Production of cytokines in human whole blood after incubation with the nucleocapsid protein of the NL63 Coronavirus

by

Aasiyah Chafekar

Thesis Submitted in Fulfillment of the Requirements for the Degree

MSc

Medical Biosciences

Faculty of Natural Sciences

University of the Western Cape

Supervisor: Prof B.C. Fielding

Department of Medical Biosciences, University of the Western Cape

November 2012

KEYWORDS

Human coronavirus NL63, severe acute respiratory syndrome (SARS), nucleocapsid protein, bacterial expression of proteins, cytokine production, interleukin 6, interferon-gamma, whole blood cultures

ABSTRACT

The *Coronaviridae* family consists of RNA viruses within the order *Nidovirales*. The family is classified into two genera, namely the corona- and toroviruses. Coronaviruses are enveloped, single stranded, positive sense RNA viruses with genomes ranging between 27-32kb in size. The 5' two-thirds of the genome encodes for the 1a/b polyprotein, while the 3' one-third of the genome encodes for the structural proteins that mediate viral entry into the host cell. These structural proteins include the spike (S), envelope (E), membrane (M) and nucleocapsid (N) proteins.

The nucleocapsid protein is expressed at high levels within an infected cell. Studies have shown that this protein plays a key regulatory role in different cellular pathways, including the inhibition of interferon production and the up-regulation of the AP1 signal transduction pathway, amongst others. Also, the N protein is vital in the formation of the ribonucleocapsid core by binding to the viral RNA during virion assembly. The focus of this study is the immune response in whole blood cultures to the presence of human coronavirus (HCoV) NL63 N protein.

To characterise the stimulation of the immune activity against HCoV-NL63 N in blood cultures, the HCoV-NL63 N gene was expressed in a bacterial system. In this pilot study, GST-tagged N constructs were then purified and used to treat whole blood cultures from three volunteers. ELISAs were used to measure the cytokine response in these treated whole blood cultures. Results showed that the nucleocapsid protein has an inflammatory response on whole blood cultures. These results have generated vital information in the potential function of the HCoV-NL63 N protein on the immune system. It is suffice to say that the HCoV-NL63 N protein is able to elicit an effective inflammatory response within the host cell. Future studies

into the cellular pathways affected by the HCoV-NL63 N protein will clarify its exact role in stimulating the host immune system.

UNIVERSITY of the
WESTERN CAPE

\

DECLARATION

I, Aasiyah Chafekar, declare that this thesis, *Production of cytokines in human whole blood after incubation with the nucleocapsid protein of NL63-CoV* hereby submitted to the University of the Western Cape for the degree of *Magister Scientiae* (MSc) has not previously been tendered by me for a degree at this or any other university or institution, that it is my own work in design and in execution, and that all materials contained herein have been duly acknowledged.

Date: _____

Signature: _____

UNIVERSITY *of the*
WESTERN CAPE

ACKNOWLEDGEMENT

Research Output from this thesis

Oral and Poster Presentations:

1.	A Chafekar, M Berry, E Pool and BC Fielding, Human coronavirus NL63 nucleocapsid protein stimulates an immune response in whole blood culture, 22nd IUBMB & 37th FEBS Congress (2012), Seville, Spain, 2-9 September 2012 (**POSTER Presentation**).

2.	A Chafekar, M Berry, E Pool and BC Fielding, Detection of serum antibodies to the human coronavirus-NL63 nucleocapsid protein, 39th Annual Conference of The Physiology Society of Southern Africa (PSSA), 28-31 August 2011 (**ORAL Presentation**).

Abstracts published in International Journals:

1.	Chafekar, A., Berry, M., Pool, E. and Fielding, B. C. (2012) Human coronavirus NL63 nucleocapsid protein stimulates an immune response in whole blood culture, 22nd IUBMB & 37th FEBS Congress, Seville, Spain, 4-9 September 2012, (Perham, R., Ed.) **FEBS Journal**, Federation of European Biochemical Societies, p 559.

2.	Chafekar, A., Berry, M. S., Pool, E., and Fielding, B. C. (2012) Human coronavirus-NL63 nucleocapsid protein induced immune response in human whole blood cultures, 39th Conference of the Physiology Society of Southern Africa, Cape Town, South Africa, 29-31 August 2011, (Fielding, B., Dietrich, D., and Van der Horst, G., Eds.) **Scientific Research and Essays** 7, Academic Journals, p 22-23.

TABLE OF CONTENTS

Chapter I: Literature Review

Chapter II: Cloning and expression of HCov-NL63 proteins

Chapter III: Immune response to HCoV-NL63 N in whole blood culture

LIST OF TABLES

UNIVERSITY *of the*
WESTERN CAPE

LIST OF FIGURES

OD$_{450}$ nm

UNIVERSITY *of the*
WESTERN CAPE

LIST OF APPENDIXES

Appendix III: Cloning into the pGEX-4T-2expression vector for further studies using the IPTG inducible system.

Chapter I

The Human Coronavirus HCoV-NL63:A Review

1.1 Background

Coronaviruses form a genus of the *Coronaviridae* family, with the name "Coronavirus" derived from the Latin word *corona*, meaning "crown". When viewed under an electron microscope, a ring of small bulbous structures can be observed in the viral envelope under an electron microscope, and this gives the virus its characteristic appearance and its name (Lin, Feng et al. 2008; Smuts 2008; Woo, Huang et al. 2010).

1.2 Discovery

The history of research into human coronaviruses began in 1965 when Tyrrell and Bynoe found that a B814 virus could be passaged in human embryonic tracheal organ cultures. Although unable at the time to grow the virus in tissue culture, they would confirm B814's presence by inoculating the medium intra-nasally into human volunteers, resulting in a significant number of the volunteers exhibiting common cold-like symptoms (Tyrrell and Bynoe 1965; Tyrrell and Bynoe 1966).
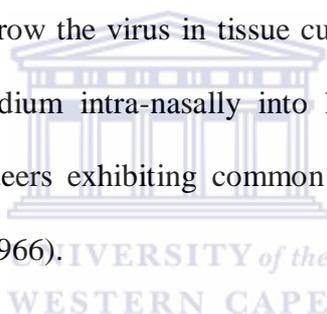
At a later stage, Hamre and Procknow (1966) isolated a virus, obtained from medical students with the common cold, in tissue culture. This virus would later be named Human Coronavirus 229E (HCoV-229E). HCoV-229E, together with the B814 virus, was found to be ether sensitive, presumably requiring a lipid coat for infectivity to occur. By using a similar technique to Tyrell, McIntosh et al (1967) were able to recover several other ether sensitive viruses in organ cultures, the viruses were termed "OC" to indicate that they were grown in organ cultures. Within the same time frame Tyrrell and Bynoe found particles in the electron microscopy of organ cultures infected with the B814 virus to be similar in morphology to the infectious bronchitis virus of chickens, as well as to the mouse hepatitis virus (MHV) and the transmissible gastroenteritis virus of swine.

In 1975 these viruses were officially classified into a new genus, *coronaviruses.* Unfortunately, many clinical samples collected in the 1960s that proved to be positive for coronavirus-like particles were subsequently lost. In the three decades after the initial discovery, extensive research attention was paid almost exclusively to the study of Hamre's HCoV-229E and McIntosh's HCoV-OC43 viruses as they were the easiest to isolate and study at the time. Despite this, it soon became clear that there were other coronavirus strains circulating in human populations.

1.3 Characteristics and taxonomy

Collectively *Coronaviridae*, *Arteriviridae* and *Roniviridae* belong to the order *Nidovirales* as all the viruses in this order produce a 3′ co-terminal nested (Latin name *nidus* meaning "nest") set of subgenomic mRNA′s during infection (Cavanagh, Brian et al. 1993; Woo, Huang et al. 2010). The *Coronaviridae* family is comprised of two genera, one being the actual Coronaviruses (CoV) and the other the Toroviruses. Originally Coronaviruses were classified according to the antigenic cross-reactivity into three antigenic groups. However, as the Coronavirus genomes were sequenced the original antigenic groups were converted into genetic groups, based on their nucleotide sequences (Liu and Sun 2008; Woo, Huang et al. 2010).

