

**CLONING OF HUMAN CORONAVIRUS NL-63 ORF3, M AND E GENES FOR  
ANTIBODY PRODUCTION**

**By**

**Randall Graeme Fisher**

A thesis submitted in fulfilment of the requirements for the degree of  
Magister Scientiae in the Department of Medical Biosciences, University of the  
UNIVERSITY *of the*  
Western Cape. PE

**SEPTEMBER 2010**

## DECLARATION

I declare that: '**Cloning of Human Coronavirus NL-63 ORF3, M and E genes for Antibody Production**' is my own work, that it has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Randall Graeme Fisher

September 2010



Signed.....UNIVERSITY of the  
WESTERN CAPE

# TABLE OF CONTENTS

DECLARATION.....	ii
TABLE OF CONTENTS .....	iii
LIST OF FIGURES.....	vi
LIST OF ABBREVIATIONS .....	viii
LIST OF ABBREVIATIONS .....	viii
ABSTRACT .....	xi
CHAPTER 1 .....	1
Introduction .....	1
Epidemiology of Human Coronaviruses .....	2
Identification and Isolation of HCoV-NL-63 .....	3
Phylogeny and Toxicology of HCoV-NL-63.....	4
Genomic Structure and Transcription Regulation of HCoV-NL-63 .....	5
Replicase Enzymes 1a and 1b.....	8
Spike Protein .....	9
Open Reading Frame 3.....	10
Envelope Protein .....	11
Membrane Protein.....	12
Nucleocapsid Protein.....	13
Viral Entry and Replication.....	13
Seasonal Incidence and Prevalence .....	16
Co-infections with NL-63 .....	18
Clinical Presentation.....	19
Laboratory Diagnosis and Detection .....	20
Detection of viral RNA .....	20
Scope of this thesis.....	22
CHAPTER 2 .....	25
Methods and Materials.....	25
Bioinformatic Analysis of M, E and ORF3 .....	25
cDNA Synthesis.....	25
Primer Design and Gene Amplification.....	26
Gel Purification .....	28
Ligation into Sequencing Vector and Transformation into JM109 cells .....	28
Colony Picking and Plasmid Extraction.....	30
Sgf1 and Pme1 Restriction Digest of Constructs .....	31

Ligation into Bacterial Expression Vector.....	31
Glycerol Stocks and Gene Insert Validation.....	32
Pilot Study: Expression of MΔN-GST Fusion Protein.....	33
Expression of Viral GST Fusion Proteins by Autoinduction .....	34
Expressed Protein Quantification .....	35
Validation of GST Fusion Protein Expression .....	35
Western Blot analysis .....	35
Antigen Generation.....	36
Balb/c Mice .....	37
Antigen Preparation.....	37
Concentration of Antigens .....	38
Inoculation of Balb/C Mice.....	38
Validation of Antibody Production.....	39
Indirect Viral Protein Specific ELISA .....	40
Optimization of ELISA's Protocol.....	41
<b>CHAPTER 3 .....</b>	<b>42</b>
Results and Discussion.....	42
In silico Analysis .....	42
Reverse Transcription and Gene Specific PCR .....	45
Ligation into the pGEM Vector for Sequencing .....	47
Transformation into the JM109 Competent E. coil .....	47
Recombinant Vector Confirmation and Gene Sequencing .....	50
Sgf I and Pme I Restriction Enzyme Digest of pGEM Constructs.....	51
pFLEXI Ligation and Transformation.....	52
pFLEXI Recombinant Plasmid Verification.....	54
Pilot Study: Expression of MΔN-GST Fusion Protein.....	55
Expression of Viral GST Fusion Protein by Autoinduction .....	56
Coomassie Stain and Western Blot Analysis .....	58
Antigen Preparation.....	63
Concentration of Antigens .....	65
Validation of Antibody Production.....	66
Viral Protein specific ELISA.....	69
Optimized, Indirect, Viral Protein Specific ELISA .....	71
Optimized Indirect Viral Protein Specific Competition ELISA.....	72
<b>CHAPTER 4 .....</b>	<b>73</b>
Conclusion .....	73
<b>REFERENCES .....</b>	<b>78</b>

APPENDIX ..... CVI  
ACKNOWLEDGMENTS ..... CXI



## LIST OF FIGURES

Figure 1: Schematic overview of the VIDISCA method .....	4
Figure 2: Genomic structure of HCoV-NL-63 .....	7
Figure 3: Formation of the Pseudo Hairpin Knot.....	14
Figure 4: NL-63 ORF3 Hydrophilicity and Antigenicity plot.....	42
Figure 5: NL-63 M Hydrophilicity and Antigenicity plot .....	44
Figure 6: Gene specific PCR of ORF3 $\Delta$ N and M $\Delta$ N.....	46
Figure 7: Gene specific PCR of E .....	46
Figure 8: Transformation into JM109 – Positive control.....	48
Figure 9: No Transformation into JM109 – Negative control .....	49
Figure 10: Transformation into JM109 – Background control .....	49
Figure 11: Transformation into JM109 – ORF3 $\Delta$ N.....	50
Figure 12: Transformation into JM109 – M $\Delta$ N .....	50
Figure 13: Transformation into JM109 – E.....	50
Figure 14: <i>Sgf</i> I and <i>Pme</i> I digestion of ORF3 $\Delta$ N and M $\Delta$ N-pGEM construct.....	52
Figure 15: <i>Sgf</i> I and <i>Pme</i> I digestion of E-pGEM construct.....	52
Figure 16: Transformation into KRX – Negative control .....	53
Figure 17: Transformation into KRX- ORF3 $\Delta$ N-GST .....	53
Figure 18: Transformation into KRX- M $\Delta$ N-GST.....	54
Figure 19: Transformation into KRX- E-GST .....	54
Figure 20: <i>Sgf</i> I and <i>Pme</i> I digest of ORF3 $\Delta$ N and M $\Delta$ N from pFLEXI constructs .....	55
Figure 21: <i>Sgf</i> I and <i>Pme</i> I digest of E pFLEXI constructs.....	55
Figure 22: 12 hour Pilot study - Coomassie stained SDS PAGE of expressed M $\Delta$ N-GST fusion protein.....	56
Figure 23: Coomassie stain displaying optimal ORF3 $\Delta$ N-GST fusion protein expression at 12 hours.....	60
Figure 24: Coomassie stain showing an optimal concentration of expressed M $\Delta$ N-GST fusion protein at 10 hours post induction. ....	61

Figure 25: Coomassie stain depicting optimal expression of E-GST fusion protein 16 hours post induction. ....	61
Figure 26: Western Blot showing an optimal concentration of ORF3ΔN-GST expressed fusion protein, 12 hours post induction .....	62
Figure 27: Western Blot displaying an optimal concentration of MΔN-GST expressed fusion protein, 10 hours post induction .....	62
Figure 28: Western Blot of E-GST expressed fusion proteins displaying an optimal protein concentration at 16 hours post induction.....	63
Figure 29: Bradford Assay Standard Curve (620nm).....	64
Figure 30: Concentration of GST-Fusion viral protein post purification .....	64
Figure 31: Bradford Assay Standard Curve (620nm).....	65
Figure 32: Concentration of GST-Fusion Viral Protein Post Purification .....	66
Figure 33: Western Blot using 1:500 Mouse anti-GST-Fusion 1° Antibodies .....	69
Figure 34: Absorbance observed for the indirect ELISA.....	70
Figure 35: Absorbance observed for the optimized, indirect ELISA .....	71
Figure 36: Absorbance observed for the optimized, indirect, competition ELISA.....	72
Figure 37: Sequencing vector pGEM T-easy .....	CVI
Figure 38: Positive control vector pGEM-3Z .....	CVI
Figure 39: Bacterial expression vector pFLEXI-pFN2A (GST) .....	CVII
Figure 40: Sequence verification of ORF3ΔN gene in pGEM.....	CVIII
Figure 41: Sequence verification of MΔN gene in pGEM .....	CIX
Figure 42: Sequence verification of E gene in pGEM .....	CX

## LIST OF ABBREVIATIONS

5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside	: X-gal
Adenosine Triphosphate	: ATP
Angiotensin-Converting Enzyme	: ACE
Amino Acid	: aa
Amino Terminus	: N-terminus
Amplified Fragment Length Polymorphism	: AFLP
Approximately	: ~
Base Pairs	: bp
Bovine Serum Albumin	: BSA
Carboxyl Terminus	: C-terminus
Clonal Deoxyribonucleic Acid	: cDNA
Coronavirus	: CoV
Cytopathic Effect	: CPE
Degrees Celsius	: °C
Deoxyribonucleic Acid	: DNA
Deoxyribonucleotide Triphosphate	: dNTP
Dithiothreitol	: DTT
Dulbeccos' Modified Phosphate Buffered Saline	: DPBS
Enzyme Linked Immunosorbent Assay	: ELISA
Escherichia Coli	: <i>E.coli</i>
Ethylenediaminetetraacetic acid	: EDTA
<i>Et alii</i> (And others)	: <i>et al.</i>
Envelope	: E
Gravity	: g
Glutathione-S-Transferase	: GST
Heptad Repeat	: HR
Horseradish Peroxidase	: HRP
Hours	: Hrs





Human Coronavirus	: HCoV
Hydrochloric Acid	: HCL
<i>Id est</i> (That is)	: i.e.
In an artificial environment	: <i>in vitro</i>
Isopropyl-beta-D-thiogalactopyranoside	: IPTG
Kilo Base Pairs	: kbp
Kilo Daltons	: kDa
Liter	: L
Luria Bertani	: LB
Magnesium Chloride	: MgCl <sub>2</sub>
Magnesium Hydroxide	: MgOH
Membrane	: M
Messenger Ribonucleic Acid	: mRNA
Micro Grams	: μg
Micro Liters	: μl
Micro Molar	: μM
Milliamps	: mA
Milliliter	: ml
Millimolar	: mM
Minute	: min
Nano Grams	: ng
Nanometer	: nm
Nano Molar	: nM
National Center for Biotechnology Information	: NCBI
Nonstructural Protein	: NSP
Nucleocapsid	: N
Nucleotides	: nt
Outside the living organism	: <i>ex vivo</i>
Papain-like Protease	: PLP



Percent	: %
Percentage Hydrogen	: pH
Performed on computer	: <i>in silico</i>
Phosphate Buffered Saline	: PBS
Pico Molar	: p.mol
Polymerase Chain Reaction	: PCR
Polyvinylidene Fluoride	: PVDF
Real Time Polymerase Chain Reaction	: RT-PCR
Revolutions per minute	: rpm
Ribonuclease inhibitor	: RNAsin
Ribonucleic Acid	: RNA
Room Temperature	: R/T
See below	: <i>vide infra</i>
Second	: sec
Sever Acute Respiratory Syndrome	: SARS
Sodium Dodecyl Sulfate	: SDS
Spike	: S
Sulfuric Acid	: H <sub>2</sub> SO <sub>4</sub>
Transcription Regulatory Sequence	: TRS
Ultraviolet	: UV
Units	: U
Untranscribed Region	: UTR
Open Reading Frame 3	: ORF3
Virus-like Particle	: VLP
Volts	: V
Volume per Volume	: v/v
Weight per Volume	: w/v
Within a living organism	: <i>in vivo</i>



## **ABSTRACT**

Human Coronavirus NL-63 is a respiratory virus with a high incidence rate, causing mild respiratory infections in children under the age of 18. The outbreak of Severe Acute Respiratory Syndrome (SARS) in 2003 sparked increased interest into the field of coronavirology and respiratory diseases subsequently led to the discovery of this novel Human Coronavirus (HCoV) by a group of scientists in Holland.

The membrane protein (M) of NL-63 has been shown to interact with the nucleocapsid, spike and envelope proteins of the virus when expressed *ex vivo*. In contrast, the envelope protein (E) is shown to exhibit ion channel activity, interacts with the membrane protein during the formation of viral-like particles.. The functions of the open reading frame 3 (ORF3) proteins remains a mystery. Research does, however, indicate that this protein is needed for *in vivo* infectivity and pathogenesis.

Bioinformatic analysis indicates that both the ORF3 and M proteins possess at least 3' C-terminal transmembrane regions. To further characterize the biological activity of these three proteins in clinical and laboratory samples, sensitive and specific antibodies are required. Thus, the antigenic regions of ORF3, M and the entire E gene were amplified by PCR and ligated into a bacterial expression vector for expression and subsequent generation of antibodies in a mouse system. The identities of the cloned genes were confirmed by sequencing before

being expressed in an *in vitro* bacterial system. Western Blots were used to identify the expression of the 41kDa, 42kDa and 34kDa GST-tagged viral proteins which were consistent with the bioinformatically predicted protein species.

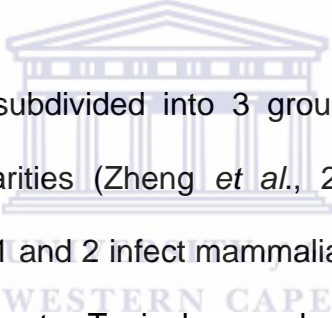
Verified fusion proteins were expressed in large quantities, quantified and concentrated for *in vivo* antibody production. Inoculation of 9 healthy, female Balb/C mice with the purified fusion proteins yielded high titers of polyclonal antibodies. Western Blotting was once again used to validate the production of the antibodies before their specificity was quantitatively measured using a modified competition ELISA.



## CHAPTER 1

### **Introduction**

The name "coronavirus" is derived from the Latin word "corona" meaning crown, as the viral morphology, under electron microscopy, appears to be crowned by a characteristic ring of small bulbous structures protruding (Spaan *et al.*, 1988). The family *Coronaviridae* forms part of the order *Nidovirales* and consists of the genera *Coronavirus* (CoV) and *Torovirus* (Plant and Dinman, 2008). Both Coronaviruses and Toroviruses infect humans and animals and are mainly associated with enteric disease (Woode, 1994).



The Coronavirus genus is subdivided into 3 groups based on their antigenic reactivity and genetic similarities (Zheng *et al.*, 2006; Fielding *et al.*, 2009). Whereas CoVs from groups 1 and 2 infect mammalian hosts, Group 3 CoVs have been found to infect avian hosts. Typical examples of Group 1 viruses include Canine coronavirus (CCV), Feline coronavirus (FeCoV), Porcine endemic diarrhea virus (PEDV) and Transmissible gastroenteritis Virus (TGEV). Bovine coronavirus (BCoV), Human coronavirus OC43 (HCoV-OC43), Mouse hepatitis virus (MHV), Porcine hemagglutinating encephalomyelitis virus (HEV), Rat coronavirus (RCV) and Human coronavirus HKU1 (Woo *et al.*, 2005a) are examples of Group 2 coronaviruses. Group 3 coronaviruses include Infectious bronchitis virus (IBV) and Turkey coronavirus (Bluecombs disease virus) (Abdul-Rasool and Fielding, 2010; Weiss and Martin, 2005).

Coronaviruses are capable of enormously varied pathogenicity and causing diseases that affect multiple organs through a variety of pathogenic mechanisms (McIntosh, 2005). Symptoms caused by coronaviruses are varied depending on the specific host species infected. In humans, these symptoms may include cough, rhinorrhoea, tachypnea, fever, abnormal breath sounds, hypoxia, severe atypical pneumonia, bronchiolitis, pharyngitis, dyspnea, desaturations, bronchospasm, vomiting, body rashes and wheezing chest (Weiss and Martin, 2005). Animal coronaviruses can cause disease in organ systems other than the respiratory tract, such as the gastrointestinal system and the central nervous system (Pyrce *et al.*, 2004).



### ***Epidemiology of Human Coronaviruses***

Human Coronaviruses (HCoVs) were first identified from clinical samples in 1965 by Tyrrell and Bynoe and was shown to cause rhinitis or the common cold (Tyrrell and Bynoe, 1965). The isolated viruses were named HCoV-229E and HCoV-OC43, were subsequently categorized as Group 1 and 2 coronaviruses, respectively (Pyrce *et al.*, 2007a; Kahn, 2006; Woo *et al.*, 2005b).

Then, in early 2003 with the outbreak of severe acute respiratory syndrome (SARS), the causative agent was identified as a third HCoV (Drosten *et al.*, 2003; Kuiken *et al.*, 2003), which was classified a distant member of the Group 2 coronaviruses (Gibbs *et al.*, 2004; Snijder *et al.*, 2003). SARS CoV, which causes acute respiratory failure, infected about 8000 people worldwide and had a mortality rate of about 10%. Subsequent to this outbreak, two additional Human

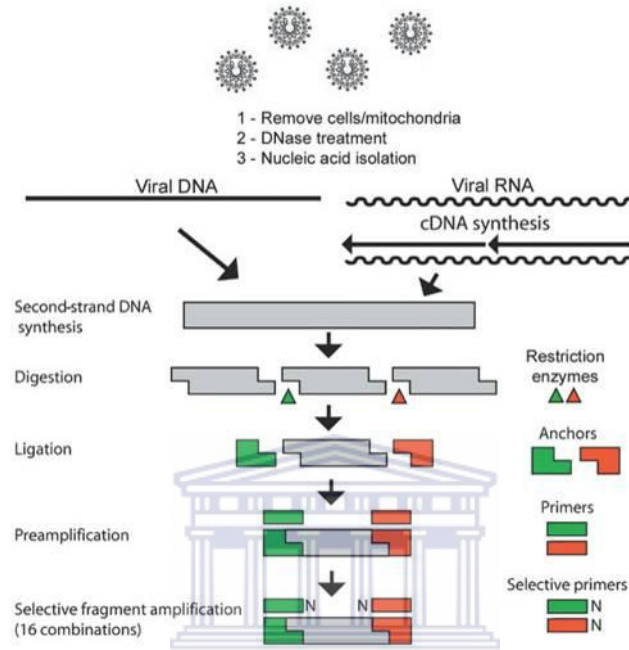
Coronaviruses were identified; NL-63 from a 7-month-old child with an upper respiratory tract infection (van der Hoek *et al.*, 2004) and HKU1 that was isolated from a 71-year-old man who presented with fever and cough (Woo *et al.*, 2005a). With the exception of SARS CoV, the human coronaviruses continuously circulate in the population especially in children.

### ***Identification and Isolation of HCoV-NL-63***

A modified cDNA-AFLP method, VIDISCA, was used to identify HCoV-NL-63 from a nasopharyngeal aspirate taken from a 7-month-old child hospitalized with bronchiolitis, conjunctivitis, and fever (van der Hoek *et al.*, 2004). While his chest X-ray showed clear signs of infection, the young boys' aspirate tested negative for all known respiratory viruses. However, the aspirate still induced cytopathic effects (CPE) when inoculated onto monkey kidney cells. The previously unknown virus was later classified as a Group 1 coronavirus (van der Hoek *et al.*, 2004; Pyrc *et al.*, 2006b).

Since its discovery, HCoV-NL-63 has been identified in many upper respiratory tract infection samples collected in different countries, especially from young children (van der Hoek *et al.*, 2006; Chen *et al.*, 2007). This virus has now been shown to have a world-wide distribution and has been isolated in places such as Australia (Arden *et al.*, 2005), Canada (Bastein *et al.*, 2005a), Japan (Suzuki *et al.*, 2005), Belgium (Moës *et al.*, 2005), USA (van der Hoek *et al.*, 2005), France (Vabret *et al.*, 2005), Germany (van der Hoek *et al.*, 2005), China and Hong Kong (Chiu *et al.*, 2005), Italy (Canducci *et al.*, 2008), Korea (Choi *et al.*, 2006), Taiwan

(Wu *et al.*, 2008), Thailand (Dare *et al.*, 2007), South Africa (Smuts *et al.*, 2008), Switzerland (Kaiser *et al.*, 2005) and in Sweden (Koetz *et al.*, 2006), to name but a few.



**Figure 1: Schematic overview of the VIDISCA method**

Virus Discovery cDNA-AFLP (VIDISCA) utilizes virus isolated from cell culture supernatants of *in vitro* samples that displayed cytopathic effect (CPE) (de Souza Luna *et al.*, 2008). Supernatants were then treated with DNase to digest cDNA before being used in a modified cDNA Amplified Fragment Length Polymorphism (AFLP) analysis. The modified AFLP uses the restriction enzyme digestion sites in the unknown DNA sequence to attach oligonucleotide adaptors, which are then used as primer binding sites for PCR and genomic sequencing (van der Hoek *et al.*, 2004; Pyrc *et al.*, 2008).

### ***Phylogeny and Toxicology of HCoV-NL-63***

Sequence analyses of the complete genome of HCoV-NL-63 reveal that the virus is more closely related to HCoV-229E than to the other human coronaviruses (van der Hoek *et al.*, 2004) and that the genome has a mosaic structure. Phylogenetic analysis suggests that the current prevalent HCoV-NL-63 strain is the result of the genetic convergence of two sister coronavirus strains. This recombination event is



proposed to have resulted from a co-infection of a human host by a NL-63 strain – that diverged from a common HCoV-229E ancestor in the 11<sup>th</sup> century – and a prevalent HCoV-229E strain (Pryc *et al.*, 2006; Arden *et al.*, 2005; Chui *et al.*, 2005).

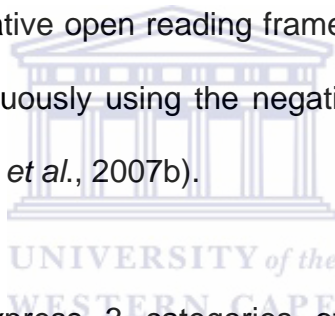
Measuring the rate at which genetic divergence occurs in a particular species requires molecular clock analysis. This type of analysis assumes a constant rate of mutation within the species and is calculated based on available field data (Bromham and Penny, 2003). Analysis conducted on coronaviruses revealed an average substitution rate of  $10^{-4}$  substitutions per year, per site (Sanchez *et al.*, 1992; Vijgen *et al.*, 2005). The substitution rate from the coronaviral molecular clock analysis was applied to available sequence data from the S gene of HCoV-229E (from known dates). Results of this analysis indicate that the most recent common ancestor between HCoV-229E and HCoV-NL-63 was indeed in the 11<sup>th</sup> century. This discovery also suggests that HCoVs have been present, and circulating, in the human race for hundreds of years (Moës *et al.*, 2005; Vabret *et al.*, 2005; Minosse *et al.*, 2008; Pryc *et al.*, 2006b).

### **Genomic Structure and Transcription Regulation of HCoV-NL-63**

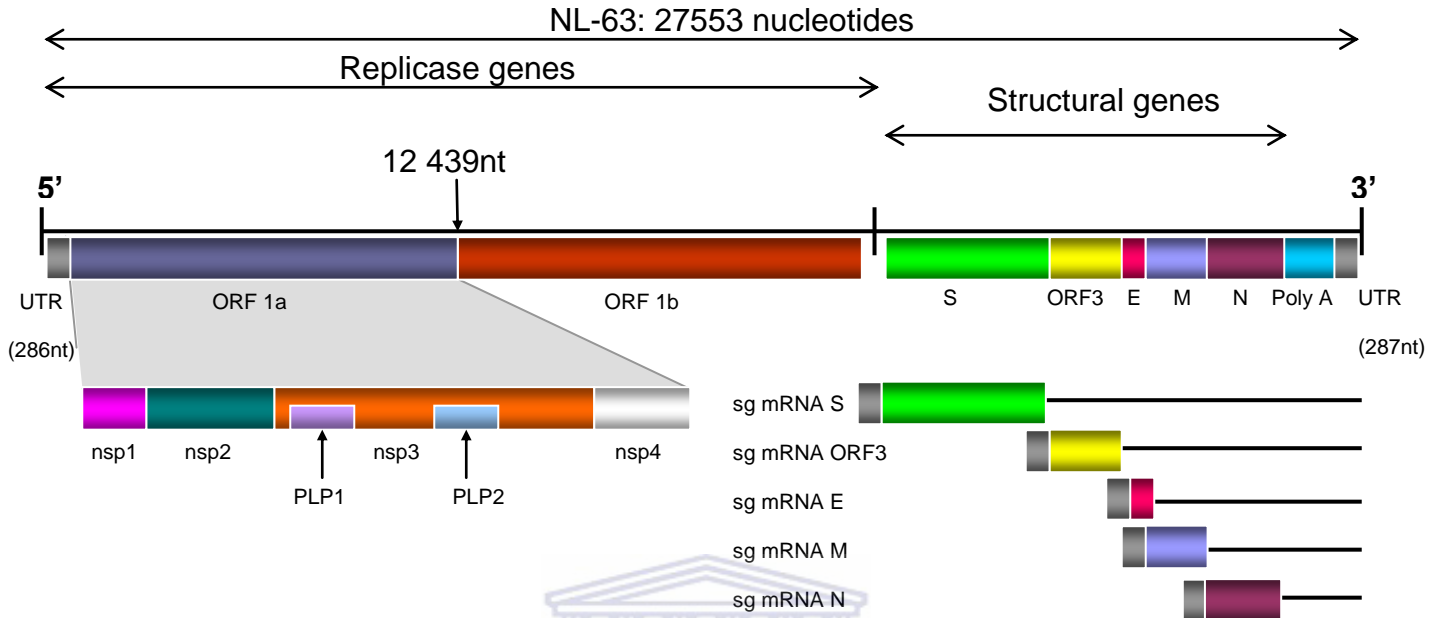
The coronavirus family consists of single stranded, plus sense, polyadenylated RNA viruses (Pryc *et al.*, 2004). The viral genome is known to have helical genetic symmetry (Siddwell *et al.*, 1983; Donaldson *et al.*, 2008; Moës *et al.*, 2005). Since coronaviruses display no DNA stage during their replication cycle

they are classified as a Group IV virus family, according to the revised Baltimore scheme of viral classification.

