DEPARTMENT OF MEDICAL BIOSCIENCES

Cloning and Expression of the M-Gene from the Human Coronavirus NL-63 in Different Expression Systems

University of the Western Cape

Lizel Lubbe

22/05/2009

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science in the University of the Western Cape, Cape Town. It has not been submitted before for any degree or examination at any other University.

_____

Lizel Lubbe

_____ th day of _____

Table of Contents

Page

List of Tables i

List of Figures ii

List of Abbreviations iv

Abstract ix

Publication of Research xii

Acknowledgements xiv

Chapter 1: *Introduction*

1.1. Background 2

1.2. Coronaviruses 5

1.2.1. Basic structure

1.2.2. Genome structure

1.2.3. Structural proteins

1.2.3.1. Spike protein (S)

1.2.3.2. Nucleocapsid protein (N)

1.2.3.3. Envelope protein (E)

1.2.3.4. Membrane protein (M)

1.2.4. Replication and life cycle

1.3. Human Coronavirus NL63 18

1.3.1. Virus cell entry

1.3.2. Culturing in permissible mammalian cell lines

1.3.3. Clinical infection, symptoms and prevalence

1.3.4. Treatment of infection

1.4. Objectives of this Thesis 27

Chapter 2: *Materials and Methods*

2.1. *In silico* analysis of HCoV-NL63 M protein 29

2.1.1. Predicative analysis of HCoV-NL63 M

2.1.2. Comparison of HCoV-NL63 M homologues

2.2. Overview of common molecular techniques used throughout Chapter 2 30

2.2.1. Reverse transcription

2.2.2. Polymerase chain reaction

2.2.2.1. Colony PCR

2.2.3. PCR purification and extraction of nucleic acid fragments from gels

2.2.4. Agarose gel electrophoresis

2.2.5. Restriction endonuclease digests

2.2.6. Ligation

2.2.7. Plasmid DNA preparation (minipreps)

2.2.8. Freezing of mammalian and insect cells

2.2.9. SDS-PAGE

2.2.10. Western Blotting

2.2.11. Coomassie blue staining

2.3. Cloning for expression in mammalian system 41
2.3.1. Plasmid construction
2.3.2. Transformation of ligations into pCMV-myc for mammalian expression
2.3.3. Plasmid DNA preparation (minipreps)
2.4. Cloning for expression in Baculovirus system 44
2.4.1. Plasmid construction
2.4.2. Transformation of ligations into pFastBac for baculovirus expression
2.4.3. Plasmid DNA preparation (minipreps)
2.4.4. Cells and viruses
2.4.5. Transfection of insect cells
2.4.6. Harvesting recombinant baculovirus P1 and P2 viral stock
2.5. Cloning for expression in bacterial system 48
2.5.1. Plasmid construction
2.5.2. Transformation of ligations into pFlexi for bacterial expression
2.5.3. Plasmid DNA preparation (minipreps)
2.5.4. Time course studies

Chapter 3: *Results and Discussion*

3.1. *In silico* characterisation of HCoV-NL63 M protein 53
3.1.1. Comparison of HCoV-NL63 homologues
3.2. Cloning for expression in mammalian system 58
3.3. Cloning for expression in Baculovirus system 63
3.4. Cloning for expression in bacterial system 70

Chapter 4: *Conclusion*

4.1. Mammalian expression 77

4.2. Baculovirus expression 78

4.3. Bacterial expression 80

References 82

Appendix A 101

Appendix B 103

Appendix C 105

List of Tables

		Page
1.	Summary of structural genes of HCoV-NL63	13
2.	Seasonal prevalence of HCoV-NL63	22
3.	Frequency of NL63 in comparison to other human coronaviruses	23
4.	Reverse transcription mix used in this study	30
5.	Reverse transcription program conditions	31
6.	PCR mixture components	32
7.	PCR program conditions	33
8.	Polymerase chain reaction primers used in this study	33
9.	Restriction endonuclease mix	36
10.	Ligation mix	37
11.	Comparison of HCoV-NL63 M amino acid sequence to homologues from selected Group I coronavirus isolates.	56

List of Figures

	Page
1. Classification of Human Coronavirus NL63	6
2. General structure of HCoV-NL63	7
3. Genome structure of HCoV-NL63	8
4. Life cycle of coronaviruses	15
5. Hydrophobicity plot	53
6. Predicted structure/topology of HCoV-NL63 M protein	54
7. Alignment of the amino acid sequences of HCoV-NL63 M and selected coronavirus M proteins.	55
8. PCR of the full-length M gene from complete HCoV-NL63 genome	58
9. Agar plate of transformation of pGEM constructs into *JM109 E.coli* competent cells	60
10. Colony PCR of M ligated into pGEM and transformed into JM109 *E. coli*	61
11. Restriction enzyme digestion of pGEM-M1 and pGEM-M4	62
12. Restriction enzyme digests of pCMV-myc constructs	62
13. 1% Agarose gel electrophoresis of PCR product generated with M-specific primers using HCoV-NL63 cDNA as template	64
14. 1% Agarose gel electrophoresis of PCR product generated from colony PCR of pGEM-M transformed into JM109 *E.coli*	65
15. *XbaI* and *EcoRI* restriction enzyme digestion of M from the three positive pGEM-M constructs identified in Figure 14	66
16. Colony PCR of the M gene from four (4) selected white colonies	67

17. Gel of PCR of recombinant bacmid 68
18. Diagrammatic representation of expected fragment sizes after gel electrophoresis of PCR of recombinant bacmid 69
19. Agarose gel of PCR of the M gene for expression in bacteria 71
20. Agarose gel of colony PCR of insert transformed into *JM109 E.coli* Coomassie stain of baculovirus expression 72
21. Agar plates of transformation of M into *KRX E.coli* cells Colony PCR of M ligated into pCMV and transformed into JM109 *E.coli* 73
22. Restriction enzyme digests of pFlexi-M with restriction enzymes *SgfI* and *PmeI* 73
23. Western Blots of GST-M protein expressed by KRX *E.coli* over a 36 hour time course 74

List of Abbreviations

A	Adenine
Aa	Amino acid
Ab	Antibody
Amp	Ampicillin
ACE	Angiotensin-converting enzyme
AcMNPV-PH	*Autographica californica* multiple nuclear polyhedrosis virus polyhedron
APS	Ammonium persulphate
BCoV	Bovine coronavirus
Bp	Base pairs
BSA	Bovine serum albumin
BV	Baculovirus
C	Celsius
C	Cytosine
cDNA	Copy DNA
CoV	Coronavirus
CPE	Cytopathic effect
Da	Dalton
DNA	Deoxyribonucleic acid

DMEM	Dulbecco’s Modified Eagle’s Medium
DMSO	Dimethyl sulfoxide
dNTP’s	Deoxyribonucleotide triphosphates
E protein	Envelope protein
EDTA	Ethylenediaminetetraacetic acid
ECoV	Equine Coronavirus
FIPV	Feline infectious peritonitis virus
FBS	Foetal Bovine Serum
G	Guanine
GP	Glycoprotein
GST	Glutathione S-transferase
HCoV	Human coronavirus
HE	Haemagglutinin esterase
HR	Heptad repeat
Hr(s)	Hour(s)
HRP	Horseradish peroxidase
HCl	Hydrochloric acid
I protein	Internal protein
IBV	Infectious bronchitis virus
IFN	Interferon

IPTG	Isopropyl-beta-D-thiogalactopyranoside
IV	Intravenous
Kb	Kilo bases
kDa	Kilo Dalton
LB	Luria Bertani
LLC	Lewis lung carcinoma
LRTI	Lower respiratory tract infections
M protein	Membrane protein
mAmp	MilliAmpere
mg	Milligram
MgCl	Magnesium chloride
MHV	Murine Hepatitis Virus
ml	Milliliter
mm	Millimeter
min	Minutes
mM	Milimolar
N protein	Nucleocapsid protein
NaCl	Sodium Chloride
nm	Nanometer
ng	Nanogram

NPA	Nasopharyngeal aspirate
nt	Nucleotide
O/N	Overnight
ORF3 protein	Open reading frame 3 protein
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEDV	Porcine epidemic diarrhea virus
PhCoV	Pheasant Coronavirus
PHEV	Heamagglutinating encephalomyelitis virus
RNA	Ribonucleic acid
Rpm	Revolutions per minute
RSV	Respiratory syncytial virus
RT	Reverse transcriptase
RT	Room temperature
s	Seconds
S protein	Spike protein
SADV	Sialodacryoadenitisvirus
SARS	Severe Acute Respiratory Syndrome
SDS	Sodium dodecyl sulphate

Sf	*Spodoptera frugiperda*
T	Thymine
Taq	*Thermus aquaticus*
TCoV	Turkey Coronavirus
TGEV	Transmissible gastro enteritis virus
TRS	Transcription regulatory sequence
Tris	Tris(hydroxymethyl)aminomethane
TBE	Tris Borate EDTA
TEMED	N'N'N'N tetramethylethylene diamine
U	Uracil
UV	Ultraviolet
V	Volt
VIDISCA	Virus discovery based on cDNA-AFLP technique
VLP	Viral like particle
WHO	World Health Organization
µl	Microliter

Abstract

Respiratory tract infections are one of the leading causes of morbidity and mortality across the world. This is especially true for infants, young children, the elderly and the immunocompromised. The strain placed on economies and health care systems of all countries by these diseases are phenomenal. Although we are familiar with various agents leading to these kinds of infections (e.g. rhino-, influenza-, parainfluenza, human metapneumo-, respiratory syncytial-, adeno- and coronaviruses), the cause of a substantial portion, (48-70%) of cases remain unidentified (Van der Hoek *et al*, 2004; Fouchier *et al*, 2004; File, 2003; Fine *et al*, 1999; Shay *et al*, 1999, Henrickson *et al* 2004; Murray *et al* 2001).

In the past, human coronaviruses have not been known to cause severe disease in humans. For this reason, little research was performed on these viruses, with research focusing on the animal coronaviruses that are of veterinary importance. However, with the outbreak of SARS in 2003, the field of human coronavirus research has received significantly more attention. Also, the subsequent identification of two additional novel human coronaviruses (NL63 and HKU1) has led to an increased awareness of the potential threat of these viruses. With the discovery of these new human coronaviruses, it has become clear that the potential for another outbreak by a yet unknown human coronavirus is a very real possibility. This has made research into the pathogenesis and the role of the various coronavirus genes in the pathogenesis of these viruses of utmost importance.

HCoV-NL63 was first discovered in January 2003 in the Netherlands. It causes upper and lower respiratory tract disease in young children, the elderly and immunocompromised individuals. The disease is also associated with croup and has

even been implicated as a possible cause of the childhood vascular ailment Kawasaki Disease. HCoV-NL63 is frequently found in combination with other respiratory viruses leading to superinfections. It is still unclear whether HCoV-NL63 is an opportunistic virus or whether it leads the way for co-infection with other respiratory viruses. This particular virus is also the only coronavirus sharing the same cellular receptor as SARS-CoV.

The virus is found all over the world and has been identified in countries like Australia (Arden et al, 2005), Japan (Ebihara *et al.,* 2005; Suzuki *et al.,* 2005), Belgium (Moës *et al.,* 2005), Hong Kong (Chiu *et al.,* 2005), Taiwan (Wu *et al.,* 2007) Korea (Choi *et al.,* 2006), Canada (Bastien *et al.,* 2005), France (Vabret *et al.,* 2005), Switzerland (Kaiser *et al.,* 2005; Garbino *et al.,* 2006), Germany (Van der Hoek *et al.,* 2005), Sweden (Koetz *et al.,* 2006) and South Africa (Smuts and Hardie, 2006).

In this study, the HCoV-NL63 genome was transcribed from RNA to DNA from which the M gene was amplified with various primers designed for use in specific expression systems. The various genes were cloned into the pGEM vector and confirmed by sequencing. The genes were now expressed in cloning vectors suited for each expression system (pFastBac for baculovirus expression, pFlexi for bacterial expression and pCMV for mammalian expression). Clones were sequenced for a second time. The recombinant clone in pFlexi was expressed in KRX cells and a 36hr time course was performed. The recombinant pFastBac clone was used to infect *Sf9* insect cells and P1 and P2 viral stocks were obtained. The recombinant pCMV clone was used to transfect *Cos1* mammalian cells.

The genome was successfully transcribed and the M gene amplified and cloned into pGEM and confirmed by sequencing. Subsequent cloning of the various M genes into pFastBac for baculovirus expression, pFlexi for bacterial expression and pCMV for mammalian expression was achieved and sequencing confirmed the presence of the inserts in frame.

pFlexi clones were successfully expressed in bacterial KRX cells with expression of the M protein in the pellet of the lysed bacterial cells. No M protein was seen in the supernatant of the lysed cells. *Sf9* insect cells were infected with the recombinant pFastBac clones and P1 and P2 viral stocks were obtained. Protein expression occurs in KRX bacterial cells with optimal expression at approximately 24 hours. The M protein expresses on the cell membrane as can be concluded from the product obtained in the pellet of the lysed bacterial cells. Very little of the expressed protein is present in the plasma of the cell as evidenced by the absence of protein in the supernatant of the lysate.

Publication of Research

Oral and poster presentations

L Lubbe, R Fisher, T Suliman and BC Fielding, Molecular biology of human coronavirus NL-63, **186th General Meeting of the Experimental Biology Group (EBG)**, University of the Western Cape, 30th October 2008. (**Invited Oral presentation)**

L Lubbe and BC Fielding, Expression studies of human coronavirus NL63 M protein, **16th Annual Faculty of Natural Sciences Research Open Day**, University of the Western Cape, 27-28 October 2008. (**Oral presentation**)

L Lubbe, S. Khan and BC Fielding, Cloning of human coronavirus NL63 M protein for expression in a baculovirus system, **Medical Research Council Research Day**, MRC Conference Centre, Parow Valley, 16-17 October 2008. (**Poster presentation**)

L Lubbe and BC Fielding, Expression studies of human coronavirus NL63 M protein, **Medical Biosciences Seminar Series**, University of the Western Cape, 11 September 2008. (**Oral presentation**)

Abstracts

L Lubbe and BC Fielding, Expression studies of human coronavirus NL63 M protein, **16th Annual Faculty of Natural Sciences Research Open Day**, University of the Western Cape, 27-28 October 2008. (**Abstract book**)

L Lubbe, S. Khan and BC Fielding, Cloning of human coronavirus NL63 M protein for expression in a baculovirus system, **Medical Research Council Research Day**, MRC Conference Centre, Parow Valley, 16-17 October 2008. (**Abstract book**)

Awards

Deutscher Akademischer Austausch Dienst (DAAD) Scholarship, 2008

Acknowledgements

I would like to thank my supervisor Prof. B. C. Fielding for all the support and advice throughout the entire project. I also take this opportunity to thank my colleagues Randall Fisher and Tasnim Suliman for all their help and encouragement. Various other people helped make this project a reality and have my heartfelt gratitude. Prof Sehaam Khan and her students (for the use of her laboratory and expertise on baculovirus protocol), Peter Witbooi (my faithful bodyguard during late night lab experiments), Hannes Ehlers (for patiently helping with hours of image editing), my parents (for endless belief and encouragement in the most frustrating of times) and last but certainly not least, Martin Immelman for all the love, patience and litres of red wine involved in a fiancé writing a thesis.

Chapter 1: *Introduction*

1.1. Background

Respiratory tract infections are one of the leading causes of morbidity and mortality across the world and according to the World Health Organization (WHO, 2003), lower respiratory tract infections (LRTI) account for up to 20% of all deaths in children under five. Not only does it cause substantial mortality in the elderly, infants and young children, but it is also the most common cause of hospitalization and antibiotics usage in immunocompromised individuals and those with chronic lung diseases. Although we are familiar with various agents leading to these kinds of infections (*e.g.* rhino-, influenza-, parainfluenza, human metapneumo-, respiratory syncytial-, adeno- and coronaviruses), the causative viral agents of a substantial portion (48-70%) of cases remains unidentified (Fouchier *et al*, 2004; Henrickson *et al* 2004; Van der Hoek *et al*, 2004; File, 2003; Murray *et al* 2001; Fine *et al*, 1999; Shay *et al*, 1999). The strain placed on economies and health care systems of all countries by these diseases are phenomenal. This is especially true for South Africa and similar countries with an increasing number of immunocompromised individuals due to the HIV/AIDS pandemic. No studies have yet been undertaken in South Africa to ascertain the effec human coronavirus NL63 (HCoV-NL63) might have on the high risk individuals.

Human coronavirus (HCoV)-229E and HCoV-OC43, both discovered in the mid 1960's (Hamre & Procknow, 1966; Tyrrell & Bynoe, 1965) are considered as the so called "common cold" viruses (Bradburne and Somerset, 1972; Bradburne *et al.,* 1967; Hamre & Procknow, 1966). It is believed that they are responsible for up to 30% of all common colds (McIntosh, 1996), often resulting in mild and self-limiting, mostly upper respiratory tract infections (McIntosh, 1996).

During the past six years three "new" human coronaviruses have been identified. In early 2003 the causative agent of severe acute respiratory syndrome

(SARS) was identified as the novel coronavirus SARS-CoV (Drosten *et al.,* 2003; Ksiazek *et al.,* 2003). This virus causes severe respiratory disease characterized by atypical pneumonia. Up to 10% of infections are fatal and the disease resulted in the loss of about 800 lives before being brought under control late in 2003 (Peiris *et al* 2003; Stadler *et al*, 2003; WHO, 2003). It is speculated that the virus was transmitted in a zoonotic event from bats to civet cats (prepared as food source in China), and ultimately to humans (Lau *et al.,* 2005; Li *et al.,* 2005).

As a result of all the attention focused on the field of coronavirology brought about by the outbreak of SARS, another two human coronaviruses were subsequently identified, namely HCoV-NL63 and HCoV-HKU1. HCoV-HKU1 is the most recently discovered human coronavirus; it was isolated and identified in January 2005 in Hong Kong from a 71 year old man with chronic pulmonary disease (Woo *et al*, 2005). The virus causes common cold-like symptoms and other acute respiratory tract infections and may also cause febrile seizures in up to 50% of cases. Typically, HCoV-HKU1 infects children, but can sometimes also infect immunocompetent adults (Woo *et al.,* 2005).

HCoV-NL63 was first isolated in January 2003 in Amsterdam, Netherlands at the Amsterdam Medical Centre. It was isolated from a nasopharyngeal aspirate (NPA) of a 7 month old hospitalized child presenting with coryza, fever, brochiolitis and conjunctivitis. The specimen from this patient tested negative for all the known respiratory viruses. It was then inoculated into a tMK cell line followed by inoculation into LLC-MK2 cells. The novel NL63 virus was identified from the supernatant using the VIDISCA method. Sequencing showed that this virus was related to the coronaviruses (with 42-81% sequence homology to HCoV-229E) but was indeed a new member of the virus coronavirus family as opposed to a recombinant member (Van der Hoek *et al.,* 2004).

In 2004, another group at the Erasmus Medical Centre in the Netherlands described this virus found in a nose swab taken from an 8 month old pneumonia patient from 1988. They cultured the virus in Vero-E6 instead of LLC-MK2 cells and identified the virus in the supernatant. They named their virus HCoV-NL. The sequence of this virus corresponded 98.8% at nucleotide level with that of the virus discovered by Van der Hoek *et al.* (2005) and it was subsequently concluded that they represent the same virus (Fouchier *et al.,* 2004). Esper *et al*. (2005) discovered a virus at the Yale University in New Haven in 2005 that had a 94-100% homology with the two previously discovered viruses. Even though they named the virus HCoV-NH this virus was also concluded to be the same virus as those discovered by Van der Hoek *et al.* (2005) and Fouchier *et al.* (2004).

Even though three groups stumbled upon this virus in the space of a year, this might in fact not be the first time that this virus was isolated. In the mid 1960's, human coronaviruses were cultured in human embryonic tracheal organ cultures. These viruses seemed to be only distantly related to HCoV-229E and HCoV-OC43 (Macintosh *et al*, 1967; Tyrrell & Bynoe, 1965). Then in the early 1980's, more human coronavirus isolates were obtained from cell or organ culture and extensively studied by Larson *et al.* (1980) and Macnaughton *et al.* (1981). Some of these isolates were serologically related to the group II HCoV-OC43 while others were serologically related to the group I HCoV-229E. Some of the group I isolates could be cultured in MRC continuous cells like HCoV-229E while others could only be cultured in human foetal tracheal and nasal organ culture cells – very much like HCoV-NL63. Unfortunately these cultured strains were lost to science before this could be further investigated. Up to this day it can only be speculated whether this was the first observed strains of HCoV-NL63, HCoV-HKU1 or a still undiscovered human coronavirus.

1.2. Coronaviruses

The order *Nidovirales* consists of the viral families *Arteriviridae*, *Roniviridae* and *Coronaviridae* (Cavanagh, 1997). The *Coronaviridae* family has been further divided into two subfamilies namely toroviruses and coronaviruses (Figure 1) (Nidovirus Symposium, 2005).

Traditionally, coronaviruses have been divided into 3 groups according to genome sequence, host preference and serological cross-reactivity (Holmes, 2001). Group 1 and 2 coronaviruses cause respiratory, central nervous system and enteric disease in humans and mammals (Guy *et al.,* 2000; Holmes and Lai, 1996). They have been isolated from mice, rats, pigs, dogs, cats, rabbits, horses and cattle (Holmes and Lai, 1996). Group 1 includes coronaviruses such as transmissible gastroenteritis virus (TGEV), porcine epidemic diarrhea virus (PEDV) and feline infectious peritonitis virus (FIPV). Also, included in Group 1 are the human coronaviruses 229E, and NL63. Group 2 includes three human coronaviruses – OC43, HKU1 and severe acute respiratory coronavirus (SARS-CoV), as well as several animal coronaviruses like murine hepatitis virus (MHV), bovine coronavirus (BCoV), porcine heamagglutinating encephalomyelitis virus (PHEV), rat sialodacryoadenitis coronavirus (SADV) and equine coronavirus (ECoV). Group 3 coronaviruses only infect avian hosts (Holmes and Lai, 1996) and have been shown to infect chickens, pigeons, geese, ducks and turkeys and include infectious bronchitis virus (IBV), turkey coronavirus (TCoV) and pheasant coronavirus (PhCoV).