1.3.1 Group I

Group I Coronaviruses contain three major viruses, HCoV-229E, Porcine Transmissible Gastroenteritis Virus (TGEV) and the Porcine Epidemic Diarrhoea Virus (PEDV), which are responsible for gastroenteritis and lower respiratory tract infections. TGEV in particular can have a severe impact on agriculture as gastroenteritis causes morbidity among pigs, and consequently high mortality rates among piglets (Enjuanes, Almazan et al. 2006). In March

2004, a novel human pathogenic coronavirus isolated from a seven-month-old baby who had bronchiolitis and conjunctivitis designated HCoV-NL63 was discovered (van der Hoek, Pyrc et al. 2004).

1.3.2 Group II

Other pathogenic coronaviruses that belong to Group II are HCoV-OC43 and the recently discovered HCoV-HKU1, which are also responsible for the common cold in humans. HCoV-OC43 has also been associated with neurological diseases like multiple sclerosis (Edwards, Denis et al. 2000). Group II is also comprised of the Murine Hepatitis Virus (MHV), which causes bronchiolitis and infects the liver and brain of mice. MHV has previously been used as a model organism to study the pathogenesis of coronaviruses (Perlman and Ries 1987).

1.3.3 Group III

The avian Infectious Bronchitis Virus (IBV) and the Turkey Coronaviruses are the only members of Group III to date (Jia, Mondal et al. 2002). IBV is one of the main causes of economic loss within the poultry industry as it causes severe bronchiolitis and nephritis (Cavanagh, Brian et al. 1993).

1.4 Emergence of SARS Coronavirus

Given the enormous diversity of animal coronaviruses, it was not surprising when in November 2002 the cause of an outbreak of atypical pneumonia in the Guangdong province of South China was identified as a coronavirus, which soon after spread globally at a quantifiable speed (Ksiazek, Erdman et al. 2003; Peiris, Lai et al. 2003). The causative agent was identified as the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV)

(Drosten, Gunther et al. 2003; Peiris, Lai et al. 2003). During the 2002-2003 outbreaks, SARS infection was reported in 29 countries with a total of 8098 patients identified and 774 deaths recorded worldwide. A highly effective infection control response halted the SARS epidemic, and SARS-CoV is currently not in circulation (Pyrc, Berkhout et al. 2007).

It is believed that SARS-CoV most likely originated from a wild animal reservoir and was transmitted zoonotically to humans via infected Himalayan palm civets, which are sold as food in live-animal markets in China (Webster 2004; Cheng, Lau et al. 2007). Curiously, previous sero-epidemiology studies conducted among food market workers in SARS epidemic regions indicated that 40% of wild animal traders and 20% of those who slaughter animals were both seropositive for SARS-CoV, even though they were all asymptomatic (Guan, Zheng et al. 2003). These findings indicate that these individuals were previously exposed through their occupations to a SARS-like virus that had resulted in an asymptomatic infection (Drosten, Preiser et al. 2003).The SARS epidemic cast a light on the world of coronavirus research, and infused a level of attention that contributed to the detection and characterisation of previously unidentified respiratory viruses. Since 2003, five new human coronaviruses have been discovered, four of which are still circulating in the human population.

Interestingly, coronaviruses as single-stranded RNA viruses have been shown to be fully equipped to adapt to rapidly changing ecological niches as they have a high substitution rate (Woo, Lau et al. 2009; Lau, Li et al. 2010). Also, homologous recombination has been observed *in vitro* and *in vivo*, implying that mutated or recombinant coronavirus variants can emerge anywhere at any time, possibly putting public health at risk as happened in case of SARS-CoV (Gibbs, Gibbs et al. 2004; Esposito, Bosis et al. 2006) .

1.5 HCoV-NL63

1.5.1 Identification and Isolation

In January 2003, a 7-month-old baby was admitted in Amsterdam hospital with coryza, conjunctivitis and fever, with a chest radiography depicting typical bronchiolitis features. A nasopharyngeal aspirate specimen was collected within five days, but all diagnostics yielded negative results. As a result, the clinical sample was inoculated onto monkey kidney cells and the cytopathic effect was observed. Initial PCR products showed that the sequence was in fact similar in sequence to other members of the *Coronaviridae* family (Brittain-Long, Nord et al. 2008).
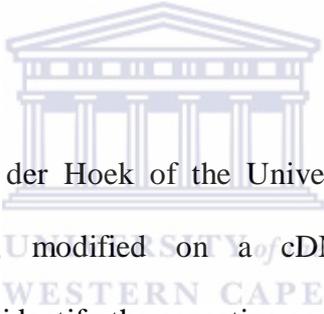
A group led by Lia Van der Hoek of the University of Amsterdam developed a VIDISCA method (Figure 1), modified on a cDNA-amplified restriction length polymorphism (cDNA-AFLP), to identify the causative agent. This technique utilises viruses that have been isolated from cell culture supernatants of *in vitro* samples that display a cytopathic effect. Supernatants are treated with DNase to remove any mitochondrial and chromosomal DNA from degraded cells that interfere with the outcome. Subsequent use of frequently-cutting restriction enzymes were used to digest cDNA, allowing this method to be specific. In so doing, "sticky end" products can be ligated into anchors for amplification with selective primers.
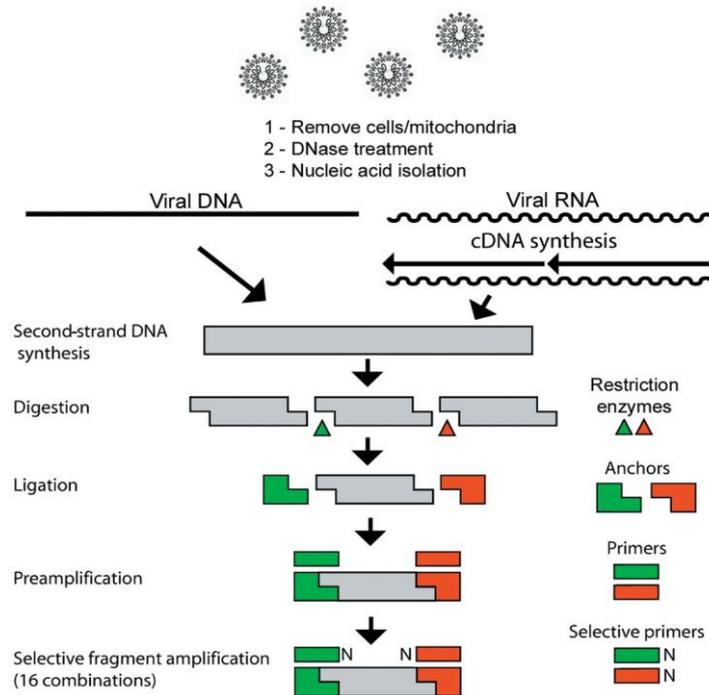
Figure 1.1: VIDISCA technique used to identify the new HCoV-NL63 (van der Hoek, Pyrc et al. 2004). VIDISCA utilizes the restriction of cDNA generated by RT-PCR of isolated viral RNA. As the restriction sites are known they, the digested fragments can be amplified with specific primers (van der Hoek, Pyrc et al. 2006)

1.5.2 Seasonality and Prevalence of infection

The first publications of HCoV-NL63 described several infected patients, indicating that infection by this coronavirus is not a rare event (van der Hoek, Pyrc et al. 2004). Patients infected with HCoV-NL63 suffer from both upper and lower respiratory tract infections, with symptoms being similar to that of HCoV-229E and HCoV-OC43 (Abdul-Rasool and Fielding 2010).