The coronaviruses have the largest genomes of known RNA viruses ranging in size from approximately 27 to 32kbp (Vabret *et al.*, 2005). The large, unstable RNA genome, coupled with the lack of a proofreading replicase enzyme, results in frequent transcriptional errors, point mutations (Pryc *et al.*, 2006b) and the increased possibility of genetic recombination (Jia *et al.*, 1995; Kusters *et al.*, 1990). The genome of HCoV-NL-63 consists of 27553 bases with a genomic organization 5'-ORF1a-ORF1b-S-ORF3-E-M-N-Poly-A tail-3' (Figure 2 below). The RNA encodes for 7 putative open reading frames (ORFs) (Pryc *et al.*, 2004), that are generated discontinuously using the negative sense strand to generate positive sense mRNAs (Pryc *et al.*, 2007b).



Coronaviruses generally express 3 categories of protein during new virion formation. The first category is expressed from ORFs 1a and 1b. These genes encode for the viral replicase enzymes that are required for viral assembly and the necessary activities for the transcription of negative stranded RNA, leader RNA, subgenomic mRNAs and progeny virion RNA as well as proteinases responsible for the cleavage of the polyprotein into functional products (Clementz *et al.*, 2010). The structural proteins are encoded by ORFs found in the last 1/3 of the 3'-end of the RNA genome (Pryc *et al.*, 2004). These proteins are expressed from 4 subgenomic mRNAs (Pryc *et al.*, 2004) and include the spike (S), envelope (E), membrane (M) and nucleocapsid (N) proteins. These proteins are essential for virion particle formation (Pryc *et al.*, 2004). The final category of expressed



**Figure 2: Genomic structure of HCoV-NL-63**

The genome is 27 553 nucleotides in length and is subdivided into replicase and structural genes. The arrow at 12 439 nucleotides indicates the position of the -1 ribosomal frame-shift that is required to fully express ORF's 1a and 1b. The figure also indicates the 5 subgenomic mRNA expressed by the structural and accessory genes, all with the same 5' UTR. ORF 1a is further subdivided into Non-structural Proteins nsp1, nsp2, nsp3 and nsp4 with the largest non-structural protein, nsp3, encoding two papain-like proteases PLP1 and PLP2.

proteins are the accessory proteins, whose genetic templates are dispersed among the structural genes. The functions of these proteins have not yet been clearly defined. Research, however, has indicated a possible involvement of these species-specific proteins in viral infectivity or modulation of pathogenesis in the natural host (de Haan *et al.*, 2002; Haijema *et al.*, 2004; Casais *et al.*, 2005; Fielding and Suliman, 2009).

***Replicase Enzymes 1a and 1b***

The genetic template for the viral replicase enzymes is found on the first 2/3 of the RNA genome N-terminal. ORF1ab is expressed as a single protein that is later cleaved by a protease (nsp3) encoded by the 5' end of protein 1ab (Pyrce *et al.*, 2007c). A -1 frame-shift is required at position 12439nt for the ribosome to successfully express 1b gene. This frame-shift is facilitated by the formation of a putative elaborate hairpin pseudoknot (Namy *et al.*, 2006), subsequently resulting in all expressed viral proteins having the same ~70nt transcription regulatory sequence at their N-terminus.

In an *in vitro* system, non-structural proteins, nsp3 and nsp4, are detectable in peri-nuclear cellular regions, 24 hours post infection with NL-63 (Chen *et al.*, 2007). Chen and colleagues have also characterized two viral papain-like proteases, PLP1 and PLP2, which play a role in the processing of polyprotein 1ab. PLP2 was also discovered to have deubiquitinating activity, but its function remains unclear.

Nsp-3 and -5 proteins are coexpressed with polyprotein 1ab and have been shown to have papain-like and serine-like protease activity. Nsp-5 is regarded as the 'main' protease, however, and is responsible for processing, by means of cleavage, of most of the 1ab polyprotein (Ziebuhr, 2005).

**Spike Protein**

The “crowned”-shape viral morphology observed is actually formed by the protruding viral spike (S) glycoproteins, which populate the surface of the virus and determine host-cell tropism (Li *et al.*, 2005; Nal *et al.*, 2005). The spike protein mediates the viral attachment to specific cellular receptors on the host-cell surface and facilitates viral entry into the cell (Zheng *et al.*, 2006; Cavanagh, 1995). It is the largest of the structural proteins and, in the case of HCoV-NL-63, is seen to be 2380 amino acids in length (Pyrce *et al.*, 2004).

Coronavirus S-protein has been classified as a class I fusion glycoprotein consisting of a globular S1 region that recognizes the ACE2 receptor on human cells and the rod-like S2 region that mediates membrane fusion through a C-terminal transmembrane fusion domain (Bosch *et al.*, 2003). Also found in the NL-63-S2 region is two heptad-repeat (HR) clusters which are larger than their correspondents found in other coronavirus species (Zheng *et al.*, 2006; Pyrc *et al.*, 2007c; Bosch *et al.*, 2004).

These HR repeat regions are highly conserved between Group II and III coronaviruses. Zheng *et al.* (2006) conducted proteolytic studies on the S2 fusion core and revealed a possible  $\alpha$ -helical domain made up of a trimer of the HR sections N57 and C42. The finalized crystal structure of this region revealed high-affinity conformations of interacting cross-sectional layers of six helices. Their

study concluded that the larger HR regions found in Group I coronaviruses, need to prime the S protein in order to activate the fusion conformation required during viral entry (Zheng *et al.*, 2006).

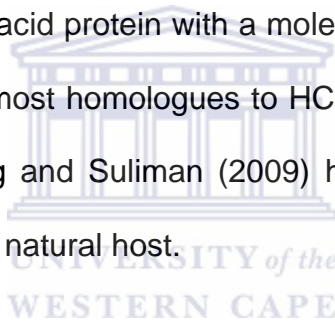
The N-terminal of the NL-63 S-protein possesses a 179 amino acid chain that has not been found in any other coronavirus. This unique section represents the most varied region of the NL-63 genome and is suspected to play a role in host immune system evasion (van der Hoek *et al.*, 2006; Pyrc *et al.*, 2007). Other research conducted on the S protein suggests that antibodies produced against the S have a neutralizing affect *in vitro* (Pyrc *et al.*, 2006). Interaction between the spike and membrane proteins has been seen during viral assembly and prevents S from being transported to the host cells membrane surface during replication (Rottier and Rose, 1987; Opstelten *et al.*, 1995). By preventing this intracellular processing and translocation of the S proteins to the infected cell surface, the membrane protein enables the virus to evade the hosts' immune system.

### ***Open Reading Frame 3***

The ORF's that encode for the accessory proteins of coronaviruses can be found distributed amongst the structural genes in the 3' 1/3 of the RNA genome (Figure 2). Group I HCoV's such as NL-63 and 229E both encode only one ORF whose genetic template can be found between the S and E genes (Franklin *et al.*, 2006). Preliminary research conducted into the functions of the accessory proteins suggested that they were non-essential for viral growth in culture (Hodgson *et al.*, 2006; Casais *et al.*, 2005; de Haan *et al.*, 2002). Later research has however

indicated that ORF's are required for *in vivo* infectivity and pathogenicity in the natural host (de Haan *et al.*, 2002; Cavanagh *et al.*, 2007; Netland *et al.*, 2007; Haijema *et al.*, 2004; Pewe *et al.*, 2006).

HCoV-NL-63 ORF3 is expressed from the fourth of the six subgenomic mRNAs produced by NL-63. This unique gene contains U-rich and A-poor pockets that suggested it to be a recently transferred gene from another genetic origin (Pyrce *et al.*, 2004). NL-63 ORF3 has 3 trans-membrane regions, is N-glycosylated and is incorporated into the completed virion (Müller *et al.*, 2010). *In silico* analysis of ORF3 predicts a 225 amino acid protein with a molecular weight of approximately 26 kilodaltons (kDa) and is most homologous to HCoV-229E ORF4 (43% identity and 62% similarity). Fielding and Suliman (2009) hypothesize that ORF3 to be linked to pathogenesis in the natural host.



### ***Envelope Protein***

E is relatively under expressed *in vitro* and accounts of approximately 2% of the progeny proteins (Pyrce *et al.*, 2004). This 78 amino acid protein has been singled out as being the only viral gene to not contain an upstream core transcription regulatory sequence (TRS). E does, however, contain its own sub-optimal variation of the TRS core sequence (AACUAUA) and a 13nt sequence identical to that of the leader sequence seen at the 5' end of the genome. Pyrc *et al.* (2007), suggests that the annealing of this secondary leader sequence to the original leader sequence may compensate for the disturbed leader-TRS/body-TRS interaction.

Coronaviral E-protein has been shown to exhibit ion channel activity but the function of these channels has yet to be determined (Wilson *et al.*, 2004). Interaction between the E and M proteins has been experimentally demonstrated (Godet *et al.*, 1992; Yu *et al.*, 1994; Lim and Liu, 2001; Maeda *et al.*, 2001) and is required for the formation of virions with normal infectivity (Kuo and Masters, 2003). Also, co-expression of coronavirus E and M proteins *in vitro*, results in the production of virus like particles (VLP's) (de Haan *et al.*, 1998; Ho *et al.*, 2004).

### ***Membrane Protein***

Interaction of the coronaviral M-protein with the spike and envelope proteins has been well characterized (Godet *et al.*, 1992; Yu *et al.*, 1994; Lim and Liu, 2001; Rottier and Rose, 1987; Opstelten *et al.*, 1995), as well as their interaction with the viral nucleocapsid proteins (Narayanan *et al.*, 2000). With its triple-membrane-spanning regions and its N-terminus endodomain, is evidently the most active viral protein expressed during virion formation (Armstrong *et al.*, 1984; Rottier *et al.*, 1986). It appears as though the progenic function of M is to incorporate the N protein into the newly formed virion during viral assembly (Narayanan *et al.*, 2003; Narayanan *et al.*, 2000; Narayanan and Makino, 2001).

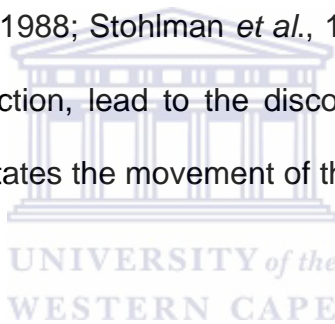
In the case of NL-63 M protein, this ~230 amino acid protein is the second most abundantly expressed viral protein, accounting for ~15% of the total expressed viral proteins during virion assembly (Pyrce *et al.*, 2004). Similarly to NL-63 ORF3,



M is also seen to be a U-rich and A-poor region, suggesting a recent gene transfer from another genomic origin (Pyrce *et al.*, 2004).

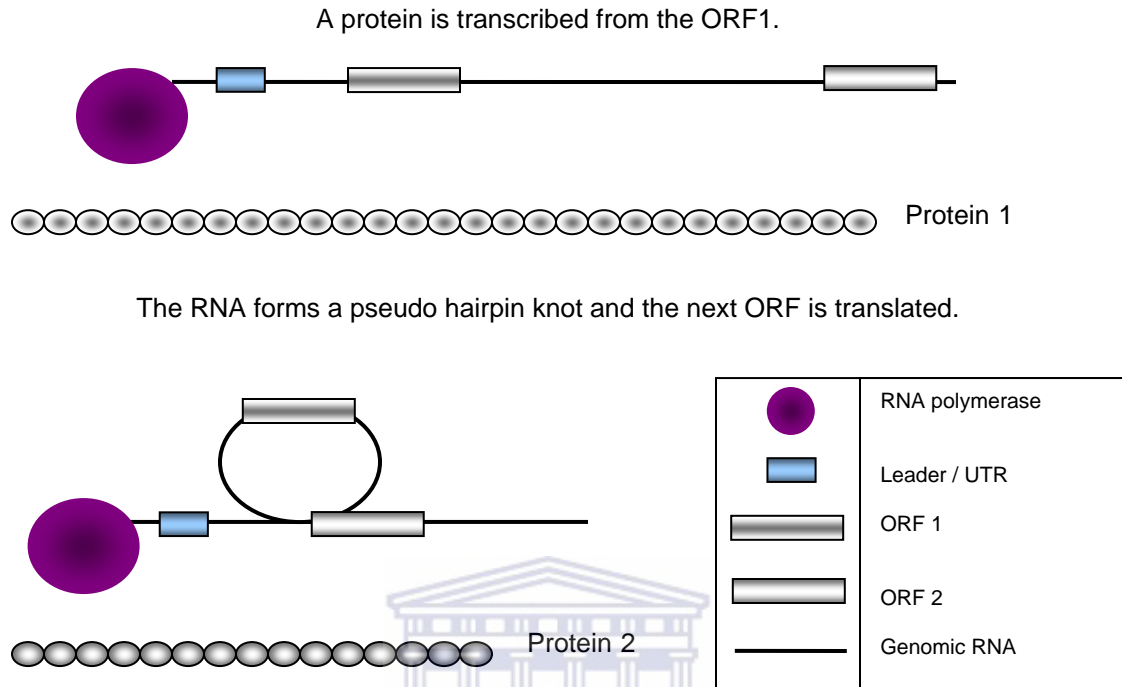
### ***Nucleocapsid Protein***

In comparison the other structural proteins, the coronavirus ~380 amino acid nucleocapsid proteins is not well characterized. It is however documented to be the most abundantly expressed viral protein, ~74% of the progeny proteins (Pyrce *et al.*, 2004). Also, coronavirus nucleocapsid proteins have been shown to interact with co-expressed N proteins, forming dimers (Yu *et al.*, 2005). Studies have shown interactions between coronaviral N-proteins and all sub-genomic and genomic RNAs (Baric *et al.*, 1988; Stohlman *et al.*, 1988). Exploration into HCoV-NL-63 N-protein-RNA interaction, lead to the discovery of the N-terminus RNA chaperone motif, which facilitates the movement of the genomic RNA into the new virion (Zúñiga *et al.*, 2006).



### **Viral Entry and Replication**

As mentioned earlier, the genomic order of the HCoV-NL-63 genome is 1ab-S-ORF3-E-M-N (Figure 2). The 5' 2/3rds (i.e.: genes 1a and 1b) of the genome encodes two polyproteins that contain all the enzymes required for RNA replication. These polyproteins are transcribed from genes 1a and 1b and require a –1 ribosomal frame shift (Figure 2) for the successful expression of gene 1b (Herold *et al.*, 1993; Giedroc *et al.*, 2000). The polyprotein that is produced by the ribosome undergoes a catalytic cleavage by the viral papain-like proteinase and a chymotrypsin-like proteinase to produce proteins 1a and 1b (Chen *et al.*, 2007).



**Figure 3: Formation of the Pseudo Hairpin Knot**

The figure above graphically depicts the formation of the pseudo hairpin knot employed by coronaviruses during viral replication. This genomic conformation results in all expressed viral proteins having the same leader sequence.

Coronaviruses replicate by using the host-cell ribosomes and metabolites to produce multiple copies of the infecting virion in the cytoplasm (Brockway *et al.*, 2003; Siddwell *et al.*, 1983; Konings *et al.*, 1988; Bond *et al.*, 1979). Genomic replication generally follows the uncoating of the viral genome inside the host cell cytoplasm (Siddwell *et al.*, 1983). The polyadenylated, positive sense RNA is read directly by the host ribosomes as mRNA and the replicase genes 1a and 1b are translated into the viral RNA dependant RNA polymerase and a viral protease (Ziebuhr *et al.*, 2000; Sawicki *et al.*, 2007). These enzymes are then employed by

the virus to produce a multitude of viral progeny in a microenvironment established in the host-cell's cytoplasm (Brockway *et al.*, 2003; Siddwell *et al.*, 1983; Konings *et al.*, 1988; Bond *et al.*, 1979).

The N protein is found to be the most abundant viral protein present during the replication cycle (Pyrce *et al.*, 2004). This is followed by M, ORF3, S and finally E. The phenomenon observed in coronavirus replication indicates that the further away from the leader sequence the gene lies, the more it is transcribed. This theory holds true for all expressed HCoV-NL-63 proteins with the exclusion of the E protein, which relatively underexpressed (Pyrce *et al.*, 2004).

The suggested explanation for this phenomenon is that the structural genes are transcribed from subgenomic mRNA during replication. Researchers suggest the formation of a pseudo hairpin knot to be responsible for each expressed protein having the same leader sequence (Herold *et al.*, 1993; Giedroc *et al.*, 2000).

Coronaviruses have a lysogenic life-cycle meaning that the virus can cause degenerative effects within a host cell without causing its death. These effects generally result in the rounding up of the infected host cell or detachment of the infected cell from the surrounding tissue. These symptoms are collectively called the cytopathic effects (Pyrce *et al.*, 2006; Herzog *et al.*, 2008). Released virions can be passed to other sites within the body by traveling through the bloodstream or even between hosts through direct contact. More commonly, transmission

occurs through contact with infected bodily fluids such as sputin and saliva (Bridges *et al.*, 2003).

The virus infects the new host via the respiratory tract and binds to the upper respiratory tract epithelium. Next, the virus gains entry into the susceptible epithelium cells through the use of the ACE 2 surface receptor. This entry into susceptible host cells is mediated by the spike protein, which is found on the virion envelope surface (Cavanagh, 1995). Once inside the cell, the virus induces the host cell to begin manufacturing the proteins necessary for virus reproduction. As well as proteins production, the virus must also direct the replication of new viral genomes. Coronaviral replication takes place with the use of nested ends, meaning that the leader sequence (found just before the start codon for of every protein) remains constant with every protein that is transcribed from genomic RNA. This is achieved by the genomic RNA bending and forming a loop, which moves the leader sequence adjacent to the start codon of the next ORF that is to be transcribed (Herold *et al.*, 1993; Giedroc *et al.*, 2000). As a result, each gene that is transcribed has the same leader sequence.

### ***Seasonal Incidence and Prevalence***

To date, HCoV-NL-63 has been shown to have a worldwide distribution. Initially, the virus was observed primarily in the winter seasons in countries with temperate climates. On the other hand, countries with extreme weather, like Canada, have also shown virus activity around January to March, although milder symptoms were reported (Bastein *et al.*, 2005a). Also, seasonal variations have been reported in China where infection with HCoV-NL-63 appeared mainly in spring

and summer (Chiu *et al.*, 2005). Additionally, Wu *et al.* (2007) reported that the virus is detectable in the autumn season in Taiwan (Wu *et al.*, 2008). A more recent study of coronaviruses in Thailand though, did not show any seasonal predilection (Dare *et al.*, 2007). Collectively, results show that the virus has no predilection to a particular season and is not affected by temperature variations as infections can occur throughout the year.

Human coronaviruses are known to be the etiology of a significant number of hospitalizations of children under the age of 18, the elderly and the immunocompromised. A one-year study conducted in Hong Kong (Chiu *et al.*, 2005) revealed the presence of human coronaviruses in 4.4% of all children hospitalized with respiratory tract infections, of which NL-63 was seen to be the most abundant (2.6%). Another respiratory study carried out in Japan revealed the presence of NL-63 in 1.2% of the 419 samples that tested negative for any other respiratory viruses (Suzuki *et al.*, 2005). Ebihara *et al.* (2005) also reported the incidence of NL-63 in 2.5% of 118 nasopharyngeal swab samples collected from kids under the age of two in Germany.

High respiratory coronavirus incidences were also noted abroad in the European regions. Documented incidences include studies conducted in France (Vabret *et al.*, 2005) and Belgium (Moës *et al.*, 2005) which positively validated the presence of HCoV-NL-63 in 9.3% of 300 samples and 2.5% of 279 samples, respectively. van der Hoek *et al.* (2005) discovered 5.2% incidence rate of NL-63 in 949 samples that they rechecked from a previous survey conducted by Forster *et al.*

(2004), to evaluate lower respiratory tract infections in children under the age of 3, in Germany.

In Australia, a respiratory survey study was conducted on 543 patients with respiratory symptoms in Melbourne (Lambert *et al.*, 2007). Eighteen of those patients tested positive for HCoV-NL-63. In the west, 3.6% of 525 patient samples tested by Bastein and colleagues in Canada (2005), were positive for NL-63 in comparison to the 8.8% of the 895 patient samples tested by Esper *et al.* in the United States (2005).

### ***Co-infections with NL-63***

Superinfections of NL-63 with other respiratory viruses appear to be quite common, often resulting in patient hospitalization due to the severity of these infections (Chui *et al.*, 2005; Minosse *et al.*, 2008; Dare *et al.*, 2007; Wu *et al.*, 2008; Canducci *et al.*, 2008; van der Hoek *et al.*, 2005; Lambert *et al.*, 2007; Kaiser *et al.*, 2005; Yoo *et al.*, 2007). The most common respiratory co-infection observed in children under 3 years of age, is a combination of Respiratory Syncytial Virus-A (RSV-A) and NL-63, whose seasonality patterns seem to overlap (van der Hoek *et al.*, 2005). Other co-infecting agents include other human coronaviruses, influenza A virus, parainfluenza virus (PIV) and human metapneumovirus (hMPV). Smuts and Hardie (2006) reported one incident of a bocavirus/NL-63 co-infection from a study screening 341 nasopharyngeal and bronchoalveolar specimens of hospitalized children in South Africa.

The HCoV-NL-63 viral load of patients infected with HCoV-NL-63 only is slightly higher than patients co-infected with NL-63 and other respiratory viruses (van der Hoek *et al.*, 2005). Professor Lia van der Hoek and colleagues proposed four theories to explain this phenomenon (2005):

1. NL-63 might be the primary infecting agent, which lowers the host's immunity, allowing for opportunistic respiratory infections. By the time diagnostic tests are done on the patient, the NL-63 infection has already been subdued.
2. The two infecting agents may be competing for receptor binding or viable cells in the host organs.
3. The secondary infection hyper activates the host's innate immune system causing it to inhibit NL-63 replication.
4. NL-63 exists as a prolonged or persistent infection resulting in low-level of virus detection.

### ***Clinical Presentation***

NL-63 has been identified in both upper and lower respiratory tract infections and the clinical presentation of these infections mimics those of seen in HCoV 229E and OC43 infections (Vabret *et al.*, 2008). Commonly, mild symptoms are observed in most patients who contract upper respiratory infections of NL-63. These symptoms generally include fever, cough and rhinorrhoea.

Lower respiratory tract infections with NL-63 display more severe symptoms such as bronchiolitis and inflammation of the bronchiole basal membranes. Although

research conducted in China did not associate HCoV-NL-63 with bronchiolitis (Chiu *et al.*, 2005), other research groups have observed these severe respiratory symptoms (Arden *et al.*, 2005; Ebihara *et al.*, 2005; Bastien *et al.*, 2005). Croup, a disease seen mostly in children, which presents with pharyngitis, sore throat and hoarseness of voice, has also been associated with NL-63 by at least two research groups (Chiu *et al.*, 2005; van der Hoek *et al.*, 2005).

The majority of NL-63 patients who present at hospitals or clinics are children and their recorded symptoms are based mainly on parental observation, which may prove to be somewhat uneducated. The lack of informed observation results in not all possible symptoms of NL-63 infections being identified. Patients examined for NL-63 infections on arrival at a hospital are generally admitted for respiratory symptoms. Hence, possible NL-63 infections of other organ systems might never be identified.

### ***Laboratory Diagnosis and Detection***

Diagnoses of HCoV-NL-63 infections are difficult to distinguish from the many other respiratory viruses that are known to co-infect with NL-63. Incorrect treatment of these misdiagnosed patients often results in increased health care expenses and increased chance of disease spread (Fox, 2007). Two methods exist for NL-63 detection from clinical samples: Detection of viral RNA and serological detection of NL-63 specific proteins.

### ***Detection of viral RNA***

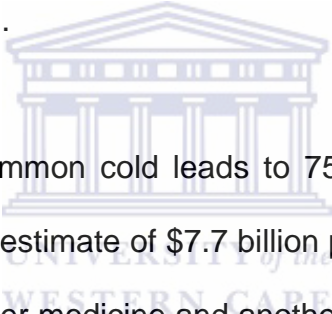


Due to its high specificity and sensitivity, the gold standard diagnostic test for the detection of coronavirus RNA is the reverse transcriptase polymerase chain reaction (Fielding and Tan, 2007). RT-PCR and other nucleic-acid-based tests are also seen as the preferred methods for detecting NL-63 from clinical samples. ORF1ab, N and S genes are generally the targets for these nuclear diagnostic tests (Arden *et al.*, 2005; Chiu *et al.*, 2005; Moës *et al.* 2005; Minosse *et al.*, 2008; Bastien *et al.*, 2005; Dare *et al.*, 2007).

*In vitro* coronavirus studies indicate that N is the most abundantly present viral protein in infected cells, which theoretically makes it a perfect target for PCR assays. This theory has however not been clearly proven in clinical studies (Cheng *et al.*, 2007). One of the main advantages of RT-PCR is that it can be performed on both fresh and frozen nasopharyngeal aspirate samples and viral culture can be used as confirmation of NL-63 infection (Kistler *et al.*, 2007; Shirmizu *et al.*, 2005).

***Scope of this thesis***

The future of drug development against RNA viruses is expected to have a significant indirect and direct impact on human health-related quality of life. RNA viruses include in excess of 350 different major human pathogens and are responsible for many emerging and reemerging diseases. These pathogens are responsible for not only very frequent benign diseases (which nonetheless have a huge economic impact, *vide infra*) but also for millions of deaths each year in both industrialized and developing countries. Even though many treatments exist for infections caused by DNA-based viruses, few efficient antiviral drugs against RNA viruses have been developed.