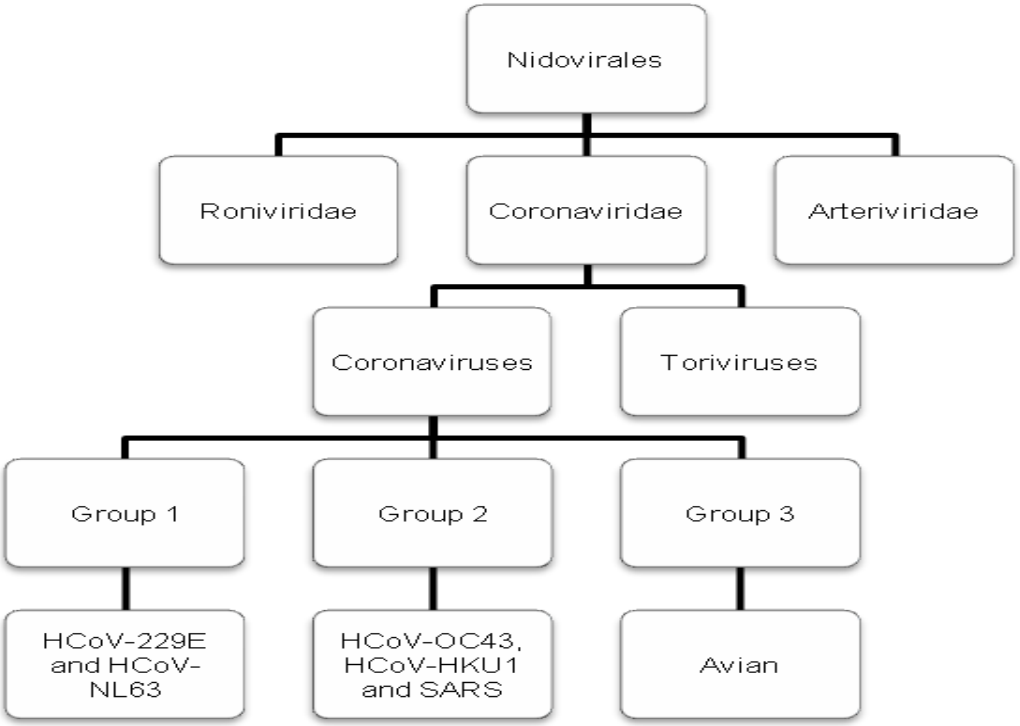


Figure 1. Classification of Human Coronavirus-NL63 (Nidovirus Symposium, 2005)

1.2.1. Basic structure

The name coronavirus was first used in 1968 and originates from the crown like appearance of these viruses when viewed under an electronmicroscope (Tyrrell *et al*, 1968). They are round/pleiomorphic viruses roughly 80-120nm in diameter (Figure 2). They are enveloped RNA viruses with the largest genomes found in all RNA viruses – 27-33 kb (Masters, 2006). Within the viral membrane there is a helical capsid, which is made up of RNA complexed with the nucleocapsid (N) protein. The membrane also contains three more proteins - spike (S), the membrane (M) and the envelope protein (E). These four proteins are the essential proteins and are found in all coronaviruses. There are also some non-essential proteins that are characteristic to certain groups. Some group 2 species has a unique protein known as the heamagglutinin esterase (HE) as well as the internal (I) protein (Pyrc *et al.,* 2004; Siddell, 1983). SARS has the ORF3a protein – an additional structural protein.

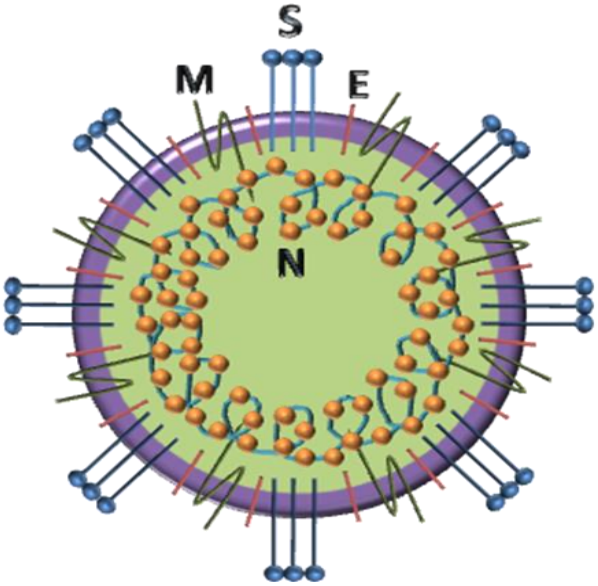


Figure 2. General structure of HCoV-NL63

1.2.2. Genome structure

Coronaviruses have linear, non-segmented, positive-sense, single stranded RNA genomes. The genomes of about 30 kb are capped and polyadenelated. The 5' end has an untranslated region of 286 nucleotides and the 3' end has an untranslated region of 287 nucleotides (Van der Hoek *et al.,* 2004).

Coronavirus genomes consist of two distinct regions, with a typical gene arrangement (Figure 3). The 5' two thirds of the genome codes for the proteins needed for RNA synthesis. The 3' third of the genome codes for 5 structural proteins in the conserved order, 5'-membrane (M)-spike (S)-nucleocapsid (N)-envelope (E)-3'. Interspersed among the structural genes coronavirus genomes contain genes that encode for accessory proteins. These non-essential accessory genes vary greatly in position and number in the different coronavirus species. The function of these proteins is not yet fully understood, but they are thought to contribute to virus pathogenicity, possibly interfering with the innate immune response of the host.

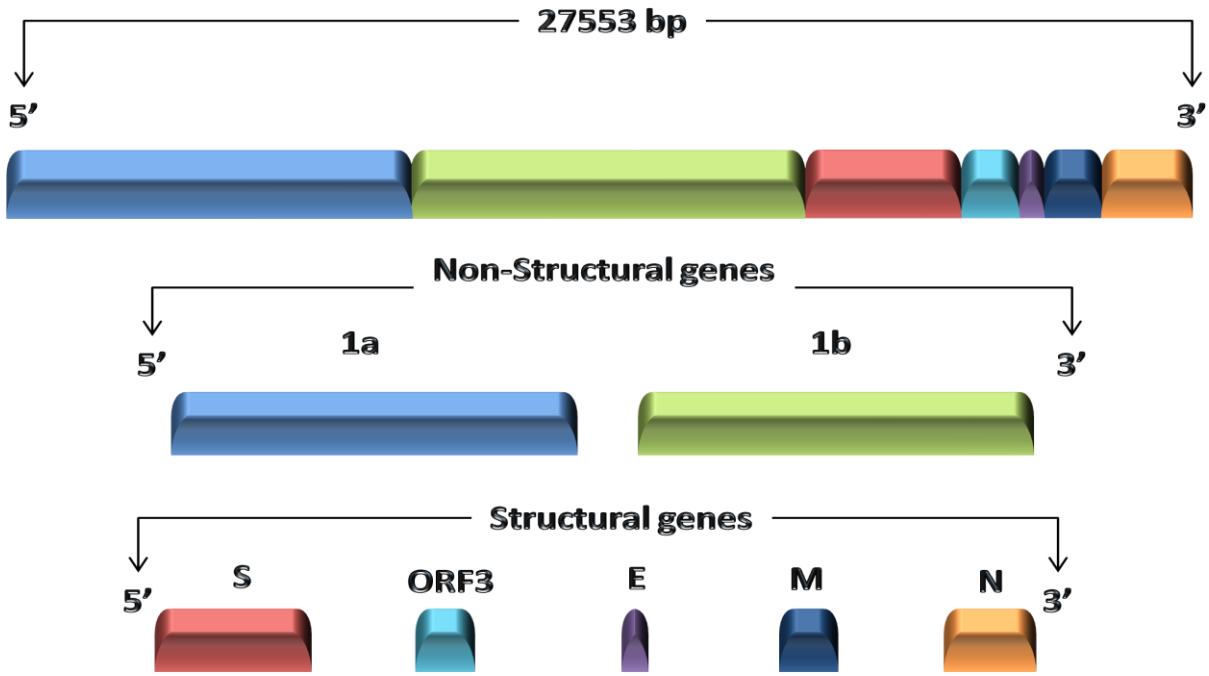


Figure 3. Genome structure of HCoV-NL63

1.2.3. Structural proteins

1.2.3.1 Spike protein (S)

The S protein is a type I glycoprotein that protrudes from the virus surface and provides the virus with its characteristic crown-like appearance (Bosch *et al.,* 2003). During processing of the protein in the Golgi apparatus, the spike is cleaved into the S1 and S2 subunit each approximately 90 kDa (Sturman and Holmes, 1977). In the globular head of the S1 protein there is a receptor binding domain located in the first 330 amino acids (Kubo *et al.,* 1994). The S2 subunit is anchored in the membrane and has 2 heptad repeat regions (HR1 and HR2) (Van der Hoek *et al.,* 2006) which are longer than coronaviruses of group II and III (Bosch *et al.,* 2004). This extra length is due to the 14 amino acid insert present in 2 extra helical turns. After infection of the cell, the HR1 and HR2 polypeptides will fuse into a stable 6 helix bundle (Pyrc *et al.,* 2006). S2 also has a transmembrane domain. It is thought that this protein has a role in infection of target cells, cell to cell spread and may also play a role in the replication cycle of the virus. Even though the S protein of NL63 is very similar to other type 1 glycoproteins, it differs in the aspect of not needing to be cleaved into the two parts to be able to infect a cell (de Haan *et al.,* 2004). Another unique feature is the 179 amino acid domain it has at the N terminal part of the spike. This sequence contains several glycosylation sites and is thought to play a role in the evasion of the host's immune response (Van der Hoek *et al.,* 2006). The virus attaches to a cellular receptor via its S protein. This causes a conformational change in the cell allowing fusion of the viral and cellular membranes (Matsuyama and Taguchi, 2002; Saeki *et al.,* 1997).

1.2.3.2. Nucleocapsid protein (N)

The N protein has a role in pathogenesis as well as transcription. Without this protein, the virus cannot be effectively recovered from infectious cDNA clones (Yount *et al.,* 2003; Yount *et al.,* 2002). While the N protein has not been studied in depth for the NL63, other coronaviruses of all 3 groups have the N protein localized in the nucleolus and cytoplasm. This could be an indication that the protein may delay or arrest the cell cycle in the G_2/M phase. It most likely achieves this by inhibiting cytokines (Wurm *et al.,* 2001).

1.2.3.3. Envelope protein (E)

The E protein of some coronaviruses can form virus like particles when expressed with or without the M protein and therefore plays a vital part in the assembly and infectivity of the virus (Kuo and Masters, 2003; Vennema *et al.,* 1996). This protein can also induce apoptosis in certain cells via the caspase dependant pathway (An *et al.,* 1999). This apoptosis may be the cause of the lymphopaenia commonly found in SARS positive patients (Young *et al.,* 2005). SARS-CoV E proteins have been shown to have cation selective ion channel activity of which the role has yet to be fully established. It has been speculated that the channel could be important for the budding, morphogenesis and assembly of the virus (Wilson *et al.,* 2004).

1.2.3.4. Membrane protein (M)

The M protein is the most abundant of the coronavirus membrane proteins. It is a type III glycoprotein consisting of three parts, a short ectodomain at the N-terminal, a triple membrane spanning region and a cytoplasmic tail (Bond *et al.,* 1979). It has been hypothesized that the membrane of the coronavirus consists of a

dense matrix of laterally interacting M proteins (de Haan *et al.,* 2000). The proteins are found in complexes in the Golgi apparatus when they are expressed on their own (Krijnse Locker *et al.,* 1995).

The M protein has been shown to be N-glycosylated (de Haan *et al.,* 2003). Some coronavirus M proteins, like that of TGEV, have proven to have interferogenic activity. It was also shown that if there are mutations in the ectodomain of the M protein causing impairment of the N-glycosylation, the interferogenic activity of these viruses will be down regulated (Laude *et al.,* 1992). Studies on MHV have shown that the glycosylation state of the M protein does not affect the virus' ability for *in vitro* replication. It does however, have an effect on the protein's ability to replicate *in vivo* in the liver, as well as its ability to induce IFN-α *in vitro* (de Haan *et al.,* 2003). It also has an effect on the interaction with host cells.

As for the general function of the M protein, it is believed to play a role in the assembly of the virion as well as in interaction with host cells and therefore infectivity. The formation of new virions is dependent on two processes. The one is the assembly of the nucleocapsid and the other, the assembly of the envelope. The M protein plays an essential role in both these processes by directing the formation of the envelope as well as providing the basis for the attachment of the nucleocapsid to enable budding (de Haan *et al.,* 2000).

The interaction of the M protein with other proteins has been studied by using the method of production of viral-like particles (VLP's). By expressing the M protein and any other protein together, the minimal interactions required for the formation of the particles resembling the original virus, can be determined. (Bos *et al*; 1996, Vennema *et al*; 1996). The M gene is vital in the formation of these VLP's. This is proven by the fact that the ability of a virus to produce these VLP's is majorly affected with mutation of the M gene (de Haan *et al*; 1998). This is true even when

there are only two amino acids removed from the carboxy terminal end of the protein. A mutation like this in fact proved to be lethal (Kuo *et al*; 2002).

The M protein does not only interact with the host molecules but also with other proteins within the virus itself. The interaction of the M protein with the E protein has been well documented and is believed to be the most important step in the viral assembly process. When the genes for these two proteins were expressed in cells, VLP's were formed. These particles were nearly identical in size and shape to the original viruses. In some studies it was shown that the M protein interacts with the N protein *in vitro* as well as *in vivo* (Kuo and Masters, 2002; Escors *et al*, 2001; Narayanan and Makino, 2001). This interaction is key to the assembly of the virus as it lacks the structural matrix protein of other enveloped viruses (He *et al*, 2004). Further down the line in the process of virion assembly, the M protein also interacts with the RNA packaging signal to form the mature packaged virus particle (Narayanan *et al*, 2003). The interaction between the M protein and the S protein also seems to be an essential step in the assembly of the virus (Rottier, 1995).

The uniqueness of the M protein is further highlighted by its ability to induce antibody production in hosts either during infection or inoculation with attenuated viral vaccines. This takes place in spite of the fact that the greater part of this protein is buried within the viral membrane (Wesseling *et al*, 1993; Vennema *et al*, 1991; Pulford and Britton, 1991; Laude *et al*, 1986; Fiscus and Teramoto, 1987). This ability is useful for the production of vaccines and utilization as a target for diagnostic tests. This immunogenicity is due to two immunodominant epitopes situated one each at the N terminal and the C terminal of the M protein as shown by He *et al* (2005) for the SARS virus. They found that the epitope at the N terminal was very immunogenic when introduced into patients but less so when introduced into immunized animals. The opposite was true for the epitope at the C terminal. This

phenomenon highlights the difference in immunogenic properties of the epitopes in immunization and infection. It is believed that the N terminal epitope is the major immunogenic epitope (He *et al*, 2005).

The HCoV-NL63 structural genes, their sizes, functions and locations are summarised in Table 1.

Table 1. Summary of structural genes of HCoV-NL63

Gene	Protein	Size (approx kDa)	Function	Location
M	Membrane	26	Virus assembly, host interactions, interferogenic activity	Membrane
E	Envelope	10	Virus assembly, apoptosis, cation selective ion channel activity	Membrane
ORF3	ORF3	26	Unknown	Membrane
N	Nucelocapsid	42	Pathogenesis and transcription	Cytoplasm
S	Spike	150-200	Infection of target cells, cell to cell spread, role in replication, evasion of host immune response	Membrane

1.2.4. Replication and life cycle

It is commonly known that RNA viruses have extremely high mutation rates – up to one mutation per replication cycle (Moya *et al.,* 2004; Drake and Holland, 1999). This leads to the quasispecies concept that says that RNA viruses can exist as a dynamic distribution of variants within a closely related, yet non-identical genome centered on a master sequence (Domingo, 2002). HCoV-NL63 also has a high frequency of recombination because of its unique method of replication. This is due to the switching of random templates during RNA replication. This mechanism is called the "copy choice" mechanism (Herrewegh *et al.,* 1998; Lai *et al.,* 1985). These features allow high mutation rates enabling these viruses to adapt to new circumstances and hosts very rapidly.

When the virus meets the target cell, it attaches to the ACE2 receptor with its spike protein. The spike is divided into S1 and S2 parts. The S1 part is situated at the amino terminal end and contains the receptor binding domain. The S2 part is situated at the carboxy terminal end of the spike and has the fusion peptide, a membrane spanning region and two heptad repeat regions – HR1 and HR2 (Van der Hoek *et al.,* 2006). This binding of the spike with the ACE 2 molecule causes a conformational change to occur in the spike protein leading to the fusion of the viral and target cell membranes (Figure 4). The nucleocapsid of the virus can now be released into the target cell (Matsuyama and Taguchi, 2002; Saeki *et al.,* 1997).

The first event after release of the viral nuclear material is replication of the viral genome. The 1a and 1b genes at the 5' end of the genome are translated into their respective polypeptides pp1a and pp1ab directly from the genomic RNA. The translation of pp1ab can only occur if there is a -1 ribosomal frameshift. This occurs via a slippery sequence in a potential RNA pseudoknot in the RdRp gene. The knot has an 11 base pair stem forming a hairpin with an 8 nucleotide loop.

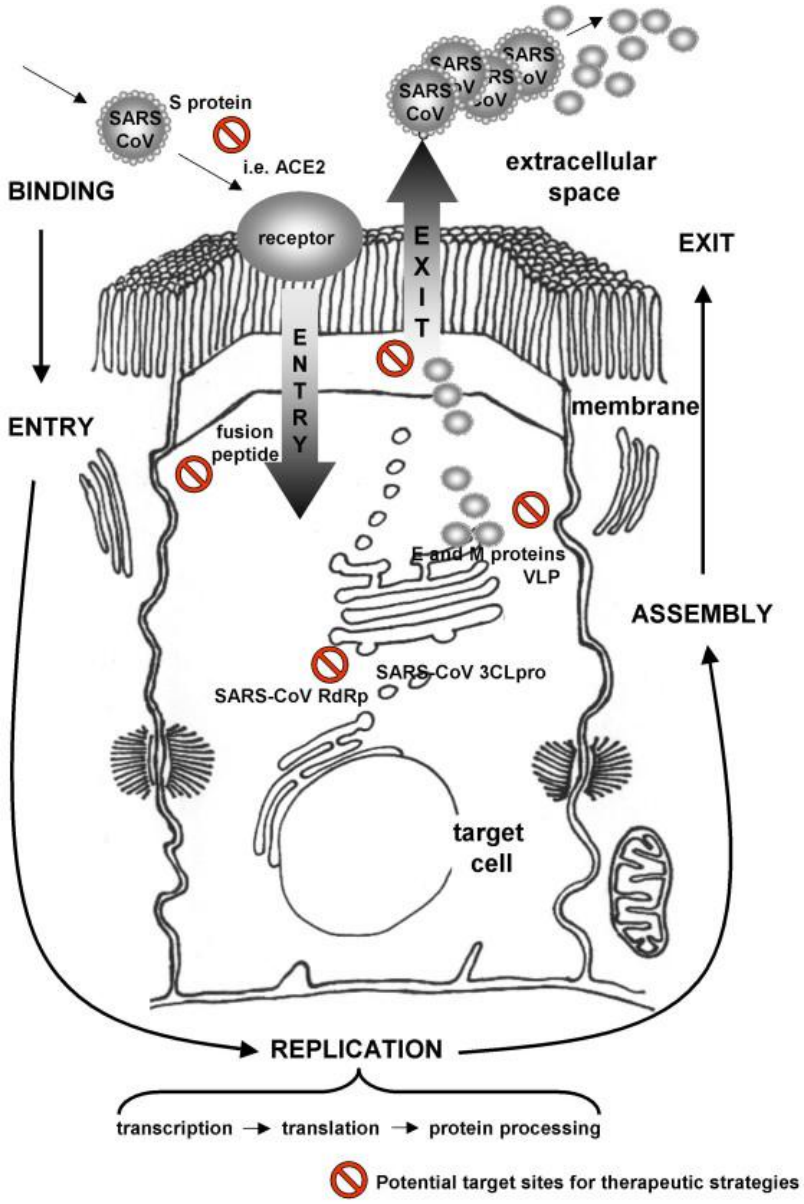


Figure 4. Life cycle of coronaviruses (Groneberg *et al., 2005*)

A sequence 167 nucleotides downstream of this pseudoknot form five base pairs with the loop to complete the structure. The slippery sequence is a sequence of seven nucleotides (UUUAAAC) slightly upstream of the pseudoknot.

The expression of the regulatory proteins is controlled by posttranslational enzymatic processing. A multisubunit is constructed under the influence of various

enzymes that cleave the 1a and 1b protein. This subunit is responsible for replication and transcription. Open reading frame (ORF) 1a codes for the papain-like protease PL^{pro} and a serine-like protease M^{pro}. These proteases cause cleavage of the 1a/1b polypeptide. Various other enzymes are coded for and produce to help with regulatory functions. ORF1b codes for a RNA dependant RNA polymerase (RdRp) and helicase as well as a putative 3'-5' exonuclease (ExoN), poly(U)-specific endoribonuclease (XendoU) and putative S-adenosylmethionine-dependant ribose 2'-*O*-methyltransferase. These enzymes may also play a role in viral metabolism and in interference of the host cell metabolism.

For the creation of the structural proteins, a set of subgenomic RNAs are made. These include subgenomic RNAs for the spike, membrane, envelope, nucleocapsid and ORF3. All of these subgenomic RNAs have an untranslated leader sequence of 72 base pairs. These sequences are identical and are derived from the amino terminal end of the genomic RNA. The exact mechanism whereby these subgenomic RNAs are created is not entirely clear but it is thought to involve an inimitable method of discontinuous transcription.

During the negative strand synthesis there is a halt at the transcription regulatory sequence (TRS). There is a base pairing between the cis-acting TRS in the 5' end of the viral genome and the TRSs upstream of each of the other ORFs. This leads to a chain transfer event taking place whereby the nascent RNA chain switches from the regulatory sequence on the negative strand to one situated in the 5' untranslated region of the genome. Antileader sequences are added to the 3' ends of all the negative strand RNAs (AACUAAA). These antileader sequences can now serve as the templates for the synthesis of mRNAs and are highly conserved in all of these subgenomic RNAs. The only exception to this rule is in the subgenomic RNA of the E protein. It is thought that this protein has an unfavourable antileader

sequence of AACUAUA. For this reason this amount of subgenomic mRNA is less than for the mRNAs of the other proteins with normal antileaders. The most profuse subgenomic mRNA is that of the N protein and has inverse proportionality between the distance of a gene to the 3' untranslated region and its RNA expression level. Translation of the structural proteins of the virus can now take place. This happens mostly exclusively from the 5' ORF.

Posttranslationally, the membrane and envelope proteins move to the Golgi intracellular membranes and are situated close to the endoplasmic reticulum Golgi intermediate compartment. This is the proposed site from where the virus will eventually bud off. There are various proteins involved in the process budding but even in their absence M and E combined will still form virus-like particles. The spike will locate itself on both the plasma and the intracellular membranes from where it will interact with M during virus assembly. N protein will complex with the genomic RNA again to form the original helical structures and will also interact with the M protein. Vesicles are formed into which the virus can bud. These vesicles will move to the surface of the cell from which it will be released.