Previous literature has indicated that HCoV-OC43 and HCoV-229E are predominantly transmitted over the winter season in temperate climate countries (Hendley,

Fishburne et al. 1972). Interestingly, HCoV-NL63 shows the same preference for the winter season, as studies conducted in Belgium, Australia, Japan, Canada, Germany and France have also indicated that the highest prevalence of this virus is within the winter season (Arden, Nissen et al. 2005; Bastien, Anderson et al. 2005; Ebihara, Endo et al. 2005; Moes, Vijgen et al. 2005; Vabret, Dina et al. 2008). Collectively these reports show that, although discovered in The Netherlands, HCoV-NL63 is not unique to this region but has a global distribution. In a Japanese study on 118 children under the age of 2 who were admitted into hospital with bronchiolitis, 3 were revealed to be positive for HCoV-NL63 infection (Ebihara, Endo et al. 2005). A second study further revealed five HCoV-NL63 infections among 419 specimens collected from children with respiratory infections(Suzuki, Okamoto et al. 2005).
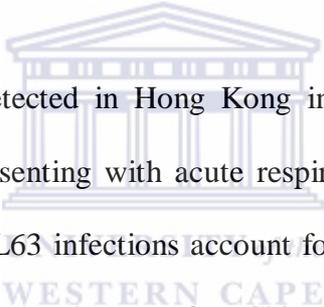
HCoV-NL63 was also detected in Hong Kong in a prospective clinical study on children under the age of 18 presenting with acute respiratory tract infections. This study estimated that annually HCoV-NL63 infections account for the hospitalisation of more than 200 children for every 100 000 under the age of 6 (Chiu, Chan et al. 2005; Lau, Woo et al. 2006). A population study in Germany found that 49 out 919 children were positive for HCoV-NL63 (Forster, Ihorst et al. 2004; Konig, Konig et al. 2004; van der Hoek, Sure et al. 2005). Studies conducted all indicate that the proportion of positive tests to be highest in the 0-5 year age group. HCoV-NL63 infection is also frequently observed in patients who are immunocompromised due to HIV infection or chronic obstructive pulmonary disease (Fouchier, Hartwig et al. 2004; Arden, Nissen et al. 2005; Bastien, Anderson et al. 2005; Moes, Vijgen et al. 2005).

HCoV-NL63 infections have been detected in 1.0-9.3% of cases of respiratory tract infections in children, 75% of them aged 2.5 to 3.5 years, the children found to be sero-

positive for the virus (Dijkman, Jebbink et al. 2008; Fielding 2011). Coronaviruses epidemiology is poorly understood mainly because these viruses, including HCoV-NL63, are for a number of reasons, not routinely tested for. Firstly, the respiratory sites chosen for clinical sampling all produced different results obtained (Kleines, Scheithauer et al. 2007) and, secondly, the difference in the sensitivity of the various diagnostic assays routinely used (Fielding 2011). As a result, the development of sensitive and specific screening methods for human coronaviruses is important in understanding the epidemiology of these viruses.

1.5.3Disease association

1.5.3.1 Croup

It has been shown that HCoV-NL63 contributes significantly to lower respiratory tract infections (LRTI). LRTI's have been globally shown to be the leading cause of morbidity in children less than 5 years of age (van der Hoek, Sure et al. 2005; Kuypers, Martin et al. 2007; Dominguez, Robinson et al. 2009). Croup is a common manifestation of LRTI in children, and is assumed to be caused by a respiratory virus, with para-influenza and influenza A frequently implicated (Klig 2004). Croup is the inflammation of the trachea and is characterised by pharyngitis, sore throat and hoarseness of voice, and also produces a "barking cough" that usually worsens at night (van der Hoek, Sure et al. 2005; Han, Chung et al. 2007; Leung, Li et al. 2009; Sung, Lee et al. 2010).

A prospective population-based study in Germany has reported an association between HCoV-NL63 and croup (Konig, Konig et al. 2004). It was found that 43% of the HCoV-NL63 positive patients without any secondary infection suffered from croup, whereas there was a prevalence of only 6% in HCoV-NL63 negative patients. The findings of this German study therefore suggest that a real link between HCov-NL63 and croup does exist

(van der Hoek, Sure et al. 2005; Weiss and Navas-Martin 2005; Sung, Lee et al. 2010). In fact, the possibility of developing croup is 6.6 times higher in HCoV-NL63 infected children, with the virus considered the second most common causative agent associated with the hospitalisation of children with croup (Fielding 2011)

1.5.3.2 Kawasaki disease

Of particular interest, HCoV-NL63 has also been associated with Kawasaki Disease (Esper, Shapiro et al. 2005). Kawasaki disease is an acute self-limiting vascular disease affecting children (Shimizu, Shike et al. 2005), and may result in aneurysms in the coronary arteries (Esper, Shapiro et al. 2005). It is the most common cause of acquired heart disease in children, although the aetiology remains unknown.

Early epidemiological evidence and confirmation has shown that the respiratory syndrome precedes the symptoms of Kawasaki disease, thereby suggesting that a respiratory virus is the likely causative agent (Bell, Brink et al. 1981; McIntosh 2005). Kawasaki disease has been shown to have a seasonal pattern that is similar to a coronavirus infection, both of which peak in the winter season (Yanagawa, Nakamura et al. 1986).

A study by Rowley et al has also shown that the Kawasaki disease will elicit a strong IgA response in the respiratory tract as well as other organs, indicating a possible entry of microbes through the respiratory tract (Rowley, Shulman et al. 2000). A case-controlled study conducted by Esper et al, investigated the association between Kawasaki disease and HCoV-NL63. Following PCR of samples collected, HCoV-NL63 was detected in 72 % of the patients that presented with Kawasaki disease. 4.5% of the control group tested positive for HCoV-NL63, thereby suggesting an association does in fact exist. This study was met with

scepticism as no corroborating studies have as yet confirmed these results (Belay, Erdman et al. 2005; Ebihara, Endo et al. 2005; Shimizu, Shike et al. 2005), Ebihara and colleagues used this method of -PCR and found that only 1 out of the 48 samples tested from patients suffering with Kawasaki disease were positive for HCoV-NL63 (Ebihara, Endo et al. 2005).

1.5.4 Co-infection

HCoV-NL63 is often found in combination with a second respiratory virus, such as Influenza A (Chiu, Chan et al. 2005) respiratory syncytial virus, para-influenza-3 or human meta-pneumovirus (A. Koetz 2006; Lambert, Allen et al. 2007; Kaplan, Dove et al. 2008). Although the viral load of HCoV-NL63 in co-infected patients is lower than a single infection with HCoV-NL63 (Debiaggi, Canducci et al. 2012), co-infected patients are more likely to be hospitalised, indicating the severity of this kind of infection (Abdul-Rasool and Fielding 2010).
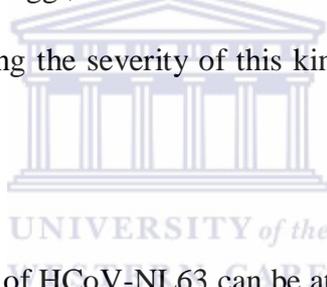
The decrease in viral load of HCoV-NL63 can be attributed to that fact that the innate immune response of the second respiratory virus inhibits the progression of HCoV-NL63, and also competes for the same target receptor in the respiratory tissues. It is also likely that HCoV-NL63 lowers the immune response, and in so doing creates a suitable environment for a subsequent infection with another virus.

It has also been proposed that HCoV-NL63 promotes bacterial co-infection by increasing the adherence capabilities of *Streptococcus pneumonia* to human airway epithelial cells (Golda, Malek et al. 2011). It has believed that respiratory viruses like HCoV-NL63 facilitate bacterial colonisation as they damage the epithelium cells found on the respiratory tract, creating a new binding site for the bacteria on the epithelial surface.

1.5.5 Genome structure

A common feature of all coronaviruses is the 27-32kb positive sense single-stranded RNA genome, making it the biggest RNA virus. The genome is capped at the 5' end and poly-denylated at the 3' end (Pyrc, Berkhout et al. 2007). The genome associated with HCoV-NL63 is organised in the following gene order: 5'-replicase 1a-1b-S-ORF3-E-M-N-3'. HCoV-NL63 has three open reading frames with the ORF3 serving as an additional accessory protein between the S and E genes (Dijkman, Jebbink et al. 2008).

The 5' two-thirds of the genome that encodes for the 1a/b poly-protein has a stable nucleotide count, while the 3' one-third of the genome encoding the structural proteins has an increased C-count, resulting in the GC ratio of HCv-NL63 being very low in comparison to other coronaviruses. The two ORF3's 1a and 1b encode proteins necessary for RNA replication and make up two thirds of the genome, sharing an overlapping region of 40-60 nucleotides. During translation of the 1a and 1b poly-proteins, there is a secondary RNA pseudo-knot structure that will initiate a ribosomal frame shift, triggering the ribosome to move one nucleotide upstream to encode for a completely new protein (Pyrc, Jebbink et al. 2004; Sawicki, Sawicki et al. 2007). This frame shift is initiated when the elongating ribosome is stalled by a hepta-nucleotide sequence that is located upstream of the RNA pseudo-knot structure, leading to the completion of the poly-proteins 1a and 1b(Giedroc, Theimer et al. 2000; Namy, Moran et al. 2006; Sawicki, Sawicki et al. 2007).