In the United States, the common cold leads to 75 to 100 million clinical visits yearly at a conservative cost estimate of \$7.7 billion per annum. Americans spend \$2.9 billion on over-the-counter medicine and another \$400 million on prescription medicines for relief of symptoms (Garibaldi, 1985; Fendrick *et al.*, 2003). More than 30% of patients who consulted a physician received an antibiotic prescription, which has implications for antibiotic resistance from misuse of such drugs (Fendrick *et al.*, 2003). Also, an estimated 22 to 189 million school days are missed annually due to the common cold which resulted in parents missing about 126 million workdays to stay home to care for their children. When added to the 150 million workdays missed by employees suffering from the common cold, the total economic impact of cold-related work loss exceeds \$20 billion per year (Garibaldi, 1985; Fendrick *et al.*, 2003). Clearly, the development of drugs against

RNA virus infections is expected to have a significant impact on human health-related quality of life.

Among the most significant viral infectious agents are coronaviruses (10- 15% of all common colds in the world). Coronaviruses are not well characterized. Until the discovery of the severe acute respiratory syndrome coronavirus (SARS-CoV) in 2003 (Drosten *et al*, 2003), coronaviruses were known to cause mainly mild respiratory infections in the general populace and various complications in immunocompromised patients (Hsueh and Yang, 2003). This changed after the outbreak of SARS-CoV. Since the identification of SARS-HCoV two additional human coronaviruses have been identified from clinical samples. Although these newly identified coronaviruses are usually not lethal, both of these causes respiratory and other infections that results in economic loss.

For studying pathogenicity and replication of viruses in the host or animal models, it is essential to generate viral protein-specific antibodies. Also, the genome of HCoV-NL-63 is not well characterized and the functions of the proteins transcribed by its genome, with a few exceptions, are not understood. Therefore the generation of specific and sensitive antibodies would allow the study of the different viral proteins *in vitro* as well as *in vivo*. Therefore, the **main objectives of this thesis** are:

1. RT-PCR amplification of the HCoV-NL-63 genes encoding for the structural proteins M and E and the accessory (group-specific) protein ORF3 from viral RNA.

2. Cloning of the amplification products into vectors for expression in a bacterial system.
3. Expression and purification of the viral-fusion proteins.
4. Generation of polyclonal antibodies specific to the viral-fusion proteins.
5. Validation of the specificity and sensitivity of the antibodies.



## **CHAPTER 2**

### **Methods and Materials**

#### ***Bioinformatic Analysis of M, E and ORF3***

*In silico* analysis of HCoV-NL-63 proteins M, E and ORF3 was done with Invitrogen's peptide selector application, which mathematically predicts and graphically displays the most antigenic region of the queried amino acid sequence (<http://peptideselect.invitrogen.com/peptide/>). This application calculates the position of the most probable antigenic regions by taking into account the specific amino acid charge, possible post-translational modification and theoretical protein tertiary and quaternary structure and represents it in a graphical presentation. In this study, HCoV-NL-63 ORF3 and M proteins were calculated to contain three theoretical membrane-spanning regions in their 5' halves. Thus, the antigenic 3' halves were selected to be analyzed experimentally. These 3' regions will be referred to as  $\Delta N$  regions from here on. The E protein however, contained no putative membrane spanning regions.

#### ***cDNA Synthesis***

HCoV-NL-63 RNA (a kind gift from Prof. L van der Hoek, Holland) was used as template to synthesize 1<sup>st</sup> strand cDNA. RNA was extracted from a fifth-passage virus (strain Amsterdam 1) obtained from a clinical sample. Reverse transcription was carried out using a mixture of 2 $\mu$ l RNA template ( $10^8$ - $10^9$  copies/ml), 1 $\mu$ l of Oligo (dT)<sub>15</sub> primer (100 $\mu$ M stock), 1 $\times$  incubation buffer, 2 $\mu$ l of dNTP mix (10mM

stock), 20U RNasin<sup>®</sup> Ribonuclease inhibitor, 15U of AMV Reverse Transcriptase and 4 $\mu$ l MgCl<sub>2</sub> (25 mM) and was made up to a total volume of 20  $\mu$ l in nuclease-free water, according to the manufacturer's specifications (Promega). The reaction was heated at 42°C for 60 min, 95°C for 5 min and then cooled to 0°C for 5 min to deactivate the enzyme.

### ***Primer Design and Gene Amplification***

PCR primers were designed to amplify the putative antigenic regions of ORF3 and M and the entire E gene and have an appropriate melting temperature and restriction sites for ligation into the expression vector. Briefly, all forward primers designed included a 5' restriction enzyme recognition site for *Sgf* I (for the in-frame, unidirectional ligation of the amplified gene into the bacterial expression vector) and 20 to 25 nucleotides identical to the 5' end of the target gene. All reverse primers were created to include a *Pme* I restriction enzyme recognition site and a 20 to 25-nucleotide sequence, which was the reverse complement of the 3' end of the target gene (Table 1). Forward and reverse primers were engineered to have a melting temperature between 55°C and 65°C.

100 $\mu$ M of the designed primers were used to amplify the hydrophilic portions of both the M (nt 300-681) and ORF3 (nt 330-678) genes and the entire E gene from the generated 1<sup>st</sup> cDNA. PCR reactions were carried out in total volume of 25  $\mu$ l using 2.5U of GoTaq DNA Polymerase, 0.5 $\mu$ l dNTP mix (10mM stock), 2 $\mu$ l MgCl<sub>2</sub> (25mM stock), 2.5 $\mu$ l 10X MgCl<sub>2</sub>-free GoTaq Flexi buffer and 2 $\mu$ l cDNA according to the manufacturer's specifications (Promega).

Amplification was carried out in a Multi Gene™ II Thermocycler (LabNet) using the following optimized conditions: Initial denaturation at 95°C for 3 min, followed by 30 cycles of: denaturation at 95°C (45 sec), annealing at 50°C (60 sec) and elongation at 72°C (60 sec). A final extension step at 72°C for 15min was included, before a final hold at 4°C.

**Table 1: Primers used for amplification of HCoV NL-63 ORF3ΔN, MΔN and E genes**

<b>Gene</b>	<b>Primer Sequence</b>
ORF3ΔN-Forward	5'-TGCGGCGATCGCCAGACTTTGGCGCCGTGTTAAAAC-3'
<b>ORF3ΔN-Reverse</b>	5'-TGTGG <b>TTTAAACT</b> TTAGATTAAATGAAGCAACTTCTC-3'
MΔN-Forward	5'-GGCCGCGATCGCCGGCTATTATGCCTATCTCTATAAAAA-3'
MΔN-Reverse	5'-GCC <b>TTTAAACT</b> CAATTAATCGAAGGAACATCTTCGTA-3'
E-Forward	5'-GCGCGCGATCGCCATGTTCCCTTCGATTAATTG-3'
E-Reverse	5'-GCGC <b>TTTAAACT</b> TTAGACATTTAGTACTTCAGC-3'

N.B.: *Sgf* I cut sites are underlined and *Pme* I cut sites are **bolded**

On completion of the PCR, 5µl of a 6X blue/orange DNA loading-dye (0.03% (v/v) bromophenol blue, 0.03% (v/v) xylene cyanol FF, 0.4% (v/v) orange G, 15% (v/v) Ficoll 400, 10mM Tris-HCl (pH 7.5) and 50mM EDTA (pH 8.0)) was added to each sample before they were loaded into a 1% (w/v) agarose gel (0.1% w/v Ethidium Bromide). 5µl of a 100bp DNA marker (Promega) was also loaded onto the gel before the amplified gene products were separated by electrophoresis at 90V.

### ***Gel Purification***

Amplicons were gel-purified with the Wizard SV Gel and PCR Clean-Up system (Promega). Amplification products were excised from the agarose gel under short-wave length UV illumination and transferred to a pre-weighed, sterile 1ml Eppendorf tube to determine the weight of the gel piece. The gel was then solubilised by adding an equal volume of Membrane Binding solution (4.5M guanidine isothiocyanate and 0.5M potassium acetate (pH 5.0)), followed by incubation at 60°C with constant vortexing to ensure complete dissolution. Next, the dissolved DNA was gravity filtered through a positively charged, nucleopore membrane by centrifugation. This membrane bound the DNA on the basis of its negative charge. The membrane was then washed with 1200µl of Membrane Wash solution (10mM potassium acetate (pH 5.0), 80% (v/v) ethanol and 16.7µM EDTA (pH 8.0)) to dissolve all unwanted salts and purify the bound DNA. Bound DNA was eluted in 30µl of nuclease-free water.

### ***Ligation into Sequencing Vector and Transformation into JM109 cells***

pGEM T-Easy was the selected sequencing vector to verify the inserted viral genes (Figure 37 in Appendix). The vector contains various characteristics that were deemed desirable for this stage of the study, including:

- A Lac-Z operon that encodes the enzyme  $\beta$ -galactosidase, which is able to digest the substrate X-gal in the presence of IPTG (Isopropyl-beta-D-thiogalactopyranoside), producing a blue dye which in turn produces a blue colony when plated. During a successful ligation, this gene is disrupted, which results in the growth of a white colony.



- M13 sequencing primer-binding site for cost effective sequencing of vector-insert constructs.
- Ampicillin resistance gene to which allows for the growth of successfully transformed bacterium on ampicillin containing selective, semisolid media.
- And 5' and 3' T-overhangs that allow for the easy ligation of adenosine appended PCR products into the vector at the multiple cloning site.

Ligation into the pGEM T-Easy vector was achieved by incubating 6 units of T<sub>4</sub> DNA Ligase, 50ng of cut pGEM T-Easy vector and 7.5µl of the purified PCR products at 4°C overnight in reaction buffer (300mM Tris-HCl (pH 7.8 at 25°C), 100mM MgCl<sub>2</sub>, 100mM Dithiothreitol and 10mM ATP). The resultant constructs were transformed into JM109 chemically competent *E. coli* cells by incubating 5µl of the constructs with 50µl of the recently thawed competent cells at 42°C for 45sec. The transformation reaction was then placed on ice for 2min before the cells were allowed to recover for 1 hour at 37°C in 1ml of antibiotic-free LB liquid media.

Each liter of LB liquid media (broth) used in this experiment consisted of 10g Bacto-tryptone, 5g yeast extract and 10g Sodium Chloride. 800ml of distilled water was then added to the dry ingredients before being thoroughly mixed and pH adjusted to 7.5 using Sodium Hydroxide. Finally, the media volume was adjusted to 1L before being autoclave sterilized.

Control measures taken during the transformation procedure included a positive transformation control (pGEM-3Z vector (Figure 38 in Appendix) to measure transformation efficiency), a background control (Ligated vector without viral insert) and 50µl of JM109 competent cells to verify the efficacy of the Ampicillin (100µg/µl) in the plating-media. Also included in the semisolid media was 1mM of the Lac-Z operon inducer (IPTG) and 50µg/ml of the β-galactosidase substrate (X-gal) to facilitate blue/white colony selection.

### ***Colony Picking and Plasmid Extraction***

After a 14-hour incubation at 37°C, 4 white colonies were picked from each of the experimental plates using a sterile pipette tip, and inoculated into 10ml of LB media containing 100µg/ml Ampicillin. These inoculated broths were then incubated at 37°C overnight with continuous shaking (150rpm). The now slightly turbid cultures displayed atypical optical densities of ~0.3 (600nm) in comparison to the expected ~2.0 absorbance (600nm). Culture volumes were thus adjusted by 4 fold to comply with the optimal biomass input requirements, as suggested by the supplier (Promega).

In compensation for the low culture concentrations observed, 20ml of the inoculated cultures were centrifuged at 2800rpm to pellet the suspended cells before being resuspended in 250µl of cell resuspension solution (50mM Tris-HCl (pH 7.5), 10mM EDTA and 100µg/ml RNase A). The cells were then lysed in 350µl cell lysis solution (0.2M NaOH and 1% (w/v) SDS) and 10µl alkaline protease before 450µl of Neutralizing reagent (1.32M potassium acetate (pH 4.8))

was added. The cell debris was pelleted at 16000xg and the cell lysates were filtered through a nucleopore membrane at 16000xg to bind the negatively charged DNA and elute unwanted compounds. 1ml of membrane wash solution (80mM potassium acetate, 8.3mM Tris-HCl (pH 7.5) and 40 $\mu$ M EDTA) purified the DNA by filtering it through the membrane at 16000xg and removing all unwanted salts. The plasmid construct was then eluted in 30 $\mu$ l of nuclease free water and placed on ice until further use. 10 $\mu$ l of the extracted plasmid DNA products were sent to Inqaba Biotech (Pretoria, SA) for the purpose of gene sequencing.

#### ***Sgf1 and Pme1 Restriction Digest of Constructs***

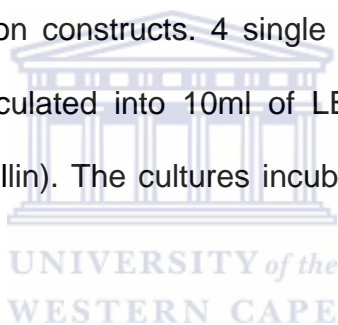
2 $\mu$ l of the 10X concentrated *Sgf* I and *Pme* I enzymes blend (Promega) was used to digest 10 $\mu$ l of the various isolated pGEM-viral-insert constructs. The reaction was then incubated at 37°C for 2Hrs in a 1X enzyme digest buffer (50mM Tris-HCl (pH 7.9 at 37°C), 250mM NaCl, 50mM MgCl<sub>2</sub>, 5mM Dithiothreitol and 0.5mg/ml acetylated BSA) before it was heated to 65°C to deactivate the restriction enzymes. The samples were separated on a 1% (w/v) agarose gel at 90 V before the genes of interest were gel purified as previously described. The products were then eluted in 30 $\mu$ l of nuclease free water.

#### ***Ligation into Bacterial Expression Vector***

Subsequently, 5 $\mu$ l of the eluted 'sticky-ended' viral genes were ligated with 50ng of the bacterial expression vector, pFLEXI-GST (pFN2A Glutathione-S-transferase tagged) (Figure 39 in Appendix). The overnight ligation took place at 4°C with the use of 12 units of T4 DNA ligase enzyme in 1X reaction buffer

(300mM Tris-HCl (pH 7.8 at 25°C), 100mM MgCl<sub>2</sub>, 100mM Dithiothreitol and 10mM ATP). The new expression constructs were transformed into 50µl of KRX chemically competent *E. coli* cells (Promega) at 42°C for 45 sec before being incubated on ice for 2 min. KRX was selected as the bacterial expression cell line due to the absence of genomic proteases expressed from the bacterial genome.

GST-tagged construct-containing KRX cells were spread-plated on semisolid LB-agar containing 100µg/ml Ampicillin. Overnight incubation of the plated samples at 37°C, resulted in the appearance of multiple white colonies containing the inserted viral-gene expression constructs. 4 single colonies were selected from each sample plate and inoculated into 10ml of LB broth with the appropriate antibiotic (50µg/ml of Ampicillin). The cultures incubated at 37°C, overnight, with constant shaking (150rpm).



### ***Glycerol Stocks and Gene Insert Validation***

Glycerol stocks were made by homogenizing 250µl of the incubated cultures with 750µl molecular grade glycerol and storing them at -80°C. A scraping of each glycerol stock was inoculated into the 10ml of LB broth containing the Ampicillin. Another cell scraping of each sample was inoculated into 10ml of broth containing an antibiotic to which the bacteria were susceptible (30µg/ml of Kanamycin was used for GST vectors) in order to validate the efficacy of the antibiotic resistant gene found in each vector.

Plasmid DNA was extracted from the first group of overnight cultures and restriction enzyme digested using *Sgf I* and *Pme I* to validate the presence of the viral insert in each construct (as previously described). Validated glycerol stocks were stored at  $-80^{\circ}\text{C}$ .

***Pilot Study: Expression of MΔN-GST Fusion Protein***

An overnight culture of MΔN-GST was diluted 1:100 in 250ml fresh LB media containing 50μg/ml Ampicillin (for GST vectors). The inoculated culture was incubated at  $37^{\circ}\text{C}$  with continuous shaking (150rpm) until it reached an optical density of  $\sim 0.2$  (600nm). Once optimal density (OD of 0.2 at 600nm) was reached, the culture was titrated with 0.1% L-Rhamnose (w/v) to induce protein expression. Every 2Hrs post induction a 10ml sample of the culture was collected and stored at  $4^{\circ}\text{C}$  until needed. On completion of the incubation period, the collected 10ml samples were centrifuged at 2800rpm to pellet the bacteria.

The pellets were resuspended in 3.6ml fresh LB media before adding 400μl of 10X concentrate FastBreak Cell Lysis reagent (Promega) and vortexing. Samples were incubated at room temperature for 30min with gentle shaking before being centrifuged at 5000xg to pellet cell debris. 200μl of the cell lysates were extracted and diluted 1:1 in SDS Laemmli's loading buffer (4% (w/v) SDS, 20% (v/v) Glycerol, 10% (v/v) β-mercaptoethanol, 0.004% (v/v) Bromphenol blue and 0.125M Tris HCL (Fluka)) preceding boiling at  $95^{\circ}\text{C}$  for 3min, in a dry bath. The expressed proteins were separated on a 15% (w/v) polyacrylamide SDS-PAGE (at 20mA) and Coomassie stained to visualize the expressed proteins and substantiate when optimal protein expression levels were achieved.

***Expression of Viral GST Fusion Proteins by Autoinduction***

Various modifications were attempted to reduce the latency of protein expression observed in the standard protein expression. One of these attempts was to modify the expression broth by adding 0.05% filter sterilized Glucose (v/v) and 0.1% L-Rhamnose (w/v), to assist the cultures in reaching optimum biomass before expression was induced. This autoinduction protocol was described by the manufacturer (Promega, USA) and was adopted as the preferred expression protocol used for the remainder of this body of work.

Overnight starter cultures (Glycerol stock scraping inoculated into 10ml LB broth with 50µl/ml Ampicillin.) were diluted 1:100 in 250ml fresh LB media containing Ampicillin (50µg/ml), 0.05% filter-sterilized Glucose (w/v) and 0.1% L-Rhamnose (w/v). Inoculated cultures were incubated at 37°C for 24Hrs with continuous shaking (150rpm). Every 2Hrs a 10ml sample of the culture was collected and stored at 4°C until needed. On completion of the incubation period, the collected samples were centrifuged at 2800rpm to pellet the bacterial cells.

The pellets were resuspended in 3.6ml fresh LB media before adding 400µl of 10X concentrate FastBreak Cell Lysis reagent (Promega) and vortexing. Samples were incubated at room temperature for 30min with gentle shaking before being centrifuged at 5000xg to pellet cell debris. 200µl of the cell lysates were extracted and the protein concentration quantified using a Bradford total protein assay (BioRad).

***Expressed Protein Quantification***

The Bradford assay required 5µl of the cell lysates to be added to 250µl Bradford 1X dye reagent (at room temperature) in a 96 well clear bottom plate. Absorbance was read at 590nm and the concentration of the six samples calculated by substituting their specific OD's into the equation of the standard curve. The samples were divided into 30mg aliquots and diluted 1:1 in SDS Laemmli's loading buffer (4% (w/v) SDS, 20% (v/v) Glycerol, 10% (v/v) β-mercaptoethanol, 0.004% (v/v) Bromphenol blue and 0.125M Tris HCL (Fluka)) preceding boiling at 95°C for 3min, in a dry bath.

***Validation of GST Fusion Protein Expression***

15mg of each of the three GST-fusion viral proteins were separate by SDS-PAGE gel electrophoresis at 20mA/gel for 74min. On completion one of the SDS-PAGE gels (10cm X 8cm) was visualized by soaking it in Coomassie Blue Stain (45% (v/v) Methanol, 10% (v/v) Glacial acetic acid, 45% (v/v) distilled water and 3g/l Coomassie Brilliant Blue R250 powder (Fluka)) overnight, with gentle rocking, followed by destaining in Coomassie Destain solution (30% (v/v) Methanol, 10% (v/v) Glacial acetic acid and (v/v) 60% Distilled water) for 12Hrs with rocking.

***Western Blot analysis***

The remaining duplicate SDS-PAGE gel was transferred by Western Blot onto PVDF (Polyvinylidene Fluoride) membrane (BioTrace) in a submersion, transfer system (Hoofer: Mighty Small Transfer) at 100V, for 90min, with the use of a

cooling system. The resultant membrane was soaked in Ponceau S (0.1% (w/v) Ponceau S powder (BioLab), 5% (v/v) acetic supplemented to 1L with distilled water) to verify protein transfer and for consistent protein-concentration loading verification, before being blocked in blocking solution (3% (w/v) fat free powder milk (BioRad) in phosphate buffered saline (PBS) containing 0.05% (v/v) Tween-20). Once blocked, the blot was exposed to a 1:5000 dilution of rabbit anti-GST primary antibody (Santa Cruz Biotechnology), in 3% the milk, overnight at 4°C with continuous rolling.

Following the overnight incubation, the blot was washed three times in 1X PBS containing 0.05% (v/v) Tween-20 for 15min per wash. The secondary antibody, goat anti-rabbit conjugated with Horse Radish Peroxidase (Santa Cruz Biotechnology), was diluted 1:5000 in the 3% milk and exposed to the blot for one hour at room temperature on a roller mixer. The blot was once again washed in the PBS solution before being exposed to the HRP chromogenic substrate (KPL) at room temperature for ~10min. Blots were photographed using a Canon (G6 Powershot) camera on a white background.

### ***Antigen Generation***

Protein production now validated, the viral GST-fusion proteins were once again expressed in KRX bacteria under the influence of the inducer L-Rhamnose. The 75ml cultures (LB broth + 100µg/ml Ampicillin) were then pelleted at 5000rpm for 10min before being reconstituted in 36ml of saline (0.7% (w/v) NaCl), bringing culture concentration closer to the required absorbance of 2 (OD 600nm). Next,



4ml of the 10X FastBreak bacterial cell lysis solution (Promega) was added to the cultures, which were incubated at R/T for 30min, with rocking.

The lysed bacteria were centrifuged at 8000rpm to pellet cell debris and isolate the soluble cytoplasmic proteins. The lysates containing the expressed proteins were subsequently gravity filtered through 10mg-capacity, GST-affinity, agarose-based column (Thermo Scientific) to isolate the expressed GST-fusion proteins. On completion of filtration, the fusion proteins were eluted in protein elution buffer (Pierce Proteomics) in preparation for inoculation of the Balb/c mice.

### ***Balb/c Mice***

The 11 Balb/c mice used for antibody production (3 mice per antigen (9 in total) + 2 control mice) were obtained from the University of Cape Town, Special Pathogen-free Animal Unit. Mice were housed in separate cages with freely available water and animal feed, supplied by the Medical Research Council South Africa. Cages were cleaned and maintained on a weekly basis.

### ***Antigen Preparation***

The eluted GST-fusion ORF3ΔN, MΔN and E viral proteins were quantified using a Bradford assay (BioRad). 250µl of 1X BioRad Bradford dye reagent (room temperature) was added to 5µl of the eluted proteins and 5µl of BioRad Bradford bovine serum albumin standards (at concentrations of 2.5mg, 2.0mg, 1.5mg, 1.0mg, 0.75mg, 0.5mg, 0.25mg and 0.0mg/ml) in a 96 well microtiter plate. The absorbance (at 620nm) obtained for the Bradford standards were used to

construct a concentration trend straight line graph. The absorbencies of the eluted samples were substituted into the equation of the above mentioned straight line graph to obtain the concentration of the samples.

### **Concentration of Antigens**

Protein samples were then concentrated using 5kDa molecular-weight-cut-off (MwCO), polysulfone membrane filters (Separations Biotech) and quantified again with a Bradford assay. 10ml of the GST protein elutes were centrifuge-filtered through the polysulfone filters at 10000rpm for 10min. 100 $\mu$ g aliquots of the isolated viral ORF3 $\Delta$ N-,  $\Delta$ N- and E-GST fusion proteins were supplemented with DPBS (Dulbecco's modified Phosphate Buffered Saline) up to 150 $\mu$ l, before emulsification with an equal volume of Freund's complete adjuvant (500 $\mu$ g/ml *Mycobacterium butyricum* suspended in 15% (v/v) Mannide Monooleate and 85% (v/v) Paraffin oil) (Life Technologies).

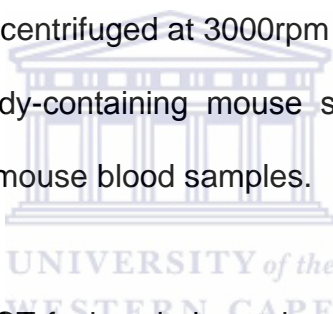
### **Inoculation of Balb/C Mice**

Approximately 33 $\mu$ g (100 $\mu$ l) of each antigen solution was injected into each mouse intra-peritoneally, which constituted the primary inoculation. Each of the GST-fusion antigens were injected into 3 mice resulting in 9 mice being used for antibody production and 2 mice being reserved as a negative control (*i.e.* these mice were not inoculated). Mice were inoculated and left for 3 weeks before being boosted with 33 $\mu$ g of each antigen emulsified in Freund's incomplete adjuvant (15% (v/v) Mannide Monooleate and 85% (v/v) Paraffin oil) (Life Technologies), every 2 weeks.

One week after the 4<sup>th</sup>, 5<sup>th</sup> and 6<sup>th</sup> booster injections, 15µl of blood was harvested from each mouse. Blood was obtained by clipping off ~2mm of tissue from the tip of the mouse's tail and extracting blood from the transverse tail vein. The harvested blood was resuspended in 1.85ml sterile PBS. All procedures for the use of these laboratory animals were approved by the University of the Western Cape Animal Research Ethics Committee and carried out in accordance with the regulations and guidelines as stipulated by this committee.