1.3. Human Coronavirus NL63 (HCoV-NL63)

1.3.1. Virus cell entry

As HCoV-NL63 shares 56% amino acid homology with its closest relative 229E (Fouchier *et al.*, 2004; Pyrc *et al.*, 2004; Van der Hoek *et al.*, 2004), one would assume that the virus will use the same receptor (aminopeptidase N/CD 13) for entry (Delmas *et al*, 1992; Yeager *et al*, 1992). It has, however, been shown that NL63 does not use this receptor for cell entry at all (Hofmann *et al.,* 2005). The reason for this being that the NL63 virus has a unique 179 bp sequence at the N-terminus (Van der Hoek *et al*, 2004). With the use of a technique employing pseudovirions, Hofmann *et al* (2005) showed that this virus is able to use the zinc metalloprotease angiotensin-converting enzyme (ACE) 2, a characteristic it shares with the SARS coronavirus (Li *et al.,* 2003). The ACE2 receptor and the ACE protein are homologues of each other, both of which plays a vital role in controlling the renin-angiotensin system (Rice *et al*, 2004, Danilczyk *et al*, 2003, Crackower *et al*, 2002). The ACE2 molecule is found on the arterial and venous endothelium, epithelium of the small intestine, arterial smooth muscle cells and on the surface of ciliated cells of human nasal and tracheobronchial epithelium (Sims *et al.,* 2005; Hamming *et al.,* 2004). In the lung this system plays an important role in acute lung injury when it is activated (Imai *et al.,* 2005). ACE is responsible for the activation of the system by the cleavage of the angiotensin I into angiotensin II. ACE 2 on the other hand protects the lungs by inactivating angiotensin II and therefore halting the damage caused by ACE. During SARS infection, it is thought that the virus partially inhibits the expression of ACE2 and that this in part the reason for the severe lung failure (Kuba *et al.,* 2005). Considering that the action of SARS and NL63 both has the same effect on this system, it would appear incongruent that HCoV-NL63 has such a

mild pathogenicity in comparison. This might suggest that this binding and inhibition of expression of ACE 2 is not the only event in determining the severity of these two viruses and thus warrants further investigation (Van der Hoek *et al.,* 2006).

1.3.2. Culturing in permissible mammalian cell lines

The receptor usage of a virus also inevitably determines which cell lines and hosts the virus can use for replication (Kuo *et al*, 2000; Holmes, 1996). To determine which cells are suitable for HCoV-NL63, Schildgen *et al* (2006) tested six cell lines that are routinely used for viral culture. The lines they tested included Human lung fibroblasts (MRC-5: ATCC-CCL-171), rhabdomyosarcoma cells (A-204; ATCC-ACC-250), Madin-Darby-Canine-kidney cells (MDCK: ATCC-CCL-34), monkey kidney cells (Vero-B4 cells: ATCC-ACC-33; LLC-MK2 cells: ATCC-CCL7), and laboratory adapted non-ATCC-listed monkey kidney cell line monkey stable (MS). Supernatant of the cell culture were sampled and the RNA was extracted and virus yield determined by polymerase chain reaction (PCR). The cells were also monitored for cytopathic effect (CPE).

The cell lines found to be permissive for HCoV-NL63 were the monkey kidney cell lines LLC-MK2 and Vero-B4. The titers of the virus increased up to 4000 fold in these cell lines. The rise in titers of the other cell lines was negligible. All cell lines showed CPE except the human lung fibroblasts. The fact that these cells (MDCK, MS and RD) showed CPE but no increase in viral titers may be due to the inoculation of NL63 being toxic to these cells initially. It could also be that these cell lines are simply not able to secrete the virus particle into the supernatant and therefore appear not to have a rise in titers. Regardless of the reason, these cells are not seen as being sustainable for the culture of HCoV-NL63.

Other studies have shown that there are other cell lines that are also permissive for infection from this virus. An example would be the human hepatocellular carcinoma cell line Huh-7, another characteristic NL63 shares with SARS (Hofmann *et al.,* 2005).

The question might arise as to why this virus was not previously discovered if it can then be cultured in routine laboratory cell lines. The reason being that the only parameter being routinely checked for the presence of viruses in cell lines is the observation of CPE. The CPE observed for NL63 appears to be identical to picorna viruses and may therefore have been observed before but have been mistaken for these known viruses (Schildgen *et al.,* 2006)

Even though coronaviruses can be cultured, they do not always grow very well and they are not easy to isolate and study. The failure to previously detect these viruses is due to unreliable methods that are time consuming and not sensitive enough. The unavailability of proper lower respiratory tract specimens further hampered the discovery of these new viruses. The most common methods currently being used for detection of these viruses include viral culture and direct immunofluorescence and more recently molecular methods like reverse transcription (RT) PCR (Vabret *et al*, 2001). Even though viral culture has been considered the gold standard, it can only detect small numbers of viruses and results depend greatly on quality of sampling, storage, transport and type of cells used. It also takes several weeks and can therefore not be used as a diagnostic tool. Direct immunofluorescence results are more rapid and have been shown to improve patient outcome with less antibiotic use and shorter hospital stays (Barenfanger, 2000; Adcock *et al*, 1997; Woo *et al*, 1997). Vijgen *et al* have now developed a new one-step real-time quantitive reverse transcriptase PCR assay for HCoV OC43 and 229E that is even more reliable, sensitive and time efficient for detection of these

respiratory viruses (Vijgen *et al*, 2005). Recently Moës *et al* adapted this technique to detect HCoV-NL63 (Moës *et al*, 2005).

1.3.3. Clinical infection, symptoms and prevalence

This virus causes not only upper respiratory tract infection but also lower respiratory tract infection and can even cause pneumonia (Bastien *et al.,* 2005; Vabret *et al.,* 2005). General symptoms of infection with NL63 can include any of the following: fever, cough, coryza, rhinorrhoea, pharyngitis, sore throat, hoarseness, bronchitis, bronchiolitis, pneumonia, and croup (Arden *et al.,* 2005; Bastien *et al.,* 2005; Chiu *et al.,* 2005; Moës *et al*., 2005; Vabret *et al.,*2005; Fouchier *et al.,* 2004; Konig *et al.,* 2004; Forster *et al.,* 2004).

Since its discovery in the Netherlands, various research groups have found this virus in countries all over the world. These countries include Australia (Arden *et al*, 2005), Japan (Ebihara *et al.,* 2005; Suzuki *et al.,* 2005), Belgium (Moës *et al.,* 2005), Hong Kong (Chiu *et al.,*2005), Taiwan (Wu *et al.,* 2007) Korea (Choi *et al.,* 2006), Canada (Bastien *et al.,* 2005), France (Vabret *et al.,* 2005), Switzerland (Garbino *et al.,* 2006; Kaiser *et al.,* 2005), Germany (Van der Hoek *et al.,* 2005), and Sweden (Koetz *et al.,* 2006). This shows that this virus occurs all over the world in tropical, subtropical and temperate countries (Table 2). The virus has shown a seasonal prevalence especially in the temperate countries where it preferred the colder climate in winter, a characteristic it shares with HCoV-229E and HCoV-OC43 (Arden *et al.* 2005; Bastien *et al.,* 2005; Ebihara *et al.,* 2005; Moës *et al.,* 2005; Vabret *et al.,* 2005; Van der Hoek *et al.,* 2005 Hendley *et al.,*1972). In the subtropical and tropical areas the virus was more prevalent in the early spring (Chiu *et al.,* 2005).

Table 2. Seasonal prevalence of HCoV-NL63

Country	Positive samples	Percentage	Seasonality
Australia	16/766	2.1	Winter
Belgium	7/279	2.3	Winter
Canada	45/1765	2.5	Winter
France	28/300	9.3	Winter
Germany	49/949	5.2	Winter
Hong Kong	15/587	2.6	Spring
Japan	5/419	1.2	Winter
Switzerland	6/82	7.0	Winter

As stated previously, HCoV-NL63 affects the young and the elderly but it has been shown in all these studies that the prevalence for infection is highest in children under five years old. It also more frequently infects the elderly and immunocompromised adults and often leads to hospitalization due to severity.

Many studies have been performed (Table 3) to ascertain how frequent NL63 infection is in comparison with other human coronaviruses commonly known to cause respiratory disease (Chiu *et al.,* 2005; Kaiser *et al. ,*2005; Moës *et al.,* 2005; Vabret *et al.,* 2005; Van der Hoek *et al.,* 2005). Based on the results, it is clear to see that this virus causes a substantial amount of infections when compared to its

relatives in other groups. Moës *et al* (2005) showed that of all the human coronaviruses it is NL63 and OC43 that causes the most hospitalization.

Table 3. Frequency of NL63 in comparison to other human coronaviruses

Country	Specimens	% NL63	% 229E	% OC43
Belgium	309	2.3	0.3	4.7
France	300	9.3	0	2.2
Hong Kong	587	2.6	0.3	1.5
Netherlands	63	4.7	0	7.7
Switzerland	82	7.3	3.7	6.1

There have been many reports and much speculation about the association of the human coronaviruses with other diseases. Some of these associations include hepatitis, multiple sclerosis and newborn enteric diseases. Esper *et al* (2005) claimed in one of their studies that human coronavirus NL63 might be associated with Kawasaki disease, one of the most common causes of childhood vasculitis (Burns and Glode, 2004). They tested respiratory tract specimens of eleven Kawasaki Disease patients and found eight tested positive for HCoV-NL63. This seems like a significant discovery as the cause of this disease has eluded scientists for many years and many claims as to the aetiology have been made in the past. Unfortunately it seems that this study might have been biased and many follow-up studies have been done – none of which could substantiate this claim (Chang *et al.,* 2006; Belay *et al.,* 2005; Ebihara *et al.,* 2005; Shimizu *et al.,* 2005).

It has however been proven that NL63 is associated with croup, bronchiolitis and pneumonia (Arden *et al.,* 2005; Bastien *et al.,* 2005; Chiu *et al.,* 2005; Moës *et al.,* 2005; Vabret *et al.,* 2005; Forster *et al.,* 2004; Konig *et al.,* 2004; Fouchier *et al.,* 2004). In fact, HCoV has been shown to be one of the three leading viruses to cause croup (laryngotracheitis), the other two being influenza and parainfluenza. Croup is characterized by a loud barking cough that may be exacerbated at night. In this same study it was found that 50% of patients positive for the virus also had croup (Forster *et al.,* 2004; Konig *et al.,* 2004).

There have also been many studies that did not reflect these same statistics (Arden *et al.,* 2005; Bastien *et al.,* 2005; Moës *et al.,* 2005; Vabret *et al.,* 2005 Fouchier *et al.,* 2004) but the reason for this was most likely the patient selection. Depending on the sample group, patients with croup may not have been included. As croup patients are very rarely admitted to hospital or subjected to disease investigation, studies that were hospital based would not necessarily have had samples from these patients (Van der Hoek *et al.,* 2006). Nonetheless, it can still be said with relative certainty that HCoV-NL63 is one of the most important causes for this disease.

1.3.3.1. Co-infections with other respiratory viruses

Many groups have reported that the occurrence of infection with HCoV-NL63 combined with another respiratory virus is not at all a rare event. In fact, Van der Hoek *et al* (2005) showed that this occurs in up to 50% of cases. They have also found that the doubly infected patients are more likely to be hospitalized, indicating the severity of this kind of superinfection. NL63 has been found in conjunction with influenza A (Chiu *et al.,* 2005), respiratory syncytial virus (RSV) (Arden *et al.,* 2005; Esper *et al.,* 2005; Van der Hoek *et al.,* 2005), parainfluenza (Arden *et al.,* 2005; Van

der Hoek *et al.,* 2005) and hMPV (Arden *et al.,* 2005; Esper *et al.,* 2005). The viral load of NL63 is lower in these coinfected patients than that of patients infected only with NL63 (Chiu *et al.,* 2005; Van der Hoek *et al.,* 2005). There are three possible explanations for this phenomenon. Firstly, the NL63 might be the initial infection that breaks down the immune system enough for a second infection to occur. By the time this second infection shows symptoms, the NL63 infection might have already been brought under control by the host immune system. A second option is that the two viruses may be in competition for the same receptor, and, lastly, the immune response triggered by the second respiratory virus may cause inhibition of NL63 (Van der Hoek *et al.,* 2006).

1.3.4. Treatment of infection

Even though the severity of disease caused by HCoV-NL63 is generally mild and self-limiting, it has to be taken into account that this virus has been known to cause more serious disease. HCoV-NL63 infection quite often leads to hospitalization and was the cause of death in at least one case (Bastien *et al.,* 2005). It is therefore important to develop an antiviral treatment for this virus. Some headway has been made in this field. There are various known inhibitors for some coronaviruses, all of which work by reducing replication of the virus (Pyrc *et al.,* 2006; Stiehm *et al.,* 1987). To date there are four ways to successfully control infection with NL63: pyrimidine nucleoside analogues, HR2 domain derived peptides, small interfering RNAs, IV immunoglobulin and protease inhibitors. The pyrimidine nucleoside analogues B-D-N^4-hydroxycytidine and 6-azauridine was shown by Pyrc *et al.* (2006) to inhibit NL63 on a transcriptional level. The precise mode of action for this inhibition is still unclear.

As mentioned previously, HR1 and HR2 form a six-helix bundle that enables the fusion of the virus and the target cell. This fusion can be inhibited by peptides derived from the HR2 region by interacting with the HR1 and thus blocking the formation of the natural HR1/HR2 complex (Pyrc *et al.,* 2006; Bosch *et al.,* 2004).

Low amounts of two types of small interfering RNA, targeting HCoV-NL63, made cells resistant to infection as shown by Pyrc *et al* (2006). RSV (Hemming *et al.,* 1987), Kawasaki Disease (Stiehm *et al.,* 1987) and various other respiratory diseases has been shown to be treated successfully by the use of intravenous immunoglobulin. This proved to be the case for HCoV-NL63 as well (Pyrc *et al.,* 2006). A broadspectrum antiviral agent can be developed containing a protease inhibitor acting against M^{pro}. This will be effective against several coronaviruses as this is a very highly conserved portion of the genome and would thus be found in most coronaviruses. N3, an inhibitor designed by Yang *et al*, has indeed shown to have broad spectrum inhibition of M^{pro} enzymes (Yang *et al.,* 2005).

1.4. Objectives of this Thesis

Human coronavirus NL-63 is an emerging virus and has, as yet, not been characterized in great detail. No molecular studies on the structural proteins from HCoV-NL63 have been published. Although the main features of infection and replication of the other coronaviruses have been studied in detail the past few decades, the mechanism of infection and pathogenicty of HCoV-NL63 needs to be understood. In light of this, it will be endeavored to partly characterise the HCoV-NL63 M gene by cloning it into various expression vectors and then expressing it in various cellular systems including bacterial cells, insect cells and mammalian cells. The resulting knowledge gained from these experiments could then in future be used for further studies into the characterization of this and other human coronaviruses.

Chapter 2: *Materials and Methods*

2.1. *In silico* analysis of HCoV-NL63 M protein

2.1.1 Predictive analysis of HCoV-NL63 M

TMHMM Server v. 2.0 (Krogh *et al.*, 2001) was used to search for transmembrane regions in the M sequence.

2.1.2. Comparison of HCoV-NL63 M homologues

The putative amino acid sequence of HCoV-NL63 M was compared to sequences in the GenBank database at the National Centre for Biotechnology by using the Basic Blast Search Server (Altschul *et al.*, 1990). Identified HCoV-NL63 M-homologues were subsequently aligned with CLUSTAL X v 1.81 (Thompson *et al.*, 1997) and viewed with GENEDOC software (Nicholas and Nicholas, 1997).

2.2. Overview of common molecular biology techniques used throughout Chapter 2 (Refer to Appendix C for all solutions made up for study)

2.2.1 Reverse transcription

Reverse transcription is a process where single stranded RNA is reverse transcribed into complementary DNA (cDNA) with the help of reverse transcriptase. RT PCR was performed by using the kit from Promega. Table 4 shows the reverse transcription mix and Table 5 the conditions used in this study.

Table 4: Reverse transcription mix used in this study

Component	Volume
MgCl (25mM)	4,0 µl
10 x Buffer	2,0 µl
dNTP's (10mM)	2,0 µl
RNAsin	0,5 µl
ANIV Reverse Transcriptase (25u/µl)	0,6 µl
Oligo (dT) primer (100µM)	1,0 µl
HCoV-NL63 RNA (Amsterdam 1)	2,0 µl
Nuclease free water	7,9 µl
Final volume	20 µl

Table 5: Reverse transcription program conditions

	Time	Temperature
Reverse transcription	60 min	42°C
Deanaturation	5 min	95°C
Extension	5 min	0°C

2.2.2. Polymerase chain reaction (PCR)

PCR is an *in vitro* enzymatic amplification of defined DNA sequences to produce a high yield of amplified target DNA. This happens under the influence of specifically designed primers and a thermostable DNA polymerase of *Thermus aquaticus* (Taq). The process consists of three distinct steps. The first is denaturation of double stranded DNA followed by annealing of the primers to their complementary sequences on the template. The last step is extension by incorporation of nucleotides under the influence of Taq polymerase. With numerous repetitions of this set of steps, the number of copies of the target sequence rises exponentially. Tables 6-8 shows the PCR mixture components, program conditions and PCR primers used in this study.

2.2.2.1. Colony PCR

The principle for a colony PCR is exactly the same as for a normal PCR except that the template DNA is taken directly from a bacterial colony growing on an agar plate.

A sterilised bacterial loop was touched to the desired colony and was then used to inoculate the PCR mix discussed above. The program stays the same bar the extension of the initial denaturation time to 4 minutes.

Table 6. PCR mixture components

Component	**Volume**	
	PCR	**Negative Control**
Template	4 µl	0 µl
dNTP (10mM)	0,5 µl	0,5 µl
MgCl (25mM)	1,5 µl	1,5 µl
10 x buffer	2,5 µl	2,5 µl
Forward primer (100µM)	1,0 µl	1,0 µl
Reverse primer (100µM)	1,0 µl	1,0 µl
Taq (5u/µl)	0,5 µl	0,5 µl
Nuclease free water	14 µl	18 µl
Final volume	25 µl	25 µl

Table 7. PCR program conditions

		Time	Temp
	Initial denaturation	3 min	95°C
30 cycles	Denaturation	45 s	95°C
	Annealing	60 s	45°C
	Elongation	60 s	72°C
	Final elongation	15 min	72°C

Table 8. Polymerase chain reaction primers used in this study (Refer to Appendix A for vector maps)

Primer name	Restriction sites for cloning	Vector(s) used for cloning	Construct name
B2Bfor1	*Not*I/SaI	pFastBac	XS2/XN5
B2Brev1	*Xba*I	pFastBac	
CMV20for2	*EcoR*I	pFlag	pFlag6.2
CMV20rev2	*Xba*I	pFlag	
M-FF-for	*Smf*I/*Pme*I	pFlexi	Flexi2

2.2.3. PCR product purification and extraction of nucleic acid fragments from gels

DNA was extracted from electrophoresis gels with Promega's Wizard SV Gel and PCR Clean-Up System. The same kit was also used to purify PCR products. The desired band was cut out of the electrophoresis gel and its weight determined. 10 µl of Membrane Binding Solution was added for every 10 mg of the gel slice. The mixture was vortexed and incubated at 60°C until the gel was dissolved. The tube was centrifuged and contents transferred to a SV Minicolumn placed in a Collection tube. DNA was bound to the membrane by centrifugation and washed twice with Membrane Wash Solution diluted with 95% ethanol. DNA was eluted of the membrane with 30 µl nuclease free water. The same kit was also used to purify PCR products by adding equal amount of the Membrane Binding Solution to the PCR product and then following the steps as outlined above.

2.2.4. Agarose gel electrophoresis

Agarose gel electrophoresis is a technique used for the separation of nucleic acids based on charge, size and conformation. DNA is negatively charged and will migrate away from the negative pole through the agarose gel to the positive pole when the gel is placed in buffer in an electrical field. The molecules will separate out with the larger and more folded molecules moving slower through the gel and settling closer to the origin than the smaller less folded molecules that are able to move faster through the gel. Markers with proteins at known sizes are run on all gels as reference for determination of size of test DNA.

1-2% agarose gels were made with 1 x Tris Borate EDTA (TBE) made up from 10x TBE with distilled water. Gels were run in TBE at between 50 V and 80 V. 5-25 µl of DNA was loaded in each well after addition of 5 µl of 6x sample loading buffer.

Blue/Orange Loading Dye, (Promega) contains 0.4% orange G, 0.03% bromophenol blue, 0.03% xylene cyanol FF, 15% Ficoll® 400, 10 mM Tris-HCl (pH 7.5) and 50 mM EDTA (pH 8.0). The dyes allow for tracking of migration of nucleic acids through the gel as the xylene cyanol FF migrates at approximately 4 kb, bromophenol blue at approximately 300 bp and orange G at approximately 50 bp in a 0.5–1.4% agarose gel in 0.5X TBE. 1 kb and 100 bp DNA ladders in 1x Blue/Orange loading dye (Promega) were run on all gels. Ethidium bromide (10 mg/ml) from Promega was added to agarose gel before it set. This dye intercalates with DNA running through the gel and fluoresces under ultra violet (UV) light enabling visualisation of DNA bands.

2.2.5. Restriction Endonuclease digests

Restriction endonuclease digestion works on the principle of each enzyme recognising its unique palindrome and cutting it in a specific way to yield either blunt or sticky ends that can in turn be ligated into a designated site of a vector. This method was used for digestion of DNA fragments for cloning into vectors, to confirm presence of inserts in ligated and transformed products and for linearising vectors. All enzymes were obtained from Promega and generally the same protocol was adhered to. All reactions were incubated at 37°C for an hour followed by the addition of 0.5 µl of the related enzyme and incubated for an additional hour. Products were separated and retrieved from agarose gels after electrophoresis. Table 9 shows the restriction endonuclease mixture used for digests.

Table 9. Restriction endonuclease mix

Component	Volume
Nuclease free water	2,3 µl
Buffer	2 µl
BSA	0,2 µl
DNA	15 µl
Enzyme	0,5 µl
Final volume	20 µl

2.2.6. Ligation

DNA fragments of interest can be cloned into plasmids which can be transformed into competent bacterial cells. These bacterial cells will replicate the plasmids with its own DNA creating multiple clones. After replication, plasmids can be removed for further experimentation. For ligation, the ligation mix was made up (as showed in Table 10) and incubated overnight at 4°C. Only for ligation of DNA into KRX cells for bacterial expression, was the ligation reaction left at room temperature for one hour.