1.5.6 Transcription and Replication

In order to express the proteins that are necessary for replication, the coronaviruses use a regulatory post-translation proteolytic process. The poly-proteins 1a and b within HCoV-NL63 are cleaved by viral proteases to form a multi subunit protein complex that subsequently plays a role in replication and transcription. HCoV-NL63 has two proteinases in

the 5'region of the 1a poly-protein, the first being a papain-like proteinase that expresses the non-structural 3 gene (nsp 3) that will then catalytically cleave the 1a and b polyprotien at 3 different locations. The second proteinase is expressed by the nsp5 gene, which contains serine-like domains that will further allow for the cleavage of a non-structural protein at the 1a and b poly-protein.

The 3' end of the genome encodes for the structural proteins that participate in a budding process, thereby incorporating them into the viral particle. In replication, six mRNAs are produced, including a complete genomic RNA, and "nested" 5 sub-genomic RNA's are produced (Pyrc, Berkhout et al. 2007). These 5 sub-genomic mRNAs each have an identical portion of 72 nucleotides and encode for the S, ORF3, E, M and N structural proteins necessary for viral assembly (Pyrc, Jebbink et al. 2004).
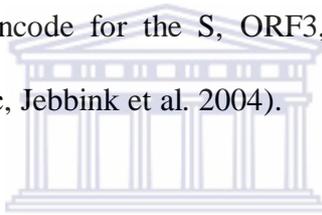
HCoV-NL63 uses a discontinuous replication mechanism, which requires the RNA-dependent RNA polymerase to pause after the transcription regulation sequence (TRS) of a specific gene is generated during negative strand synthesis. This delay allows the negative strand to be copied into positive strand mRNAs. The TRS is common to all sub-genomic mRNAs, with the exception of the E protein (Pyrc, Jebbink et al. 2004; Pyrc, Berkhout et al. 2007; Sawicki, Sawicki et al. 2007). RNA expression levels of these sg-mRNA's increases in the 3'end direction in relation to their genomic position; as a result the sg-mRNA for the N protein is found to have the highest level of expression compared to the sg-mRNA of the S protein, which has the lowest. An exception does occur with the E proteins sg-mRNA, which is lower than expected. This is due to the fact that the E protein has a sub-optimal TRS core of AACUAUA instead of AACUAAA. This point mutation in the TRS core sequence may account for the low levels of expression of the E protein in coronaviruses.

1.5.7   Cell Tropism and receptor usage

Infection of the host specific target cells occurs in coronaviruses via the binding of the S protein to the cellular receptors. The spike protein found in HCoV-NL63 is a class 1 fusion protein similar to that found in the Influenza virus HE protein and the HIV-1 envelope glycoprotein (van der Hoek, Pyrc et al. 2006). The viral replication strategy in coronaviruses begins with the binding of the spike protein to cellular receptors on the target cells. Initially it was believed that all group 1 coronaviruses, including the HCoV-229E, TGEV, PRCoV, feline and canine coronaviruses, all utilise the CD13 (aminopeptidase) receptor for infection in the target cell (Delmas, Gelfi et al. 1992; Yeager, Ashmun et al. 1992; Tresnan, Levis et al. 1996).
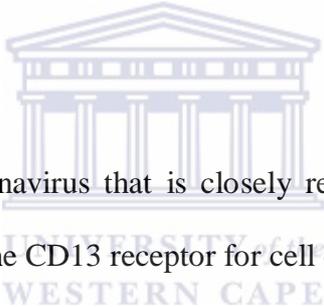
However, a group 1 coronavirus that is closely related to HCoV-229E the HCoV-NL63 has been found to not use the CD13 receptor for cell entry (Hofmann, Pyrc et al. 2005). This impact of receptor specificity could be due to that fact that the HCoV-NL63 S protein has a unique 179 amino acid sequence at this N terminal that is not homologous to any other coronavirus proteins (Hofmann, Simmons et al. 2006). In addition HCo-V-NL63 and SARS-CoV have both been shown to replicate in monkey kidney cells (Fouchier, Hartwig et al. 2004; Hofmann and Pohlmann 2004), and although from different coronavirus groups, further infect Huh-7 cells: a human hepatocellular carcinoma cell line (Hofmann, Pyrc et al. 2005) This analysis, coupled with receptor usage, has revealed that the HCoV-NL63-like SARS-CoV are the only coronaviruses that use the angiotensin-converting enzyme-II (ACE2) receptor to gain entry into the target cells (Li, Moore et al. 2003). This is puzzling as the HCoV-NL63 shares no amino acid identity with SARS-CoV. Hoffmann et al, showed that

despite a high homology between HCoV-NL63 and HCoV-229E, they form distinct surfaces at the C-terminal and thereby recognise the ACE 2 and CD13 receptors, respectively.

1.6   Structural proteins

1.6.1   Spike (S) protein

The HCoV-NL63 S protein is a type 1 glycoprotein that is found on the viral surface, giving the virus its "crown-like" structure (Bosch, van der Zee et al. 2003). The S protein is produced in the Golgi bodies, resulting in the cleavage of the S protein into two S1 and S2 subunits, each approximately 90kDa in size (Sturman and Holmes 1977). The N terminal of the S protein (S1) contains a receptor-binding domain located in the first 330 amino acids that is able to recognise the ACE2 receptor on the host target cell (Kubo, Yamada et al. 1994). The C terminal is found in the membrane region, together with two heptad regions (HR1 and HR2) and the fusion protein (van der Hoek, Pyrc et al. 2006).

The primary function of the spike protein is to mediate the viral attachment to host-specific cellular receptors, as well as facilitate viral entry into the host cell with antibodies, produced *in vitro* against the S protein having a neutralising effect (Pyrc, Bosch et al. 2006)The S and M protein have also been shown to interact during the assembly of the virus. This will prevent the translocation of the S protein to the host cells membrane surface enabling the virus to evade the host immune system instead of being processed intracellularly (Opstelten, Raamsman et al. 1995).

1.6.2 Envelope (E) protein

As mentioned previously, expression levels of the E protein remain low, making it a poorly characterised protein, this the result of the suboptimal variation of the TRS core sequence compared to other sub-genomic mRNA's. The E protein is an integral membrane protein (Yu, Bi et al. 1994; Hurst, Kuo et al. 2005), playing a critical role in viral assembly (Vennema, Godeke et al. 1996; Hurst, Kuo et al. 2005; DeDiego, Alvarez et al. 2007) and virus replication (Yang, Xiong et al. 2005).

Studies have shown that the co-expression of CoVs M and E protein in cell lines will result in the formation of a virus-like particle (Vennema, Godeke et al. 1996; Baudoux, Carrat et al. 1998; Corse and Machamer 2000; Huang, Yang et al. 2004; Bosch, de Haan et al. 2005; DeDiego, Alvarez et al. 2007), thereby playing a vital role on viral morphology. The E protein has also been shown to induce apoptosis (Weiss and Navas-Martin 2005) via the caspase-dependent pathway, as seen with the E protein of the Mouse Hepatitis Virus. In MHV is as been show that inhibition occurs via the over-expression of the anti-caspase protein Bcl2 (An, Chen et al. 1999; Yang, Xiong et al. 2005).