### ***Validation of Antibody Production***

The resuspended blood was centrifuged at 3000rpm for 10min to separate the red blood cells from the antibody-containing mouse serum. ~1400µl of sera was isolated from each of the 11 mouse blood samples.



Approximately 10µg of the GST-fusion viral proteins were separated by SDS-PAGE gel electrophoresis at 20mA/gel for 74min. The GST-fusion proteins were transferred by Western Blot onto PVDF (Polyvinylidene Fluoride) membrane (BioTrace) in a submersion, transfer system (Hooper: Mighty Small Transfer) at 100V, for 90min, with the use of a cooling system. The resultant membrane was soaked in Ponceau S (0.1% (w/v) Ponceau S (BioLab), 5% (v/v) acetic acid supplemented to 1L with distilled water) to verify protein transfer and for consistent protein-concentration loading verification, before being blocked in blocking solution (3% (w/v) fat free powder milk (BioRad) in phosphate buffered saline (PBS) containing 0.05% (v/v) Tween-20). Next, the blot was cut into 3

strips thus isolating each GST-fusion antigen and 1 negative control on a separate piece of PVDF (Polyvinylidene fluoride) membrane.

The sera from the first bleed, containing the mouse polyclonal antibodies against the GST-fusion viral proteins, were diluted 1:500 in 5ml of blocking solution (3% (w/v) fat-free milk (BioRad) in 1xPBS containing 0.05% (v/v) Tween-20). The individual blots were exposed overnight to the various mouse polyclonal antibodies at 4°C, with continuous rolling. Following the overnight incubation, the three blots were rinsed 3 times in 1X PBS (containing 0.05% (v/v) Tween-20) of 15min. The secondary antibody, goat anti-mouse polyclonal conjugated with Horse Radish Peroxidase (Santa Cruz Biotechnology), was diluted 1:5000 in the blocking solution and exposed to the blots for one hour at room temperature with continuous rolling. The blots were once again rinsed in the PBS solution before being exposed to the HRP chromogenic substrate (KPL), at room temperature for ~10min. Blots were photographed using a Canon (G6 Powershot) camera on a white background.

As a result of the Western Blot, it was evident that the mouse antibodies were able to detect the GST-fused viral proteins. Detection of untagged viral proteins was imperative for the validation of specific antibody-antigen recognition. In an attempt to elucidate specific antibody activity, an ELISA (enzyme-linked immunosorbent assay) was executed.

### ***Indirect Viral Protein Specific ELISA***

A flat bottom 96 well plate (Nunc-Immuno plate, MaxiSorp) was coated with 50µl/well (10µg) of GST-fused viral protein in PBS and incubated overnight, at 4°C. After 3 washings in wash buffer (autoclaved phosphate buffered saline and 0.05% (v/v) Tween-20), non-specific binding sites were blocked with 1% (v/v) BSA (Bovine serum albumin) made up in PBS for 1 hour at R/T. A 50µl solution containing the diluted primary antibody (1:500 anti-GST-viral protein) constituted in PBS and added to all wells. This preceded a 1-hour incubation period at R/T, with orbital shaking. Subsequently, the plate was washed 3 times and 50µl of goat anti-mouse-horseradish peroxidase conjugate (1:5000 in PBS) was added to each well. Following an incubation period of 1 hour (at R/T with orbital shaking), the plate was washed 3 times in sterile PBS with 0.05% (v/v) Tween-20. 50µl of substrate solution (Tetramethylbenzidine solution) was added to all wells and incubated for 10-15min at R/T. The chromogenic reaction was then stopped with the addition of 50µl/well of stop solution (1M H<sub>2</sub>SO<sub>4</sub>) and the absorbance was read spectrophotometrically at 450nm (Original Multiskan EX, Type 355, Thermo Electron Corporation, Shanghai, China). The results of this ELISA displayed universal non-specificity of the antibodies resulting in protocol optimization.

### ***Optimization of ELISA's Protocol***

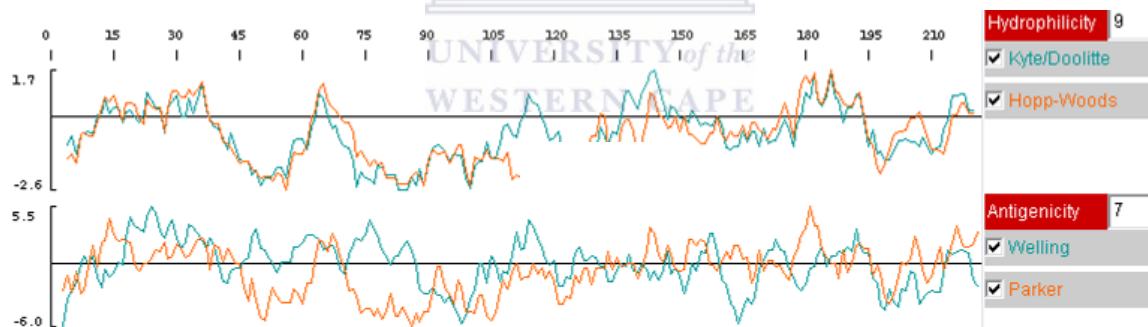
In an attempt to optimize antibody specificity, 0.05% (v/v) Tween-20 was added to the blocking solution and the assay diluents for the primary antibody. The addition of a random GST-tagged antigens (10µg/well) to the primary antibody solution, resulting in an indirect viral protein specific competition ELISA, as a final attempt to further decrease the high GST affinity and non-specific binding observed.

## CHAPTER 3

### Results and Discussion

#### *In silico Analysis*

The *in silico* analysis of the human coronavirus NL-63 ORF3 and M genes were carried out using Invitrogen's peptide selector tool (<http://peptideselect.invitrogen.com/peptide/>). Whereas the hydrophilic state of the proteins was analyzed using the Kyte-Doolittle (1982) and Hopp-Woods (1981) selection criteria, the protein antigenicity was analyzed according to the Welling and Parker methods.



**Figure 4: NL-63 ORF3 Hydrophilicity and Antigenicity plot**

The above figure depicts increased hydrophilicity and antigenicity in the 3' half of the HCoV-NL-63, ORF3 gene. Peaks protruding above the X-axis indicate the regions with increased hydrophilicity or antigenicity.

Based on the results shown in Figures 4 and 5, the 3'-halves of the ORF3 (from 330nt to 678nt) and M (from 298nt to 681nt) genes were selected as the most

immunogenic regions of the protein sequences based on amino acid polarity and membrane-surface orientation.

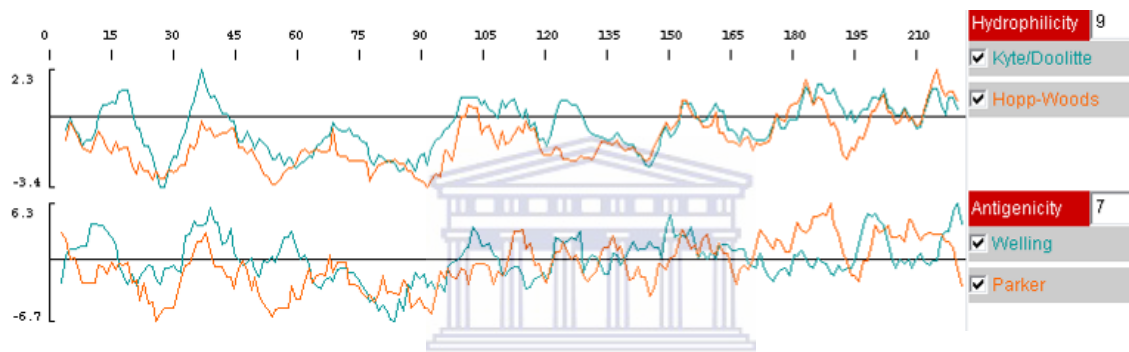
The Kyte-Doolittle method used to calculate the antigenicity of proteins considers the molecular structure of each amino acid and assigns each one a hydrophilicity value between negative 4.6 (Hydrophilic) and positive 4.6 (Hydrophobic). The program then averages the hydrophilicity scores of every 9 amino acids to provide the user with a hydrophilicity plot for every section of the protein (Kyte and Doolittle, 1982). The Hopp-Wood method utilizes a similar method of calculating protein solubility with the exception of the suggested average window size being 7 amino acids (Hopp and Woods, 1981).

The Welling model for calculating antigenicity is based on comparing the functionality of the input amino acid sequences to that of known amino acid peptide epitopes with known antigenicity (Welling *et al.*, 1985). Epitope similarity between the query protein and the known proteins suggests antigenic similarity. Similar to Welling's method, the Parker method of calculating antigenicity also compares peptides to known antigenic strings of amino acids. However, the Parker method differs in that it accommodates the effect of various protein side chains when calculating antigenicity (Parker *et al.*, 1994).

The graphs indicated 3 hydrophobic regions in the 5' halves of the ORF3 and M genes. For the purpose of antibody generation, the expressed tagged proteins would have to be soluble and not membrane bound in order to be isolated for

inoculation. This discovery led to the truncation of the undesired region in the 5' halves of the proteins.

For the purpose of this study primers were designed to amplify the 3'-portions ORF3 and M genes. Primers were also designed to amplify the entire E gene. All primers were designed to incorporate restriction enzyme sites for unidirectional, in-frame cloning of the inserts into vectors for expression in a bacterial system.



**Figure 5: NL-63 M Hydrophilicity and Antigenicity plot**

The figure above indicates the increased antigenicity in the 3' half of the HCoV-NL-63 M gene according to the Welling and Parker methods. The Kyte-Doolittle and Hopp-Woods tests also showed more stable hydrophilicity in that same region of the gene.

Forward primers consisted of a nucleotide sequence to append the restriction enzyme cut site for *Sgf* I, 1 or 2 additional nucleotides to allow for in-frame ligation of the viral genes into the expression vector, and 24 to 27 nucleotides complementary to the start of the section of DNA (ORF3- 330 to 355nt; M- 298 to 225nt and E- 1 to 25nt) to be amplified on the complementary DNA strand. The reverse primer consisted of a nucleotide sequence to append the *Pme* I restriction enzyme site as well as 24 to 27 nucleotides, complementary end of the gene to

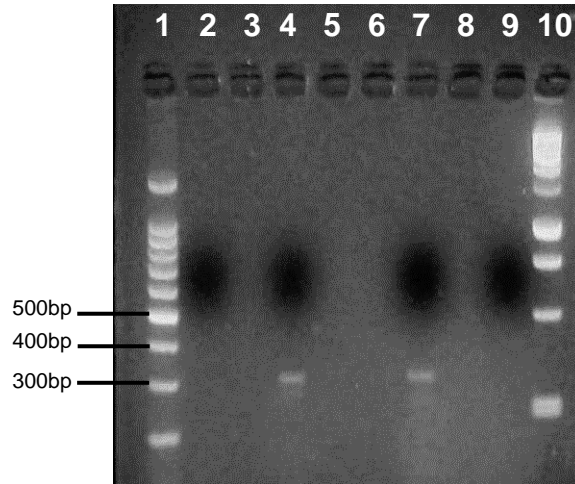


be amplified (ORF3- 654 to 678nt; M- 654 to 681nt and E- 177 to 201nt). These primers were then synthesized by Inqaba Biotech (Table 1, Page 27).

### ***Reverse Transcription and Gene Specific PCR***

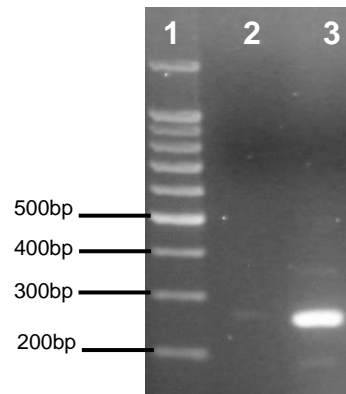
Template viral RNA was extracted from infected Vero E6 cells (African Green Monkey kidney epithelial cells) and given to us as a kind gift from Professor Lia van der Hoek in Amsterdam. The RNA was used for first strand cDNA synthesis by reverse transcription using random primers. To ensure solubility of bacterially expressed viral recombinant proteins, only the predicted hydrophilic portions of proteins were amplified by polymerase chain reaction.

Prior experiments have indicated difficulty when expressing insoluble viral proteins (Kondo *et al.*, 2008). Therefore, polymerase chain reaction amplification of the hydrophilic portions of M and ORF3 (ORF3 $\Delta$ N and M $\Delta$ N) and the complete E gene was done. The amplification products observed at ~350bp (Figure 6, Lanes 4 and 7, respectively) corresponded to the predicted size as determined by the *in silico* analysis, *i.e.* ORF3 (~360nt) and M (~390nt). The amplified product for the E gene, with an estimated size of ~250nt (Figure 7, Lane 3), corresponded to the predicted size of the *in silico* analysis. The accurate size of the amplicons is not a definitive indication of amplification of the genes of interest however, and it was decided to confirm the identity of the products by sequence verification. Therefore, products were purified and cloned for sequencing.



**Figure 6: Gene specific PCR of ORF3 $\Delta$ N and M $\Delta$ N**

Figure 6 depicting a 100bp DNA marker in lane 1, the ORF3 $\Delta$ N negative control (water) seen in lane 2, the amplified ORF3 $\Delta$ N gene (~360nt.) in lane 4, the amplified M $\Delta$ N gene (~390nt.) in lane 7 and the M $\Delta$ N negative control (water) seen in lane 9. Lane 10 contains a 1Kbp DNA marker.



**Figure 7: Gene specific PCR of E**

The figure above depicts a 100bp DNA marker in lane 1, the E negative control (water) in lane 2 and the amplified E gene (~250nt.) in lane 3.

### ***Ligation into the pGEM Vector for Sequencing***

Amplicons were gel purified to remove the ethidium bromide molecules and any salts that could interfere with the ligation process. The adenosine appended 'sticky ended' purified genes were now ready to be ligated into pGEM T-Easy vectors, which were subsequently screened by blue-white colony selection. The pGEM T-Easy was used for three reasons:

- This vector contains up-stream and down-stream M13 sequencing primer binding site, making it relatively easy and cost effective to sequence inserted products;
- The 5' and 3' ligation sites found in the vector contain 'T' overhangs for easy ligation of 'blunt ended' PCR products into the vector;
- The vector allows for Blue/White colony selection via the presence of a  $\beta$ -galactidase.

Selected white colonies were screened by genetic sequencing as discussed elsewhere.

### ***Transformation into the JM109 Competent E. coli***

pGEM constructs were transformed into JM109 chemically competent *E. coli*. Once the JM109 cells were heat shocked into accepting the pGEM constructs, cells containing unligated plasmids digest the substrate (X-gal) to produce a blue dye via the action of the  $\beta$ -galactidase gene. The presence of this die would result in a blue colony growing on the LB agar plate. Ligated plasmids with their

disrupted  $\beta$ -galactidase genes produced no dye and would thus results in white colonies.

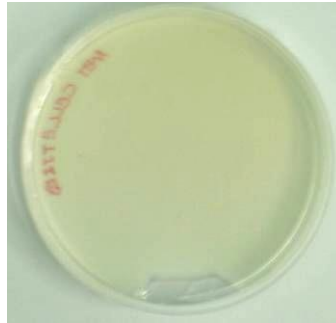
The positive transformation control plate, displaying only white colonies (Figure 8), indicated the activity of the LacZ operon inducer (IPTG), the  $\beta$ -galactidase enzyme substrate X-gal and the efficacy of the pGEM-3Z vector transformation procedure. Figure 9 depicts the negative control transformation plate that displayed no growth and thus validated the antibiotic included in the semisolid media. The blue colonies observed on the background control plate (Figure 10) indicated the efficiency of the ligation reaction. These colonies contain religated pGEM vector, insert deficient with reconstructed LacZ operons which produced blue colonies.

The blue and white colonies displayed on the experimental plates (Figures 11, 12 and 13) showed a number of successfully ligated and transformed pGEM-viral gene constructs. The correct functioning of the experimental controls validated the white colonies observed on the experimental plates.



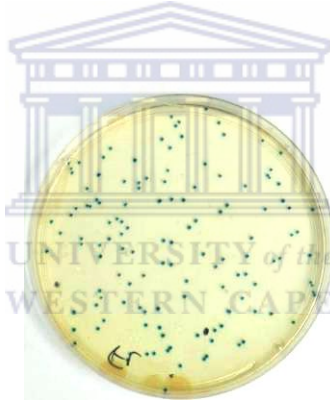
**Figure 8: Transformation into JM109 – Positive control**

The numerous white colonies observed on the positive control plate, containing 100µg/ml ampicillin, indicate a successful transformation of the pGEM-3Z plasmid without an insert, into the JM109 competent *E.coli* cells.



**Figure 9: No Transformation into JM109 – Negative control**

The absence of the untransformed *E.coli* growth on the negative control plate (LB agar and 100µg/ml ampicillin) indicates the efficacy of the antibiotic used in the transformation experiments.



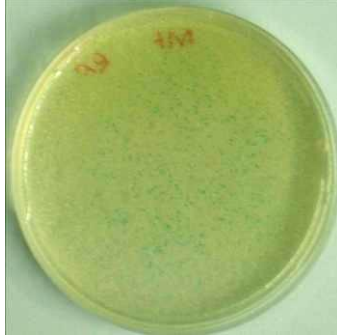
**Figure 10: Transformation into JM109 – Background control**

The blue colonies seen on the background control plate (LB agar and 100µg/ml ampicillin) validate the ligation procedure. The colonies seen here contain competent *E.coli* transformed with self-ligated pGEM t-Easy vector with no insert.

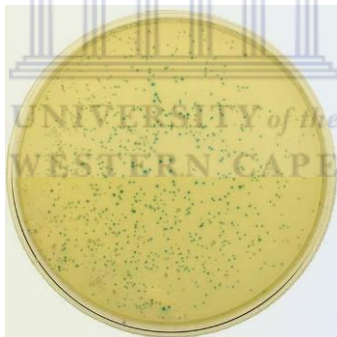


**Figure 11: Transformation into JM109 – ORF3ΔN**

Both blue and white colonies are observed on this transformation plate (LB agar with 100µg/ml ampicillin). The white colonies seen here contain properly ligated pGEM t-Easy vectors with the viral ORF3 gene inserted, successfully transformed into the JM109 bacteria.

**Figure 12: Transformation into JM109 – MΔN**

Both blue and white colonies are observed on this transformation plate (LB agar with 100µg/ml ampicillin). The white colonies seen here contain properly ligated pGEM t-Easy vectors with the viral M gene inserted, successfully transformed into the JM109 bacteria.

**Figure 13: Transformation into JM109 – E**

Mostly blue colonies are observed on this plate (LB agar with 100µg/ml ampicillin). The white colonies seen here contain properly ligated pGEM t-Easy vectors with the viral E gene inserted, successfully transformed into the JM109 bacteria.

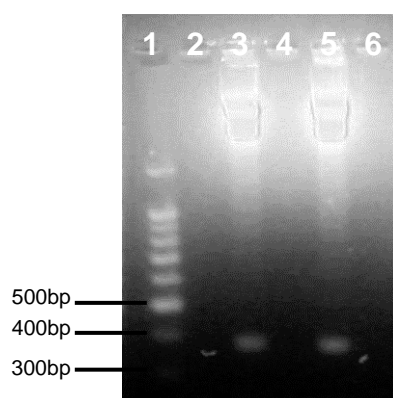
***Recombinant Vector Confirmation and Gene Sequencing***

10µl of the purified viral gene-pGEM constructs were sent to Inqaba Biotech for gene sequencing using the M13 universal sequencing primers. Figures 40, 41 and 42 indicate the sequenced ORF3ΔN, MΔN and E genes in the pGEM vector.

Positive gene identification was made by firstly aligning the forward and reverse sequences obtained from the Inqaba Biotech. Next, the consensus sequence was compared to the 3' half of the full-length ORF3 and M genes and the full length E gene available online from NCBI (NL-63-ORF3: AY697419, NL-63-M: AY697422 and NL-63-E: AY697421). The homologies observed between the original sequences on the database and the experimentally determined sequences, validated the specificity of the designed primers for gene amplification.

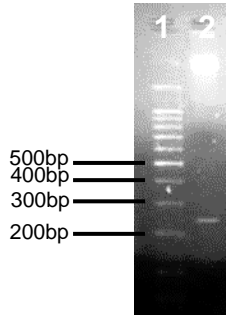
### ***Sgf I and Pme I Restriction Enzyme Digest of pGEM Constructs***

The restriction digestion of the ORF3 $\Delta$ N, M $\Delta$ N and E-pGEM constructs resulted in the generation of 'sticky ends' that was ligated into the compatible sites of the GST bacterial expression vector. Whereas Figure 15 depicts the digested and sequence-verified viral E gene (~230bp), Figure 14 indicates the successful digestion of the sequenced ORF3 $\Delta$ N (~350bp) and M $\Delta$ N (~380bp) genes. The overhangs produced by this digestion allowed for the in-frame ligation of the viral genes into the bacterial expression vector, resulting in the production of the correct viral protein.



**Figure 14: *Sgf* I and *Pme* I digestion of ORF3ΔN and MΔN-pGEM construct**

Figure 14 above depicts a 100bp DNA marker lane 1, the released MΔN gene (~381nt) in lane 3 and the released ORF3ΔN gene (~348nt.) seen in lane 5.

**Figure 15: *Sgf* I and *Pme* I digestion of E-pGEM construct**

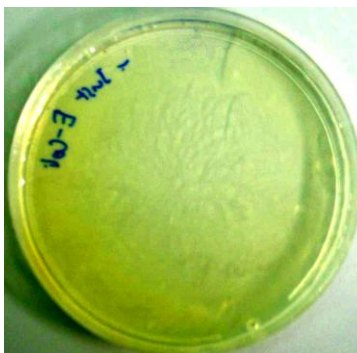
The figure above displays a 100bp DNA marker in lane 1 and the digested and released E gene (~234nt) in lane 2.

***pFLEXI* Ligation and Transformation**

Sequence-verified pGEM-constructs were digested with *Sgf* I and *Pme* I and cloned into compatible sites of pFLEXI vector pFN2A (GST). The use of infrequent cutting restriction enzymes *Sgf* I and *Pme* I ensured the unidirectional cloning of the inserts into the vector. The pFLEXI vector also contains a lethal barnase gene that is disrupted during the process of ligation of inserts into the vector. The disruption of the barnase gene results in only properly ligated and transformed cells being able to grow in culture, therefore serving as its own positive control. Additionally, the vector contains two incompatible 5'- and 3'-end overhanging to prevent re-ligation of the vector with itself. The combinations of these characteristics result in the transformation experiment only requiring a negative control to validate the antibiotic used (Figure 16) and duplicate experimental plates for repeatability (Figures 17, 18 and 19). Thus, the single



white colonies observed on the experimental plates were a result of successful ligation and subsequent transformed of competent cells.



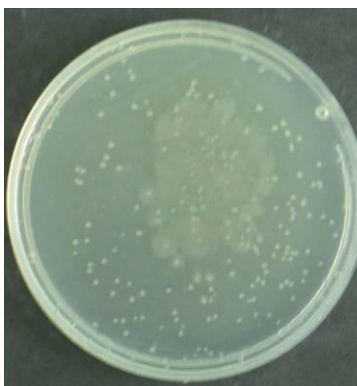
**Figure 16: Transformation into KRX – Negative control**

Here untransformed KRX bacteria were spread on an LB agar plate containing 100µg/ml ampicillin. The absence of bacterial growth validates the efficacy of the antibiotic utilized.



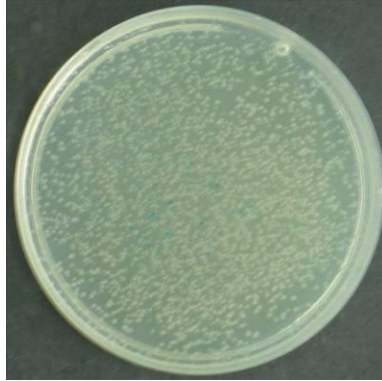
**Figure 17: Transformation into KRX- ORF3ΔN-GST**

Sparse growth observed on this plate indicates sub-optimal transformation of the ORF3ΔN-Flexi construct into the KRX bacteria. The LB agar used in this transformation contained 100µg/ml ampicillin.

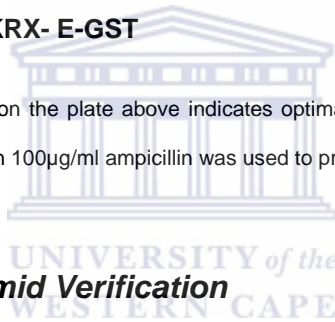


**Figure 18: Transformation into KRX- MΔN-GST**

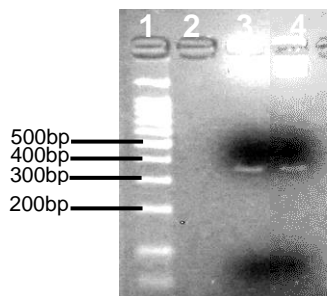
Sub-optimal growth was observed for the growth of the MΔN-Flexi-transformed KRX bacteria. 100µg/ml ampicillin was added to the LB agar in order to discourage growth of any possible contaminant bacteria.

**Figure 19: Transformation into KRX- E-GST**

The multitude of white colonies observed on the plate above indicates optimal ligation and transformation of the E-Flexi construct into the KRX bacteria. Once again 100µg/ml ampicillin was used to prevent bacterial contaminant growth.