Table 10. Ligation mix

Component	Volume
Purified DNA	16 µl
Ligation buffer	2 µl
T4 DNA ligase (3u/µl)	1 µl
Vector	1 µl
Final volume	20 µl

2.2.7. Plasmid DNA preparation (Minipreps)

Plasmid DNA of overnight cultures (approximately 14 hours) were purified with the Promega Wizard Plus SV Minipreps DNA Purification kit. Briefly, 5 ml of culture was pelleted by centrifugation. The supernatant was discarded and 250 µl Cell Lysis Solution was added and incubated until the solution cleared. Alkaline Protease was added to inactivate any endonucleases or proteins released during lysis that could have an adverse effect on the DNA being isolated. Neutralization Solution was added and suspension was cleared by centrifugation. Cleared lysate was transferred to a spin column and DNA bound to the membrane by centrifugation. The DNA was washed twice with Column Wash Solution and eluted from the membrane with 30 µl nuclease free water.

2.2.8. Freezing of mammalian and insect cells

Cos 1 and Sf9 cells were frozen down at regular intervals to ensure constant supply of cells. This was achieved by adding 1 ml DMSO to 2 ml FBS and 7 ml of cells suspended in DMEM. Cells were then frozen at -80°C in cryovials.

2.2.9. SDS-PAGE

When determining the molecular weight of a protein, general electrophoresis will not suffice. This is because the charge of all the proteins running on the SDS-PAGE gel has to be the same. To achieve this, SDS is added to the gel. SDS is an anionic detergent that will linearise proteins by denaturation. It then binds to the proteins to give them a uniform charge. The migration of the proteins in the gel will now only be dependent on size. B-mercpatoethanol is also added to the samples to aid in the reduction of disulphide bond in the proteins. The proteins run in a gel matrix that is formed by the polymerisation of acrylamide and the crosslinkage of the N'N-methylene bisacrylamide with ammonium persulphate (APS) and N'N'N'N tetramethylethylene diamine (TEMED) as catalysts.

15% SDS-PAGE gels were made up using the kit (Sigma). Discontinuous gels were made up to ensure all proteins start at the same point before separating. SDS sample buffer was added to all samples and samples were boiled at 95°C for 2 minutes before loading 5-25 µl of the sample per well. Coloured protein markers were also loaded as size reference. Gels were run at 40 mA for approximately 90 minutes. Gels were now either stained with Coomassie brilliant blue or used for Western Blotting.

2.2.10. Western Blotting

Western blotting refers to the electrotransfer of proteins from a SDS-PAGE gel onto a nitrocellulose membrane. Transferred proteins are now detectable by using labelled antibodies specific for the relevant protein. Substrate is added for detection purposes.

Before use, the nitrocellulose membrane was soaked in pure methanol and rinsed with distilled water. The membrane, SDS-PAGE gel, sponges and blotting paper was soaked in 1x running buffer for 15 minutes at RT. The sandwich was assembled in the holders and placed in the tank with 1x transfer buffer. Transfers were run at 40V for 6 hours at 4°C. Membranes were stained with Poncheau Red until bands appeared. This serves as confirmation that protein transfer was successful. After rinsing with distilled water, membranes were blocked from non-specific antibody binding by washing in a 3% fat free milk solution for 30 minutes. The membrane was then incubated overnight in the primary antibody (1:5000 in 3% milk solution) on rollers at 4°C. Membranes were washed in PBS (with 0.5% Tween) for 45 minutes and then exposed to secondary antibody (1:2000 in 3% milk solution) for 1 hour at room temperature on rollers. The membrane was washed in PBS (with 0.5% Tween) for 45 minutes and exposed to a colour reaction.

2.2.11. Coomassie blue staining

Coomassie Blue R250 is a stain that binds non-specifically to virtually all proteins. It is a two step process whereby the SDS-PAGE gel is over-stained with the staining solution turning the entire gel blue. This is now followed by a destaining step which removes excess stain revealing proteins bands on a clear gel. SDS-PAGE gels were left in the staining solution to stain overnight (12-14hrs). Stained gels were destained with destaining solution until excess Coomassie was removed from the gels and a clear background was revealed. The destaining solutions were replenished several times during the process. Destained gels were dried using a gel drying kit from Promega.

2.3. Cloning for expression in mammalian system

2.3.1. Plasmid construction

The sequence encoding HCoV-NL63 M gene was amplified by PCR using cDNA generated by RT-PCR from genomic RNA as a template (kind gift from L Van der Hoek, Holland). The oligonucleotide primers used for the amplification of M were forward 5'-gc**gaattc**ATGTCTAATAGTAGTGTGCCTC-3' (sense, *Eco*RI site in bold and translational start codon underlined) and for reverse 5-gc**tctaga**TaAGATTAAATGAAGCAATTCTC-3' (antisense, *Xba*I site in bold – no stop codon). PCR products were subjected to agarose gel electrophoresis.

Following gel extraction and purification of the appropriate bands, the product was ligated into pGEM vector using the Promega pGEM-T Easy Vector Systems. Ligation reactions were transformed into competent JM109 *Escherichia coli (E.coli)* cells and transformation reactions were plated onto LB agar plates. Plates were incubated overnight at 37°C and white colonies were selected from the plates. Bacterial cells from white colonies were used as template for colony PCRs to confirm presence of insert. Luria Bertani (LB) broth containing Ampicillin (100 mg/ml) was inoculated with the white colonies and grown overnight at 37°C with agitation and miniprepped the following day. Constructs were digested with *Xba1* and *EcoR1* to confirm the presence of the gene of interest. Digestions were visualised on agarose electrophoresis gels. The full-length M gene was DNA sequenced (Inqaba Biotech) and the insert was verified by comparison to the existing sequence for the corresponding gene on Genbank (AY697422). For ligation protocols, refer to 2.2.6.

2.3.2. Transformation of ligations into pCMV-myc vector for mammalian expression

Once the DNA fragment of interest is successfully ligated into a vector, it can be transformed into competent cells. Competent cells are bacterial cells capable of taking up naked fragments of DNA into its cytoplasm. With replication of the cell, these fragments will be replicated with the cell's native DNA.

4 µl of each ligation reaction was added to 50 µl of the competent cells. This was incubated on ice for twenty minutes, heat shocked at 42°C for 45 seconds, followed by another incubation on ice for 2 minutes. 950 µl of LB medium, prewarmed to 37°C, was added to the reaction and it was incubated for 90 min at 37°C with agitation. 200 µl was spread plated onto LB agar plates (10 g/L NaCl, 5 g/L bacto tryptone, 5 g bacto yeast extract, 5 g/L bacteriological agar) containing 100 µl IPTG (100 mM), 20 µl X-Gal (50 mg/ml) and 10 µl Ampicillin (100 mg/ml). The plates were incubated overnight at 37°C. Appropriate colonies were picked off plates with a sterile bacteriological loop and inoculated into LB broth containing Ampicillin (100 mg/ml). Cultures were grown up overnight at 37°C with agitation.

2.3.3. Plasmid DNA preparation (Minipreps)

Plasmid DNA of overnight cultures (approximately 14 hours) were purified with the Promega Wizard Plus SV Minipreps DNA Purification kit. Briefly, 5 ml of culture was pelleted by centrifugation, the supernatant discarded and 250 µl Cell Lysis Solution was added and incubated until the solution cleared. Alkaline Protease was added to inactivate any endonucleases or proteins released during lysis that could have an adverse effect on the DNA being isolated. Neutralization Solution was added and suspension was cleared by centrifugation. Cleared lysate was transferred

to a spin column and DNA bound to the membrane by centrifugation. The DNA was washed twice with Column Wash Solution and eluted from the membrane with 30µl nuclease free water. The products was DNA sequenced (Inqaba Biotech) and the insert was verified by comparison to the existing sequence for the corresponding gene on Genbank (AY697422)

2.4. Cloning for expression in Baculovirus system

2.4.1. Plasmid construction

The sequence encoding HCoV-NL63 M gene was amplified by PCR using cDNA generated by RT-PCR from genomic RNA as a template (kind gift from L Van der Hoek, Holland). The oligonucleotide primers used for the amplification of M were forward 5'-gc**gaattc**aATGTCTAATAGTAGTGTGCCTC-3' (sense, *Eco*RI site in bold and translational start codon underlined) and for reverse 5-gc**tctaga**TTAGATTAAATGAAGCAATTCTC-3' (antisense, *Xba*I site in bold and stop codon underlined). PCR products were subjected to agarose gel electrophoresis.

Following gel extraction and purification of the appropriate bands, the product was ligated into pGEM vector using the Promega pGEM-T Easy Vector Systems. Ligation reactions were transformed into competent JM109 *E.coli* cells and transformation reactions were plated onto LB agar plates. Plates were incubated overnight at 37°C and white colonies were selected from the plates. Bacterial cells from white colonies were used as template for colony PCRs to confirm presence of insert. LB broth containing Ampicillin (100 mg/ml) was inoculated with the white colonies and grown overnight at 37°C with agitation and miniprepped the following day. Constructs were digested with *Xba1* and *EcoR1* to confirm the presence of the gene of interest. Digestions were visualised on agarose electrophoresis gels. The full-length M gene was DNA sequenced (Inqaba Biotech) and the insert was verified by comparison to the existing sequence for the corresponding gene on Genbank (AY697422). For ligation protocol refer to 2.2.6.

2.4.2. Transformation of ligations into pFastBac for baculovirus expression

1 ng of ligated DNA reaction was added to 50 µl of DH10Bac competent cells. This was incubated on ice for 30 minutes, heat shocked at 42°C for 45 seconds, followed by another incubation on ice for 2 minutes. 950 µl of SOC medium was added to the reaction and it was incubated for 4 hrs at 37°C with agitation (225 rpm). Cells were serially diluted with SOC medium to 10^{-1}, 10^{-2} and 10^{-3}. 100 µl of each of these dilutions were spread onto LB agar plates containing 50 µg/ml Kanamycin, 7 µg/ml Gentamicin, 10 µg/ml tetracycline, 40 µg/ml IPTG and 100µg/ml X-gal. Plates were incubated at 37°C for 48 hrs. White colonies were selected from the plates containing 100 – 200 colonies (blue colonies also selected to act as negative controls). Appropriate colonies were selected and restreaked onto fresh plates containing all of abovementioned antibiotics, IPTG and X-gal. Plates were incubated overnight at 37°C. Liquid cultures were set up from appropriate confirmed single white colonies (blue colonies picked as negative controls). LB broth with all abovementioned antibiotics were incubated overnight at 37°C with agitation (225 rpm).

2.4.3. Plasmid DNA preparation (Minipreps)

Plasmid DNA of overnight cultures (approximately 14 hours) was purified with the Promega Wizard Plus SV Minipreps DNA Purification kit. Briefly, 5 ml of culture was pelleted by centrifugation. The supernatant was discarded and 250 µl Cell Lysis Solution was added and incubated until the solution cleared. Alkaline Protease was added to inactivate any endonucleases or proteins released during lysis that could have an adverse effect on the DNA being isolated. Neutralization Solution was

added and suspension was cleared by centrifugation. Cleared lysate was transferred to a spin column and DNA bound to the membrane by centrifugation. The DNA was washed twice with Column Wash Solution and eluted from the membrane with 30 µl nuclease free water. The products were then subjected to agarose gel electrophoresis or colony PCR for confirmation of desired insert. The products was DNA sequenced (Inqaba Biotech) and the insert was verified by comparison to the existing sequence for the corresponding gene on Genbank (AY697422)

2.4.4. Cells and viruses

Sf9 cells were cultured in *Sf*9 medium with addition of Penicillin/Streptomycin (50 units/ml penicillin and 50 µg/m streptomycin). Cells were incubated at 27°C in 250 ml flasks. Cells were split every 2-3 days depending on confluency. Cells were split by mechanically dislodging it by repeatedly spraying medium over the growth with a pipette. When cells were released into the medium, it was split into 2 or 3 clean flasks containing fresh media. Flasks were incubated at 27°C in order for cells to attach and reach confluency. The recombinant baculoviruses, containing the coding sequences of M, were generated by using the Bac-to-Bac™ system (Invitrogen). Propagation of the recombinant virus was performed according to standard procedures.

2.4.5. Transfection of Insect cells

9 x 5^{10} Sf9 cells were seeded out into 35 mm 6 well plates containing 2 ml of Sf 900 II Sfm media with penicillin/streptomycin at 0.5X final concentration (50 units/ml penicillin and 50 µg/m streptomycin). Cells were left to attach for 60 minutes. Transfection solutions were prepared according to protocol and left to incubate for 45 minutes at room temperature. Cells were gently washed with SF 900 II Sfm without antibiotics and media was aspirated. Lipid-DNA complexes were layered over cells and incubated for 5 hours at 27°C. The transfection mixture was removed, 2 ml of Sf 900 II Sfm media with antibiotics were added, and the plates were incubated for 72 hours at 27°C. Virus was harvested after 72 hours incubation.

2.4.6. Harvesting recombinant baculovirus P1 and P2 viral stock

Supernatant from infected Sf9 cells were aspirated and clarified by centrifugation for 5 minutes at 500 x *g*, and clarified virus-containing supernatant was removed (P1 viral stock). Six well plates were prepared with 2 x 10^6 cells per well. Cells were left to attach for 1 hr at room temperature. 500 µl of P1 viral stock was added to each well and incubated for 48 hrs at 27^oC in a humidified incubator. Supernatant was collected and clarified by centrifugation (500 x 5 *g* for 5 minutes). These P2 viral stocks were stored at -80^0C for future use.

2.5. Cloning for expression in bacterial system

2.5.1. Plasmid construction

The sequence encoding HCoV-NL63 M gene was amplified by PCR using cDNA generated by RT-PCR from genomic RNA as a template (kind gift from L Van der Hoek, Holland). The oligonucleotide primers used for the amplification of M were forward 5'- M-FF-for- GCGC**GCGATCGC**CATGTCTAATAGTAGTGTGC (sense, SmfI site in bold and translational start codon underlined) and for reverse 5-gc**tctaga**TTAGATTAAATGAAGCAATTCTC-3' (antisense, *Xba*I site in bold and stop codon underlined). PCR products were subjected to agarose gel electrophoresis.

Following gel extraction and purification of the appropriate bands, the product was ligated into pGEM vector using the Promega pGEM-T Easy Vector Systems. Ligation reactions were transformed into competent JM109 *E.coli* cells and transformation reactions were plated onto LB agar plates. Plates were incubated overnight at 37°C and white colonies were selected from the plates. Bacterial cells from white colonies were used as template for colony PCRs to confirm presence of insert. LB broth containing Ampicillin (100 mg/ml) was inoculated with the white colonies and grown overnight at 37°C with agitation and miniprepped the following day. Constructs were digested *EcoR1* to confirm the presence of the gene of interest. Digestions were visualised on agarose electrophoresis gels. The full-length M gene was DNA sequenced (Inqaba Biotech) and the insert was verified by comparison to the existing sequence for the corresponding gene on Genbank (AY697422)

2.5.2. Transformation of ligations into pFlexi for bacterial expression

The circular Flexi vector was digested with Flexi Enzyme Blend (*SgfI* and *PmeI*) to yield a linear vector into which the gene of interest could be cloned (Promega). The reaction was incubated at 37°C for 30 minutes after which it was heated to 65°C for 20 minutes to inactivate the restriction enzymes. For ligation method refer to 2.2.6.

Once the DNA fragment of interest is successfully ligated into a vector, it can be transformed into competent cells. Competent cells are bacterial cells capable of taking up naked fragments of DNA into its cytoplasm. With replication of the cell, these fragments will be replicated with the cell's native DNA. Constructs for bacterial expression was transformed into competent KRX *E.coli* cells from which expression studies were carried out. All competent cells were acquired from Promega and a general protocol was followed for all transformations.

4 µl of each ligation reaction was added to 50 µl of the competent cells. This was incubated on ice for twenty minutes, heat shocked at 42°C for 45 seconds, followed by another incubation on ice for 2 minutes. 950 µl of LB medium, prewarmed to 37°C, was added to the reaction and it was incubated for 90 min at 37°C with agitation. 200 µl was spread plated onto LB agar plates (10 g/L NaCl, 5 g/L bacto tryptone, 5g bacto yeast extract, 5 g/L bacteriological agar) containing 100 µl IPTG (100 mM), 20 µl X-Gal (50 mg/ml) and 10 µl Ampicillin (100 mg/ml). The plates were incubated overnight at 37°C. Appropriate colonies were picked off plates with a sterile bacteriological loop and inoculated into LB broth containing Ampicillin (100 mg/ml). Cultures were grown up overnight at 37°C with agitation.

2.5.3. Plasmid DNA preparation (Minipreps)

Plasmid DNA of overnight cultures (approximately 14 hours) were purified with the Promega Wizard Plus SV Minipreps DNA Purification kit. Briefly, 5 ml of culture was pelleted by centrifugation. The supernatant was discarded and 250 µl Cell Lysis Solution was added and incubated until the solution cleared. Alkaline Protease was added to inactivate any endonucleases or proteins released during lysis that could have an adverse effect on the DNA being isolated. Neutralization Solution was added and suspension was cleared by centrifugation. Cleared lysate was transferred to a spin column and DNA bound to the membrane by centrifugation. The DNA was washed twice with Column Wash Solution and eluted from the membrane with 30 µl nuclease free water. The products was DNA sequenced (Inqaba Biotech) and the insert was verified by comparison to the existing sequence for the corresponding gene on Genbank (AY697422)

2.5.4.Time course studies

150 µl of the overnight cultures were added to 150 ml LB broth with 10% glucose, 20% Rhamnose and Ampicillin (100 mg/ml). Flasks were incubated at 37°C with agitation. A time course was performed by sampling 10 ml of the growing culture at regular intervals – every two hours for the first twelve hour followed by twelve hourly intervals up to 36 hours. The bacteria in the samples were lysed using Promega FastBreak reagent. 10 ml of overnight cultures were spun down to pellet cells. 6,4 ml of supernatant was discarded and cells were resuspended. 400 µl of FastBreak Cell Lysis Solution (Promega) was added to resuspended cells and was left to lyse at room temperature for 10 minutes. Samples were centrifuged to pellet

the cell debris. 300 µl of SDS sample buffer was added to 20 µl of the pellet. Similarly, 100 µl of SDS sample buffer was added to 100 µl of the supernatant. Samples were boiled at 94°C for 10 minutes before running on SDS-PAGE gels. The gels were then subjected to Western Blotting. The primary antibody used was rabbit-anti-GST (1:5000) followed by the secondary antibody of goat-anti-rabbit (1:2000). The antibody is Horseradish peroxidase (HRP)-conjugated and therefore a HRP substrate containing detection agent was used to detect the protein on the blots with a chromagenic reaction. SDS-PAGE gels were also subjected to Coomassie staining.

Chapter 3: *Results and Discussion*

3.1. *In silico* characterisation of HCoV-NL63 M protein

The HCoV-NL63 genome produces at least six distinct subgenomic (sg) mRNAs, which encodes for six (6) potential ORFs. M is expressed from an ORF on monocistronic sg mRNA 5 (Pyrc *et al*., 2004). The 681 bp ORF encodes for a putative 25.2 kDa protein, 227 amino acids in length. A previous study of the RNA expression levels of the HCoV-NL63 sg mRNAs shows that excluding N, M has the highest expression level (Pyrc *et al*., 2004). In the current study, TMHMM Server v. 2.0 (Krogh *et al.*, 2001) and Kyte-Doolittle hydrophobicity plots (Figure 5) were used to search for transmembrane regions in the HCoV-NL63 M protein sequence. Results showed that similarly to other coronavirus M proteins, HCoV-NL63 M was a triple spanning membrane protein. Three transmembrane regions were found at amino acid position 20-36, 75-97 and 129-151 (Figure 6). Predictive analysis also showed that the N-terminal tail faced towards the extracellular space and the short 76aa C-terminus tail towards the cytosolic side.

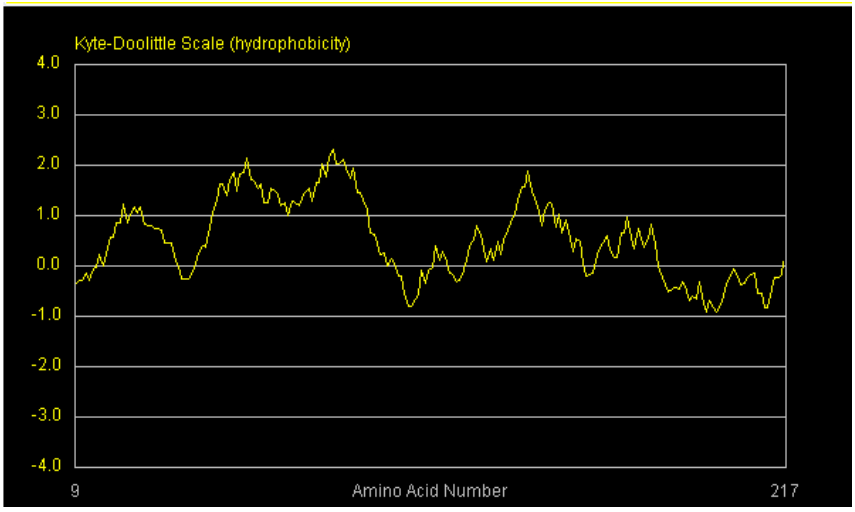


Figure 5. Hydrophobicity plot showing 3 distinct regions >1.6 on the Kyte-Doolittle scale. Peaks above 1.6 on the plot indicate probable membrane spanning regions.

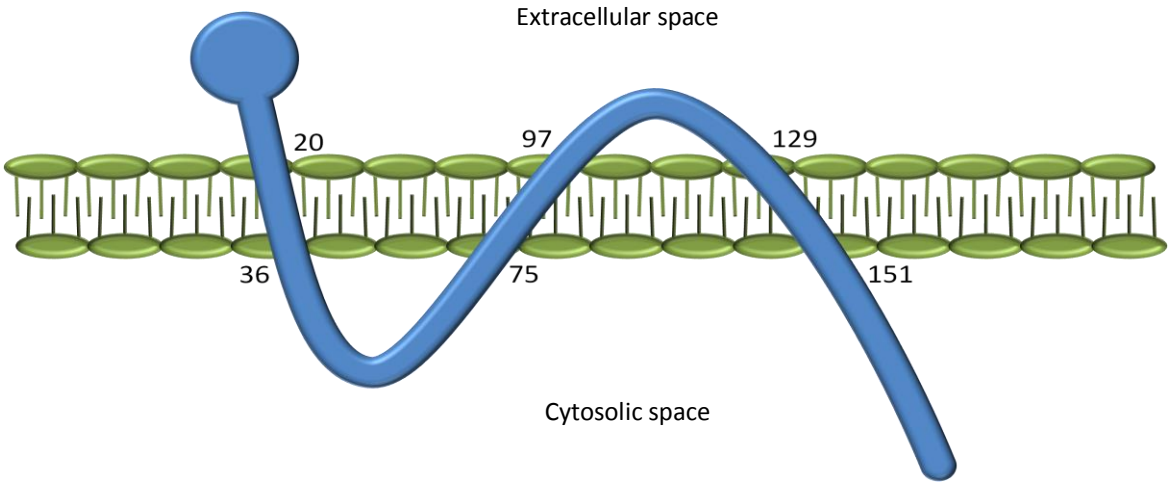


Figure 6. Predicted structure/topology of HCoV-NL63 M protein. Analysis revealed a putative ~25.2 Da protein with three membrane spanning regions at amino acid positions 20-36, 75-97 and 129-151.