1.6.3 Open Reading Frame 3 (ORF3) protein

ORF3 is one of several accessory genes belonging to HCoV-NL63, which encodes the open reading frame-3 protein. The function of ORF3 protein in HCoV-NL63 is largely unknown; however, several homologues of the protein have been studied. The ORF3 protein of HCoV-NL63 is homologous to ORF3 proteins of the other group 1b coronaviruses, namely PEDV, Bt-CoV/512/2005 (Tang, Zhang et al. 2006) and the ORF4 protein of HCoV-229E.(Dijkman, Jebbink et al. 2006). The ORF3 protein has also been found to play a role in the infectivity and pathogenicity of porcine epidemic diarrhoea virus (PEDV) (group 1b) and transmissible

gastroenteritis virus (TGEV) of group 1a (Kim, Hayes et al. 2000). Studies into the use of truncated forms of the ORF3 protein after serial passages in culture resulted in viruses with decreased pathogenicity (Woods 2001; Song, Yang et al. 2003). Piglets that were inoculated with the wild type PEDV developed severe diarrhoea (Andries and Pensaert 1981), while those inoculated with the highly cell-passaged form of the virus did not exhibit such symptoms.

TGEV, which is a group 1a coronavirus, as mentioned above, is associated with vomiting, diarrhoea, fever and even fatality in pigs. This is caused by infection and destruction of the epithelium of the small intestine. It was found that serially passaged TGEV led to a modified ORF3 gene that was smaller than the corresponding wild-type, which exhibited a decrease in virulence. Piglets inoculated with the modified TGEV had a lower morbidity upon exposure to the wild-type than the control group. Hence ORF3 provides a lucrative candidate for vaccine development. The deletion of accessory genes of feline infectious peritonitis virus (FIPV) and Murine Hepatitis Virus (MHV) have not been fatal to the virus in cell culture; it has, however, resulted in attenuation *in vivo*. (Herrewegh, Vennema et al. 1995; de Haan, Masters et al. 2002; Haijema, Volders et al. 2004). This is consistent with the observations of the group 1b ORF3 protein, but has not been found to be homologous with HCoV-NL63 ORF3.

Extensive research has been done on the ORF3 protein of SARS, and although SARS ORF3a/b does not possess a high nucleotide homology with HCoV-NL63 ORF3, its location is consistent with the HCoV-NL63 homologues, i.e. between the S and E genes. (Tan, Barkham et al. 2005). The SARS ORF3 protein has been demonstrated to localise in both intracellular and plasma membranes (Song, Yang et al. 2003; Ito, Mossel et al. 2005; Yuan,

Li et al. 2005; Oostra, de Haan et al. 2006). Based on these findings, similarities between ORF3 of SARS and NL63 may be drawn

1.6.4 Membrane (M) protein

The M protein is a membrane glycoprotein that consists of three integral parts, a short N terminal, a triple membrane-spanning region and a cytoplasmic tail (Bond, Leibowitz et al. 1979). The M protein is proposed to exist independently in the endoplasmic reticulum and the Golgi complex (Lim and Liu 2001). It has been shown to interact with the envelope protein in the budding compartment of the host cell. Also, the M protein in Group 2 coronaviruses has been shown to form a complex with both the hemagglutinin esterase and spike protein, respectively. Interaction with the nucleocapsid protein is believed to participate in RNA packaging into the virus (Escors, Ortego et al. 2001; Narayanan and Makino 2001; Kuo and Masters 2002; Bosch, de Haan et al. 2005).
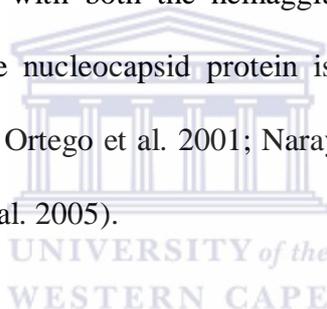
In the process of virion assembly, the M protein also interacts with the RNA- packaging signal to form the mature packaged virus particle (Narayanan, Chen et al. 2003). The interaction between the M protein and the S protein also seems to be an essential step in the assembly of the virus (He, Leeson et al. 2004; Bosch, de Haan et al. 2005). The uniqueness of the M protein is further highlighted by its ability to induce antibody production in hosts either during infection or during inoculation with attenuated viral vaccines. This takes place even though the greater part of the protein is buried within the viral membrane (Pulford and Britton 1991; Laude, Gelfi et al. 1992; Wesseling, Godeke et al. 1993; Vennema, Godeke et al. 1996).

1.6.5 Nucleocapsid (N) protein

1.6.5.1 Structure and composition

During the process of cell entry, virus particle assembly and release, the coronavirus nucleocapsid protein is produced abundantly within the infected target cell (Krokhin, Li et al. 2003). The N protein's primary function is the formation of the ribonucleocapsid helical core, which occurs when the N protein binds to a specific assembly signal on the viral RNA (He, Dobie et al. 2004; Surjit, Liu et al. 2004; Luo, Chen et al. 2005; Yu, Gustafson et al. 2005; Luo, Chen et al. 2006). Through this binding action with viral RNA dimers are formed as a result of SR-rich motifs. (Nelson, Stohlman et al. 2000; Fan, Ooi et al. 2005; Tan, Fang et al. 2006; Chen, Chang et al. 2007). As virus assembly continues, the ribonucleocapsid core will bind to the C-terminal of the membrane glycoprotein inside the endoplasmic reticulum, which allows immature virions to develop (He, Leeson et al. 2004; Hurst, Koetzner et al. 2009) and which are imperative for an effective viral replication cycle to occur.
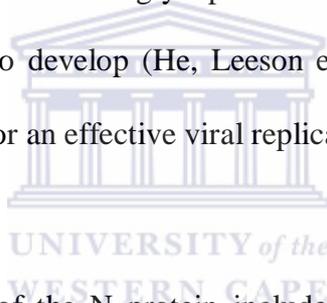
Other molecular aspects of the N protein include self-dimerisation (Surjit and Lal 2008) RNA binding capabilities and the regulation of transcription, translation and replication (Baric, Nelson et al. 1988; Almazan, Galan et al. 2004; Verheije, Hagemeijer et al. 2010). Aside from being easily expressed in a bacterial system (Tan, Fang et al. 2006), the N protein is highly immunogenic and proves to be a suitable candidate for the detection of infection via the development of ELISA. The immunogenic responses of the N protein will be discussed in further detail in subsequent chapters. The N protein has also been shown to undertake post-translation modifications like phosphorylation (Zakhartchouk, Viswanathan et al. 2005), acetylation (Dunker, Lawson et al. 2001) and sumolytion (Li, Xiao et al. 2005).

1.6.5.2 Localisation

As mentioned in the preliminary section of this chapter, *Coronaviruses* and *Arteriviruses* are placed in the order *Nidovorales* (He, Dobie et al. 2004). The N proteins of these viruses have been shown to localise in the cytoplasm (Rowland, Chauhan et al. 2005; Surjit, Kumar et al. 2005) nucleus and nucleolus (Hiscox, Wurm et al. 2001; Wurm, Chen et al. 2001; Timani, Liao et al. 2005; Reed, Dove et al. 2006). The potential function of the HCoV-NL63 N protein in regulating cell processes via its cellular localisation is currently unknown. Studies in other coronaviruses of the movement of the N proteins from the cytoplasm to the nucleoplasm have found that movement is dependant and initiated by nuclear localisation signals (NLS's) that are found on the surface of the N protein The basic amino acids, lysine and arginine, enrich these NLSs, causing them to conform to three types of motifs: either monopartitespat4 and pat7, or a bipartite motif (Rowland, Chauhan et al. 2005). The movement of N proteins from the nucleus is mediated by nuclear export signals (NESs) (You, Dove et al. 2005). The passive transport of ions and molecules is carried out by these NESs, allowing for exportation between the cytoplasm and nucleoplasm to occur (Wurm, Chen et al. 2001).

1.6.5.3 Activation of IL-6 expression through NF-Kb transcription factor

Upon examination of the lung lesions of patients suffering from an acute stage of SARS-CoV, high serum levels of IL-6 were revealed (Zhang, Wu et al. 2007). This indicates that a high inflammatory response has been brought about by the presence of the SARS-CoV in the respiratory tract. IL-6 is a pleotrophic cytokine that provides the host with a defence mechanism against a foreign substance (Dosch, Mahajan et al. 2009). Aside from mounting an immunological inflammatory response, IL-6 also plays a crucial role in bone metabolism, neoplasia, ageing and anti-immune diseases (Akira, Taga et al. 1993; Papanicolaou, Wilder et al. 1998).