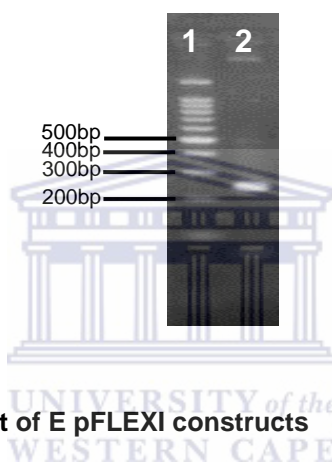
***pFLEXI Recombinant Plasmid Verification***

White colonies were selected from plates and recombinant plasmids were confirmed by restriction digestion analysis with *Sgf* I and *Pme* I for the correct insert size. This step is essential to ensure that the genes are still present and intact in the KRX expression bacteria before viral protein expression is induced. The bands seen at ~350bp (ORF3ΔN), ~380bp (MΔN) in Figure 20, and the band seen at ~230bp (E) in Figure 21, positively indicate the presence of the selected genes. Thus it is assured that the expressed proteins will be GST fusion proteins.



**Figure 20: *Sgf* I and *Pme* I digest of ORF3 $\Delta$ N and M $\Delta$ N from pFLEXI constructs**

Figure 20 above displays a 100bp DNA marker in lane 1, the released ORF3 $\Delta$ N gene (~350nt) in lane 3 and the released M $\Delta$ N gene (~380nt) in lane 4.

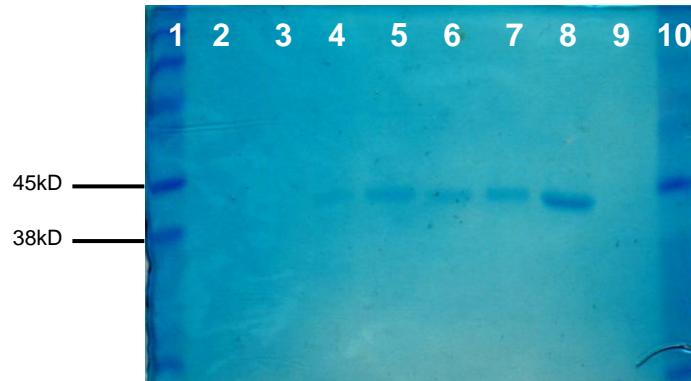


**Figure 21: *Sgf* I and *Pme* I digest of E pFLEXI constructs**

The above figure depicts a 100bp DNA marker in lane 1 and the released E gene (~230nt) seen in lane 2.

### ***Pilot Study: Expression of M $\Delta$ N-GST Fusion Protein***

A pilot study was conducted for two reasons: Firstly, to verify the efficacy of the proposed expression protocol and secondly, to gauge the time taken to reach optimal expressed protein concentrations. M $\Delta$ N-GST was cultured to an optical density of ~0.2 (600nm) and induced with L-Rhamnose. Theoretically, expression should occur within 2Hrs of induction as previously seen (Tan *et al.*, 2004; Keng *et al.*, 2005). As seen in the Coomassie stain (Figure 22), optimal levels of protein concentration are only seen as late as 12Hrs post induction.



**Figure 22: 12 hour Pilot study - Coomassie stained SDS PAGE of expressed MΔN-GST fusion protein**

The figure above displays the supernatant of bacterial expression of MΔN-GST. Rainbow SDS protein marker is seen in lanes 1 and 10. 2hr, 4hr, 6hr, 8hr, 10hr and 12hr (lanes 3 to 8) supernatant showing optimal expression of MΔN-GST (~40KD.) is seen in lane 8 (at 12Hrs).

It was evident from the results of the pilot study that protein expression only reached optimal levels 12hrs after induction which was much later than the expected 2 to 4 hrs post induction. As a result of this late induction, the expression procedure was modified by supplementation of the expression broth with 0.05% (w/v) glucose and 0.1% (w/v) L-Rhamnose at the inoculation stage.

### ***Expression of Viral GST Fusion Protein by Autoinduction***

Previous studies have shown that antibodies targeting coronaviral proteins expressed in bacteria are able to neutralize the virus in culture (Zhang *et al.*, 2004; Zhou *et al.*, 2004; Keng *et al.*, 2005). N and S coronaviruses proteins produced in bacteria have been utilized to generate *in vivo* antibodies that are able to detect NL-63 from clinical samples (Dijkman *et al.*, 2008; Shao *et al.*,

2007; Hofmann *et al.*, 2005). Cross-reactivity observed between antigenic epitopes of coronavirus species does however hinder accurate diagnosis of infection (Chan *et al.*, 2005). Our aim was to produce protein-specific antibodies that could ultimately be used for detection of virus and viral products in laboratory and clinical samples. To produce these antibodies, the E gene and the antigenic regions of the ORF3 and M genes were expressed as GST-tagged fusion proteins in a bacterial system.

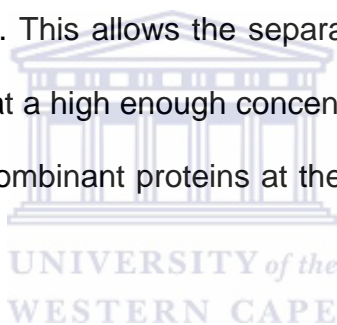
In our system, recombinant protein production was induced by the addition of L-Rhamnose to the culture media. Also, the supplementation of glucose to the media allowed the *E. coli* cultures sufficient time to reach the optimal required cell density (i.e.: the equilibrium phase of bacterial growth) for effective metabolism of the L-Rhamnose and subsequent optimal induction of protein expression. As the bacteria preferably metabolize the glucose, L-Rhamnose metabolism, which induces protein expression, is delayed until all glucose in the medium is depleted. This cocktail of supplements resulted in the induction of expression approximately 6 to 10Hrs after the culture was inoculated (Figures 23, 24 and 25).

Expressed recombinant fusion proteins were extracted from the KRX bacterial cells and separated on 15% (w/v) SDS-PAGE gel, which resolves the negatively charged proteins through the acrylamide matrix based on the basis of their molecular weight. The observed molecular weights of the recombinant viral proteins (excluding the ~27kDa of the GST tag) was similar to the predicted MW of ORF3ΔN (~14kDa), MΔN (~15kDa) and E-GST (6kDa). It is important to

remember that the observed molecular weight of the recombinant proteins (ORF3: ~42kDa, M: ~42kDa and E: ~36kDa) consists of the viral protein and the N-terminal GST-tag.

### ***Coomassie Stain and Western Blot Analysis***

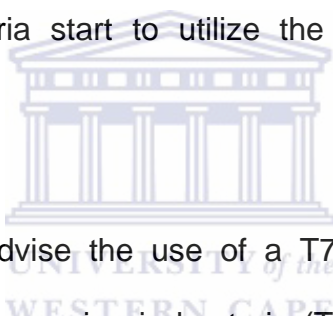
To determine the optimal expression conditions (time and temperature), a time course expression study was conducted over a 12-hour period for the ORF3-GST and M-GST mutant proteins; 36-hour time course was done for E-GST. The Coomassie Blue stain is used to stain proteins in polyacrylamide gels without permanently staining the gel. This allows the separated proteins to be visualized in the gel provided they are at a high enough concentrations. Staining of the SDS-PAGE gels revealed the recombinant proteins at their expected size (Figures 23, 24 and 25).



It was however, also observed that the induction of protein expression was significantly later than expected. Whereas previous studies commonly report the expression of GST-coronaviral recombinant proteins as early as 2Hrs after induction in a bacterial system (Tan *et al.*, 2004; Keng *et al.*, 2005), the delayed response we observed in this study could possibly be attributed to viral proteins toxicity. Even though the expression procedure was carried out according to the manufacturer's instructions (Promega, USA), optimal protein concentrations were only observed after 10Hrs post induction with L-Rhamnose and Glucose supplementation. This late induction indicated a prolonged exponential bacterial

growth phase before the bacteria reach their equilibrium-phase and protein production could begin.

The presence of an unstable promoter sequence in an expression vector has been seen to result in low levels of toxic protein expression during both the initial and lag phases as observed in this study (Giacalone *et al.*, 2006). The background toxic expressions result in reduced bacterial growth and thus a delayed onset of bacterial equilibrium phase. As full scale induction only occurs at the final phase of bacterial growth, when the glucose in the media has been metabolized and the bacteria start to utilize the L-Rhamnose, latent protein expression is observed.



Even so, researchers do advise the use of a T7 promoter to allow for tight regulation of toxic protein expression in bacteria (Tabor and Richardson. 1985). Coupled with the use of a rare inducer, such as L-Rhamnose, the tightly regulating T7 promoter should have resulted in more stringently controlled expression.

A variety of procedures were explored to reduce the delay of protein expression induction. One of these methods included expression of proteins at 27°C, with continuous shaking (150rpm), to reduce basal toxic protein expression during the bacterial growth phase. Another method attempted to decrease the latency of protein expression was the addition of 1mM IPTG to the expression broth. The addition of IPTG more stringently control expression and reduces background

toxic protein expression levels. Both attempts to reduce the delayed expression were unsuccessful.

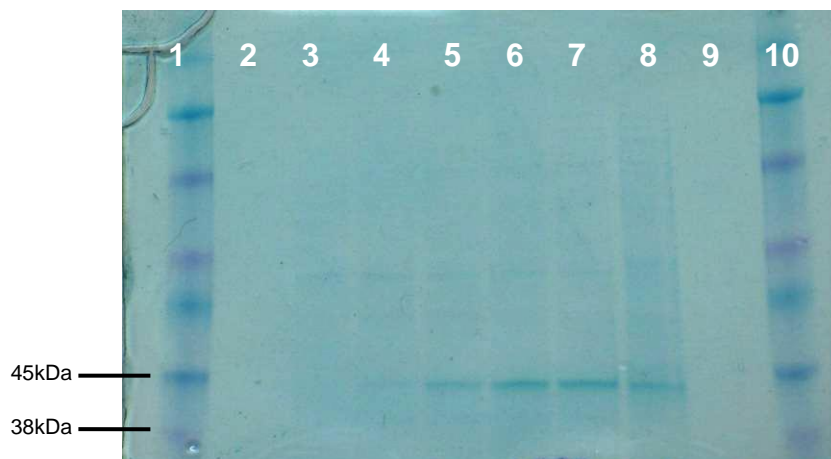
The time course experiments however revealed the highest titer of ORF3 proteins 12Hrs after induction and at 10Hrs for M protein. The highest titers of E-GST proteins were observed after 16Hrs. The expression times with the highest protein titers were used in the final large-scale expression to obtain the greatest antigen concentration for the mouse inoculations.



**Figure 23: Coomassie stain displaying optimal ORF3ΔN-GST fusion protein expression at 12 hours**

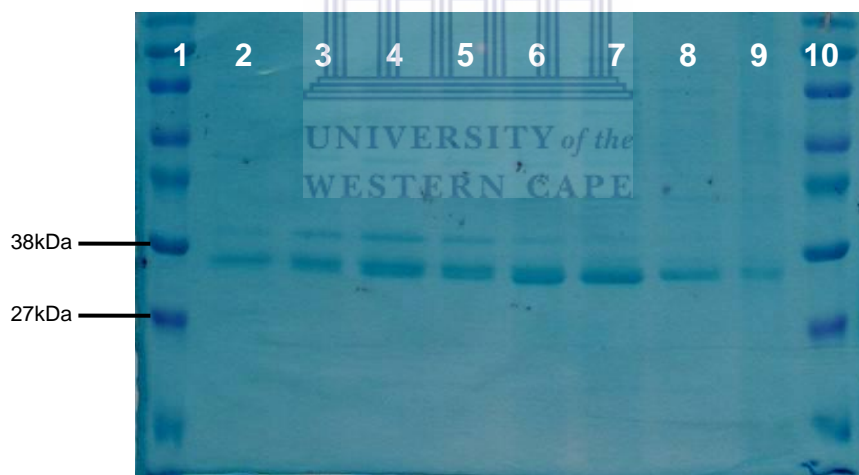
The Coomassie stained gel above depicts the optimal concentration of bacterially expressed ORF3ΔN-GST (~41kDa) at 12 hours post induction (lane 8). A rainbow SDS protein marker is seen in lanes 1 and 10, and the 2, 4, 6, 8 and 10 hour supernatant see in lanes 3 to 9.





**Figure 24: Coomassie stain showing an optimal concentration of expressed MΔN-GST fusion protein at 10 hours post induction.**

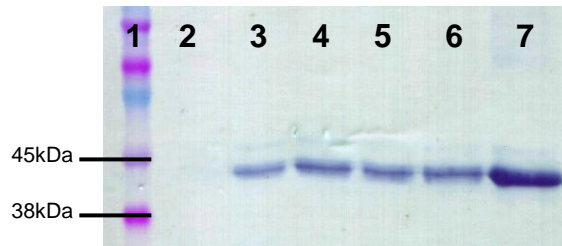
The Coomassie stained gel above displays the optimal concentration of bacterially expressed of MΔN-GST (~42kDa) at 10 hours post induction (lane 7). A rainbow SDS protein marker is seen in lanes 1 and 10, and the 2, 4, 6, 8 and 12 hour supernatant see in lanes 3, 4, 5, 6 and 8 respectively



**Figure 25: Coomassie stain depicting optimal expression of E-GST fusion protein 16 hours post induction.**

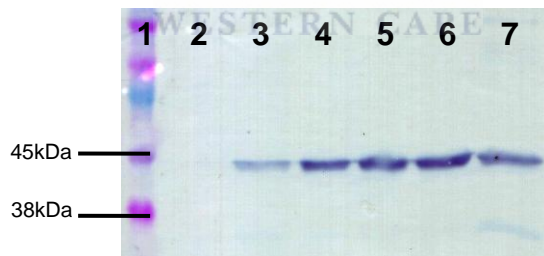
The Coomassie stained gel above depicts the optimal concentration of bacterially expressed of E-GST protein (~34kDa) at 16 hours post induction (lane 6). A rainbow SDS protein marker is seen in lanes 1 and 10. Lanes 2, 3, 4, 5, 7, 8 and 9 contain the supernatants of the 8, 10, 12, 14, 18, 24 and 36 hour samples, respectively.

Western blotting verified the presence of the fusion proteins with the use of GST specific antibodies (Figures 26, 27 and 28). The specific binding of the antibodies at the correct proteins validated the presence of the tagged proteins.



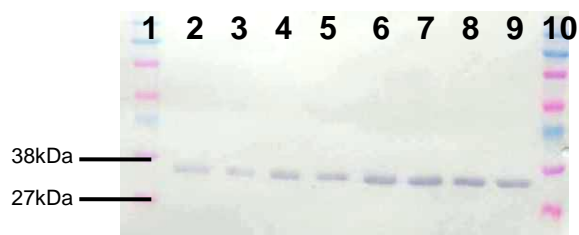
**Figure 26: Western Blot showing an optimal concentration of ORF3 $\Delta$ N-GST expressed fusion protein, 12 hours post induction**

Rainbow SDS protein marker is seen in lane 1. Lanes 3 to 7 displays the 4, 6, 8, 10 and 12 hour supernatant, showing an optimal concentration of expressed ORF3 $\Delta$ N-GST (~41kDa) at 12 hours (lane 7).



**Figure 27: Western Blot displaying an optimal concentration of M $\Delta$ N-GST expressed fusion protein, 10 hours post induction**

Rainbow SDS protein marker is seen in lane 1. Lanes 3 to 7 displays the 4, 6, 8, 10 and 12 hour supernatant, showing an optimal concentration of expressed M $\Delta$ N-GST (~42kDa) at 10 hours (lane 6).



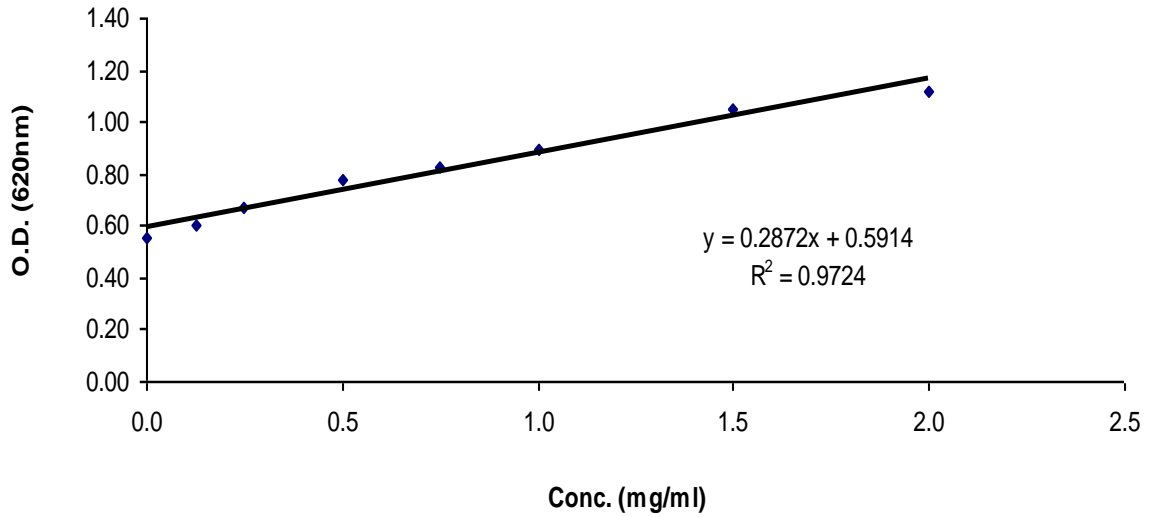
**Figure 28: Western Blot of E-GST expressed fusion proteins displaying an optimal protein concentration at 16 hours post induction.**

The rainbow SDS protein marker is seen in lanes 1 and 10. Lanes 2 to 9 consists of the 8, 10, 12, 14, 16, 18, 24 and 36 hour supernatant samples, respectively. Lane 6 shows the optimal concentration of the expressed E-GST fusion protein (~34kDa) at 16 hours post induction.

### ***Antigen Preparation***

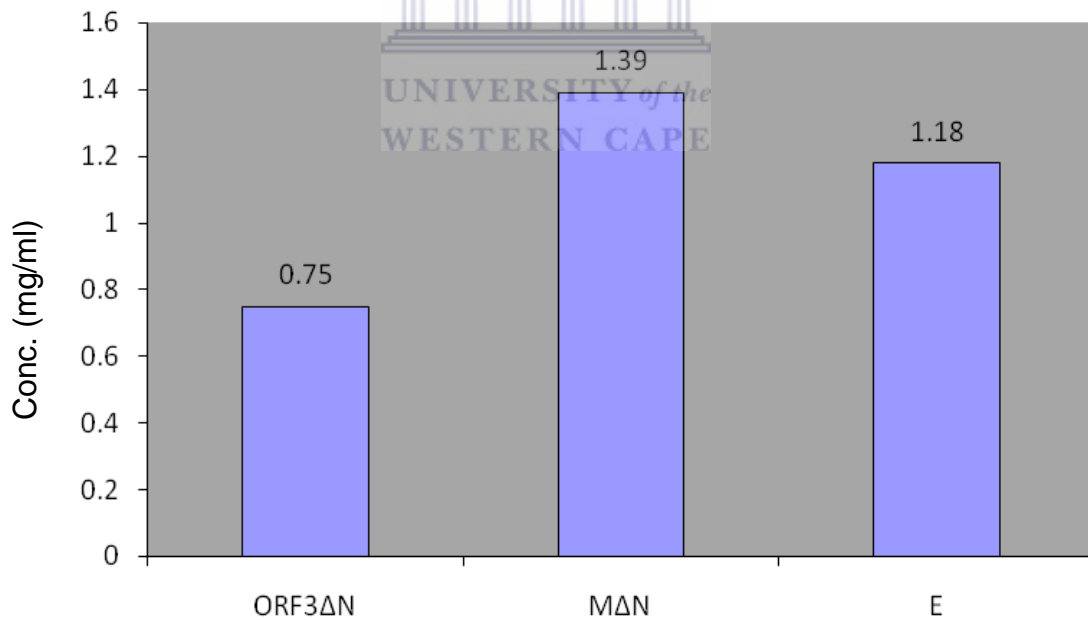
Once the proteins were isolated from the expression bacteria and filtered through the purification columns, they were quantified using a Bradford assay. Figure 29 indicates the standard curve as obtained from the Bradford Assay kit standards.

The specific protein concentrations of the GST-fusion proteins ORF3ΔN, MΔN and E (Figure 30) were obtained using the Bradford standard curve equation from Figure 29.



**Figure 29: Bradford Assay Standard Curve (620nm)**

This figure above indicates the optical densities (Y-axis) of the Bradford assay standards in relation to their concentration which ranges from 0mg/ml to 2mg/ml, as seen on the X-axis.



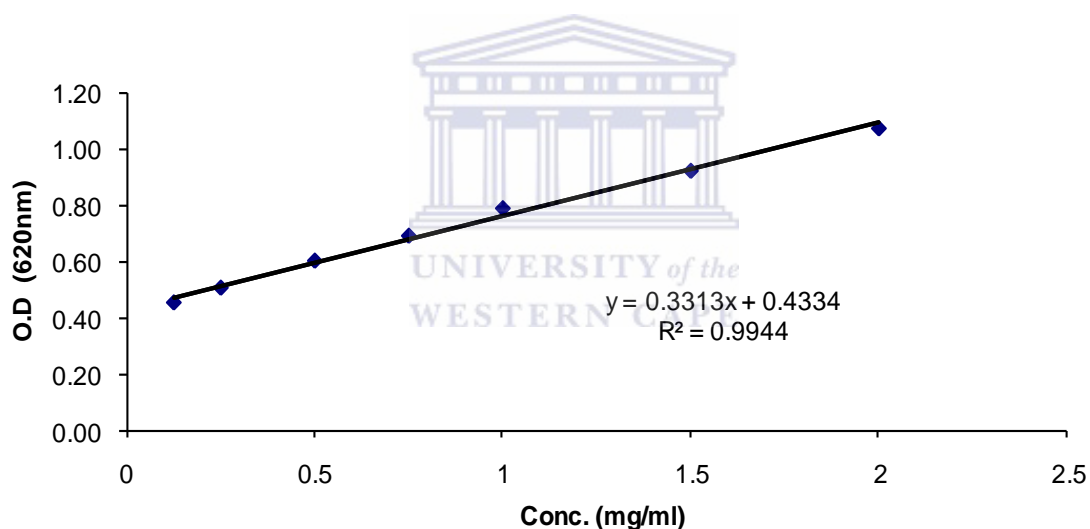
**Figure 30: Concentration of GST-Fusion viral protein post purification**

Figure 30 above depicts the concentrations of the GST-fusion viral proteins as calculated by substituting the optical densities of the purified protein samples into the equation obtained from the Bradford standard curve in Figure 29.

Results indicated that the antigens MΔN and E were at concentrations greater than 1mg/ml, which was optimal for the inoculation of the mice. The ORF3ΔN antigen was however, seen to be at a concentration of 0.75mg/ml and needed to be concentrated for optimal affectivity.

### **Concentration of Antigens**

GST-fusion proteins were concentrated by centrifuge-filtering through 5kD molecular weight cut off filters. The resultant samples were once again assayed to determine their protein concentrations as indicated in Figures 31 and 32 below.

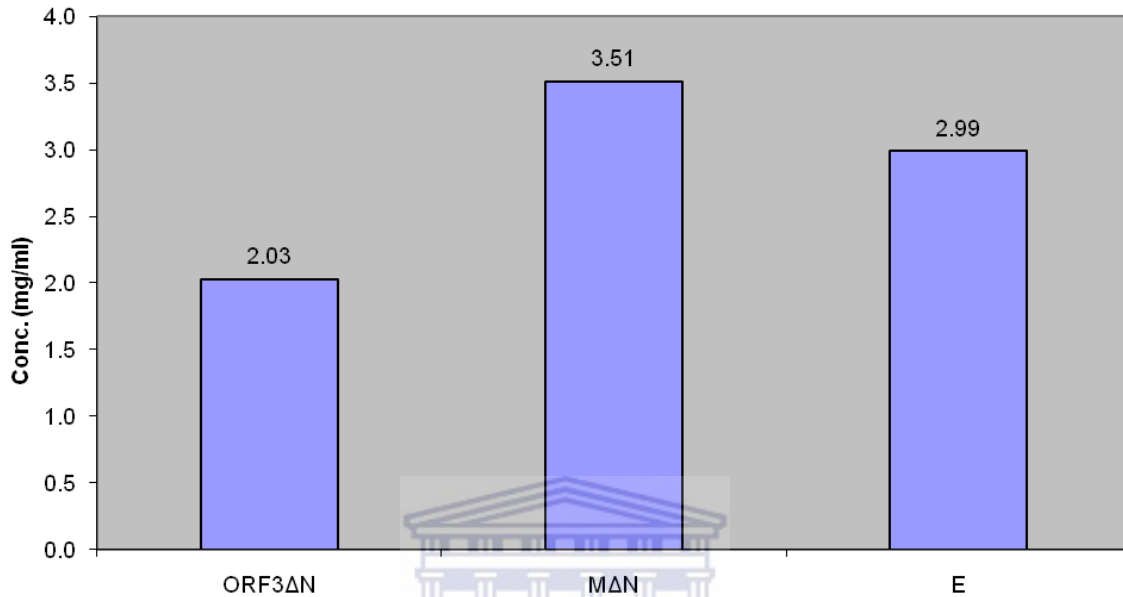


**Figure 31: Bradford Assay Standard Curve (620nm)**

Figure 31 indicates the optical densities (Y-axis) of the Bradford assay standards in relation to their concentration which ranges from 0mg/ml to 2mg/ml, as seen on the X-axis.

The optical densities of the concentrated GST-fusion proteins were substituted into the formula of the standard curve generated by the graph in Figure 31. The resultant concentration is seen in Figure 32 below. It was evident from the

increased optical densities that the proteins had a significantly higher concentration.