3.1.1. Comparison of HCoV-NL63 M homologues

The putative amino acid sequence of HCoV-NL63 M was compared to sequences in the GenBank database at the National Centre for Biotechnology by using the Basic Blast Search Server (Altschul *et al.,* 1990). Identified HCoV-NL63 M-homologues were subsequently aligned with CLUSTAL X v 1.81 (Thompson *et al*., 1997) and viewed with GENEDOC software (Nicholas and Nicholas, 1997). The M protein of HCoV-NL63 was found to be more similar to M proteins of other Group I coronaviruses (Figure 7).

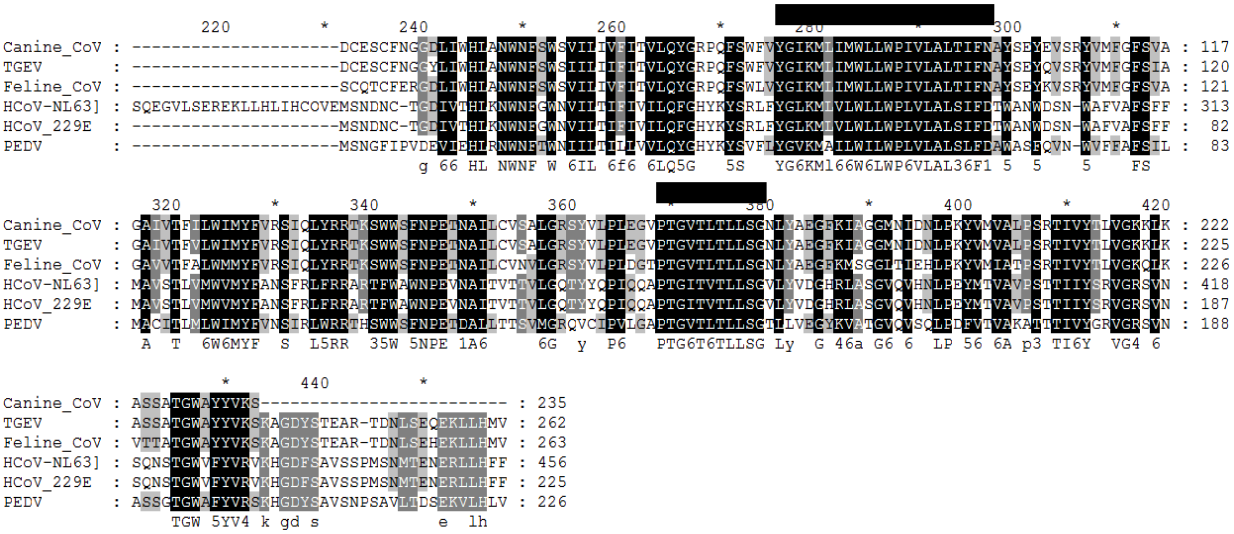


Figure 7. Alignment of the amino acid sequences of HCoV-NL63 M and selected coronavirus M proteins. Identified homologues were aligned; shading indicates conserved regions and gaps were introduced to align sequences. Conserved regions (>10 aa) indicated by solid bars and consensus sequence shown below the aligned sequences. Three levels of shading shown.

The M amino acid sequences aligned (Figure 7 and Table 11) were (HCoV-NL63), Human coronavirus NL63 Amsterdam 1 M; (HCoV-229E), Human coronavirus 229E CCU T935; (PEDV) Porcine epidemic diarrhoea virus DX; (TEGV), Transmissible gastroenteritis virus BW021898B; (CCoV), Canine coronavirus type I Elmo/02; and (FCoV), Feline coronavirus FCoV/NTU2/R/2003.

Table 11. Comparison of HCoV-NL63 M amino acid sequence to homologues from selected Group I coronavirus isolates. Identity values (%) are shown in bold and similarity values (%) are shown in italics.

	TGEV	**FCoV**	**HCoV-NL63**	**HCoV-229E**	**PEDV**
CCoV	**84%** *87%*	**74%** *82%*	**20%** *34%*	**34%** *56%*	**40%** *58%*
TGEV		**81%** *91%*	**22%** *36%*	**37%** *60%*	**46%** *63%*
FCoV			**22%** *35%*	**37%** *59%*	**41%** *63%*
HCoV-NL63				**49%** *49%*	**27%** *39%*
HCoV-229E					**55%** *80%*

The amino acid sequence of HCoV-NL63 M was shown to be most similar to human coronavirus 229E M (Table 11; 49% identity and 49% similarity). The comparative analysis showed two (2) regions ≥ 10 amino acids that were well conserved between the M homologues (Figure 7). The functions, if any, of these regions are not known, but could warrant further research.

Coronavirus M proteins are N-glycosylated, triple-spanning membrane proteins (de Haan *et al.,* 2003). The TGEV M protein has been shown to have interferogenic activity. Mutations in the ectodomain causing disruption of N-glycosylation, down-regulates this interferogenic activity (Laude *et al.,* 1992). The coronavirus M protein also plays a role in virion assembly, as well as in interaction with host cells and therefore infectivity. The formation of new virions is dependent on two processes; the one is the assembly of the nucleocapsid and the other, the assembly of the envelope. The M protein plays an essential role in both these processes by directing the formation of the envelope, as well as providing the basis for the attachment of the nucleocapsid to enable budding (de Haan *et al.,* 2000).

3.2. Cloning for expression in mammalian system

The vector used for the cloning of M for expression in a mammalian cell system, was pCMV-Myc Mammalian Expression Vector (Clontech). The pCMV-Myc vector is able to express N-terminal c-Myc tagged. This tag is known to be decidedly immunogenic. The vector contains an Ampicillin resistance gene for selection of successfully ligated and transformed constructs in competent *E.coli*, a multiple cloning site into which the gene of interest can be cloned, and a SV40 polyadenalation signal. Under the influence of the promoter *P*CMV IE from cytomegalovirus, the proteins will be expressed at high levels in mammalian cells.

The initial step involved the generation of 1st strand cDNA from viral RNA using RT-PCR. Next, PCR amplification was used to amplify the HCoV-NL63 M gene from the cDNA using specific primers. PCR primers were designed with *Eco*RI and *Xba*I restriction sites for unidirectional cloning of the insert into the vector, and also to clone the insert “in-frame” with the N-terminal myc-tag nucleotide sequence in the vector.

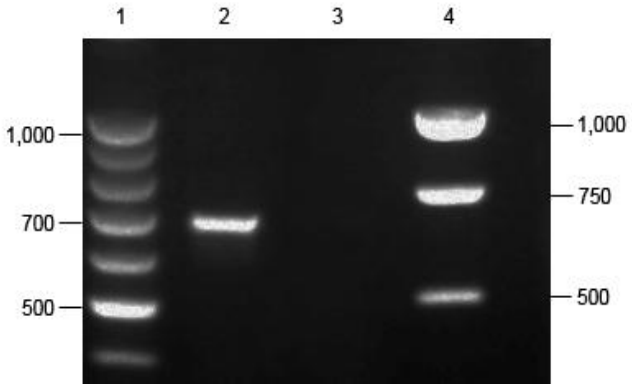


Figure 8. PCR of the full-length M gene from complete HCoV-NL63 genome. Lane 1: 100bp DNA ladder; Lane 2: Amplification of M from cDNA; Lane 3: Negative PCR control; Lane 4: 1Kb DNA ladder.

Lane 2 (Figure 8) showed the amplification product (~700 bp), that corresponded to the expected size of the M gene (~681bp). The result showed that the viral RNA genome was successfully transcribed from RNA to DNA during the reverse transcription step. Lane 3 showed the negative PCR control, *i.e.* no cDNA added to the PCR reagent mixture; no amplification products were visible in this lane indicating that no contamination took place during the preparation or running of the PCR. Next, the product was ligated into the pGEM vector to generate pGEM-M, followed by transformation into competent *JM109 E. coli* cells for blue-white colony selection. The presence of white colonies on the test plate indicated successful ligation of the gene of interest into the plasmid (pGEM) as recombinant clones will disable bacterial cells from digesting the IPTG in the presence of X-gal. This is due to the disruption of the β-galactosidase gene by the insertion of the gene of interest into the vector. The unsuccessfully ligated plasmids will enable the bacterial cells to use the IPTG as substrate and form the blue colonies in the presence of the X-Gal as the β-galactosidase gene was not disrupted.

Transformed plates of pGEM-M showed various white colonies (Figure 9a) of which four (4) were picked and used as template in a colony PCR to tentatively confirm the presence of the full-length M gene (Figure 11) based on size of the insert. The absence of growth on the negative control plate (Figure 9b) confirmed the ampicillin used in the plates was active. This affirms that all colonies present on the test plate contained plasmids that were effectively transformed into the competent JM109 *E.coli* cells; this is true because the plasmid confers Ampicillin resistance to the bacteria. The bacterial cells will therefore not be able to grow in the presence of the Ampicillin on the plate unless it contains the plasmid. The presence of the M gene in pGEM was confirmed by colony PCR amplification of products ~700bp in

size, corresponding to the expected size of 681bp for M (Figure 10, Lanes 3-6). No bands are visible in this lane indicating that no contamination took place during the setup or running of the PCR. This serves as proof that the bands seen in the test lanes are valid positive results and not contamination.

Cloning into pGEM was done to facilitate the confirmation of the nucleotide composition of the amplicon by sequencing (INQABA Biotech). When colonies were screened for the correct insert size by colony PCR (pGEM-M1, pGEM-M2, pGEM-M3 and pGEM-M4), these colonies were also inoculated into LB broth containing ampicillin and cultured overnight for plasmid extraction. Since all colonies were tentatively positive for insert by colony PCR, two constructs (pGEM-M1 and pGEM-M4) were selected for cloning M into the mammalian expression vector.

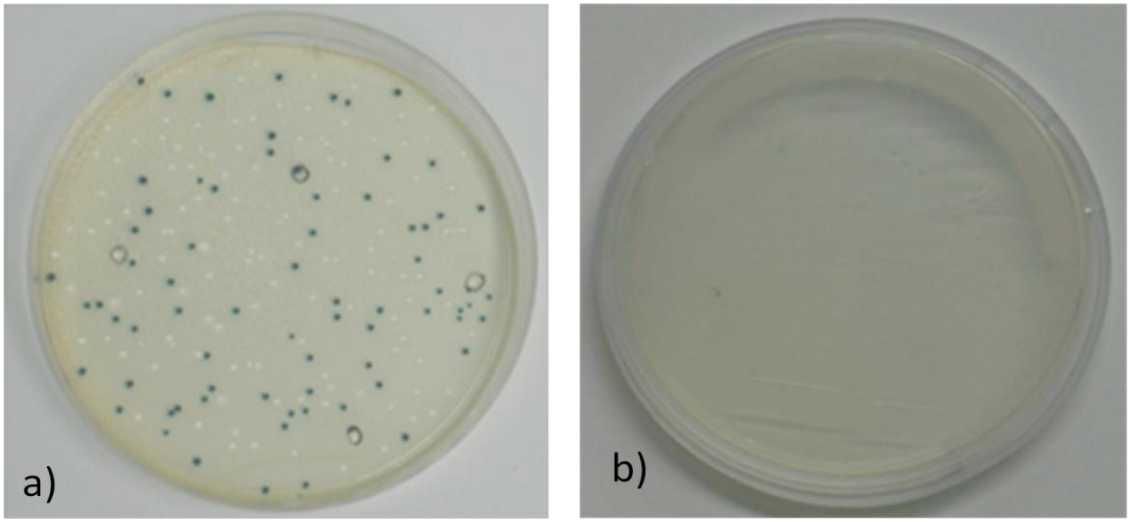


Figure 9. Agar plate of transformation of pGEM constructs into *JM109 E.coli* competent cells. (a) JM109 cells transformed with pGEM-M plated onto LB plates containing ampicillin; (b) Negative control. JM109 plated onto LB plates containing ampicillin.

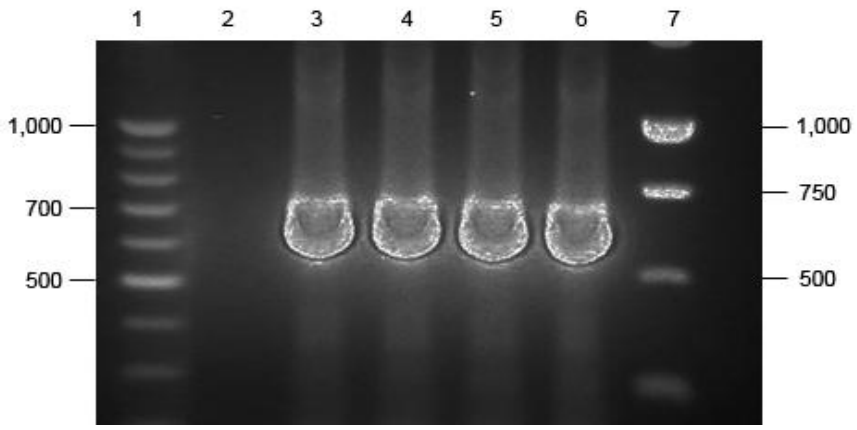


Figure 10. Colony PCR of M ligated into pGEM and transformed into JM109 *E. coli*. Lane 1: 100bp DNA ladder; Lane 2: Negative PCR control; Lanes 3-6 – Colony PCR of white colonies; Lane 7: 1kb DNA ladder.

Plasmid DNA was extracted from these cultures and subsequently restriction enzyme digested with *XbaI* and *EcoRI* to confirm insert size (Figure 11). In the meantime, pGEM-M1 and pGEM-M4 DNA were submitted for sequence verification (results not shown); sequencing results showed nucleotide sequences of the amplicons that were identical to the HCoV-NL63 M sequence on the NCBI database. For the purposes of this project, both inserts were selected for cloning into the pCMV-myc vector to generate pCMV-myc-M1 and pCMV-myc-M4.

Briefly, inserts digested with *Xba*I and *EcoR*I from pCMV-myc-M1 and pCMV-myc-M4 were gel purified and ligated into the compatible sites of pCMV-myc vector. This was followed by transformation into *JM109 E.coli* supercompetent cells (Promega) for blue-white selection. Several white colonies were picked from the agar plates and were used as template in a colony PCR to confirm correct insert size (data not shown).

One colony PCR-positive for pCMV-myc-M1 and pCMV-myc-M4 were selected for overnight growth in LB broth with ampicillin. Plasmid DNA was extracted

and subjected to a restriction enzyme digestion with *Xba*I and *Eco*RI to further confirm the presence of the gene of interest. Lanes 1 and 2 (Figure 12) each showed two bands; the lower band at ~700 bp was that of the M gene that was cut out of the vector and the top band represents the vector at approximately ~4 kb. These samples were therefore confirmed to have the insert cloned into the vector. Lane 3 was the negative control. There were no bands present in the lane, indicating that no contamination took place during the digestion and that the results were valid.

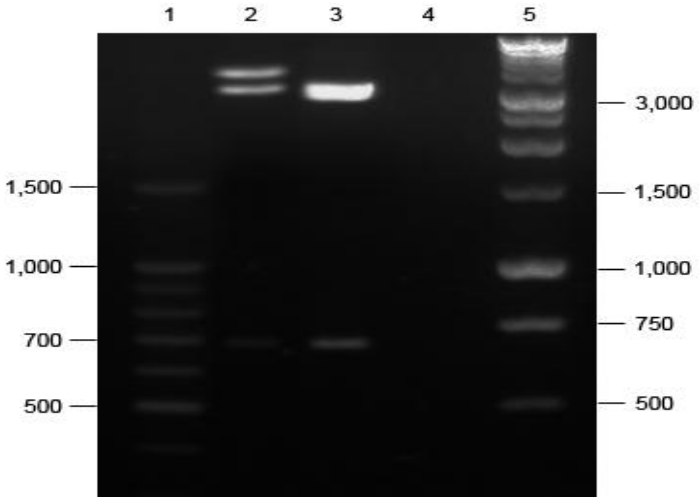


Figure 11. Restriction enzyme digestion of pGEM-M1 and pGEM-M4. Lane 1: 100bp DNA ladder; Lanes 2: pGEM-M1; Lane 3: pGEM-M4; Lane 4: Negative control; Lane 5: 1kb DNA ladder

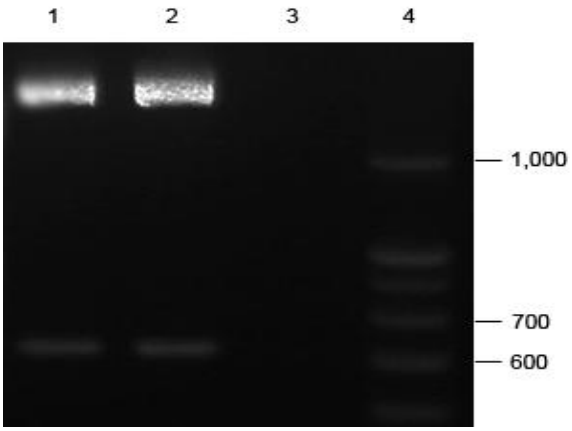


Figure 12. Restriction enzyme digests of pCMV-myc constructs. Lane 1: pCMV-myc-M1; Lane 2: pCMV-myc-M4; Lane 3: Negative control; Lane 4: 100bp DNA ladder.

3.3. Cloning for expression in Baculovirus system

Baculovirus expression technologies have made significant advances in the past 25 years, leading to them becoming more popular and quite versatile for the production of recombinant proteins (Freisen and Miller, 2001; Kost and Condreay 1999; Miller, 1997; Possee, 1997; O'reilly *et al*., 1992; Luckow and Summers, 1988). In these systems, a recombinant protein is created, processed, modified and targeted to its appropriate subcellular location as it would have been in the original cell (Kitts *et al*., 1990). These proteins are expressed at high levels and are also functionally similar to the original parent proteins. A very wide variety of genes have been expressed in this way including genes from viruses, mammals, fungi and plants (King and Possee, 1992; O'reilly *et al*., 1992; Luckow, 1991; Murhammer, 1991; Maeda, 1989; Luckow and Summers, 1988;). In the present study, the Bac-2-Bac system was used for cloning of the full-length HCoV-NL63 M gene for expression in insect cells.

The Bac-to-Bac cloning system has a number of features distinguishing it from other baculovirus cloning systems. One of the most attractive being the significantly reduced time (<2 weeks) it takes to create purified recombinant genes. This is because the recombinant DNA is never mixed with the original parent DNA and therefore there is no need for repeated plaque purification procedures. This separation of recombinant and non-recombinant DNA is achieved by the insertion of the lacZ gene into the vector allowing for blue white colony selection on agar plates. Single and multiple genes can easily be integrated into baculoviruses using commercially available plasmids and can be used for Gateway cloning. These characteristic make the Bac-to-Bac system the most efficient system for multiparallel gene expression. The simplicity and relative speed of the system allows for greater

throughput with the amplification of multiple viruses at a time (Hunt, 2005). This system is also suited for expression of proteins for structure and function studies (Invitrogen).

In this study, the first step involved PCR amplification of the HCoV-NL63 M gene from the cDNA generated earlier. Lane 3 (Figure 13) showed the amplification product (~700 bp), that corresponded to the expected size of the M gene (~681bp). The result showed that the viral RNA genome was successfully transcribed from RNA to DNA during the reverse transcription step. Lane 2 depicted the negative control; no amplification products were visible in this lane indicating that no contamination took place during the setup or running of the PCR.

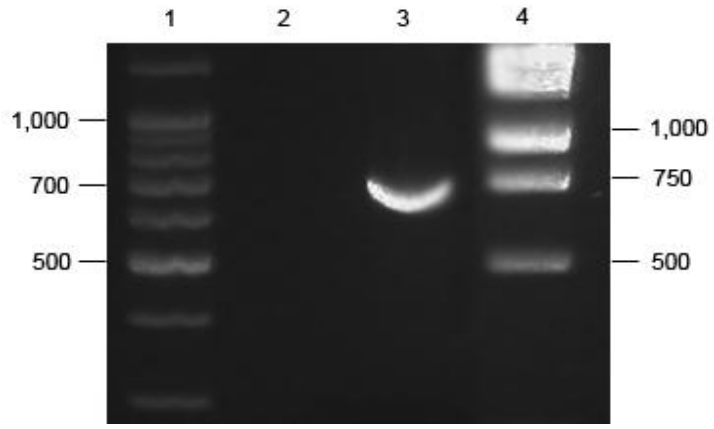


Figure 13. 1% Agarose gel electrophoresis of PCR product generated with M-specific primers using HCoV-NL63 cDNA as template. Primers incorporated restrictions enzyme sites for cloning into Bac-2-Bac system. Lane 1: 100bp DNA ladder; Lane 2: Negative control; Lane 3: PCR of M; Lane 4: 1kb DNA ladder.

In order to verify the nucleotide sequence of the amplicon, the product was then ligated into the pGEM vector to generate pGEM-M, followed by transformation into competent *JM109 E. coli* cells. Transformed plates of pGEM-M showed various white colonies (results not shown) that were picked and used as template in a colony

PCR to tentatively confirm the presence of M. The presence of the M gene in pGEM was confirmed by amplification products ~700bp in size, corresponding to the expected size of 681bp for M (Figure 14, Lanes 2-4).

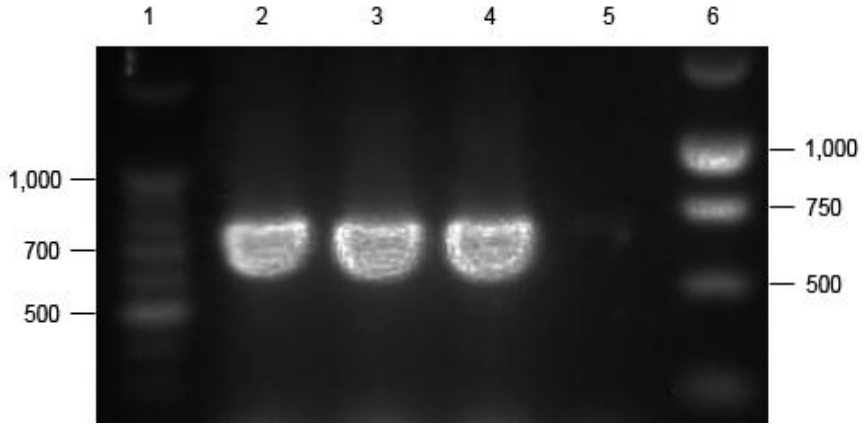


Figure 14. 1% Agarose gel electrophoresis of PCR product generated from colony PCR of pGEM-M transformed into JM109 *E.coli.* Lane 1: 100bp DNA marker; Lane 2: pGEM-M1; Lane 3: pGEM-M2; Lane 4: pGEM-M3; Lane 5: Negative PCR control; Lane 6: 1kb DNA ladder.