The activation of IL-6 within the host cell is dependent on the activity of many transcription factors like IRF-1, AP-1, C/EBP, SP1 and NF-Kb. In the presence of viruses, bacteria and mitogens in the cytoplasm, phosphorylation and degradation of an inhibitor to which the NF-Kb transcription factor is bound will allow the NF-Kb to translocate to the nucleus. To understand the molecular mechanisms involved in an inflammatory response, Zhang et al, (2007) have shown that the N protein stimulates the expression of IL-6. They further investigated if the NF-Kb transcription factors, along with the N protein, are necessary for the expression of IL-6. Other studies also revealed that NF-Kb is necessary for the N protein to activate IL-6 expression, with both NF-Kb and the N proteins having a synergistic effect (Zhang, Wu et al. 2007).

1.6.5.4 Inhibition of IFN-

The production of a type 1 interferon within the host cell is important in the innate immune response to a viral pathogen (Cheung, Poon et al. 2005; Frieman, Heise et al. 2008). The toll-like and RIG-1 receptor dependent pathways are employed to activate Type 1 interferons, which takes place via the detection of ds or ss RNA viral genome during replication or the lysis of infected cells. After interferon synthesis, it is released from the cells and binds to a receptor that will initiate the JAK-STAT pathway, and express interferon-stimulating genes that will allow the host system to combat infection. (Kopecky-Bromberg, Martinez-Sobrido et al. 2007). SARS-CoV infection does not produce an interferon response, indicating that it has developed a mechanism to overcome this. Studies have shown that the ORF3b, ORF6, N (Frieman, Heise et al. 2008) and M (Siu, Kok et al. 2009) proteins of the SARS-CoV can suppress interferon responses.

1.6.5.5 Activation of the AP1 pathway

As seen with many other proteins, the nucleocapsid protein has also been shown to enter the nucleolus during the interphase of the cell cycle (Hiscox, Wurm et al. 2001). This may be a result of its phosphorylation after translation has occurred (Smith and Kotwal 2002; Reed, Howell et al. 2007). This movement allows the N protein to interact with the signal transduction pathways and thus evade the host immune system.

An example of this interaction was observed in the activator AP1 pathway, which is involved in the regulation of cell proliferation, differentiation and apoptosis (Shaulian and Karin 2002). There are two molecules homodimers and heterodimers that make up the AP-1 signal transduction pathway, and a study by He and colleagues has proposed that, due to these dimers, the AP-1 signal transduction pathway is activated by the SARS-CoV N protein in a dose-dependent manner (He, Leeson et al. 2003).
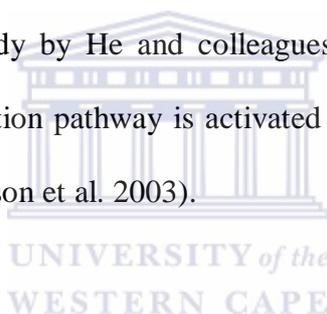
1.7 Aim of this study

1. RT-PCR amplification of HCoV-NL63 genes encoding for the structural proteins M, E, and the accessory ORF3 protein from viral RNA.

2. Cloning of these amplified products into pFLEXI vector to allow expression into a bacterial system.

3. Cloning of amplification products into a pGEX-4T-2 vector for bacterial expression.

4. Expression and purification of these structural proteins, including previously cloned and verified SARS-N and HCoV-NL63 N proteins.

Studies have indicated that the SARS-N protein will elicit an immune response in various *in vitro* systems, causing the release of inflammatory cytokines and inhibiting interferon production. As a result, is it plausible to infer that the HCoV-NL63 protein will cause a similar response? Therefore our hypothesis is that the inflammatory cytokine IL-6 is stimulated by these recombinant viral proteins (SARS-N and HCoV-NL63 N) in whole blood cell (WBC) culture systems. Our alternate hypothesis is that these viral proteins do not elicit an immune response.

1.8 References

A. Koetz, P. N., M. Linde´n, 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.pdf>."

Abdul-Rasool, S. and B. C. Fielding (2010). "Understanding Human Coronavirus HCoV-NL63." The open virology journal **4**: 76-84.

Akira, S., T. Taga, et al. (1993). "Interleukin-6 in biology and medicine." Adv Immunol **54**: 1-78.

Almazan, F., C. Galan, et al. (2004). "The nucleoprotein is required for efficient coronavirus genome replication." J Virol **78**(22): 12683-12688.

An, S., C. J. Chen, et al. (1999). "Induction of apoptosis in murine coronavirus-infected cultured cells and demonstration of E protein as an apoptosis inducer." J Virol **73**(9): 7853-7859.

Andries, K. and M. Pensaert (1981). "Vomiting and wasting disease, a coronavirus infection of pigs." Adv Exp Med Biol **142**: 399-408.

Arden, K. E., M. D. Nissen, et al. (2005). "New human coronavirus, HCoV-NL63, associated with severe lower respiratory tract disease in Australia." J Med Virol **75**(3): 455-462.

Baric, R. S., G. W. Nelson, et al. (1988). "Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription." J Virol **62**(11): 4280-4287.

Bastien, N., K. Anderson, et al. (2005). "Human coronavirus NL63 infection in Canada." J Infect Dis **191**(4): 503-506.

Baudoux, P., C. Carrat, et al. (1998). "Coronavirus pseudoparticles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes." J Virol **72**(11): 8636-8643.

Belay, E. D., D. D. Erdman, et al. (2005). "Kawasaki disease and human coronavirus." J Infect Dis **192**(2): 352-353; author reply 353.

Bell, D. M., E. W. Brink, et al. (1981). "Kawasaki syndrome: description of two outbreaks in the United States." N Engl J Med **304**(26): 1568-1575.

Bond, C. W., J. L. Leibowitz, et al. (1979). "Pathogenic murine coronaviruses. II. Characterization of virus-specific proteins of murine coronaviruses JHMV and A59V." Virology **94**(2): 371-384.

Bosch, B. J., C. A. de Haan, et al. (2005). "Spike protein assembly into the coronavirion: exploring the limits of its sequence requirements." Virology **334**(2): 306-318.

Bosch, B. J., R. van der Zee, et al. (2003). "The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex." J Virol **77**(16): 8801-8811.

Brittain-Long, R., S. Nord, et al. (2008). "Multiplex real-time PCR for detection of respiratory tract infections." J Clin Virol **41**(1): 53-56.

Cavanagh, D., D. A. Brian, et al. (1993). "The Coronaviridae now comprises two genera, coronavirus and torovirus: report of the Coronaviridae Study Group." Adv Exp Med Biol **342**: 255-257.

Chen, C. Y., C. K. Chang, et al. (2007). "Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA." J Mol Biol **368**(4): 1075-1086.

Cheng, V. C., S. K. Lau, et al. (2007). "Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection." Clin Microbiol Rev **20**(4): 660-694.

Cheung, C. Y., L. L. Poon, et al. (2005). "Cytokine responses in severe acute respiratory syndrome coronavirus-infected macrophages in vitro: possible relevance to pathogenesis." J Virol **79**(12): 7819-7826.

Chiu, S. S., K. H. Chan, et al. (2005). "Human coronavirus NL63 infection and other coronavirus infections in children hospitalized with acute respiratory disease in Hong Kong, China." Clin Infect Dis **40**(12): 1721-1729.

Corse, E. and C. E. Machamer (2000). "Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles." J Virol **74**(9): 4319-4326.

de Haan, C. A., P. S. Masters, et al. (2002). "The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host." Virology **296**(1): 177-189.

Debiaggi, M., F. Canducci, et al. (2012). "The role of infections and coinfections with newly identified and emerging respiratory viruses in children." Virol J **9**(1): 247.

DeDiego, M. L., E. Alvarez, et al. (2007). "A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo." J Virol **81**(4): 1701-1713.

Delmas, B., J. Gelfi, et al. (1992). "Aminopeptidase N is a major receptor for the entero-pathogenic coronavirus TGEV." Nature **357**(6377): 417-420.

Dijkman, R., M. F. Jebbink, et al. (2008). "Human coronavirus NL63 and 229E seroconversion in children." J Clin Microbiol **46**(7): 2368-2373.

Dijkman, R., M. F. Jebbink, et al. (2006). "Human coronavirus 229E encodes a single ORF4 protein between the spike and the envelope genes." Virol J **3**: 106.