**Figure 32: Concentration of GST-Fusion Viral Protein Post Purification**

Figure 32 above depicts the concentrations of the GST-fusion viral proteins as calculated by substituting the optical densities of the concentrated purified protein samples into the equation obtained from the Bradford standard curve in Figure 31.

### ***Validation of Antibody Production***

The production of polyclonal antibodies begins with the immunization of the mice with the fusion viral proteins (produced in *E. coli*) suspended in Freund's complete adjuvant. The preferred route of inoculation in this case was subcutaneously (intrapertoneally) as opposed to intramuscularly as described by Hendriksen and Hau (2003). The function of the immune system is to eliminate non-self or foreign objects from the body (Kidd, 2003) and the exposure of host's immune system to

the viral fusion antigens sets off an immune response (von Behring and Kitasato, 1890).

The production of specific antibodies can be divided into 3 phases namely: the inductive phase, the effector phase, and the establishment of immunological memory (McCullough *et al.*, 1997). The innate immune system consists of Natural killer cells, mast cells, eosinophils, basophils and the phagocytic or antigen presenting cells (APC) which include macrophages, neutrophils and dendritic cells. These immune cells are the first to encounter the foreign antigen in the mouse's body and initiate the immune response (Randall *et al.*, 2008).

The antigen is internalized and processed by the APC's and displayed on the surface of the APC along with major histocompatibility complex molecules (MHC) which attract the antigen receptors on the T-cells (thymus-derived lymphocytes). The contact between the antigen, the MHC molecules and the T-cell activates the T-cells. This primary exposure, antigen processing and T-cell activation constitute the inductive phase.

When dealing with a protein antigen, the T-cell will employ the help of a T-helper (Th) cell. This action marks the beginning of the effector phase. The activated Th cells become sensitized to growth factors such as the cytokine interleukins 1 and 2 (IL1 and IL2) that are released by the distressed cells at the site of the immunization. These cytokines will eventually cause the B-lymphocytes to proliferate and differentiate into antibody producing plasma cells. Every single

plasma cell is genetically programmed to produce a specific antibody that will only recognize a single antigen epitope. These antibodies, also called immunoglobulins can be divided into 5 basic classes: IgA, IgD, IgE, IgG and IgM each with their own specific ranges of activity.

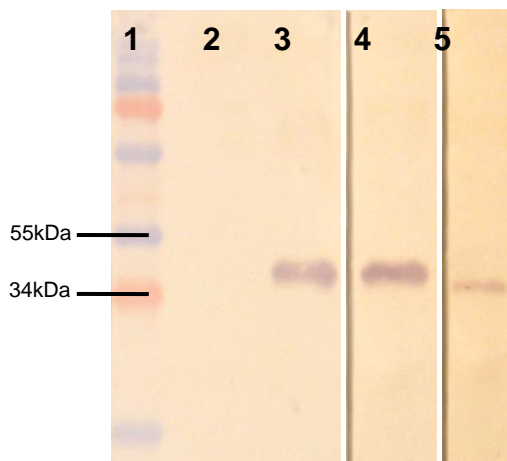
An antigen epitope usually comprises of between 5 and 10 amino acids. Hence, the viral-fusion tag antigens would consist of quite a number of antigen epitopes, and would require antibodies from more than one stimulated B-cell. Thus, the immunological response to such large antigens requires a combination of hundreds of monoclonal antibodies against the various epitopes.

The induction of memory phase relies on the prior exposure to antigens of pathogens (Ladics, 2007; Vos and Loveren, 1998). Immune system memory relies on the differentiation of the B-cells and makes the response to secondary exposures, more rapid. In the case of the bi-weekly booster immunizations, the IgG concentration would greatly outweigh that of IgM and the immunological response is more effective (McCullough *et al.* 1997).

After the second inoculation, the mice were bled to determine the presence of antibodies. Harvesting of antibodies is carried out by the collection of blood serum via the lateral tail bleed method as described by Loeb and Quimby in 1999. The polyclonal immunoglobulins found in the blood plasma are separated from the blood cells and diluted in phosphate buffered saline and used for Western Blotting



and ELISAs. The Western Blot seen in Figure 33 below verifies the presence of antibodies specific to GST-fusion viral proteins at a concentration of 1:500.



**Figure 33: Western Blot using 1:500 Mouse anti-GST-Fusion 1<sup>o</sup> Antibodies**

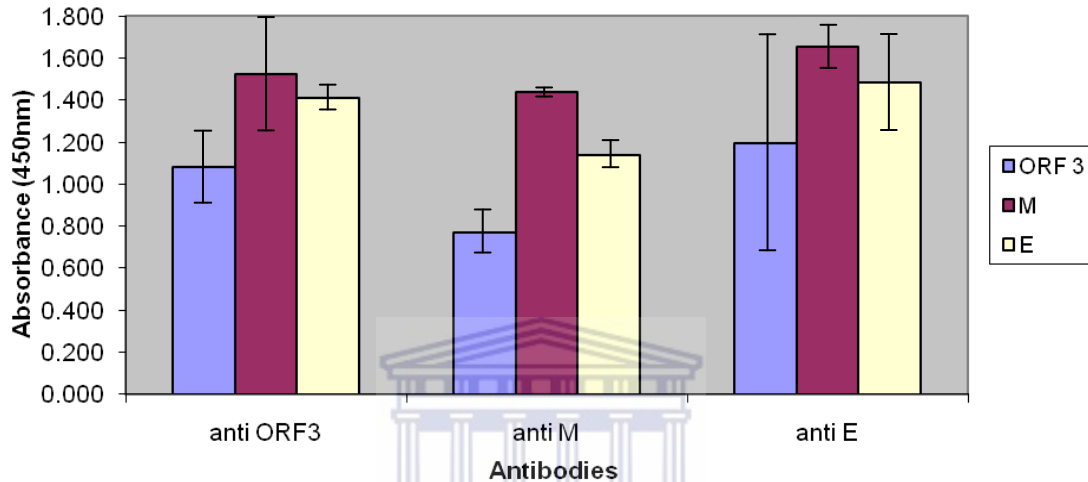
This figure depicts a positive Western Blot for the presence of the anti GST-fusion mouse antibodies produced as a result of the inoculations with the fusion antigens. Rainbow SDS page marker Lane 1; MΔN-GST (~42kDa.) Lane 3; ORF3ΔN-GST (~41kDa.) Lane 4; and E-GST (~34kDa) in Lane 5. (Negative control lanes not shown.)

The GST-tagged proteins isolated were seen to be of the expected molecular weight (GST~25kDa + ORF3~7kDa / M~11kDa or E ~9kDa) when in comparison to the protein marker (ProSieve). The activity of the produced antibodies was thus validated. Antibody specificity was yet to be determined.

### ***Viral Protein specific ELISA***

An ELISA was conducted to determine the specificity of the produced antibodies. Theoretically, each antibody has been produced with an affinity to both GST and the viral antigens. Hence this indirect ELISA was used to investigate GST affinity and viral protein specificity.

The results depicted in Figure 34 below suggest poor specificity of the antibodies to their corresponding viral antigens. It can thus be deduced that the high immunogenicity of the GST tag fused to the viral protein far outweighed that of the viral protein itself.



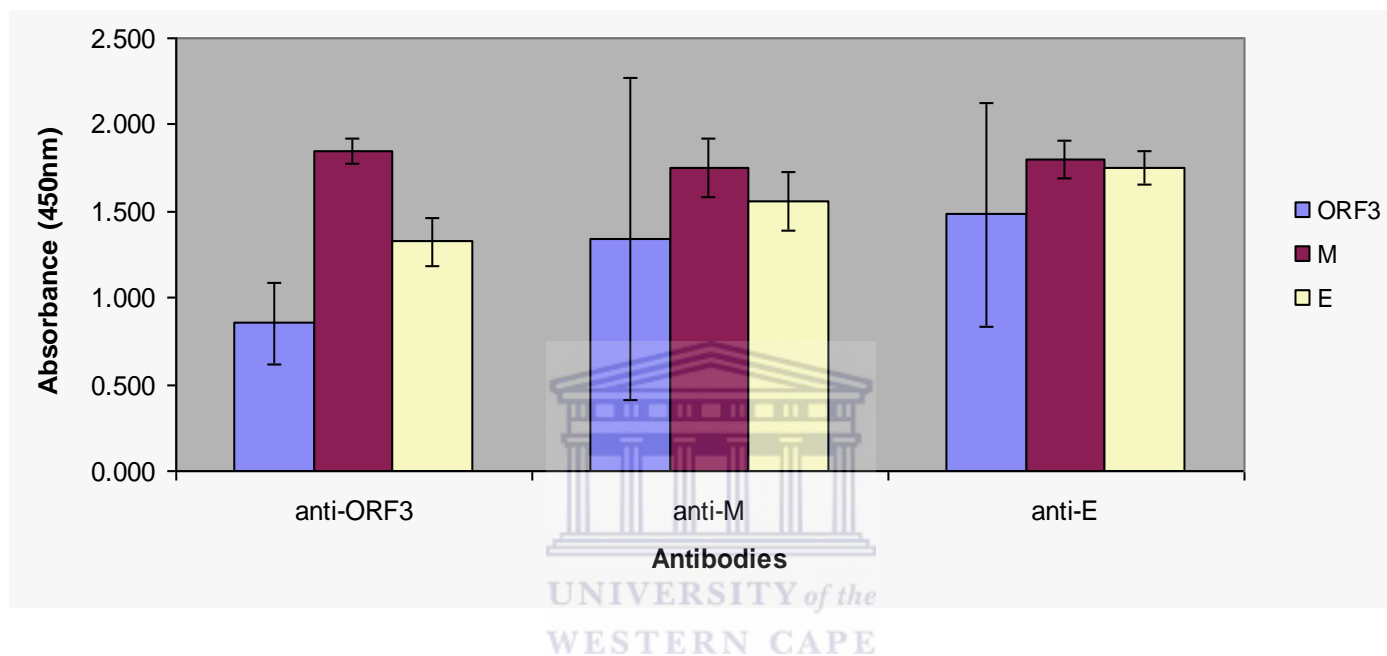
**Figure 34: Absorbance observed for the indirect ELISA**

The above graph indicates the average absorbance observed for the wells coated with the antigens GST-ORF3 (shaded blue), GST-M (shaded maroon) and GST-E (shaded yellow). No definitive antibody specificity is observed.

There are various scientifically proven solutions to counteract the great immunogenicity of the GST fusion tag. Cleavage of the fusion tag from the viral protein using a TEV (Tobacco Etch Virus (Promega)) protease would eliminate the strong GST affinity seen in the Western Blots and ELISAs and thus validate the specificity of the produced antibodies. Despite numerous attempts, we were unable to successfully perform this experiment. The indirect ELISA procedure was thus modified to render antibody binding more specific.

### ***Optimized, Indirect, Viral Protein Specific ELISA***

The previous ELISA was repeated with the addition of 1% BSA to all solutions in an attempt to reduce non-specific, GST binding. The graph below (Figure 35) demonstrates the results of this modification to standard procedure.



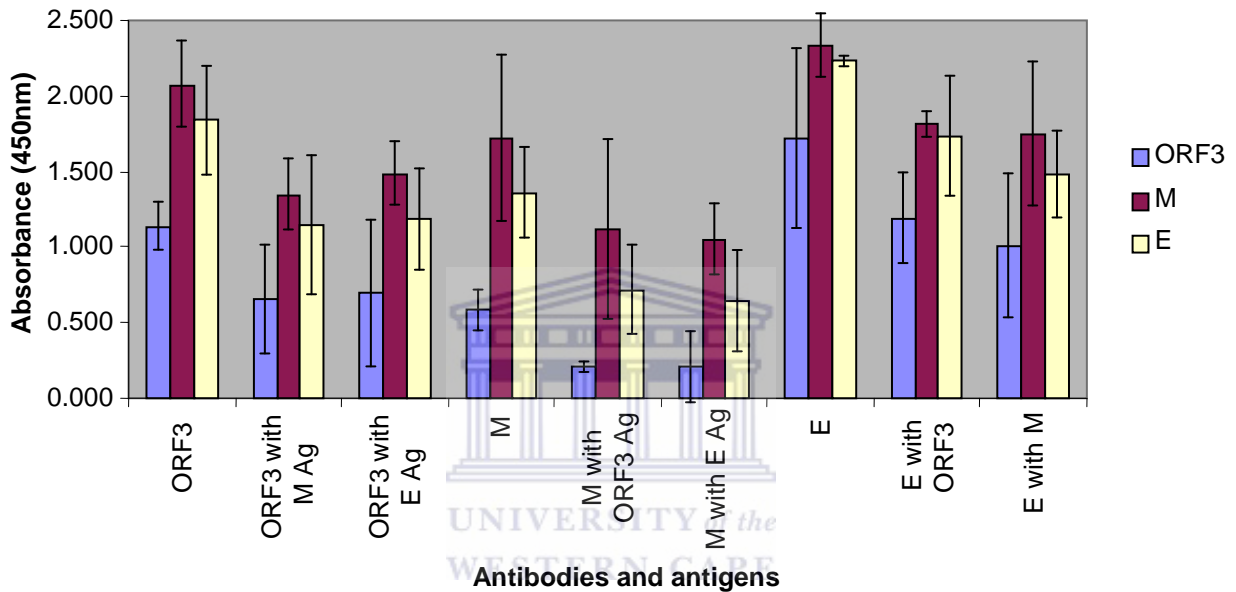
**Figure 35: Absorbance observed for the optimized, indirect ELISA**

Figure 35 above indicates the average absorbance observed for the wells coated with the antigens GST-ORF3 (shaded blue), GST-M (shaded maroon) and GST-E (shaded yellow). No significant antibody specificity was observed.

From the data obtained it is evident that the addition of BSA to the ELISA solutions had minimal effect on the non-specific binding to GST. Results still indicated minimal binding affinity to the targeted viral proteins. The ELISA procedure was once again modified to achieve the desired result.

### Optimized Indirect Viral Protein Specific Competition ELISA

Random, GST-tagged proteins were incubated with the primary antibodies to neutralize GST affinity. The remaining active antigen binding sites would thus be specific for the targeted viral proteins. The results of this ELISA, seen below (Figure 36), indicate no significant levels of antibody-viral protein specificity.



**Figure 36: Absorbance observed for the optimized, indirect, competition ELISA**

The above graph indicates the average absorbance observed for the wells coated with the antigens GST-ORF3 (shaded blue), GST-M (shaded maroon) and GST-E (shaded yellow) and competed with the antigens listed in the X-axis.

## **CHAPTER 4**

### **Conclusion**

Acute respiratory illnesses (ARIs) are a major health concern worldwide. Worldwide, ARIs affect all age groups and remain the leading cause of infant and young children mortality, accounting for approximately 2 million deaths annually (Kieny and Girard, 2005; Mizgerd, 2006). These deaths exert an economic burden on healthcare systems, especially in third world countries. Even though rhinoviruses, influenza viruses, parainfluenza viruses, respiratory syncytial viruses, adenoviruses and coronaviruses have been shown to be most frequently associated with respiratory infections, a significant proportion of these respiratory tract infections have no known cause. Recently, two new coronaviruses have been implicated in a number of these infections (van der Hoek, *et al.*, 2004; Woo *et al.*, 2005).

It has previously been reported that human coronaviruses account for a significant number of hospitalizations for children under 8 years of age, accounting for 4.4% of all admissions for acute respiratory infections. Interestingly, in a recent study of human coronavirus infections, the newly discovered human coronavirus NL-63 was the most common coronavirus identified (Chiu *et al.*, 2005). Initial clinical data indicate that HCoV-NL-63 causes acute respiratory disease in young infants and in immunocompromised adults and was detected in 1.3%-8.8% of respiratory

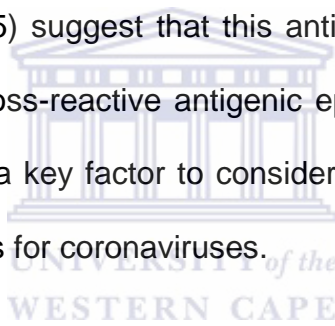
samples tested (Arden *et al.*, 2005; Bastien *et al.*, 2005; Chiu *et al.*, 2005; Ebihara *et al.*, 2005; Suzuki *et al.*, 2005; Vabret *et al.*, 2005; Chung *et al.*, 2007; Han *et al.*, 2007) and can also present as asthma exacerbation, febrile seizures and high fever (Chiu *et al.*, 2005).

NL-63 contains a single strand of positive sense, poly adenylated RNA strand, which resembles the host cell's mRNA. After the viral entry and uncoating, sections 1a and 1b on the 5' end of the RNA genome are directly translated into the viral enzymes required for viral progeny by the host cells cytoplasmic ribosomes. The self-encoded viral replicase enzymes (RNA dependant RNA Polymerase and viral protease) produced are deficient in mammalian cells and require a -1 ribosomal frame shift to express gene 1b, in frame (Herold *et al.* 1993; Giedroc *et al.* 2000). The action of these enzymes brings about the formation of new virions through the cytoplasmic production of viral proteins. The virus however, does not encode a proofreading viral enzyme and transcription and translational errors and point mutations are common with such huge, unstable RNA genomes.

Enzyme linked immunosorbent assays (ELISA) have previously been used to detect NL-63 specific antibodies (Dijkman *et al.*, 2008; Shao *et al.*, 2007). These N-specific ELISA assays were used to screen serum samples from children under the age of 20 (Shoa *et al.*, 2007). These ELISAs were able to detect maternally acquired N-directed antibodies (Dijkman *et al.*, 2008; Shao *et al.*, 2007), which normally decrease within the first 4-5 months of life (Shao *et al.*, 2007). Another study conducted by Hofmann *et al.* (2005), positively identified NL-63-S protein in

almost all samples isolated from children under the age of eight (Hofmann *et al.*, 2005). Combined, these findings suggest that NL-63 infections are commonly acquired during the early stages of life.

Serological cross-reactivity has been observed between SARS-CoV and other HCoV's including NL-63 (Chan *et al.*, 2005). This phenomenon has previously been observed specifically when using immunofluorescence and complement fixation assays (McIntosh *et al.*, 1974; Monto and Rhodes, 1977). Even ELISA's specific to recombinant N proteins of SARS-CoV detected false positives (Woo *et al.*, 2004). Chan *et al.* (2005) suggest that this antigenic cross-reactivity is as a result of the presence of cross-reactive antigenic epitopes of the coronaviruses. Antigenic cross-reactivity is a key factor to consider when attempting developing serological diagnostic assays for coronaviruses.



Here, we generated cDNA from HCoV-NL-63 RNA and used to amplify the E gene as well as the hydrophilic regions of ORF3 and M for expression in a bacterial system. These hydrophilic regions were chosen to increase the solubility of the fusion proteins expressed in the bacterial system for improved ease of purification. SDS-PAGE and Western-Blotting techniques, using anti-GST antibodies, were used to confirm the presence of the over expressed GST-fusion proteins, GST-M, GST-ORF3 and GST-E. The fusion proteins were soluble and high levels were detected in the lysate portion of our bacterial cultures.

These expressed fusion proteins served as the antigens, which were presented to the mouse immune system, in an attempt to the generate antibodies for use in

basic research, as well as for the screening of clinical samples for the presence of HCoV-NL-63.

The procedures for *in vivo* antibody production are both invasive and time consuming and require the use of immunopotentiating products (adjuvants) to enhance the immune response (Hendriksen and Hau 2003). These undesirable effects are however necessary for the production of polyclonal antibodies (Amyx, 1987) and result in animal experiments being tightly regulated by the universities animal handling and ethics committees.

In this study, gene amplification and primer specificity has been validated by the correct band sizes seen for both the ORF3 and M gene mutants and the native E gene in ethidium bromide gel electrophoresis. The genes were later verified by genomic sequencing and *in silico* analysis proved them to be homologous to the online genbank sequences. These validated genes were subsequently ligated into a bacterial expression vector and later verified by restriction enzyme digestion. Western Blot analysis of bacterially expressed GST-tagged viral-fusion proteins revealed high concentrations of tagged proteins at the expected molecular weights in the bacterial lysates.

Expressed viral-fusion proteins were purified by gravity filtration through GST-affinity agarose columns before being quantified by Bradford assay and concentrated for *in vivo* antibody production in mice. Two weeks after inoculation, mice were administered booster shots to increase blood-antibody titers.



Antibodies were harvested from the lateral tail vein and suspended in sterile PBS. *In vivo* antibody sensitivity was validated by Western Blotting before specificity determination was attempted using a modified indirect competition ELISA.

To date, the antibodies generated in this project are seen to be active and capable of detecting GST tagged viral proteins. Specificity of the antibodies is yet to be determined and will require further experimentation and resources. Once specificity has been validated, the ability of these antibodies to differentiate between viral species present in clinical samples may be hindered by the huge similarities observed between HCoV NL-63 and HCoV 229E.



---

## **REFERENCES**

**Abdul-Rasool, S., and B.C. Fielding.** 2010. Understanding Human Coronavirus HCoV-NL-63. *Open Virol. J.* **4**:76-84.

**Arden, K.E., M.D. Nissen, T.P. Sloots, and I.M. Mackay.** 2005. New human coronavirus, HCoV-NL-63, associated with severe lower respiratory tract disease in Australia. *J. Med. Virol.* **75**:455–462.

**Armstrong, J., H. Niemann, S. Smeekens, P. Rottier, and G. Warren.** 1984. Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. *Nat.* **308**:751–752.

**Amyx, H.L.** 1987. Control of animal pain and distress in antibody production and infectious disease studies. *J. Am. Vet. Med. Assoc.* **191**:1287-1289.

**Baric, R.S., G.W. Nelson, J.O. Fleming, R.J. Deans, J.G. Keck, N. Casteel, and S.A. Stohlman.** 1988. Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. *J. Virol.* **62**:4280-4287.

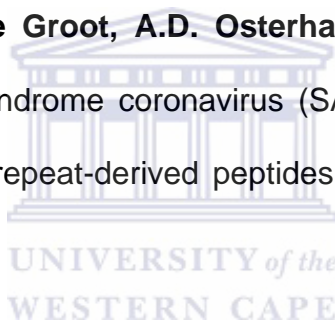
**Bastien, N., K. Anderson, L. Hart, P. Van Caesele, K. Brandt, D. Milley, T. Hatchette, E.C. Weiss, and Y. Li.** 2005a. Human coronavirus NL-63 infection in Canada. *J. Infect. Dis.* **191**:503–506.

---

**Bastien, N., J.L. Robinson, A. Tse, B.E. Lee, L. Hart, and Y. Li.** 2005b. Human coronavirus NL-63 infections in children: a 1-year study. *J. Clin. Microbiol.* **43**:4567-4573.

**Bond, C.W., J.L. Leibowitz, and J.A. Robb.** 1979. Pathogenic murine coronaviruses. II. Characterization of virus-specific proteins of murine coronaviruses JHMV and A59V. *Viol.* **94**:371–384.

**Bosch, B.J., B.E. Martina, R. Van Der Zee, J. Lepault, B.J. Haijema, C. Versluis, A.J. Heck, R. De Groot, A.D. Osterhaus, and P.J. Rottier.** 2004. Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. *Proc. Natl. Acad. Sci. USA* **101**:8455–8460.



**Bosch, B.J., R. van der Zee, C.A.M. de Haan, and P.J.M. Rottier.** 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J. Virol.* **77**:8801–8811.

**Bridges, C.B., M.J. Kuehnert, and C.B. Hall.** 2003. Transmission of Influenza: Implications for Control in Health Care Settings. *Clin. Infec. Dis.* **37**:1094–1101.

**Brockway, S.M., C.T. Clay, X.T. Lu, and M.R. Denison.** 2003. Characterization of the expression, intracellular localization, and replication complex association of

---

the putative mouse hepatitis virus RNA-dependent RNA polymerase. *J. Virol.* **77**:10515–10527.

**Bromham, L., and D. Penny.** 2003. The modern molecular clock. *Nat. Rev. Genet.* **4**:216-224.

**Canducci, F., M. Debiaggi, M. Sampaolo, M.C. Marinozzi, S. Berre, C. Terulla, G. Gargantini, P. Cambieri, E. Romero, and M. Clementi.** 2008. Two-year prospective study of single infections and co-infections by respiratory syncytial virus and viruses identified recently in infants with acute respiratory disease. *J. Med. Virol.* **80**:716-723

**Casais, R., M. Davies, D. Cavanagh, and P. Britton.** 2005. Gene 5 of the avian coronavirus infectious bronchitis virus is not essential for replication. *J. Virol.* **79**:8065-8078.

**Cavanagh, D., R. Casais, M. Armesto, T. Hodgson, S. Izadkhasti, M. Davies, F. Lin, I. Tarpey, and P. Britton.** 2007. Manipulation of the infectious bronchitis coronavirus genome for vaccine development and analysis of the accessory proteins. *Vac.* **25**:5558-5562.

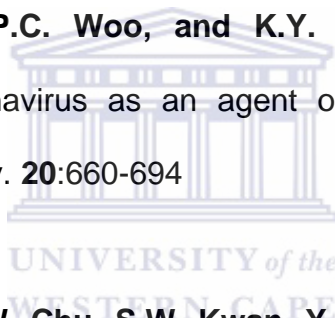
**Cavanagh, D.** 1995. The coronavirus surface glycoprotein. p. 73–113. In S.G. Siddell (ed.), *The Coronaviridae*. Plenum Press, New York, N.Y.

---

**Chan, K.H., V.C. Cheng, P.C. Woo, S.K. Lau, L.L. Poon, Y. Guan, W.H. Seto, K.Y. Yuen, and J.S. Peiris.** 2005. Serological responses in patients with severe acute respiratory syndrome coronavirus infection and cross-reactivity with human coronaviruses 229E, OC43, and NL-63. *Clin. Diagn. Lab. Immunol.* **12**:1317-1321.