At the same time, colonies selected for colony PCR (pGEM-M1, pGEM-M2 and pGEM-M3) were inoculated into LB broth containing ampicillin and cultured overnight. Since all colonies were tentatively positive for insert by colony PCR, plasmid DNA was extracted from all three cultures and subsequently restriction enzyme digested with *XbaI* and *EcoRI* to confirm insert size (Figure 15). All three pGEM-M constructs were submitted for sequence verification (results not shown) and pGEM-M2 was selected for cloning into the pFastBac vector to generate pFastBac-M.

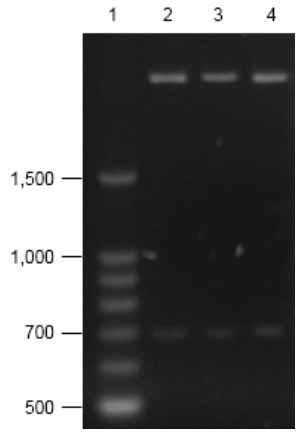


Figure 15. *XbaI* and *EcoRI* restriction enzyme digestion of M from the three positive pGEM-M constructs identified in Figure 14. Lane 1: 100bp DNA ladder; Lane 2: pGEM-M1 restriction enzyme digest; Lane 3: pGEM-M2 restriction enzyme digest; Lane 4: pGEM-M3 restriction enzyme digest; Restriction enzyme digests of pGEM-M constructs.

The *XbaI* and *EcoRI*-digested and gel purified M2 insert was now ligated into the pFastBac vector to generate pFastBac-M and was then transformed into *DH10Bac* competent *E.coli* cells. In our system, the expression of M would be regulated by the *Autographica californica* multiple nuclear polyhedrosis virus (AcMNPV) polyhedron (PD). The DH10Bac *E.coli* cells contained a bacmid (baculovirus shuttle vector) with a mini-*att*Tn7 target site and a helper plasmid (pMON7124). The bacmid (bMON14272) is 136kb and contains a low-copy number mini-F replicon, Kanamycin resistance marker and a segment of DNA coding for the LacZα peptide from a PUC-based cloning vector. This is the site into which the bacterial transposon Tn7 attachment site (mini-attTn7) has been inserted. The helper plasmid (13.2kb) provides the Tn7 transposition function *in trans* (Barry, 1988) and encodes the necessary transposase as well as conferring tetracycline resistance. When the pFastBac vector has been transformed into the DH10Bac cells, transposition occurs (under influence of transposase supplied by helper plasmid)

between the mini-*att*Tn7 target site on the bacmid and the mini-Tn7 element on the pFastBac vector to form the recombinant bacmid.

Plates of pFastBac-M transformed *DH10Bac* cells showed various white colonies, of which four (4) were picked with a toothpick and used as template in a colony PCR with M-specific primers to confirm the presence of the M gene (Figure 16). At the same time, the selected colonies were cultured overnight and recombinant bacmid DNA extracted.

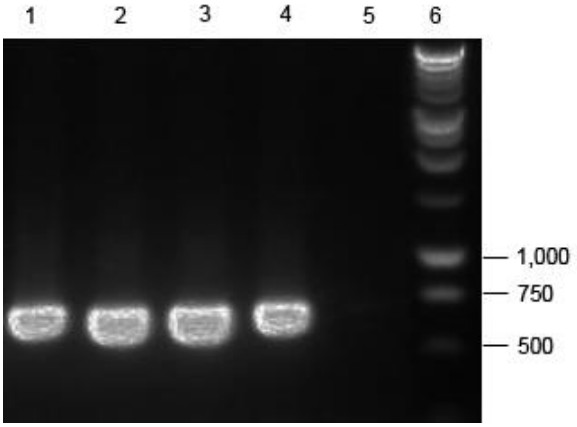


Figure 16. Colony PCR of the M gene from four (4) selected white colonies. Lanes 1–4: Tests; Lane 5: Negative PCR control; Lane 6: 1Kb DNA ladder.

The bacmid will propagate in the *E.coli* as a large plasmid and will lend the bacteria resistance to Kanamycin as well as complementation to a LacZ deletion on the chromosome. This last characteristic enables the cells to form blue colonies in the presence of a chromagenic substrate such as X-gal or Bluo-gal, together with IPTG. As it is difficult to analyse the recombinant bacmid by the more conventional method of restriction by endonucleases, the presence of the recombinant gene was confirmed by PCR and agarose gel electrophoresis (Figure 17). Lane 2 showed the complete product, amplified with both the M13 forward and reverse primers, *i.e*. the

gene of interest (681bp), plus the bacmid (273 bp) and transposed vector (1980 bp) at 2934 bp. This was the expected size as shown in Figure 18.

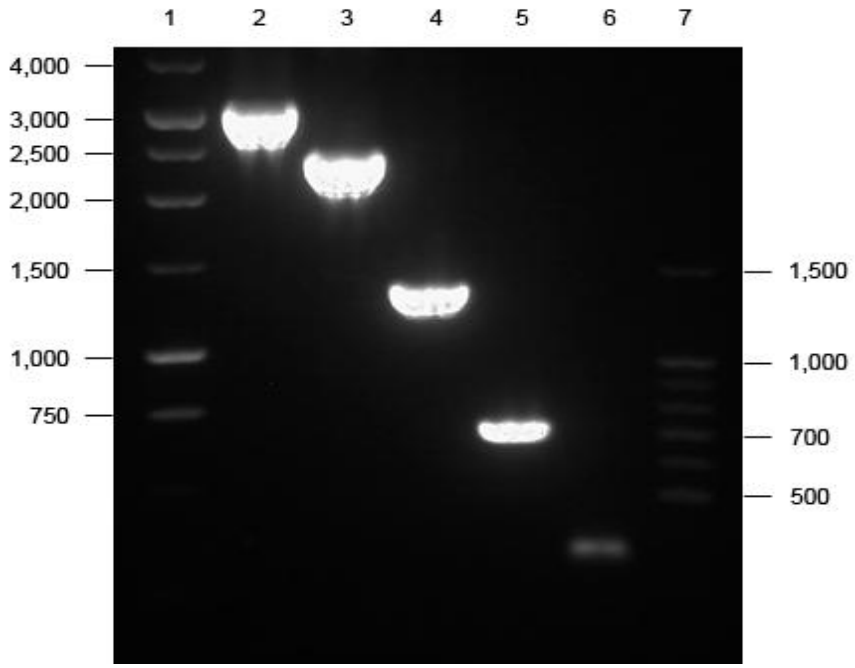


Figure 17. Gel of PCR of recombinant bacmid. Lane 1: 1kb DNA ladder; Lane 2: Product amplified with M13-forward and -reverse primers; Lane 3: Product amplified with the M13-forward primer and the gene-specific reverse primer; Lane 4: Product amplified with the M13-reverse primer and the gene-specific forward primer; Lane 5: Product amplified with gene specific primers; Lane 6: Bacmid only; Lane 7: 100bp marker.

Lane 3 showed the product when amplified with the M13 forward primer and the gene specific reverse primer yielding a product of 2331bp; this represented the bacmid + pFastBac + M gene. Lane 4 showed the product amplified with the M13 reverse primer and the gene specific forward primer yielding a product of approximately 1284bp; this represented bacmid + pFastBac + M gene. Lane 5 showed the insert amplified with gene specific primers at 681bp. Finally, Lane 6 showed the bacmid only at 300bp (Figure 17 and 18). From these results it can be concluded that the M gene was successfully cloned into the pFastBac vector and consequently into the bacmid and was now ready to be expressed in *Sf9* cells for

expression studies. The recombinant bacmid was used to infect *Spodoptera frugiperda* Sf9 insect cells using Cellfectin (Invitrogen). P1 recombinant baculovirus stock was obtained and cells were re-infected to generate P2 viral stock as discussed in the Materials and Methods section. Both P1 and P2 viral stocks were frozen and stored at -80^{0}C for future use. The titer of these viral stocks will be determined in future. However, M-specific antibodies are currently being raised as part of another project. These antibodies will be used to screen our P1 and P2 viral stocks for expression of full length HCoV-NL63 M protein by Western Blot and immune- detection.

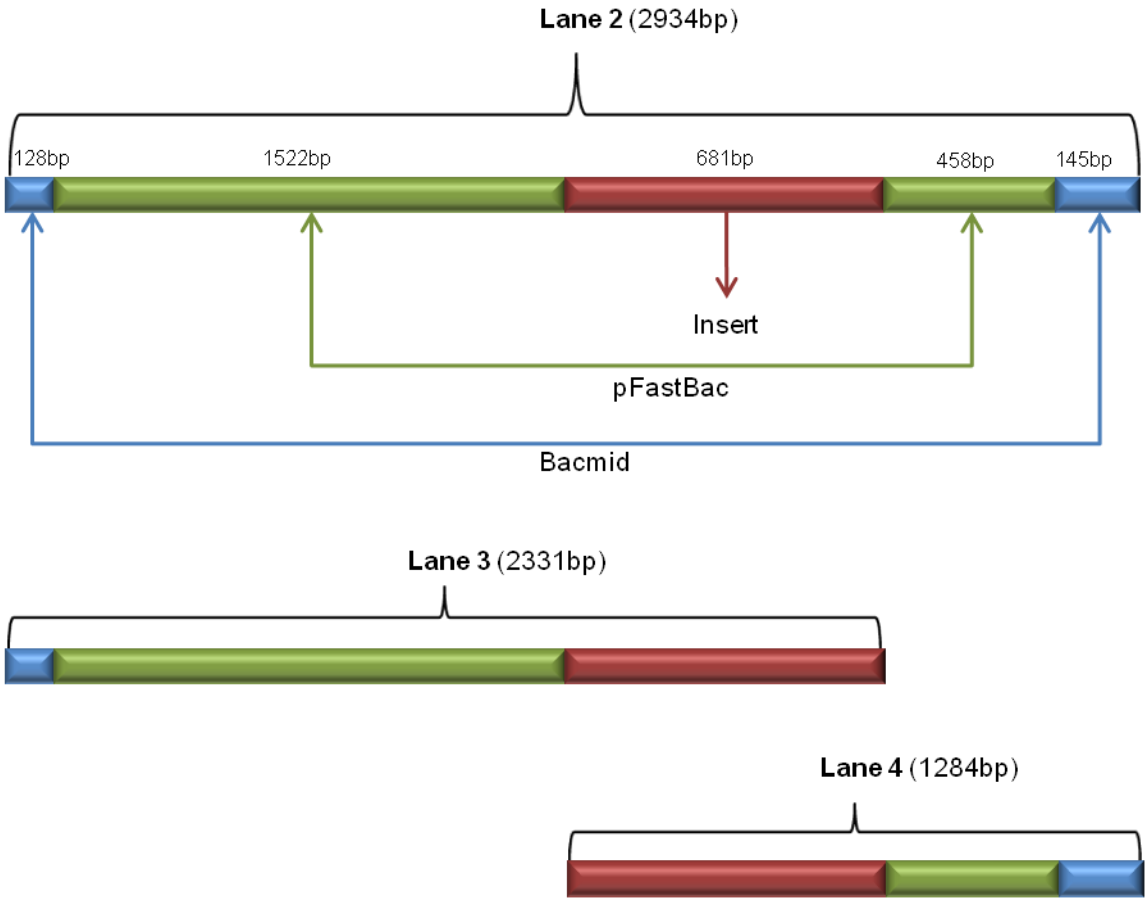


Figure 18. Diagrammatic representation of expected fragment sizes after gel electrophoresis of PCR of recombinant bacmid.

3.4. Cloning for expression in Bacterial system

Many vectors are available within the Flexi range. For this study, the pFN2A (GST) vector was chosen for protein expression. This vector is especially designed to express amino terminal GST-fusion proteins in bacteria. The GST tag can be used to detect and even purify expressed protein. The tag can also be cleaved off the expressed protein with TEV protease. For protein expression, this vector requires the use of a competent *E.coli* strain that expresses T7 RNA polymerase like KRX *E.coli*.

KRX *E.coli* (Promega) was the cloning strain of choice for various reasons. Some of these include the lack of common nucleases present in other *E.coli* strains, partially defective restriction systems, mutations to prevent proteolysis of the overexpressed protein and mutations to limit undesirable recombinations. The system is compatible with various commercially available vectors that makes use of the T7 promoter (e.g. Flexi, as in the case of this study), making it extremely user friendly. Another advantage is that recombinant protein expression can be precisely controlled. This is due to the presence of a T7 RNA polymerase gene which is driven by the rhamnose promoter *rha*PBAD. This type of T7 RNA polymerase-based system enjoys wide use because the promoter is totally independent of the *E.coli* RNA promoters, and the fact that the T7 polymerase's elongation rate is about 5x that of the *E.coli* RNA polymerase. The *rha*BAD gene is replaced by the T7 RNA polymerase gene so that the *rha*BAD promoter will control its expression. The promoter is activated by rhamnose and suppressed by glucose. Adding controlled amounts of these substances to the growth medium therefore lend excellent control of recombinant protein production by the researcher.

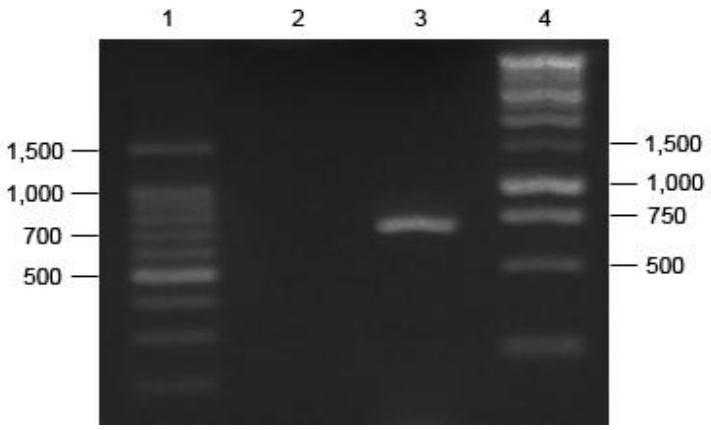


Figure 19. Agarose gel of PCR of the M gene for expression in bacteria. Lane 1: 100bp DNA ladder; Lane 2: Negative PCR control; Lane 3: PCR of full-length M gene; Lane 4: 1kb DNA ladder.

The gene of interest (M gene) was amplified from the 1st copy DNA with gene specific primers for unidirectional cloning into the vectors. Primers were designed to append the *SgfI* and *PmeI* restriction enzyme sites to the amplification product. These enzymes are the most infrequent cutting enzymes in human cDNA (98%) and also rarely cut in other organisms. The *SgfI* cut site is upstream of the start codon of the protein coding region. This enables expression of untagged proteins as well as N terminal tagged proteins. The *PmeI* site contains the stop codon for the protein coding region and adds a Valine residue to the carboxy terminus of the protein. The reading frame and the orientation of the insert are maintained by these two enzyme cut sites. The amplification product was separated on a 1% Agarose gel (Figure 19) by electrophoresis; Lane 1 and 4 showed the 100bp and 1kb marker respectively. Lane 3 showed the amplification product at ~700bp, which corresponded to the expected size of the M gene at ~681bp. First, the PCR product was ligated into the pGEM vector (for ease of sequencing) and transformed into competent *JM109 E.coli* cells. Following incubation overnight, two white colonies were picked and used as template in a colony PCR (Figure 20). At the same time, these colonies were

cultured overnight, plasmid DNA extracted by the mini-preparation method and sequenced to determine the nucleotide sequence (INQABA Biotech). Both construct sequences were determined to be identical to the HCoV-NL63 M gene sequence at the NCBI database (results not shown) and therefore one was selected for all subsequent cloning.

Next the plasmid DNA was enzyme digested with *SgfI* and *PmeI* and the insert was gel purified (Promega) and ligated into compatible sites of the pFlexi vector, generating pFlexi-M. This construct was transformed into KRX *E. coli* cells and plated on LB agar plates containing ampicillin (Figure 21).

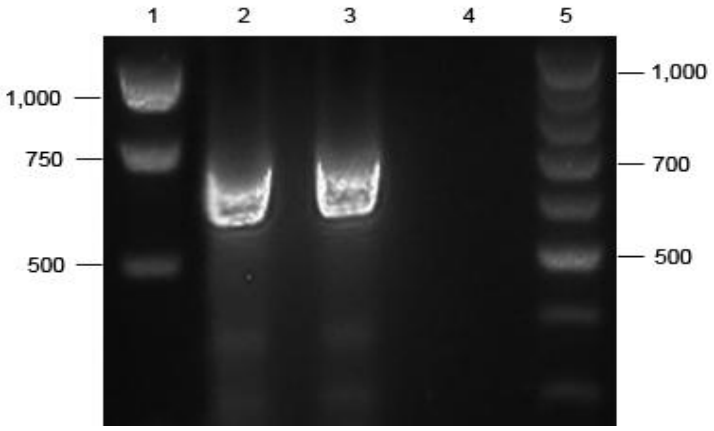


Figure 20. Agarose gel of colony PCR of insert transformed into *JM109 E.coli.* Lane 1: 1kb DNA ladder; Lane 2 and 3: Colony PCRs; Lane 4: Negative PCR control; Lane 5: 100bp DNA ladder.

All colonies on plate (b) (Figure 21) should be successfully ligated and transformed recombinants as the flexi vector contained the lethal barnase gene which was replaced by the M. Additionally, the vector also conferred antibiotic resistance to the competent cells which would otherwise be unable to grow in the presence of the Ampicillin that was added to the plates. Two colonies were picked and grown overnight in LB broth containing ampicillin. Plasmid DNA was extracted and

constructs pFlexi-M1 and pFlexi-M2 was digested with *SgfI* and *PmeI* to confirm the presence of full-length M.

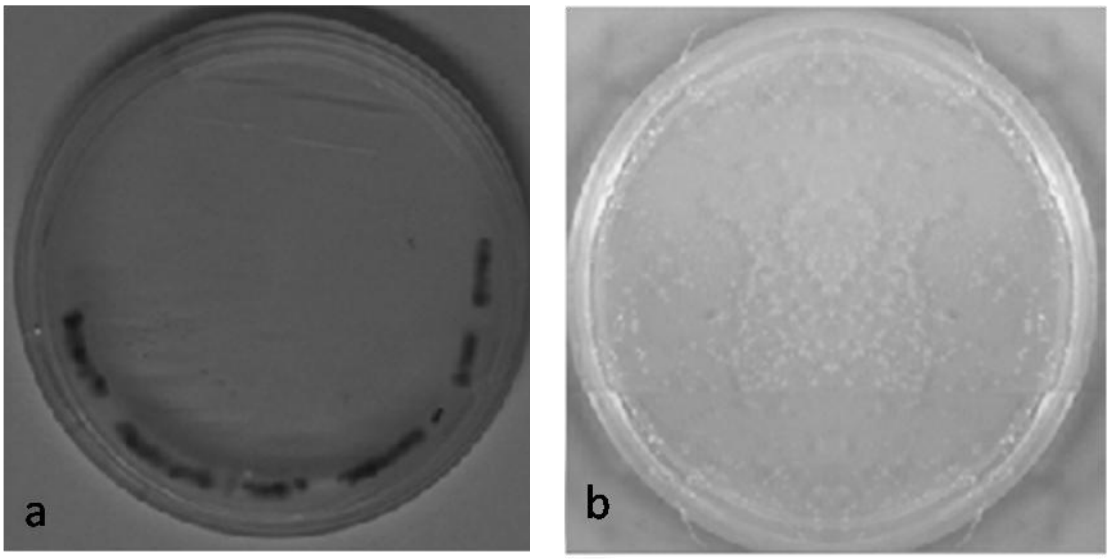


Figure 21. Agar plates of transformation of M into *KRX E.coli* cells. (a) Negative control; (b) Test plate.

The gene of interest was successfully digested from the vector as can be seen by the bands at approximately 700 bp and the vector at approximately 3kb (Figure 22, lanes 2 and 3).

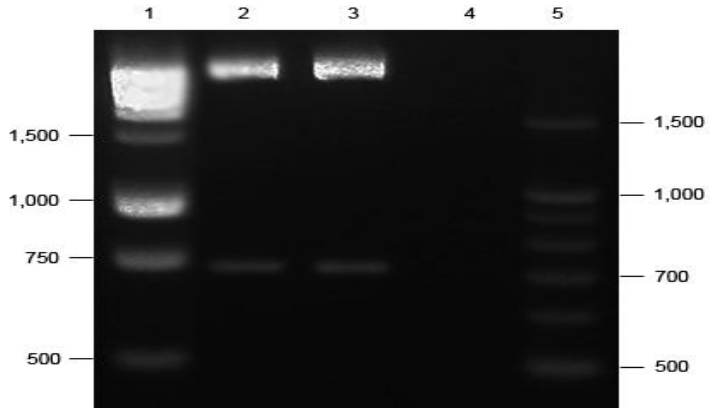


Figure 22. Restriction enzyme digests of pFlexi-M with restriction enzymes *SgfI* and *PmeI*. Lane 1: 1kb DNA ladder; Lanes 2: pFlexi-M1 enzyme digest; Lane 3: pFlexi-M2 enzyme digest; Lane 5: 100bp DNA ladder.