Dominguez, S. R., C. C. Robinson, et al. (2009). "Detection of four human coronaviruses in respiratory infections in children: a one-year study in Colorado." J Med Virol **81**(9): 1597-1604.

Dosch, S. F., S. D. Mahajan, et al. (2009). "SARS coronavirus spike protein-induced innate immune response occurs via activation of the NF-kappaB pathway in human monocyte macrophages in vitro." Virus Res **142**(1-2): 19-27.

Drosten, C., S. Gunther, et al. (2003). "Identification of a novel coronavirus in patients with severe acute respiratory syndrome." N Engl J Med **348**(20): 1967-1976.

Drosten, C., W. Preiser, et al. (2003). "Severe acute respiratory syndrome: identification of the etiological agent." Trends Mol Med **9**(8): 325-327.

Dunker, A. K., J. D. Lawson, et al. (2001). "Intrinsically disordered protein." J Mol Graph Model **19**(1): 26-59.

Ebihara, T., R. Endo, et al. (2005). "Detection of human coronavirus NL63 in young children with bronchiolitis." J Med Virol **75**(3): 463-465.

Ebihara, T., R. Endo, et al. (2005). "Lack of association between New Haven coronavirus and Kawasaki disease." J Infect Dis **192**(2): 351-352; author reply 353.

Edwards, J. A., F. Denis, et al. (2000). "Activation of glial cells by human coronavirus OC43 infection." J Neuroimmunol **108**(1-2): 73-81.

Enjuanes, L., F. Almazan, et al. (2006). "Biochemical aspects of coronavirus replication and virus-host interaction." Annu Rev Microbiol **60**: 211-230.

Escors, D., J. Ortego, et al. (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.

Esper, F., E. D. Shapiro, et al. (2005). "Association between a novel human coronavirus and Kawasaki disease." J Infect Dis **191**(4): 499-502.

Esposito, S., S. Bosis, et al. (2006). "Impact of human coronavirus infections in otherwise healthy children who attended an emergency department." Journal of Medical Virology **78**(12): 1609-1615.

Fan, H., A. Ooi, et al. (2005). "The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal structure of its N-terminal domain and multimerization properties." Structure **13**(12): 1859-1868.

Fielding, B. C. (2011). "Human coronavirus NL63: a clinically important virus?" Future microbiology **6**(2): 153-159.

Forster, J., G. Ihorst, 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**(12): 709-716.

Fouchier, R. A., N. G. Hartwig, et al. (2004). "A previously undescribed coronavirus associated with respiratory disease in humans." Proceedings of the National Academy of Sciences of the United States of America **101**(16): 6212-6216.

Frieman, M., M. Heise, et al. (2008). "SARS coronavirus and innate immunity." Virus Res **133**(1): 101-112.

Gibbs, A. J., M. J. Gibbs, et al. (2004). "The phylogeny of SARS coronavirus." Arch Virol **149**(3): 621-624.

Giedroc, D. P., C. A. Theimer, et al. (2000). "Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting." J Mol Biol **298**(2): 167-185.

Golda, A., N. Malek, et al. (2011). "Infection with human coronavirus NL63 enhances streptococcal adherence to epithelial cells." The Journal of general virology **92**(Pt 6): 1358-1368.

Guan, Y., B. J. Zheng, et al. (2003). "Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China." Science **302**(5643): 276-278.

Haijema, B. J., H. Volders, et al. (2004). "Live, attenuated coronavirus vaccines through the directed deletion of group-specific genes provide protection against feline infectious peritonitis." J Virol **78**(8): 3863-3871.

Han, T. H., J. Y. Chung, et al. (2007). "Human Coronavirus-NL63 infections in Korean children, 2004-2006." Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology **38**(1): 27-31.

He, R., F. Dobie, et al. (2004). "Analysis of multimerization of the SARS coronavirus nucleocapsid protein." Biochem Biophys Res Commun **316**(2): 476-483.

He, R., A. Leeson, et al. (2003). "Activation of AP-1 signal transduction pathway by SARS coronavirus nucleocapsid protein." Biochem Biophys Res Commun **311**(4): 870-876.

He, R., A. Leeson, et al. (2004). "Characterization of protein-protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus." Virus Res **105**(2): 121-125.

Hendley, J. O., H. B. Fishburne, et al. (1972). "Coronavirus infections in working adults. Eight-year study with 229 E and OC 43." Am Rev Respir Dis **105**(5): 805-811.

Herrewegh, A. A., H. Vennema, et al. (1995). "The molecular genetics of feline coronaviruses: comparative sequence analysis of the ORF7a/7b transcription unit of different biotypes." Virology **212**(2): 622-631.

Hiscox, J. A., T. Wurm, et al. (2001). "The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus." J Virol **75**(1): 506-512.

Hofmann, H. and S. Pohlmann (2004). "Cellular entry of the SARS coronavirus." Trends Microbiol **12**(10): 466-472.

Hofmann, H., K. Pyrc, et al. (2005). "Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry." Proc Natl Acad Sci U S A **102**(22): 7988-7993.

Hofmann, H., K. Pyrc, et al. (2005). "Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry." Proceedings of the National Academy of Sciences of the United States of America **102**(22): 7988-7993.

Hofmann, H., G. Simmons, et al. (2006). "Highly conserved regions within the spike proteins of human coronaviruses 229E and NL63 determine recognition of their respective cellular receptors." Journal of Virology **80**(17): 8639-8652.

Huang, Y., Z. Y. Yang, et al. (2004). "Generation of synthetic severe acute respiratory syndrome coronavirus pseudoparticles: implications for assembly and vaccine production." J Virol **78**(22): 12557-12565.

Hurst, K. R., C. A. Koetzner, et al. (2009). "Identification of in vivo-interacting domains of the murine coronavirus nucleocapsid protein." J Virol **83**(14): 7221-7234.

Hurst, K. R., L. Kuo, et al. (2005). "A major determinant for membrane protein interaction localizes to the carboxy-terminal domain of the mouse coronavirus nucleocapsid protein." J Virol **79**(21): 13285-13297.

Ito, N., E. C. Mossel, et al. (2005). "Severe acute respiratory syndrome coronavirus 3a protein is a viral structural protein." J Virol **79**(5): 3182-3186.

Jia, W., S. P. Mondal, et al. (2002). "Genetic and antigenic diversity in avian infectious bronchitis virus isolates of the 1940s." Avian Dis **46**(2): 437-441.

Kaplan, N. M., W. Dove, et al. (2008). "Molecular epidemiology and disease severity of respiratory syncytial virus in relation to other potential pathogens in children

hospitalized with acute respiratory infection in Jordan." <u>Journal of Medical Virology</u> **80**(1): 168-174.

Kim, L., J. Hayes, et al. (2000). "Molecular characterization and pathogenesis of transmissible gastroenteritis coronavirus (TGEV) and porcine respiratory coronavirus (PRCV) field isolates co-circulating in a swine herd." <u>Arch Virol</u> **145**(6): 1133-1147.

Kleines, M., S. Scheithauer, et al. (2007). "High prevalence of human bocavirus detected in young children with severe acute lower respiratory tract disease by use of a standard PCR protocol and a novel real-time PCR protocol." <u>J Clin Microbiol</u> **45**(3): 1032-1034.

Klig, J. E. (2004). "Current challenges in lower respiratory infections in children." <u>Curr Opin Pediatr</u> **16**(1): 107-112.

Konig, B., W. Konig, et al. (2004). "Prospective study of human metapneumovirus infection in children less than 3 years of age." <u>J Clin Microbiol</u> **42**(10): 4632-4635.

Kopecky-Bromberg, S. A., L. Martinez-Sobrido, et al. (2007). "Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists." <u>J Virol</u> **81**(2): 548-557.

Krokhin, O., Y. Li, et al. (2003). "Mass spectrometric characterization of proteins from the SARS virus: a preliminary report." <u>Mol Cell Proteomics</u> **2**(5): 346-356.

Ksiazek, T. G., D. Erdman, et al. (2003). "A novel coronavirus associated with severe acute respiratory syndrome." <u>N Engl J Med</u> **348**(20): 1953-1966.

Kubo, H., Y. K. Yamada, et al. (1994). "Localization of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein." <u>J Virol</u> **68**(9): 5403-5410.