**Chen, Z., Y. Wang, K. Ratia, A.D. Mesecar, K.D. Wilkinson, and S.C. Baker.** 2007. Proteolytic Processing and Deubiquitinating Activity of Papain-Like Proteases of Human Coronavirus NL-63. *J. Virol.* **81**:6007-6018.

**Cheng, V.C., S.K. Lau, P.C. Woo, and K.Y. Yuen.** 2007. Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. *Clin. Microbiol. Rev.* **20**:660-694



**Chiu, S.S., K.H. Chan, K.W. Chu, S.W. Kwan, Y. Guan, L.L. Poon, and J.S. Peiris.** 2005. Human coronavirus NL-63 infection and other coronavirus infections in children hospitalized with acute respiratory disease in Hong Kong, China. *Clin. Infect. Dis.* **40**:1721-1729.

**Choi, E.H., H.J. Lee, S.J. Kim, B.W. Eun, N.H. Kim, J.A. Lee, J.H. Lee, E.K. Song, S.H. Kim, J.Y. Park, and J.Y. Sung.** 2006. The association of newly identified respiratory viruses with lower respiratory tract infections in Korean children, 2000-2005. *Clin. Infect. Dis.* **43**:585-592.

---

**Chung, J.Y., T.H. Han, S.W. Kim, C.K. Kim, and E.S. Hwang.** 2007. Detection of viruses identified recently in children with acute wheezing. *J. Med. Virol.* **79**(8):1238-43.

**Clementz, M.A., Z. Chen, B.S. Banach, Y. Wang, L. Sun, K. Ratia, Y.M. Baez-Santos, J. Wang, J. Takayama, A.K. Ghosh, K. Li, A.D. Mesecar, and S.C. Baker.** 2010. Deubiquitinating and interferon antagonism activities of coronavirus papain-like proteases. *J. Virol.* **84**:4619-4629.

**Cotton, R.G.H. and C. Milstein.** 1973. Fusion of two immunoglobulin in-producing myeloma cells. *Nat.* **244**:42-43.

**Dare, R.K., A.M. Fry, M. Chittaganpitch, P. Sawanpanyalert, S.J. Olsen, and D.D. Erdman.** 2007. Human coronavirus infections in rural Thailand: a comprehensive study using real-time reverse-transcription polymerase chain reaction assays. *J. Infect. Dis.* **196**:1321-1328.

**de Haan, C.A., L. Kuo, P.S. Masters, H. Vennema, and P.J. Rottier.** 1998. Coronavirus particle assembly: primary structure requirements of the membrane protein. *J. Virol.* **72**(8):6838-6850.

**de Haan, C.A., P.S. Masters, X. Shen, S. Weiss, and P.J. Rottier.** 2002. The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. *Virol.* **296**:177-189.

---

**de Souza Luna, L.K., S. Baumgarte, K. Grywna, M. Panning, J.F. Drexler, and C. Drosten.** 2008. Identification of a contemporary human parechovirus type 1 by VIDISCA and characterisation of its full genome. *Viol. J.* **5**:26-36.

**Dijkman, R., M.F. Jebbink, B. Wilbrink, K. Pyrc, H.L. Zaaijer, P.D. Minor, S. Franklin, B. Berkhout, V. Thiel, and L. van der Hoek.** 2006. Human coronavirus 229E encodes a single ORF4 protein between the spike and the envelope genes. *Viol. J.* **3**:106.

**Dijkman, R., M.F. Jebbink, N.B. El Idrissi, K. Pyrc, M.A. Muller, T.W. Kuijpers, H.L. Zaaijer, and L. van der Hoek.** 2008. Human coronavirus NL-63 and 229E seroconversion in children. *J. Clin. Microbiol.* **46**:2368-2373.

**Donaldson, E., B. Yount, A.C. Sims, S. Burkett, R.J. Pickles, and R.S. Baric.** 2008. Systematic Assembly of a Full-Length Infectious Clone of Human Coronavirus NL-63. *J. Virol.* **82**(23):11948-11957.

**Donoghue, M., F. Hsieh, E. Baronas, K. Godbout, M. Gosselin, N. Stagliano, M. Donovan, B. Woolf, K. Robison, R. Jeyaseelan, R.E. Breitbart, and S. Acton.** 2000. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin 1 to angiotensin 1-9. *Circ. Res.* **87**(5):E1-9.

**Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H.R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R.A. Fouchier, A. Berger, A.M.**

---

**Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J.C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H.D. Klenk, A.D. Osterhaus, H. Schmitz, and H.W. Doerr.** 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N. Eng. J. Med.* **348**:1967-1976.

**Ebihara, T., R. Endo, X. Ma, N. Ishiguro, and H. Kikuta.** 2005. Detection of human coronavirus NL-63 in young children with bronchiolitis. *J. Med. Virol.* **75**:463-465.

**Esper, F., C. Weibel, D. Ferguson, M.L. Landry, J.S. Kahn.** 2006. Coronavirus HKU1 Infection in the United State. *Emerg. Infec. Dis.* **12**(4):775-779.

**Esper, F., E.D. Shapiro, C. Weibel, D. Ferguson, M. L. Landry, and J. S. Kahn.** 2005. Association between a novel human coronavirus and Kawasaki disease. *J. Infect. Dis.* **191**:499–502.

**Fendrick, A.M., A.S. Monto, B. Nightengale, and M. Sarnes.** 2003. The economic burden of non-influenza-related viral respiratory tract infection in the United States. *Arch. Intern. Med.* **163**(4):487–494.

**Fielding, B.C. and Y.J. Tan.** 2007. The Singapore contribution in the battle against SARS. *Iss. Infect. Dis.* **4**:1-22.

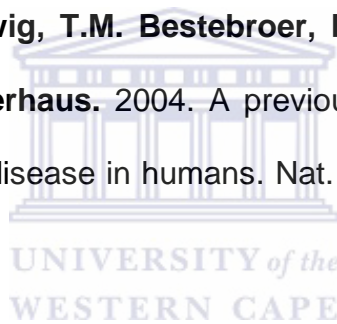


---

**Fielding, B.C. and T. Suliman.** 2009. Comparative analysis of human coronavirus-NL-63 ORF3 protein homologues. *Afri. J. Biotech.* **8**(14):3175-3178.

**Forster, J., G. Ihorst, C.H. Rieger, V. Stephan, H.D. Frank, H. Gurth, R. Berner, A. Rohwedder, H. Werchau, M. Schumacher, T. Tsai, and G. Petersen.** (2004) Prospective population-based study of viral lower respiratory tract infections in children under 3 years of age (the PRI.DE study). *Eur. J. Pediatr.* **163**:709-716

**Fouchier, R.A., N.G. Hartwig, T.M. Bestebroer, B. Niemever, J.C. de Jong, J.H. Simon, and A.D. Osterhaus.** 2004. A previously undescribed coronavirus associated with respiratory disease in humans. *Nat. Acad. Sci. U.S.A.* **101**:6212–6216.



**Fox, J.D.** 2007. Respiratory virus surveillance and outbreak investigation. *J. Clin. Virol.* **40**(1):S24-30.

**Garibaldi, R.A.** 1985. Epidemiology of community-acquired respiratory tract infections in adults. Incidence, etiology, and impact. *Am. J. Med.* **78**(6B):32-37.

**Giacalone, M.J., A.M. Gentile, B.T. Lovitt, N.L. Berkley, C.W. Gunderson, and M.W. Surber.** 2006. Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *Biotech.* **40**(3):355-364.

---

**Gibbs, A.J., M.J. Gibbs, and J.S. Armstrong.** 2004. The phylogeny of SARS coronavirus. *J. Arch. Virol.* **149**:621-624.

**Giedroc, D.P., C.A. Theimer, and P.L. Nixon.** 2000. Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J. Mol. Biol.* **298**:167-185.

**Godet, M., R. L'Haridon, J.F. Vautherot, and H. Laude.** 1992. TGEV corona virus ORF4 encodes a membrane protein that is incorporated into virions. *Virol.* **188**(2):666–675.

**Guan, Y., B.J. Zheng, Y.Q. He, X.L. Liu, Z.X. Zhuang, C.L. Cheung, S.W. Luo, P.H. Li, L.J. Zhang, Y.J. Guan, K.M. Butt, K.L. Wong, K.W. Chan, W. Lim, K.F. Shortridge, K.Y. Yuen, J.S. Peiris, and L.L. Poon.** 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Sci.* **302**:276–278.

**Haijema, B.J., H. Volders, and P.J. Rottier.** 2004. Live, attenuated coronavirus vaccines through the directed deletion of group-specific genes provide protection against feline infectious peritonitis. *J. Virol.* **78**:3863-3871.

**Hamming, I., W. Timens, M.L. Bulthuis, A.T. Lely, G.J. Navis, and H. van Goor.** 2004. Tissue distribution of ACE2 protein, the functional receptor for SARS

---

coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* **203**:631–637.

**Han, T.H., J.Y. Chung, S.W. Kim, and E.S. Hwang.** 2007. Human Coronavirus-NL-63 infections in Korean children, 2004–2006. *J Clin. Virol.* **38**:27-31.

**Harmer, D., M. Gilbert, R. Borman, and K.L. Clark.** 2002. Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme. *FEBS. Lett.* **532**(1-2):107-110.

**Hendriksen, C. and J. Hau.** (eds) 2003. Production of monoclonal and polyclonal antibodies. In: J. Hau and G. van Hoosier. *Handbook in Laboratory Animal Science. Vol I: Essential Principles and Practices.* Boca Raton: CRC Press. p391-412.

**Herold, J., and S.G. Siddell.** 1993. An 'elaborated' pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic. Acids. Res.* **21**:5838-5842.

**Ho, Y., P.H. Lin, C.Y. Liu, S.P. Lee, and Y.C. Chao.** 2004. Assembly of human severe acute respiratory syndrome coronavirus-like particles. *Biochem. Biophys. Res. Commun.* **318**(4):833-838.

---

**Hodgson, T., P. Britton, and D. Cavanagh.** 2006 Neither the RNA nor the proteins of open reading frames 3a and 3b of the coronavirus infectious bronchitis virus are essential for replication. *J. Virol.* **80**:296-305

**Hofmann, H., K. Pyrc, L. van der Hoek, M. Geier, B. Berkhout, and S. Pohlmann.** 2005. Human coronavirus NL-63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl. Acad. Sci. USA* **102**:7988–7993.

**Hopp, T.P., and K.R. Woods.** 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. USA* **78**:3824-3828.

**Hsueh, P.R., and P.C. Yang.** 2003. Severe acute respiratory syndrome (SARS) - an emerging infection of the 21st century. *J. Formos. Med. Assoc.* **102**(12):825-839.

<http://www3.niaid.nih.gov/healthscience/healthtopics/colds/>.

<http://www.vizier-europe.org/article29.html?artsuite=0>

**Jia, W., K. Karaca, C.R. Parrish, and S.A. Naqi.** 1995. A novel variant of avian infectious bronchitis virus resulting from recombination among three different strains. *Arch. Virol.* **140**:259–271.

---

**Kahn, J.S.** 2006. The widening scope of coronaviruses. *Curr. Opin. Pediatr.* **18**:42-47.

**Kaiser, L., N. Regamey, H. Roiha, C. Deffernez, and U. Frey.** 2005. Human coronavirus NL-63 associated with lower respiratory tract symptoms in early life. *Pediatr. Infect. Dis J.* **24**:1015-1017.

**Keng, C.T., A. Zhang, S. Shen, K.M. Lip, B.C. Fielding, T.H. Tan, C.F. Chou, C.B. Loh, S. Wang, J. Fu, X. Yang, S.G. Lim, W. Hong, and Y.J. Tan.** 2005. Amino acids 1055 to 1192 in the S2 region of severe acute respiratory syndrome coronavirus S protein induce neutralizing antibodies: implications for the development of vaccines and antiviral agents. *J. Virol.* **79**:3289–3296.

**Kidd, P.** 2003. Th1/Th2 balance: The hypothesis, its limitations and its implications for health and disease. *Altern. Med. Rev.* **8**(3):223-246.

**Kieny, M.P. and M.P. Girard.** 2005. Human vaccine research and development: an overview. *Vac.* **23**(50):5705-7.

**Kistler, A., P.C. Avila, S. Rouskin, D. Wang, T. Ward, S. Yagi, D. Schnurr, D. Ganem, J.L. DeRisi, and H.A. Boushey.** 2007. Pan-viral screening of respiratory tract infections in adults with and without asthma reveals unexpected human coronavirus and human rhinovirus diversity. *J. Infect. Dis.* **196**:817-825

---

**Kyte, J., and R.F. Doolittle.** 1982. A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**:105-132.

**Koetz, A., P. Nilsson, M. Linden, L. van der Hoek, and T. Ripa.** 2006. Detection of human coronavirus NL-63, human metapneumovirus and respiratory syncytial virus in children with respiratory tract infections in south-west Sweden. *Clin. Microbiol. Infect.* **12**:1089-1096.

**Kohler, G. and C. Milstein.** 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nat.* **256**:495-497.

**Kondo, N., A. Ebihara, H. Ru, S. Kuramitsu, A. Iwamoto, Z. Rao, and Z. Matsuda.** 2008. *Thermus thermophilus*-derived protein tags that aid in preparation of insoluble viral proteins. *Analyt. Biochem.* **385**: 278-285.

**Konings, D.A.M., P.J. Bredenbeek, J.F.H. Noten, P. Hogeweg, W.J.M. Spaan.** 1988. Differential premature termination of transcription as a proposed mechanism for the regulation of coronavirus gene expression. *Nuc. Acid. Res.* **16**(22):10849-10860.

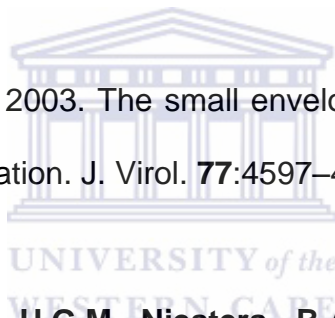
**Kuba, K., Y. Imai, S. Rao, H. Gao, F. Guo, B. Guan, Y. Huan, P. Yang, Y. Zhang, W. Deng, L. Bao, B. Zhang, G. Liu, Z. Wang, M. Chappell, Y. Liu, D. Zheng, A. Leibbrandt, T. Wada, A.S. Slutsky, D. Liu, C. Qin, C. Jiang, and J.**

---

**M. Penninger.** 2005. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nat. Med.* **11**:875–879.

**Kuiken, T., R.A. Fouchier, M. Schutten, G.F. Rimmelzwaan, G. van Amerongen, D. van Riel, J.D. Laman, T. de Jong, G. van Doornum, W. Lim, A.E. Ling, P.K. Chan, J.S. Tam, M.C. Zambon, R. Gopal, C. Drosten, S. van der Werf, N. Escriou, J.C. Manuguerra, K. Stohr, J.S. Peiris, and A.D. Osterhaus.** 2003. Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome. *Lancet.* **362**:263-270.

**Kuo, L., and P.S. Masters.** 2003. The small envelope protein E is not essential for murine coronavirus replication. *J. Virol.* **77**:4597–4608.



**Kusters, J.G., E.J. Jager, H.G.M. Niesters, B.A.M. van der Zeijst.** 1990. Sequence evidence for RNA recombination in field isolates of avian coronavirus infectious bronchitis virus. *Vac.* **8**:605-608.

**Ladics, G.S.** 2007. Use of SRBC antibody responses for immunotoxicity testing. *Meth.* **41**:9-19.

**Lambert, S.B., K.M. Allen, J.D. Druce, C.J. Birch, I.M. Mackay, J.B. Carlin, J.R. Carapetis, T.P. Sloots, M.D. Nissen, and T.M. Nolan.** 2007. Community epidemiology of human metapneumovirus, human coronavirus NL-63, and other

---

respiratory viruses in healthy preschool-aged children using parent-collected specimens. *Pediat.* **120**:e929-937.

**Li, F., W. Li, M. Farzan, and S.C. Harrison.** 2005. Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor.” *Sci.* **309**:1864-1868.

**Li, W., M.J. Moore, N. Vasilieva, J. Sui, S.K. Wong, M.A. Berne, M. Somasundaran, J.L. Sullivan, K. Luzuriaga, T.C. Greenough, H. Choe, and M. Farzan.** 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nat.* **426**:450–454.

**Lim, K.P. and D.X. Liu.** 2001. The missing link in coronavirus assembly. Retention of the avian coronavirus infectious bronchitis virus envelope protein in the pre-Golgi compartments and physical interaction between the envelope and membrane proteins. *J. Bio. Chem.* **276**:17515–17523.

**Loeb, W., and F. Quimby.** (eds.). 1999. *The Clinical Chemistry of Laboratory Animals*, Second Edition. Philadelphia, PA: Taylor and Francis.

**Maeda, J., J.F. Repass, A. Maeda, and S. Makino.** 2001. Membrane topology of coronavirus E protein. *Viol.* **281**:163–169.



---

**Mathewson, A.C., A. Bishop, Y. Yao, F. Kemp, J. Ren, H. Chen, X. Xu, B. Berkhout, L. van der Hoek, and I.M. Jones.** 2008. Interaction of severe acute respiratory syndromecoronavirus and NL-63 coronavirus spike proteins with angiotensin converting enzyme-2. *J. Gen. Virol.* **89**:2741–2745.

**McCullough, K.C., C.F.M. Hendriksen, and T. Seebeck.** 1997. *In vitro* methods in Vaccinology. In: Pastoret P-P, Blancou J, Vannier P, Verschueren C, eds. *Veterinary Vaccinology*. Amsterdam: Elsevier. p 69.

**McIntosh, K.** 2005. Coronaviruses in the Limelight. *J. Infect. Dis.* **191**:489–491.

**McIntosh, K., R.K. Chao, H.E. Krause, R. Wasil, H.E. Mocega, and M.A. Mufson.** 1974. Coronavirus infection in acute lower respiratory tract disease of infants. *J. Infect. Dis.* **130**:502-507.

**Minosse, C., M. Selleri, M.S. Zaniratti, G. Cappiello, A. Spano, E. Schifano, F.N. Lauria, G. Gualano, V. Puro, G. Campanini, G. Gerna, and M.R. Capobianchi.** 2008. Phylogenetic analysis of human coronavirus NL-63 circulating in Italy. *J. Clin. Virol.* **43**:114-119

**Mizgerd, J.P.** 2006. Lung infection--a public health priority. *PLoS Med.* **3**(2):e76.

**Moës, E., L. Vijgen, E. Keyaerts, K. Zlateva, S. Li, P. Maes, K. Pyrc, B. Berkhout, L. van der Hoek, and M. Van Ranst.** 2005. A novel pancoronavirus

---

RT-PCR assay: frequent detection of human coronavirus NL-63 in children hospitalized with respiratory tract infections in Belgium. *BMC. Infect. Dis.* **5**:6.

**Monto, A.S., and L.M. Rhodes.** 1977. Detection of coronavirus infection of man by immunofluorescence. *Proc. Soc. Exp. Biol. Med.* **155**:143-148.

**Muller, A., R.L. Tillmann, A. Muller, A. Simon, and O. Schildgen.** 2008. Stability of human metapneumovirus and human coronavirus NL-63 on medical instruments and in the patient environment. *J. Hosp. Infect.* **69**:406-408.

**Müller, M.A., L. van der Hoek, D. Voss, O. Bader, D. Lehmann, A.R. Schulz, S. Kallies, T. Suliman, B.C. Fielding, C. Drosten, and M. Niedrig.** 2010. Human Coronavirus NL-63 Open Reading Frame 3 encodes a virion-incorporated N-glycosylated membrane protein. *Viol. J.* **6**:7.

**Nal, B., C. Chan, F. Kien, L. Siu, J. Tse, K. Chu, J. Kam, I. Staropoli, B. Crescenzo-Chaigne, N. Escriou, S. van der Werf, K.Y. Yuen, and R. Altemeyer.** 2005. Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. *J. Gen. Virol.* **86**(Pt 5):1423-34.

**Namy, O., S.J. Moran, D.I. Stuart, R.J. Gilbert, and I. Brierley.** 2006. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nat.* **441**:244–247.

---

**Narayanan, K., A. Maeda, J. Maeda, and S. Makino.** 2000. Characterization of the coronavirus M protein and nucleocapsid interaction in infected cells. *J. Virol.* **74**:8127–8134.

**Narayanan, K., and S. Makino.** 2001. Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA packaging. *J. Virol.* **75**:9059–9067.

**Narayanan, K., C.J. Chen, J. Maeda, and S. Makino.** 2003. Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal. *J. Virol.* **77**:2922–2927.

**Narayanan, K., C. Huang, and S. Makino.** 2008. SARS coronavirus accessory proteins. *Vir. Res.* **133**:113-121.

**Narayanan, K., K.H. Kim, and S. Makino.** 2003. Characterization of N protein self-association in coronavirus ribonucleoprotein complexes. *Virus. Res.* **98**:131–140.

**Netline, J., D. Ferraro, L. Pewe, H. Olivares, T. Gallagher, and S. Perlman.** 2007. Enhancement of murine coronavirus replication by severe acute respiratory syndrome coronavirus protein 6 requires the N-terminal hydrophobic region but not the C-terminal sorting motifs. *J. Virol.* **81**:11520-11525.

---

**Opstelten, D.J., M.J. Raamsman, K. Wolfs, M.C. Horzinek, and P.J. Rottier.** 1995. Coexpression and association of the spike protein and the membrane protein of mouse hepatitis virus. *Adv. Exp. Med. Biol.* **380**:291-297.

**Parker, K.C., M.A. Bednarek, and J.E. Coligan.** 1994. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J. Immunol.* **152**:163-175.

**Parker, M.M. and P.S. Masters.** 1990. Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. *Viol.* **179**(1):463-468.

**Pearson, T., G. Galfri, A. Ziegler, and C. Milstein.** 1977. A myeloma hybrid producing antibody specific for an allotypic determinant on "IgD-like" molecules of the mouse. *Eur. J. Immunol.* **7**:684-690.

**Pewe, L., H. Zhou, J. Netland, C. Tangadu, H. Olivares, L. Shi, D. Look, T. Gallagher, and S. Perlman.** 2006 A SARS-CoV-specific protein enhances virulence of an attenuated strain of mouse hepatitis virus. *Adv. Exp. Med. Biol.* **581**:493-498

**Plant, E.P. and J.D. Dinman.** 2008. The role of programmed-1 ribosomal frameshifting in coronavirus Propagation. *Frnt. Biosci.* **13**:4873–4881.

---

**Pyrce, K., B. Berkhout, and L. der Hoek.** 2007a. Antiviral strategies against human coronaviruses. *Infect. Disord. Drug. Targ.* **7**:59-66.

**Pyrce, K., B. Berkhout, and L. van der Hoek.** 2007b. Identification of new human coronaviruses. *Expert. Rev. Anti. Infect. Ther.* **5**:245-253.

**Pyrce, K., B. Berkhout, and L. van der Hoek.** 2007c. The novel human coronaviruses NL-63 and HKU1. *J. Virol.* **81**:3051-3057.

**Pyrce, K., B.J. Bosch, B. Berkhout, M.F. Jebbink, R. Dijkman, P. Rottier, and L. van der Hoek.** 2006a. Inhibition of Human Coronavirus NL-63 Infection at Early Stages of the Replication Cycle. *Antimic. Agnts. Chemo.* **50**(6):2000-2008.

**Pyrce, K., M.F. Jebbink, B. Berkhout, and L. van der Hoek.** 2004. Genome structure and transcriptional regulation of human coronavirus NL-63. *Virol. J.* **1**: 7.

**Pyrce, K., M.F. Jebbink, B. Berkhout, and L. van der Hoek.** 2008. Detection of new viruses by VIDISCA. Virus discovery based on cDNA-amplified fragment length polymorphism. *Meth. Mol. Biol.* **454**:73-89.

**Pyrce, K., R. Dijkman, L. Deng, M.F. Jebbink, H. A. Ross, B. Berkhout, and L. van der Hoek.** 2006b. Mosaic structure of human coronavirus NL-63, one thousand years of evolution. *J. Mol. Biol.* **364**:964–973.

---

**Randall, L.M., F.H. Amante, Y. Zhou, A.C. Stanley, A. Haque, F. Rivera, K. Pfeffer, S. Scheu, G.R. Hill, K. Tamada, and C.R. Engwerda.** 2008. Cutting edge: Selective blockade of LIGHT-Lymphotoxin  $\beta$  receptor signaling protects mice from experimental cerebral malaria caused by *Plasmodium berghei* ANKA. *J. Immuno.* **181**(11):7458-7462.

**Rota, P.A., M.S. Oberste, S.S. Monroe, W.A. Nix, R. Campagnoli, J.P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M.H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J.L. DeRisi, Q. Chen, D. Wang, D.D. Erdman, T.C. Peret, C. Burns, T.G. Ksiazek, P.E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A.D. Osterhaus, C. Drosten, M.A. Pallansch, L.J. Anderson, and W.J. Bellini.** 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *J. Sci.* **300**:1394-1399.

**Rottier, P.J., G.W. Welling, S. Welling-Wester, H.G. Niesters, J.A. Lenstra, and B.A.M. van der Zeijst.** 1986. Predicted membrane topology of the coronavirus protein E1. *Biochem.* **25**:1335–1339.

**Rottier, P.J., J.K. Rose.** 1987. Coronavirus E1 glycoprotein expressed from cloned cDNA localizes in the Golgi region. *J. Virol.* **61**:2042–2045.

---

**Sanchez, C.M., F. Gebauer, C. Sune, A. Mendez, J. Dopazo, and L. Enjuanes.** 1992. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. *Viol.* **190**:92-105.