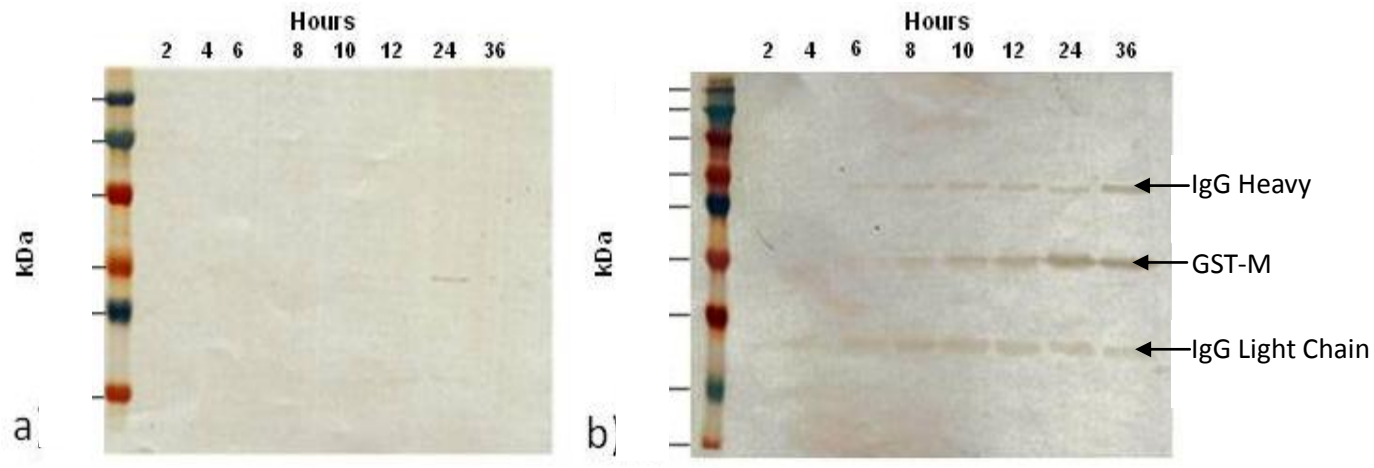


Figure 23. Western Blots of GST-M protein expressed by KRX *E.coli* over a 36 hour time course. (a) Supernatant; (b) Pellet.

Subsequently, pFlexi-M1 was selected for expression studies. A 36 hour protein expression time course was run to access the expression characteristics of the M protein in the *KRX E.coli.* Figure 23 shows the expression profile of GST-M over a 36 hour period. Cells were harvested, lysed and total proteins were separated on SDS PAGE and Western Blotted. Both pellet and supernatant samples were subjected to immunoblotting with anti-GST antibody. The blot of supernatant sample showed virtually no bands at all. This is indicative of the M protein not being present in the supernatant of the cell (Figure 23 a). The Western blot of the pellet sample showed that M expression was detected approximately 6-8 hrs after protein induction, with optimal expression levels seen at 24 hrs; expression levels then increased through to 36 hrs. The GST-M protein was detected as a 50 kDa protein; *i.e.* a 25 kDa GST tag appended to the 25 kDa M protein. The anti-GST antibody specifically detected the M fusion protein. As expected, since M is an integral membrane protein, GST-M was detected in the pellet portion of the cell lysate (Figure 23b); this consisted of all membrane structures. Due to the nature of the antibody used, the IgG Heavy and IgG light chains were also detected on the blot

(Figure 23b). M was detected as ~25 kDa protein (excluding the ~25 kDa) GST fusion tag), which corresponded to the predicted protein size. The predicted size excluded any posttranslational modification of M. It was therefore evident that the HCoV-NL63 M protein was not posttranslationally modified in this system. In order to study any modification, processing or function of M, studies need to done in a system that allows for the proper processing of the proteins, such as a mammalian and/or baculovirus system.

Chapter 4: *Conclusion*

The complete HCoV-NL63 RNA genome was transcribed to 1st strand cDNA using random poly-A primers. The HCoV-NL63 M gene was amplified from the cDNA using different gene specific primer sets for cloning into different expression vector systems. The full-length M gene was cloned into vectors for expression in bacterial, baculovirus and mammalian expression systems; all sequences were verified by double stranded sequencing of constructs.

4.1. Mammalian expression

The pCMV-Myc vector is able to express N-terminal c-Myc tagged proteins. This tag is known to be decidedly immunogenic. The vector contains an Ampicillin resistance gene for selection of ligated and transformed constructs in competent *E.coli.* The vector also contains a multiple cloning site into which the gene of interest can be cloned and a SV40 polyadenalation signal. Under the influence of the promoter *P*CMV IE from cytomegalovirus, the proteins will be expressed at high levels in mammalian cells.

1st strand cDNA was generated from viral RNA by RT-PCR. The full-length M gene was PCR amplified for unidirectional cloning into the pCMV-Myc vector “in-frame” with the c-myc tag sequence. The amplified M gene was cloned into the pGEM-T-Easy vector for sequencing. Sequence-verified inserts were digested from the pGEM vector and cloned into compatible sites of the pCMV-Myc vector.

4.2. Baculovirus expression (Refer to Appendix B)

The Bac-to-Bac system is a rapid, easy and efficient system to use for the generation of recombinant baculoviruses (Ciccarone *et al.,* 1997). The system is based on the site-specific transposition of an expression plasmid into a baculovirus shuttle vector (bacmid) propagated in *E. coli* (Luckow *et al*., 1993). The system uses the site specific transposition properties of the Tn7 transposon to simplify and enhance the process of generating recombinant bacmid DNA. Expression of the recombinant gene is under the influence of either the *Autographica californica* multiple nuclear polyhedrosis virus polyhedron (AcMNPV-PH) or p10 (AcMNPV-P10), which are genes that are highly transcribed in the late stages of infection. This leads to high expression levels in the infected cells which in turn results in proteins that are functionally similar to the wild-type proteins (King and Possee, 1992; O'reilly et al, 1992; Luckow, 1991; Murhammer, 1991; Maeda, 1989; Luckow and Summers, 1988; Miller, 1988).

Even though Bac-to-Bac is an extremely efficient system for gene expression in baculoviruses, it has to be said that it also has some disadvantages. Studies have shown that recombinant baculovirus obtained via bacmid transposition as opposed to the more traditional homologous recombination method, tends to be inherently unstable and may be spontaneously deleted from the genome upon increased passaging of the insect cells (Pijlman *et al*., 2003). This should however not create a problem in the routine research laboratory as the number of passages required for this deletion tends to be quite high.

The M gene was amplified from 1st strand cDNA for the expression in a baculovirus expression system. The full-length HCoV-NL63 M gene was cloned into the pGEM-T-Easy shuttle vector to facilitate sequencing of the insert. The sequence

verified insert was enzyme digested from the pGEM-M construct. It was then cloned into the pFastBac vector followed by transformation into competent DH10Bac *E.coli* cells. P1 and P2 viral stocks were generated and stored for future use. Currently, M-specific antibodies are being generated that will be used to screen the P1 and P2 virus stock lysates for expression of the full-length M protein.

4.3. Bacterial expression

The Flexi vector system (Promega) is an extremely suitable vector system to use for bacterial protein expression. It is a directional cloning system making use of two extremely rare restriction endonucleases *SgfI* and *PmeI*. These enzymes are the most infrequent cutting enzymes in human cDNA (98%) and also rarely cut in other organisms. The *SgfI* cut site is upstream of the start codon of the protein coding region. This enables expression of untagged proteins as well as N terminal tagged proteins. The *PmeI* site contains the stop codon for the protein coding region and adds a Valine residue to the carboxy terminus of the protein. The reading frame and the orientation of the insert is maintained by these two enzyme cut sites.

The vector also contains a barnase gene which is a gene coding for ribonuclease. This gene is lethal and needs to be replaced by the gene of interest to enable the clone to survive. This acts as a positive control for ligation as only successfully ligated clones will survive. This eliminates the need for blue/white colony selection and screening for successfully ligated colonies. Another feature used to act as a positive control is the presence of a antibiotic resistance gene (either Kanamycin or Ampicillin depending on the vector used). Only the colonies that have been successfully ligated and transformed would be able to survive on bacterial plates containing these antibiotics. This system does not need multiple amino acids to be appended to the amino- or carboxy termini of the gene of interest as are required by more traditional site specific recombination systems.

The HCoV-NL63 M gene was amplified from 1st strand cDNA, cloned into shuttle vector pGEM-T-Easy and finally cloned into the pFlexi vector to append a GST at the N-terminus of the M protein. The pFlexi-M construct was transformed into competent KRX *E.coli* cells and a 36hr time course expression profile was

performed; expression of the GST-M fusion protein was first detected from approximately 8 hours with optimum expression levels reached after 24 hours. As the majority of the GST-M proteins were seen in the pellet after cell lysis, and only a minimal portion seen in the supernatant, it was concluded that the HCoV-NL63 M protein was expressed on the membrane of the host cell.

References

1. **Adcock, P. M., G. G. Stout, M. A. Hauck, G. S. Marshall.** 1997. Effect of rapid viral diagnosis on the management of children hospitalized with lower respiratory tract infection. Pediatr. Infect. Dis. J. **16:**842-846.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** (1990). Basic local alignment search tool. J. Mol. Biol. **215:**403-410.
3. **An, S., C. J. Chen, X. Yu, J. L. Leibowitz, and S. Makino.** 1999. Induction of apoptosis in murine coronavirus-infected cultured cells and demonstration of E protein as an apoptosis inducer. J. Virol. **73:**7853-7859.
4. **Arden, K. E., M. D. Nissen, T. P. Sloots, and I. M. Mackay.** 2005. New human coronavirus HCoV-NL63 associated with severe lower respiratory tract disease in Australia. J. Med. Virol. **75:**455-462.
5. **Barenfanger, J., C. Drake, N. Leon, T. Mueller, T. Troutt.** 2000. Clinical and financial benefits of rapid detection of respiratory viruses: an outcomes study. J. Clin. Microbiol. **38:**2824-2828.
6. **Barry, G. F.** 1988. A broad host-range shuttle system for gene insertion into the chromosomes of Gram negative bacteria. Gene. **71:**75-84.
7. **Bastien, N., J. L. Robinson, A. Tse, B. E. Lee, L. Hart, and Y. Li.** 2005. Human coronavirus NL-63 infections in children: a 1-year study. J. Clin. Microbiol. **43:**4567-4573.
8. **Bastien, N., K. Anderson, L. Hart, P. van Caeseele, K. Brandt, D. Milley, T. Hatchette, E. C. Weiss, and Y. Li.** 2005. Human coronavirus NL63 infection in Canada.
9. **Belay, E. D., D. D. Erdman, L. J. Anderson, T. C. Peret, S. J. Schrag, B. S. Fields, J. C. Burns, and L. B. Schonberger.** 2005. Kawasaki disease and human coronavirus. J. Infect. Dis. **192:**352-353

10. **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. Virology **94:**371-384.
11. **Bos, E. C. W., Luytjes, H. van der Meulen, H. K. Koerten, and W. J. M. Spaan.** 1996. The production of recombinant infectious DI-particles of a murine coronavirus in the absence of helper virus. Virology **218**:52-60.
12. **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.
13. **Bosch, B. J., R. van der Zee, C. A. de Haan, and P. J. 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.
14. **Bradburne, A. F. and B. A. Somerset.** 1972. Coronative antibody titers in sera of healthy adults and experimentally infected volunteers. J. Hyg. **70:**253-244.
15. **Bradburne, A. F., M. L. Bynoe, and D. A. Tyrrell.** 1967. Effects of a “new” human respiratory virus in volunteers. Br. Med. J. **3:**767-769.
16. **Burns, J. C., and M. P. Glode.** 2004. Kawasaki syndrome. Lancet. **364:**533-544.
17. **Cavanagh, D.** 1997. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch. Virol. **142:**629-633.
18. **Chang, L. Y., B. L. Chiang, C. L. Kao, M. H. Wu, P. J. Chen, B. Berkhout, H. C. Yang, and L. M. Huang.** 2006. Lack of association between infection

with a novel human coronavirus (HCoV) HCoV-NH and Kawasaki Disease in Taiwan. J. Infect. Dis. **193:**283-286.

19. **Chiu, S. S., K. H. Chan, K. W. Chu, S. W. Kwan, Y. Guan, L. L. Poon, and J. S. Peiris.** 2005. Human coronavirus NL63 infection and other coronavirus infections in children hospitalized with acute respiratory disease in Hong Kong China. Clin. Infect. Dis. **40:**1721-1729.
20. **Choi, E. W., 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 viruses with lower respiratory tract infections in Korean children, 2000-2005. CID **43:**585-592.
21. **Ciccarone, V. C., D. Polayes, and V. A. Luckow.** 1997. Generation of Recombinant Baculovirus DNA in *E.coli* Using Baculovirus Shuttle Vector. Volume 13. U. Reischt (ed.). Humana Press Inc., New Jersey.
22. **Crackower, M. A., R. Sarao, G. Y. Oudit, C. Yagil, I. Kozieradski, S. E. Scanga, A. J. Oliveira-dos-Santos, J. da Costa, L. Zhang, Y. Pei, *et al*.** 2002. Angiotensin-converting enzyme 2 is an essential regulator of heart function. Nature. **417:**822-828.
23. **Danilczyk, U., U. Eriksson, M. A. Crackower, J. M. Penninger.** 2003. A story of two ACEs. J. Mol. Med. **81:**227-234.
24. **de Haan, C. A. M., H. Vennema, and P. J. M. Rottier.** 2000. Assembly of the coronavirus envelope: homeotypic interactions between the M proteins. J. Virol. **74:**4967-4978.
25. **de Haan, C. A. M., L. Kuo, P. S. Masters, H. Vennema, and P. J. M. Rottier.** 1998. Coronavirus particle assembly: primary structure requirements of the membrane protein. J. Virol. **72:**6838-6850.

26. **de Haan, C. A., K. Stadler, G. J. Godeke, B. J. Bosch, and P. J. Rottier.** 2004. Cleavage inhibition of the murine coronavirus spike protein by a furin-like enzyme affects cell-cell but not virus-cell fusion. J. Virol. **78:**6048-6054.

27. **de Haan, C. A., M. de Wit, L. Kuo, C. Montalto-Morrison, B. L. Haagmans, S. R. Weiss, P. J. Rottier.** 2003. The glycosylation status of the murine hepatitis coronavirus M protein affects the interferogenic capacity of the virus in vitro and its ability to replicate in the liver but not the brain. Virology **312:**395-406.

28. **Delmas, B., J. Gelfi, R. L'Haridon, L. K. Vogel, H. Sjostrom, O. Noren, H. Laude.** 1992. Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. Nat. **357**:417-420.

29. **Domingo, E.** 2002. Quasispecies theory in virology. J. Virol. **76:**463-465.

30. **Drake, J. W., and J. J. Holland.** 1999. Mutation rates among RNA viruses. Proc. Natl. Acad. Sci. USA. **96:**13910-13913.

31. **Drosten, C., S. Gunther, W. Preiser, et al.** 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1967-1976.

32. **Ebihara, T., R. Endo, X. Ma, N. Ishiguro, and H. Kikuta.** 2005. Detection of human coronavirus NL63 in young children with bronchiolitis. J. Med. Virol. **75:**463-465.

33. **Ebihara, T., R. Endo, X. Ma, N. Ishiguro, and H. Kikuta.** 2005. Lack of association between New Haven coronavirus and Kawasaki disease. J. Infect. Dis. **192:**351-352.

34. **Escors, D., J. Ortego, and L. Enjuanes.** 2001. The membrane M protein of the transmissible gastroenteritis coronavirus binds to the internal core through the carboxy-terminus. Adv. Exp. Med. Biol. **494:**589-593.

35. **Esper, F., C. Weibel, D. Ferguson, L. Landry, and J. S. Kahn.** 2005. Evidence of a novel human coronavirus that is associated with respiratory tract disease in infants and young children. J. Infect. Dis. **191:**492-498.

36. **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.

37. **File, T. M.** 2003. Community-acquired pneumonia. Lancet. **362:**1991-2001.

38. **Fine, M. J., R. A. Stone, D. E. Singer, C. M. Coley, T. J. Marrie, J. R. Lave, L. J. Hough, D. S. Obrosky, R. Schulz, E. M. Ricci, *et al*.** 1999. Processes and outcomes of care for patients with community acquired pneumonia: results from the Pneumonia Patient Outcomes Research Team (PORT) cohort study. Arch. Intern. Med. **159:**970-980.

39. **Fiscus, S. A., and Y. A. Teramoto.** 1987. Antigenic comparison of feline coronavirus isolates: evidence for markedly different peplomer glycoproteins. J. Virol. **61:**2607-2613.

40. **Forster, J. G. Ihorst, C. H. Rieger, et al.** 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.

41. **Fouchier, R. A., N. G. Hartwig, T. M. Bestebroer, B. Niemeyer, J. C. de Jong, J. H. Simon, and A. D. Osterhaus.** 2004. A previously undescribed coronavirus associated with respiratory disease in humans. Proc. Natl. Acad. Sci. USA. **101:**6212-6216.

42. **Freisen, D. P., and L. K. Miller.** 2001. Insect viruses, p. 1871-1940. *In* D. M. Knipe *et al* (eds.), Fields' Virology. Lippincott Williams and Wilkins, Philadelphia.

43. **Garbino, J., R. Sommer, A. Gerber, C. Regamey, P. Vernazza, D. Genne, P. Dur, M. Rothen, J. P. Unger, and D. Lew.** 2002. Prospective epidemiologic survey of patients with community-acquired pneumonia requiring hospitalization in Switzerland. Int. J. Infect. Dis. **6:**288-293.

44. **Garbino, J., S. Crespo, J.D. Aubert, T. Rochat, B. Ninet, C. Deffernez, W. Wunderli, J. Pache, P. M. Soccal, and L. Kaiser.** 2006. A prospective hospital-based study of the clinical impact of non-severe acute respiratory syndrome (Non-SARS)-related human coronavirus infection. CID **43:**1009-1015.

45. **Guy, J. S., J. J. Breslin, B. Breuhaus, S. Vivrette, and L. G. Smith**. 2000. Characterisation of a coronavirus isolated from a diarrheic foal. J. Clin. Microbiol. **38:**4523-4526.

46. **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.

47. **Hamre, D., and J. J. Procknow.** 1966. A new virus isolated from the human respiratory tract. Proc. Soc. Exp. Biol. Med. **121:**190-193.

48. **He, R.**, **A. Leeson, M. Ballantine, A. Andonov, L. Baker, F. Dobie, Y. Li, N. Bastien, H. Feldman, U. Strocher, S. Theriault, T. Cutts, J. Cao, T. F. Booth, F. A. Plummer, S. Tyler, and X. Li.** 2004. Characterization of protein-protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus. Vir. Res. **105:**121-125.

49. **He, Y., Y. Zhou, P. Siddiqui, J. Niu, and S. Jiang.** 2005. Identification of immunodominant epitopes on the membrane protein of the severe acute respiratory syndrome-associated coronavirus. J. Clin. Micr. **43:**3718-3726.

50. **Hemming, V. G., W. Rodriguez, H. W. Kim, C. D. Brandt, R. H. Parrott, B. Burch, G. A. Prince, P. A. Baron, R. J. Fink, and G. Reaman.** 1987. Intravenous immunoglobulin treatment of respiratory syncytial virus infections in infants and young children. Antimicrob. Agents Chemother. **31:**1882-1886.

51. **Henrickson, K. J., S. Hoover, K. S. Kehl, and W. Hua.** 2004. National disease burden of respiratory viruses detected in children by polymerase chain reaction. Pediatr. Infect. Dis. J. **23:**S11-S18.

52. **Herrewegh, A. A., I. Smeenk, M. C. Horzinek, P. J. Rottier, and R. J. de Groot.** 1998. Feline coronavirus type II strains 79-1683 and 79-1146 originate from a double recombination between feline coronavirus type I and canine coronavirus. J. Virol. **72:**4508-4514.

53. **Hofmann, H., K. Pyrc, L. van der Hoek, M. Geier, B. Berkhout, and S. Pohlmann.** 2005. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. Proc. Natl. Acad. Sci. USA. **102:**7988-7993.

54. **Holmes, K. V.** 2001. p. 1187-1203. *In* B. N. Fields, D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Strauss (eds.), Fields Virology. Lippincott-Raven Publishers, Philadelphia.

55. **Holmes, K. V. and M. M. C. Lai.** 1996. *Coronaviridae*: The viruses and their replication, p. 1075-1093. *In* B. N. Fields, D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Strauss. (eds.), Fields Virology. Lippincott-Raven Publishers, Philadelphia.

56. **Hunt, I.** 2005. From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. Prot. Expres. Purific. **40:**1-22.

57. **Imai, Y., K. Kuba, S. Rao, et al.** 2005. Angiotensin-converting enzyme 2 protects from severe acute lung failure. Nature. **436:**112-116.

58. **Kaiser, L. N. Regamey, H. Roiha, C. Deffernez, and U. Frey.** 2005. Human coronavirus NL63 associated with lower respiratory tract symptoms in early life. Pediatr. Infect. Dis. J. **24:**1015-1017.

59. **King, L. A., and R. D. Possee.** 1992. The baculovirus expression system: a laboratory guide, p. 1-229. Chapman & Hall, London.

60. **Kitts, P. A., M. D. Ayres, and R. D. Possee.** 1990. Linearization of baculovirus DNA enhances the recovery of recombinant virus expression vectors. Nucleic Acids Res, **18:**5667-5672.

61. **Koetz, A., P. Nilsson, M. Linden, L. van der Hoek, and T. Ripa.** 2006. Detection of human coronavirus NL63, human metapneumovirus and respiratory syncytial virus in children with respiratory tract infections in south-west Sweden. Clin. Microbiol. Infect. **12:**1089-1096.

62. **Konig, B. W. Konig, R. Arnold, H. Werchau, G. Ihorst, and J Forster.** 2004. Prospective study of human metapneumovirus infection in children less than 3 years of age. J. Clin. Microbiol. **42:** 4632-4635.

63. **Kost, T. A., and J. P. Condreay.** 1999. Recombinant baculoviruses as expression vectors for insect and mammalian cells. Curr. Opin. Biotechnol. **10:**428-433.

64. **Krijnse Locker, J. M., D.-J. E. Opstelten, M. Ericsson, C. Horzinek, and P. J. M. Rottier.** 1995. oligomerization of a trans-Golgi/trans-Golgi network retained protein occurs in the Golgi complex and may be part of its retention. J. Biol. Chem. **270:**8815-8821.

65. **Krogh, A., B. Larsson, G. von Heijne, and E. L. L. Sonnhammer.** 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J. Mol. Biol. **305(3):**567-580.

66. **Ksiazek, T. G., D. Erdman, C. S. Goldsmith, et al.** 2003. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1953-1966.

67. **Kuba, K., Y. Imai, S. Rao, et al.** 2005. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. Nat. Med. **11:**875-879.

68. **Kubo, H., Y. K. Yamada, and F. Taguchi.** 1994. Localization of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein. J. Virol. **68:**5403-5410.

69. **Kuo, L., and P. S. Masters.** 2002. Genetic evidence for a structural interaction between the carboxy termini of the membrane and nucleocapsid proteins of mouse hepatitis virus. J. Virol. **76:**4987-4999.

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

71. **Kuo, L., G. J. Godeke, M. J. Raamsman, P. S. Masters, P. J. Rottier.** 2000. Retargeting of Coronavirus by substitution of the spike glycoprotein ectodomain: Crossing the host cell species barrier. J. Virol. **74:**1393-1406.

72. **Lai, M. M., R. S. Baric, S. Makino, J. G. Keck, J. Egbert, J. L. Leibowitz, and S. A. Stohlman.** 1985. Recombination between non-segmented RNA genomes of murine coronaviruses. J. Virol. **56:**449-456.