**Sawicki, S.G., D. Sawicki, and S.G. Siddell.** 2007. A contemporary view of coronavirus transcription. *J. Virol.* **81**(1):20-9.

**Schwaber, J. and E.P. Cohen.** 1974. Pattern of immunoglobulin synthesis and assembly in a human-mouse somatic cell hybrid clone. *Proc. Natl. Acad. Sci. USA.* **71**: 2203-2207.

**Shimizu, C., H. Shike, S.C. Baker, F. Garcia, L. van der Hoek, T.W. Kuijpers, S.L. Reed, A.H. Rowley, S.T. Shulman, H.K. Talbot, J.V. Williams, and J.C. Burns.** 2005. Human coronavirus NL-63 is not detected in the respiratory tracts of children with acute Kawasaki disease. *J. Infect. Dis.* **192**:1767-1771.

**Shao, X., X. Guo, F. Esper, C. Weibel, and J.S. Kahn.** 2007. Seroepidemiology of group I human coronaviruses in children. *J. Clin. Virol.* **40**:207-213.

**Siddell, S., H. Wege, and V. Ter Meulen.** 1983. The biology of coronaviruses. *J. Gen. Virol.* **64**, 761-776.

---

**Sloots, T.P., P. McErlean, D.J. Speicher, K.E. Arden, M.D. Nissen, and I.M. Mackay.** 2006. Evidence of human coronavirus HKU1 and human bocavirus in Australian children. *J. Clin. Virol.* 35:99-102.

**Smuts, H., and D. Hardie.** 2006. Human bocavirus in hospitalized children, South Africa. *Emerg. Infect. Dis.* 12:1457-1458

**Smuts, H., L. Workman, and H.J. Zar.** 2008. Role of human metapneumovirus, human coronavirus NL-63 and human bocavirus in infants and young children with acute wheezing. *J. Med. Virol.* 80:906-912.

**Snijder, E.J., P.J. Bredenbeek, J.C. Dobbe, V. Thiel, J. Ziebuhr, L.L. Poon, Y. Guan, M. Rozanov, W.J. Spaan, and A.E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Bio.* 331:991-1004.

**Spaan, W., D. Cavanagh, and M.C. Horzinek.** 1988. Coronaviruses: Structure and Genome Expression. *J. Gen. Virol.* 69:2939-2952.

**Stohlman, S.A., R.S. Baric, G.N. Nelson, L.H. Soe, L.M. Welte, and R.J. Deans.** 1988. Specific Interaction between Coronavirus Leader RNA and Nucleocapsid Protein. *J. Virol.* 62(11):4288-4295.



---

**Suzuki, A., M. Okamoto, A. Ohmi, O. Watanabe, S. Miyabayashi, and H. Nishimura.** 2005. Detection of human coronavirus-NL-63 in children in Japan. *Pediatr. Infect. Dis. J.* **24**:645–646.

**Tang, J.W., and R.C.W. Chan.** 2004. Severe acute respiratory syndrome (SARS) in intensive care units (ICUs): limiting the risk to healthcare workers. *Cur. Anaest. Crit. Care.* **15**:143-155.

**Tabor, S., and C.C. Richardson.** 1985. A bacteriophage T7 RNAPolymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl. Acad. Sci. USA.* **82**(4):1074 – 1078.

**Thiel, V., J. Herold, B. Schell, and S.G. Sidell.** 2001. Infectious RNA transcribed *in vitro* from cDNA copy of the human coronavirus genome cloned in vaccinia virus. *J. Gen. Virol.* **82**:1273-1281.

**Tyrrell, D.A., and M.L. Bynoe.** 1965. Cultivation of a novel type of common-cold virus in organ cultures. *Br. Med. J.* **5448**:1467–1470.

**Vabret, A., J. Dina, E. Brison, J. Brouard, and F. Freymuth.** 2008. [Human coronaviruses.]. *Pathol. Biol.* **2**:2.

---

**Vabret, A., T. Mourez, J. Dina, L. van der Hoek, S. Gouarin, J. Petitjean, J. Brouard, and F. Freymuth.** 2005. Human Coronavirus NL-63, France. *Emerg. Infect. Dis.* **11**:1225-1229.

**van der Hoek, L., K. Pyrc, M.F. Jebbink, W. Vermeulen-Oost, R.J. Berkhout, and K.C. Wolthers.** 2004. Identification of a new human coronavirus. *J. Nat. Med.* **10**:368–373.

**van der Hoek, L., K. Sure, G. Ihorst, A. Stang, K. Pyrc, M.F. Jebbink, G. Petersen, J. Forster, B. Berkhout, and K. Uberla.** 2005. Croup is associated with the novel coronavirus NL-63. *PloS. Med.* **2**:e240.

**van der Hoek, L., K. Pyrc, and B. Berkhout.** 2006. Human Coronavirus NL-63, a new respiratory virus. *FEMS Microbiol. Rev.* **30**:760-773.

**Vijgen, L., P. Lemey, E. Keyaerts, and M. Van Ranst.** 2005. Genetic variability of human respiratory coronavirus OC43. *J. Virol.* **79**:3223-3225.

**Vos, J.G., and H. van Loveren.** 1998. Experimental studies on immunosuppression: How do they predict for man? *Toxicol.* **129**:13-26.

**Webster, R.G.** 2004. Wet markets: continuing source of severe acute respiratory syndrome and influenza? *Lancet.* **363**(9404):234-236.

---

**Weiss, S.R., and S.N. Martin.** 2005. Coronavirus Pathogenesis and the Emerging Pathogen Severe Acute Respiratory Syndrome Coronavirus. *Micro. Mol. Bio. Rev.* **69**(4):635–664.

**Welling, G.W., W.J. Weijer, R. van der Zee, and S. Welling-Wester.** 1985. Prediction of sequential antigenic regions in proteins. *FEBS Lett.* **188**(2):215-218.

**W.H.O.** 2010. Report 1: Clinical research on treatment of SARS with integrated Traditional Chinese medicine and Western Medicine. Viewed online at <http://apps.who.int/medicinedocs/en/d/Js6170e/4.html> on March 4, 2010, 11:45 am.

**Wilson, L., C. McKinlay, P. Gage, and G. Ewart.** 2004. SARS coronavirus E protein forms cation-selective ion channels. *Virology*. **330**:322–331.

**Woo, P.C., S.K. Lau, B.H. Wong, K.H. Chan, W.T. Hui, G.S. Kwan, J.S. Peiris, R.B. Couch, and K.Y. Yuen.** 2004. False-positive results in a recombinant severe acute respiratory syndrome-associated coronavirus (SARS-CoV) nucleocapsid enzyme-linked immunosorbent assay due to HCoV-OC43 and HCoV-229E rectified by Western blotting with recombinant SARS-CoV spike polypeptide. *J. Clin. Microbiol.* **42**:5885–5888.

**Woo, P.C., S.K. Lau, C.M. Chu, K.H. Chan, H.W. Tsoi, Y. Huang, B.H. Wong, R.W. Poon, J.J. Cai, W.K. Luk, L.L. Poon, S.S. Wong, Y. Guan, J.S. Peiris,**

---

**and K.Y. Yuen.** 2005a. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* **79**:884–895.

**Woo, P.C., S.K. Lau, H.W. Tsoi, Y. Huang, R.W. Poon, C.M. Chu, R.A. Lee, W.K. Luk, G.K. Wong, B.H. Wong, V.C. Cheng, B.S. Tang, A.K. Wu, R.W. Yung, H. Chen, Y. Guan, K.H. Chan, and K.Y. Yuen.** 2005b. Clinical and molecular epidemiological features of coronavirus HKU1-associated community-acquired pneumonia. *J. Infect. Dis.* **192**:1898-1907.

**Woode, G.N.** 1994. The toroviruses: bovine (Breda virus) and equine (Berne virus) and the torovirus-like agents of humans and animals, p. 581–602. *In* A. Z. Kapikian (ed.), *Viral Infections of the Gastrointestinal Tract*, 2nd ed. Marcel Dekker Inc., New York, N.Y.

**Wu, P.S., L.Y. Chang, B. Berkhout, L. van der Hoek, C.Y. Lu, C.L. Kao, P.I. Lee, P.L. Shao, C.Y. Lee, F.Y. Huang, and L.M. Huang.** 2008. Clinical manifestations of human coronavirus NL-63 infection in children in Taiwan. *Eur. J. Pediatr.* **167**:75-80.

**Yoo, S.J., E.Y. Kuak, and B.M. Shin.** 2007. Detection of 12 respiratory viruses with two-set multiplex reverse transcriptase-PCR assay using a dual priming oligonucleotide system. *Kor. J. Lab. Med.* **27**:420-427.

---

**Yu, C.L.T. Gustafson, J. Diao, B.J. W., Z. Li, J. Zhang, and J. Chen.** 2005. Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain. *J. Biol. Chem.* **280**:23280–23286.

**Yu, X., W. Bi, S.R. Weiss, and J.L. Leibowitz.** 1994. Mouse hepatitis virus gene 5b protein is a new virion envelope protein. *Viol.* **202**:1018–1023.

**Zheng, Q., Y. Deng, J. Liu, L. van der Hoek, B. Berkhout, and M. Lu.** 2006. Core structure of S2 from the human coronavirus NL-63 spike glycoprotein. *Biochem.* **45**:15205-15215.

**Ziebuhr, J.** 2005. The coronavirus replicase. *Curr. Top. Microbiol. Immunol.* **287**:57–94.



**Ziebuhr, J., E.J. Snijder, and A.E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J. Gen. Virol.* **81**:853–879.

**Zúñiga, S., I. Sola, J.L. Moreno, P. Sabella, J. Plana-Durán, and L. Enjuanes.** 2006. Coronavirus nucleocapsid protein is an RNA chaperone. *Viol.* **357**(2):215-227.

**Appendix :1 Vectors**

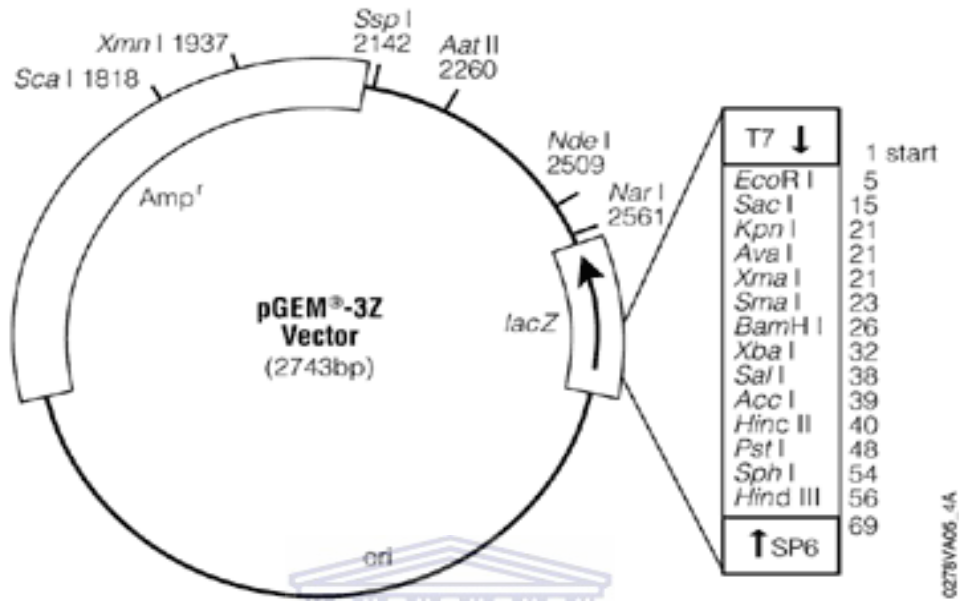


Figure 37: Sequencing vector pGEM T-easy

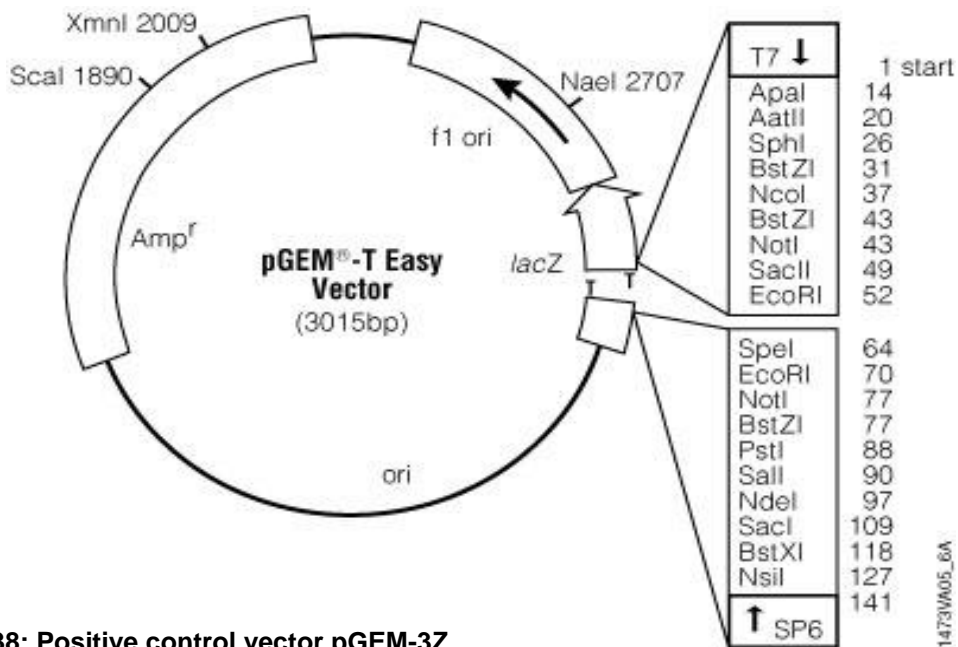


Figure 38: Positive control vector pGEM-3Z

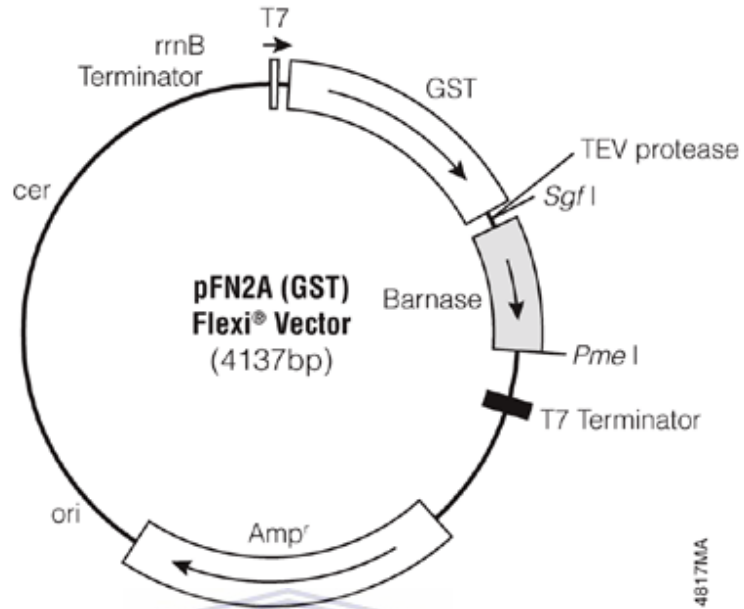
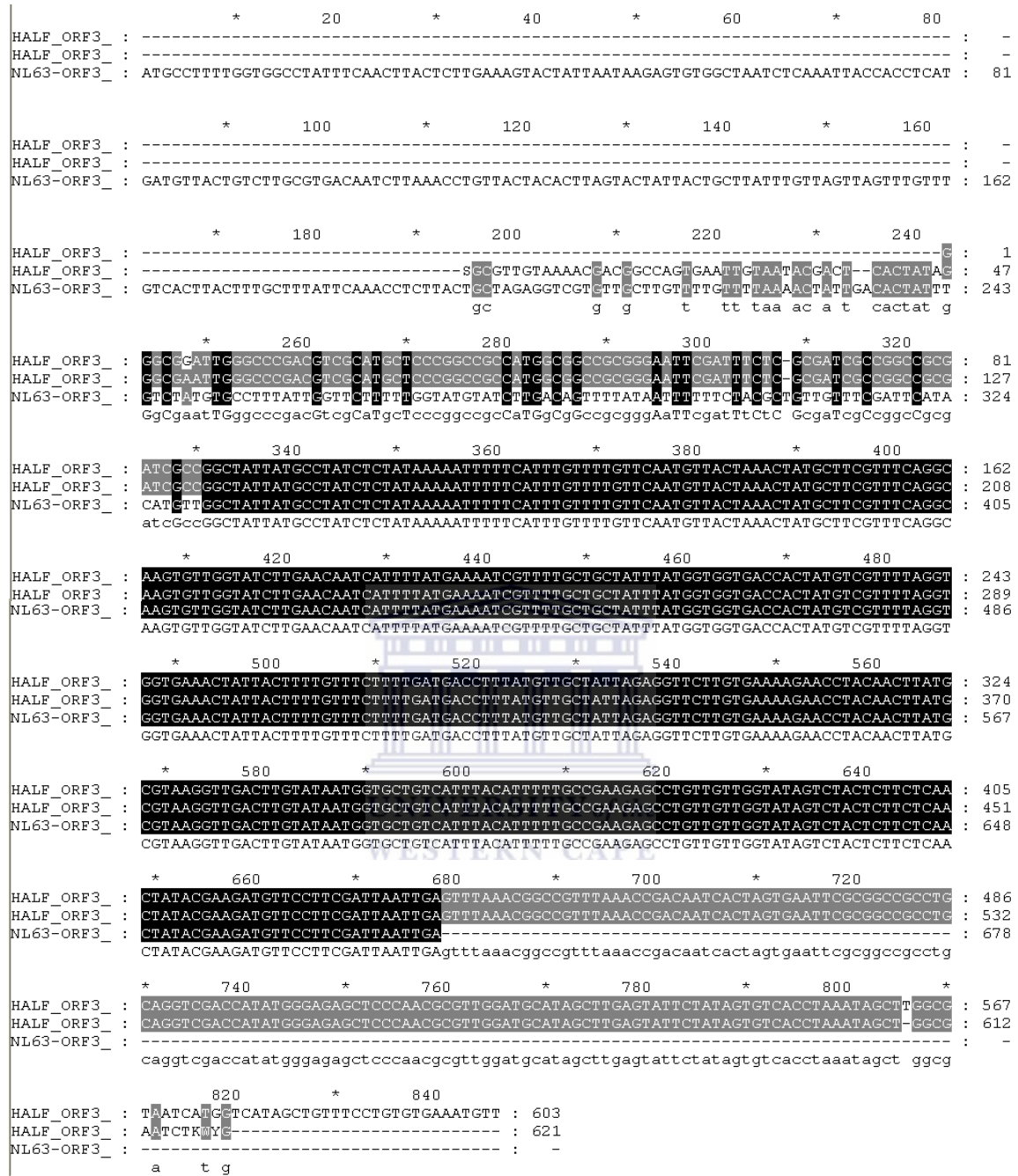


Figure 39: Bacterial expression vector pFLEXI-pFN2A (GST)





**Figure 40: Sequence verification of ORF3ΔN gene in pGEM**

Figure 40 above depicts the homology between the Human Coronavirus NL-63 (Amsterdam 1 strain) ORF3 sequence and the consensus of the forward and reverse sequences obtained from the pGEM-ORF3ΔN amplicons which were sent to Inqaba biotech for ABI sequencing.





```

E_Forward : -----*-----20-----*-----40-----*-----60-----*-----80----- : 50
NL63-E_AY6 : -----*-----20-----*-----40-----*-----60-----*-----80----- : -
E_Reverse : RGTGTAACGACGGCCAGTGAATTGTAATACGACTCACTATAAGGGCGAAATGGGCCCGACGTCGCATGCTCCCGGCCGCATC : 83
          ga tc ctat gggcga ttgggcccgcagtcgcatgctccccggccgcatg

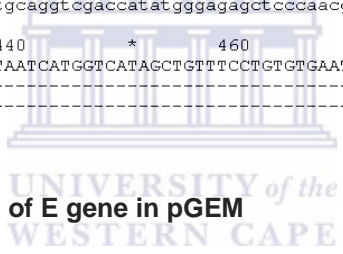
E_Forward : -----*-----100-----*-----120-----*-----140-----*-----160----- : 133
NL63-E_AY6 : -----*-----100-----*-----120-----*-----140-----*-----160----- : 50
E_Reverse : CCGGCCCGGGAAATTCGATTGGGCGCGATCGCCATGTTTCCTCGATTAATTGATGACAATGGTATTGTCCTCAATCCATTTT : 166
          gcgcccgcggaattcgattgggcgcgatcgccATGTTTCCTCGATTAATTGATGACAATGGTATTGTCCTCAATCCATTTT

E_Forward : -----*-----180-----*-----200-----*-----220-----*-----240----- : 216
NL63-E_AY6 : -----*-----180-----*-----200-----*-----220-----*-----240----- : 133
E_Reverse : ATGGCTCCTTGTTATGATATTTTTCTTTGTGTTGGCAATGACCTTTATTAACCTGATTC AATTGTTTTACTTGTCAATTT : 249
          ATGGCTCCTTGTTATGATATTTTTCTTTGTGTTGGCAATGACCTTTATTAACCTGATTC AATTGTTTTACTTGTCAATTT

E_Forward : -----*-----260-----*-----280-----*-----300-----*-----320-----*----- : 299
NL63-E_AY6 : -----*-----260-----*-----280-----*-----300-----*-----320----- : 201
E_Reverse : TTTTGTAGTAGGACATTATATCAACCAGTTTATAAAATTTTTCTTGCTTACCAAGATTATATGCAAAATAGTTTAAACACAAAAT : 332
          TTTTGTAGTAGGACATTATATCAACCAGTTTATAAAATTTTTCTTGCTTACCAAGATTATATGCAAAATAGttttaaacacaaaat

E_Forward : -----*-----340-----*-----360-----*-----380-----*-----400-----*----- : 382
NL63-E_AY6 : -----*-----340-----*-----360-----*-----380-----*-----400-----*----- : -
E_Reverse : CACTAGTGAATTCGCGGCCGCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATA : 415
          cactagtgaattcgcggccgctgcaggtcgaccat atgggagagctcccaacgcgttggatgcatagcttgagttattctata

E_Forward : -----*-----420-----*-----440-----*-----460-----*----- : 437
NL63-E_AY6 : -----*-----420-----*-----440-----*-----460-----*----- : -
E_Reverse : GTGTCACCTAAATAGCTTGGCG----- : 436
          gtgtcacctaaatagcttgg
    
```



**Figure 42: Sequence verification of E gene in pGEM**

Figure 42 above depicts the homology between the Human Coronavirus NL-63 (Amsterdam 1 strain) E sequence and the consensus of the forward and reverse sequences obtained from the pGEM-E amplicons which were sent to Inqaba biotech for ABI sequencing.

## ACKNOWLEDGMENTS

Firstly I would like to thank my Lord Jesus Christ for giving me the strength to endure my Masters' degree and preserving my sanity when all hope seemed lost.

Next, I would like to thank my Super supervisor, Proffessory Fielding, for always motivating and pushing me to work harder...even if it didn't work... your effort is GREATLY appreciated. I hold you in the deepest of respect and pray that one day you'll make it out of academia before you start to look 70...oops...too late. If ever I knew a man who was over worked and under paid... it was me... but you're hot on my heels for that trophy. Thanks again oh captain my captain.

Not far behind Proffessory on the "irritatingly inspirational and nauseatingly motivational scale" is my lovely girlfriend and majority share holding life partner Farzana Rahiman (soon to be Dr. Fuzzy). Babe, you're the bestis-sta-tist-tist. Thank you for every punch, every textbook, stapler, ring binder, every kick, shoe and cooking utensil you threw at me to get me to work... not only did it work... but I actually got ten bucks for your left shoe on Gumtree. I love you and thanks for the continued support and fantastic cooking.

Special mention must go to all my awe inspiring friends: Dr. T. Govender (Mr. Banan jammer), Ryan and Renétjie, Riaan Cedras (Faggie), Mu'azzam (Batman) and family and Adiel (E-Roboobie) and Laylah (Wan-dur Woman). I say 'awe

inspiring' because I look at you guys and wonder how the hell I ended up having such bad choice in people. Thanks for the support guys... Especially you Tash.

To all the Smurfs in my village: Paranoia Smurf (Tasnimetjie), Kerk-broer Smurf (Yanga), Hairy Armpit Smurf (Michael), Lady Quagga Smurf (Tarryn-Lee), Sokkie-Sokkie Smurf (Lizel), Smurfoholic Alch (Anele), Foreign National Ninja Smurf (Aasiyah) and finally Emo-Schizophrenic Smurf (Tiza). You guys are the beast! Since the first time I really got to know all you guys... I've developed this intense passion to do a Ph.D. in genetic profiling...'cause I believe that people should know the challenges ahead if they decide to have kids...and the reality of things going completely off track genetically. As long as you guys are in science... science will always be a mystery. Thanks for all the laughs comrades.

Finally, to my family: Mom... thanks for the good hair and for always being a mom when I needed a teaspoon. I know you always care...what exactly you care for...I'm not sure anyone can figure that out...but I love you none the less. Dad... If you're reading this then mom hasn't killed me yet... thanks for always coming to the rescue. I wish everyone I knew could have you as a dad...but then you'd be totally broke and Farzana would be my sister so let's scrap that idea. Love you Dada. Stuart, Sean and Nicole... You guys rock...as a group...I think it's the cheerleader effect though (B. Stinson *et al.*, 2004). I wish you guys could all move into one big house together... so I would know what part of Cape Town to avoid ;) Love you guys!