73. **Larson, H. E., S. E. Reed, and D. A. Tyrrell.** 1980. Isolation of rhinoviruses and coronaviruses from 38 colds in adults. J. Med. Virol. **5:**221-229.

74. **Lau, S. K., P. C. Woo, K. S. Li, Y. Huang, H. W. Tsoi, B. H. Wong, S. S. Wong, S. Y. Leung, K. H. Chan, and K. Y. Yeun.** 2005. Severe acute respiratory syndrome corona-like virus in Chinese horseshoe bats. Proc. Natl. Acad. Sci. USA. **102:**14040-14045.

75. **Laude, H., J. Gelfi, L. Lavenant, and B. Charley.** 1992. Single amino acid changes in the viral glycoprotein M affect induction of alpha interferon by the coronavirus transmissible gastroenteritis virus. J. Virol. **66:**743-749.

76. **Laude, H., J. M. Capsal, J. Gelfi, S. Labiau, and J. Grosclaude.** 1986. Antigenic structure of transmissible gastroenteritis virus. I. Properties of monoclonal antibodies directed at against virion proteins. J. Gen. Virol. **67**(Pt.1):119-130.

77. **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. Nature **426:**450-454.

78. **Li, W., Z. Shi, M. Yu, et al.** 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science **310:**676-679.

79. **Luckow, V. A.** 1991. Cloning and expression of heterologous genes in insect cells with baculovirus vectors, p. 97-152. *In* A. Prokop, R. K. Bajpai, and C. Ho (ed.), Recombinant DNA technology and applications. McGraw-Hill, New York.

80. **Luckow, V. A., C. S. Lee, G. F. Barry, and P. O. Olins.** 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in *Escherichia coli*. J. Virol. **67:**4566-4579.

81. **Luckow, V. L., and M. D. Summers.** 1988. Trends in the development of baculovirus expression vectors. Bio/Technology. **6:**47-55.

82. **Macnaughton, M. R., M. H. Madge, and S. E. Reed.** 1981. Two antigenic groups of human coronaviruses detected by using enzyme-linked immunosorbent assay. Infect. Immun. **33:**734-737.

83. **Maeda, S.** 1989. Expression of foreign genes in insects using baculovirus vectors. Annu. Rev. Entomol. **34:**351-372.

84. **Masters, P. S.** 2006. The molecular biology of coronaviruses. Adv. Virus Res. **66:**193-292.

85. **Matsuyama, S., and F. Taguchi.** 2002. Receptor-induced conformational changes of murine coronavirus spike protein. J. Virol. **76:**11819-11826.

86. **McIntosh, K.** 1996. Coronaviruses. Fields virology (B. N. Fields, D. M. Knipe, and P. M. Howley, *et al.* eds), 1095 pp. Lippincott-Raven Publishers, Philadelphia.

87. **McIntosh, K., J. H. Dees, W. B. Becker, A. Z. Kapikian, and R. M. Chanock.** 1967. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. Proc. Natl. Acad. Sci. USA. **57:**933-940.

88. **Miller, L. K.** (ed.). 1997. The Baculoviruses. Plenum Press, New York.

89. **Miller, L. K.** 1988. Baculoviruses as gene expression vectors. Annu. Rev. Microbiol. **42:**177-199.

90. **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 NL63 in children hospitalized with respiratory tract infections in Belgium. BMC. Infect. Dis. **5:**6.

91. **Moya, A., E. C. Holmes, and F. Gonzalez-Candelas.** 2004. The population genetics and evolutionary epidemiology of RNA viruses.

92. **Murhammer, D. W.** 1991. The use of insect cell cultures for recombinant protein synthesis: engineering aspects. Appl. Biochem. Biotechnol. **31:**283-310.

93. **Murray, C. J. L., A. D. Lopez, C. D. Mathers, and C. Stein.** 2001. The global burden of disease 2000 project: aims, methods and data sources. *In* Global program on evidence for health policy. Geneva: World Health Organization.

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

95. **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.

96. **Nicholas, K. B., H. B. Nicholas Jr., and D. W. Deerfield.** (1997). GeneDoc: analysis and visualization of genetic variation. EMBNEW News **4:**14.

97. **O'Reilly, D. R., L. K. Miller., and V. A. Luckow.** 1992. Baculovirus expression vectors: a laboratory manual, p. 1-347. W. H. Freeman and Company, New York.

98. **Peiris, J. S., K. Y. Yuen, A. D. Osterhaus, and K. Stohr.** 2003. The severe acute respiratory syndrome. N. Engl. J. Med. **349:**2431-2441.

99. **Pijlman, G. P., J. E. van Schijndel, and J. M. Vlak.** 2003. Spontaneous excision of BAC vector sequences from bacmid-derived baculovirus expression vectors upon passage in insect cells. J. Gen. Virol. **84:**2669-2878.

100. **Possee, R. D.** 1997. Baculoviruses as expression vectors. Curr. Opin. Biotechnol. **8:**569-572.

101. **Pulford, D. J., and P. Britton.** 1991. Expression and cellular localization of porcine transmissible gastroenteritis virus N and M proteins by recombinant vaccinia viruses. Virus Res. **18:**203-217.

102. **Pyrc, K., B. J. Bosch, B. Berkhout, M. F. Jebbink, R. Dijkman, P. Rottier, and L. van der Hoek.** 2006. Inhibition of HCoV-NL63 infection at early stages of the replication cycle. Antimicrob. Agents Chemother. **50:**2000-2008.

103. **Pyrc, K., M. F. Jebbink, B. Berkhout, and L. van der Hoek.** 2004. Genome structure and transcriptional regulation of human coronavirus NL63. Virol. J. **1:**7

104. **Rice, G. I., D. A. Thomas, P. J. Grant, A. J. Turner, N. M. Hooper.** 2004. Evaluation of angiotensin-converting enzyme (ACE), its homologue ACE2 and neprilysin in angiotensin peptide metabolism. Biochem. J. **383:**45-51.

105. **Rottier, P.** 1995. The coronavirus membrane glycoprotein, p. 73-114. *In* S. G. Siddell (ed.), The *coronaviridae.* Plenum Press, Inc., New York, N.Y.

106. **Saeki, K., N. Ohtsuka, and F. Taguchi.** 1997. Identification of spike protein residues of murine coronavirus responsible for receptor-binding activity by use of soluble receptor-resistant mutants. J. Virol. **71:**9024-9031.

107. **Schildgen, O., M. F. Jebbink, M. de Vries, K. Pyrc, R. Dijkman, A. Simon, A. Müller, B. Kupfer, and L. van der Hoek.** 2006. Identification of cell lines permissive for human coronavirus NL63. J. Virol. Meth. **138:**207-210.

108. **Shay, D. K., R. C. Holman, R. D. Newman, L. L. Liu, J. W. Stout, and L. J. Anderson.** 1999. Bronchiolitis associated hospitalizations among US children, 1980-1996. JAMA. **282:**1440-1446.

109. **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. B. Talbot, J. V. Williams, and J. C. Burns.** 2005. Human coronavirus NL63 is not detected in the respiratory tracts of children with acute Kawasaki Disease. J. Infect. Dis. **192:**1767-1771.

110. **Siddell, S.** 1983. Coronavirus JHM: coding assignments of subgenomic mRNAs. J. Gen. Virol. **64(Pt1):**113-125.

111. **Sims, A. C., R. S. Baric, B. Yount, S. E. Burkett, P. L. Collins, and R. J. Pickles.** 2005. Severe acute respiratory syndrome coronavirus infection of human ciliated airway epithelia: role of ciliated cells in viral spread in the conducting airways of the lung. J. Virol. **79:**15511-15524.

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

113. **Stadler, K., V. Masignani, M. Eickman, S. Becker, S. Abrignani, H. D. Klenk, and R. Rappuoli.** 2003. SARS – beginning to understand a new virus. Nat. Rev. Microbiol. **1:**209-218.

114. **Stiehm, E. R., E. Ashida, K. S. Kim, D. J. Winston, A. Haas, and R. P. Gale.** 1987. Intravenous immunoglobulins as therapeutic agents. Ann. Intern. Med. **107:**367-382.

115. **Sturman, L. S., and K. V. Holmes.** 1977. Characterization of coronavirus. II. Glycoproteins of the viral envelope: tryptic peptide analysis. Virology **77:**650-660.

116. **Suzuki, A., M. Okamoto, A. Ohmi, O. Watanabe, S. Miyabayashi, and H. Nishimura.** 2005. Detection of human coronavirus-NL63 in children in Japan. Pediatr. Infect. Dis. J. **24:**645-646.

117. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** (1997). The ClustalX windows interface: flexible strategies for multiplesequence alignment aided by quality analysis tools. Nucleic Acids Research. **25:**4876-4882.

118. **Tyrrell, D. A. J., J. D. Almedia, D. M. Berry, C. H. Cunningham, D. Hamre, M. S. Hofstad, L. Malluci, and K. McIntosh.** 1968. Coronavirus. Nature **220:**650.

119. **Tyrrell, D. A. J., and M. L. Bynoe.** 1965. Cultivation of novel type of common-cold virus in organ cultures. Br. Med. J. **1:**1467-1470.

120. **Vabret, A., T. Mourez, S. Gouarin, J. Petitjean, and F. Freymuth.** 2003. An outbreak of coronavirus OC43 respiratory infection in Normandy, France. Clin. Infect. Dis. **36:**985-989.

121. **Vabret, A., T. Mourez, T. Dina, L. van der Hoek, S. Gouarin, J. Petitjean, J. Brouard, and F. Freymuth.** 2005. Human coronavirus NL63, France. Emerg. Infect Dis. **11:**1225-1229.

122. **Van der Hoek, L., K. Pyrc, and B. Berkhout.** 2006. Human coronavirus NL63, a new respiratory virus. FEMS Microbiol. Rev. **30:**760-773.

123. **Van der Hoek, L., K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. M. Berkhout, K. C. Wolthers, P. M. E. Wertheim-van Dillen, J. Kaandorp, J.**

Spaargaren, and B. Berkhout. 2004. Identification of a new human coronavirus. Nat. Med. **10:**368-373.

124. **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 NL63. PLoS Med. **2:**e240.

125. **Vennema, H., G.-J. Godeke, J. W. A. Rossen, W. F. Voorhout, M. C. Horzinek, D. J. E. Opstelten, and P. J. m. Rottier.** 1996. Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. EMBO J. **15:**2020-2028.

126. **Vennema, H., R. J. de Groot, D. A. Harbour, M. C. Horzinek, and W. J. Spaan.** 1991. Primary structure of the membrane and nucleocapsid protein genes of feline infectious peritonitis virus and immunogenicity of recombinant vaccinia viruses in kittens. Virology **181:**327-335.

127. **Vijgen, L., E. Keyaerts, E. Moës, P. Maes, G. Duson, M van Ranst.** 2005. Development of One-Step, Real-Time, Quantitative Reverse Transcriptase PCR Assays for Absolute Quantitation of Human Coronaviruses OC43 and 229E. J. Clin. Microbiol. **43:**5452-5456.

128. **W.H.O. Communicable Disease Surveillance and Response (CSR).** 2003. Summary of probable SARS cases with onset of illness from November 1 2002 to July 31 2003.

129. **Wesseling, J. G., G. J. Godeke, V. E. Schijns, L. Prevec, F. L. Graham, M. C. Horzinek, and P. J. Rottier.** 1993. Mouse hepatitis virus spike and nucleocapsid proteins expressed by adenovirus vectors protect mice against a lethal infection. J. Gen. Virol. **74**(Pt. 10)**:**2061-2069.

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

131. **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.** 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol. **79:**884-859.

132. **Woo, P. C., S. S. Chiu, W. H. Seto, M. Peiris.** 1997. Cost-effectiveness of rapid diagnosis of viral respiratory tract infections in pediatric patients. J. Clin. Microbiol. **35:**1579-1581.

133. **Wu, P., L. Chang, B. Berkhout, L. van der Hoek, C. Lu, C. Kao, P. Lee, P. Shao, C. Lee, F. Huang, and L. Huang.** 2007. Clinical manifestations of human coronavirus NL63 infection in children in Taiwan. Eur. J. Pediatr.

134. **Wurm, T., H. Chen, T. Hodgson, P. Britton, G. Brooks, and J. A. Hiscox.** 2001. Localization to the nucleolus is a common feature of coronavirus nucleoproteins, and the protein may disrupt host cell division. J. Virol. **75:**9345-9356.

135. **Yang, H., W. Xie, X. Xue, *et al*.** 2005. Design of wide spectrum inhibitors targeting coronavirus main proteases. PLoS Biol. **3:**e324.

136. **Yang, Y., Z. Xiong, S. Zhang, Y. Yan, J. Nguyen, B. Ng, H. Lu, J. Brendese, F. Yang, H. Wang, and X. F. Yang.** 2005. Bcl-xL inhibits T cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors. Biochem. J. **392:**135-143.

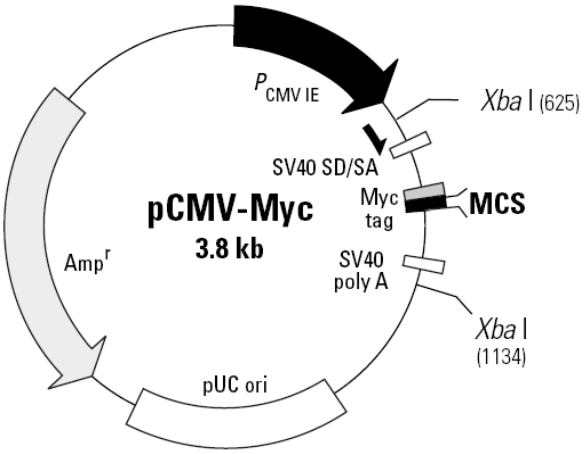
137. **Yeager, C. L., R. A. Ashmun, R. K. Williams, C. B. Cardellichio, L. H. Shapiro, A. T. Look, K. V. Holmes.** 1992. Human aminopeptidase N is a receptor for human coronavirus 229E. Nat. **357**:420-422.

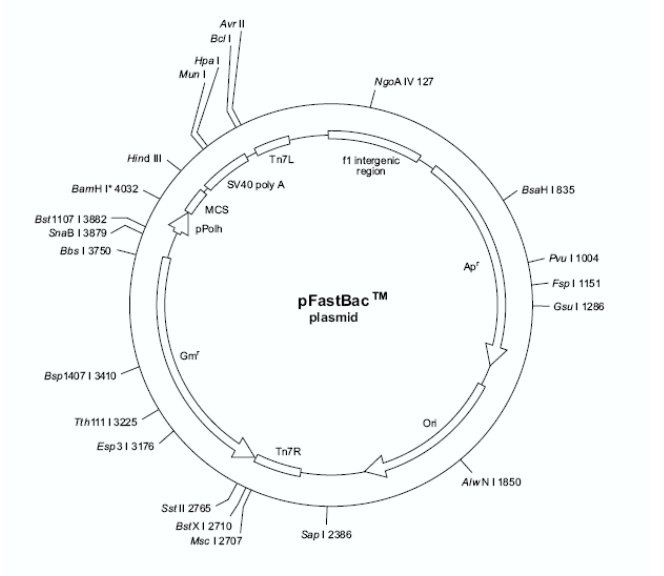
138. **Yount, B., K. M. Curtis, E. A. Fritz, L. E. Hensley, P. B. Jahrling, E. Prentice, M. R. Denison, T. W. Geisbert, and R. S. Baric.** 2003. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. Proc. Natl. Acad. Sci. USA **100:**12995-13000.

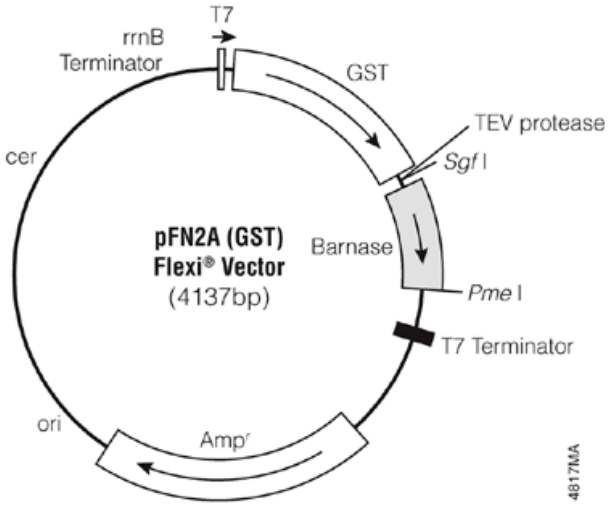
139. **Yount, B., M. R. Denison, S. R. Weiss, and R. S. Baric.** 2002. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. J. Virol. **76:**11065-11078.

Appendix A

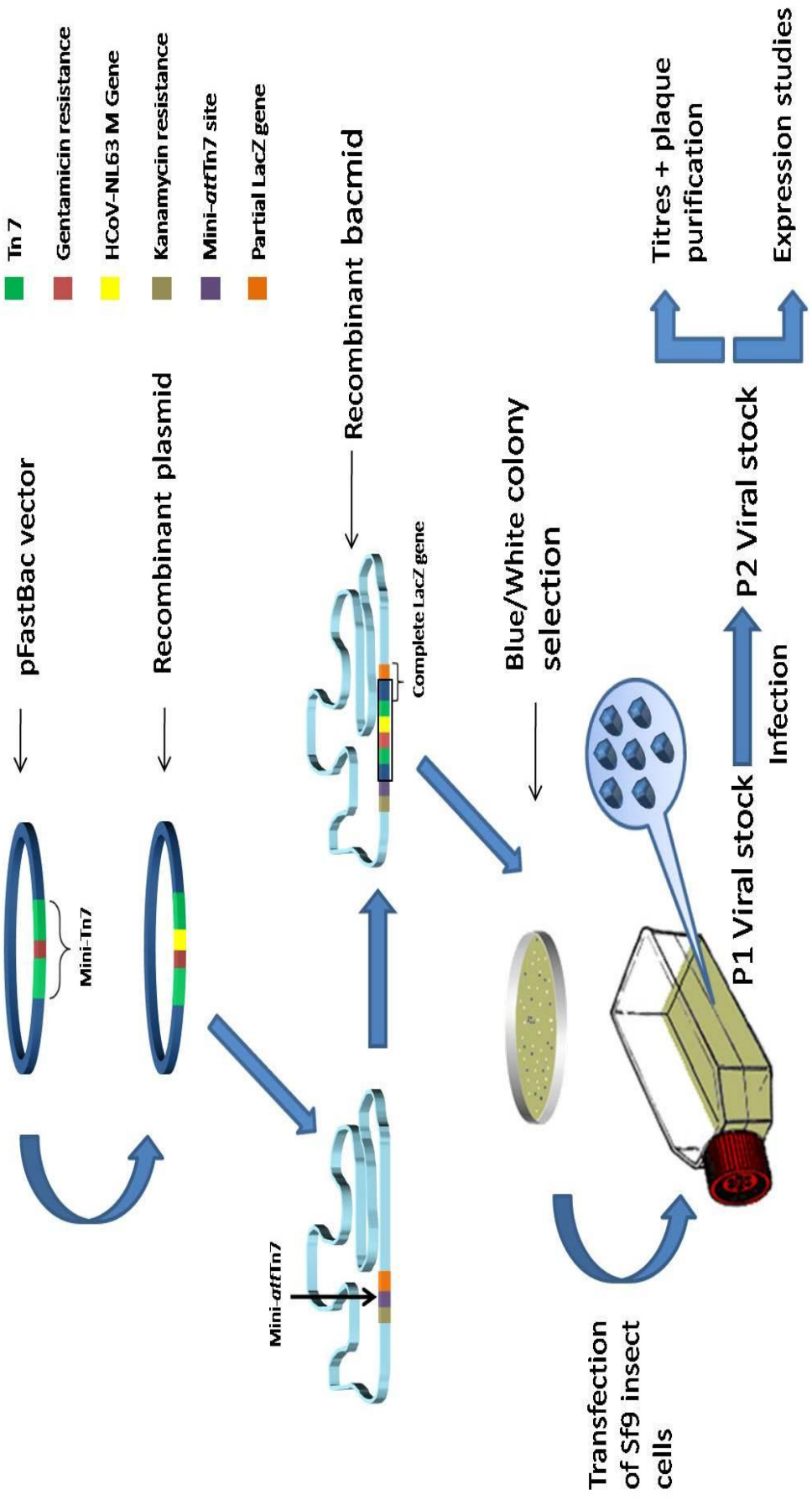
Vectors used in this study:







Appendix B


Tn 7
Gentamicin resistance
HCoV-NL63 M Gene
Kanamycin resistance
Mini-attTn7 site
Partial LacZ gene
pFastBac vector
Mini-Tn7
Recombinant plasmid
Mini-attTn7
Recombinant bacmid
Complete LacZ gene
Blue/White colony selection
Transfection of Sf9 insect cells
P1 Viral stock
Infection
P2 Viral stock
Titres + plaque purification
Expression studies

Appendix C

Solutions used:

1. **4x stacking gel buffer**

   60,5g Tris in 1000ml distilled water at a pH of 6,8.

2. **4x resolving gel buffer**

   181,5g Tris in 1000ml distilled water at a pH of 8,8.

3. **10 x running buffer**

   30g Tris, 144,0g glycine and 10,0g SDS in 1000ml distilled water at a pH 8,3

4. **10% SDS stock solution**

   10g SDS in 100ml distilled water.

5. **2x SDS-PAGE sample buffer**

2ml	4x stacking gel buffer
1,6 ml	Molecular grade glycerol
3,2ml	10% SDS
0,8ml	β-mercaptoethanol
0,4ml	1% Bromophenol blue

6. Separating/resolving gel (15%)

3,15ml	Distilled water
3,75ml	Acryl/bisacrylamide
2,5ml	4x resolving buffer
0,1ml	10% SDS
10µl	TEMED
0,5ml	APS

7. Stacking gel

2.86ml	Distilled water
0,3ml	Acryl/bis acrylamide
0,5ml	Stacking gel buffer
0,04ml	10% SDS
10µl	TEMED
0,3ml	APS

8. Transfer buffer (10x)

30,3g Tris and 144,1g glycine in 1000ml distilled water

Made up to 1x with distilled water and 20% methanol for transfers.

9. 3% Milk blotting solution

3g	Fat Free milk powder (Biorad)
100ml	PBS (with 0,5% Tween)

10. Coomassie Brilliant Blue R250 staining solution

0,25g	Coomassie Brilliant Blue R250 powder
90ml	MetOH:H2O (1:1 v/v)
10ml	Glacial acetic acid

11. Coomassie Brilliant Blue Destaining Solution

90ml	MetOH:H2O (1:1 v/v)
10ml	Glacial acetic acid
10%	Glycerol