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**(10): 4987-4999.

Kuypers, J., E. T. Martin, et al. (2007). "Clinical disease in children associated with newly described coronavirus subtypes." Pediatrics **119**(1): e70-76.

Lambert, S. B., K. M. Allen, et al. (2007). "Community epidemiology of human metapneumovirus, human coronavirus NL63, and other respiratory viruses in healthy preschool-aged children using parent-collected specimens." Pediatrics **120**(4): e929-937.

Lau, S. K., K. S. Li, et al. (2010). "Ecoepidemiology and complete genome comparison of different strains of severe acute respiratory syndrome-related Rhinolophus bat coronavirus in China reveal bats as a reservoir for acute, self-limiting infection that allows recombination events." J Virol **84**(6): 2808-2819.

Lau, S. K., P. C. Woo, et al. (2006). "Coronavirus HKU1 and other coronavirus infections in Hong Kong." Journal of Clinical Microbiology **44**(6): 2063-2071.

Laude, H., J. Gelfi, et al. (1992). "Single amino acid changes in the viral glycoprotein M affect induction of alpha interferon by the coronavirus transmissible gastroenteritis virus." J Virol **66**(2): 743-749.

Leung, T. F., C. Y. Li, et al. (2009). "Epidemiology and clinical presentations of human coronavirus NL63 infections in hong kong children." Journal of Clinical Microbiology **47**(11): 3486-3492.

Li, F. Q., H. Xiao, et al. (2005). "Sumoylation of the nucleocapsid protein of severe acute respiratory syndrome coronavirus." FEBS Lett **579**(11): 2387-2396.

Li, W., M. J. Moore, et al. (2003). "Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus." Nature **426**(6965): 450-454.

Lim, K. P. and D. X. Liu (2001). "The missing link in coronavirus assembly. Retention of the avian coronavirus infectious bronchitis virus envelope protein in the pre-Golgi compartments and physical interaction between the envelope and membrane proteins." J Biol Chem **276**(20): 17515-17523.

Lin, H. X., Y. Feng, et al. (2008). "Identification of residues in the receptor-binding domain (RBD) of the spike protein of human coronavirus NL63 that are critical for the RBD-ACE2 receptor interaction." The Journal of general virology **89**(Pt 4): 1015-1024.

Liu, Z. H. and X. Sun (2008). "Coronavirus phylogeny based on base-base correlation." International journal of bioinformatics research and applications **4**(2): 211-220.

Luo, H., J. Chen, et al. (2006). "Carboxyl terminus of severe acute respiratory syndrome coronavirus nucleocapsid protein: self-association analysis and nucleic acid binding characterization." Biochemistry **45**(39): 11827-11835.

Luo, H., Q. Chen, et al. (2005). "The nucleocapsid protein of SARS coronavirus has a high binding affinity to the human cellular heterogeneous nuclear ribonucleoprotein A1." FEBS Lett **579**(12): 2623-2628.

McIntosh, K. (2005). "Coronaviruses in the limelight." J Infect Dis **191**(4): 489-491.

Moes, E., L. Vijgen, et al. (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.

Namy, O., S. J. Moran, et al. (2006). "A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting." Nature **441**(7090): 244-247.

Narayanan, K., C. J. Chen, et al. (2003). "Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal." J Virol **77**(5): 2922-2927.

Narayanan, K. and S. Makino (2001). "Cooperation of an RNA packaging signal and a viral envelope protein in coronavirus RNA packaging." J Virol **75**(19): 9059-9067.

Nelson, G. W., S. A. Stohlman, et al. (2000). "High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA." J Gen Virol **81**(Pt 1): 181-188.

Oostra, M., C. A. de Haan, et al. (2006). "Glycosylation of the severe acute respiratory syndrome coronavirus triple-spanning membrane proteins 3a and M." J Virol **80**(5): 2326-2336.

Opstelten, D. J., M. J. Raamsman, et al. (1995). "Coexpression and association of the spike protein and the membrane protein of mouse hepatitis virus." Adv Exp Med Biol **380**: 291-297.

Papanicolaou, D. A., R. L. Wilder, et al. (1998). "The pathophysiologic roles of interleukin-6 in human disease." Ann Intern Med **128**(2): 127-137.

Peiris, J. S., S. T. Lai, et al. (2003). "Coronavirus as a possible cause of severe acute respiratory syndrome." Lancet **361**(9366): 1319-1325.

Perlman, S. and D. Ries (1987). "The astrocyte is a target cell in mice persistently infected with mouse hepatitis virus, strain JHM." Microb Pathog **3**(4): 309-314.

Pulford, D. J. and P. Britton (1991). "Expression and cellular localisation of porcine transmissible gastroenteritis virus N and M proteins by recombinant vaccinia viruses." Virus Res **18**(2-3): 203-217.

Pyrc, K., B. Berkhout, et al. (2007). "The novel human coronaviruses NL63 and HKU1." J Virol **81**(7): 3051-3057.

Pyrc, K., B. J. Bosch, et al. (2006). "Inhibition of human coronavirus NL63 infection at early stages of the replication cycle." Antimicrob Agents Chemother **50**(6): 2000-2008.

Pyrc, K., M. F. Jebbink, et al. (2004). "Genome structure and transcriptional regulation of human coronavirus NL63." Virol J **1**: 7.

Reed, M. L., B. K. Dove, et al. (2006). "Delineation and modelling of a nucleolar retention signal in the coronavirus nucleocapsid protein." Traffic **7**(7): 833-848.

Reed, M. L., G. Howell, et al. (2007). "Characterization of the nuclear export signal in the coronavirus infectious bronchitis virus nucleocapsid protein." J Virol **81**(8): 4298-4304.

Rowland, R. R., V. Chauhan, et al. (2005). "Intracellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein: absence of nucleolar accumulation during infection and after expression as a recombinant protein in vero cells." J Virol **79**(17): 11507-11512.

Rowley, A. H., S. T. Shulman, et al. (2000). "IgA plasma cell infiltration of proximal respiratory tract, pancreas, kidney, and coronary artery in acute Kawasaki disease." J Infect Dis **182**(4): 1183-1191.

Sawicki, S. G., D. L. Sawicki, et al. (2007). "A contemporary view of coronavirus transcription." J Virol **81**(1): 20-29.

Shaulian, E. and M. Karin (2002). "AP-1 as a regulator of cell life and death." Nat Cell Biol **4**(5): E131-136.

Shimizu, C., H. Shike, et al. (2005). "Human coronavirus NL63 is not detected in the respiratory tracts of children with acute Kawasaki disease." J Infect Dis **192**(10): 1767-1771.

Siu, K. L., K. H. Kok, et al. (2009). "Severe acute respiratory syndrome coronavirus M protein inhibits type I interferon production by impeding the formation of TRAF3.TANK.TBK1/IKKepsilon complex." J Biol Chem **284**(24): 16202-16209.

Smith, S. A. and G. J. Kotwal (2002). "Immune response to poxvirus infections in various animals." Crit Rev Microbiol **28**(3): 149-185.

Smuts, H. (2008). "Human coronavirus NL63 infections in infants hospitalised with acute respiratory tract infections in South Africa." Influenza Other Respi Viruses **2**(4): 135-138.

Song, D. S., J. S. Yang, et al. (2003). "Differentiation of a Vero cell adapted porcine epidemic diarrhea virus from Korean field strains by restriction fragment length polymorphism analysis of ORF 3." Vaccine **21**(17-18): 1833-1842.

Sturman, L. S. and K. V. Holmes (1977). "Characterization of coronavirus II. Glycoproteins of the viral envelope: tryptic peptide analysis." Virology **77**(2): 650-660.

Sung, J. Y., H. J. Lee, et al. (2010). "Role of human coronavirus NL63 in hospitalized children with croup." Pediatr Infect Dis J **29**(9): 822-826.

Surjit, M., R. Kumar, et al. (2005). "The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation." J Virol **79**(17): 11476-11486.

Surjit, M. and S. K. Lal (2008). "The SARS-CoV nucleocapsid protein: a protein with multifarious activities." Infect Genet Evol **8**(4): 397-405.

Surjit, M., B. Liu, et al. (2004). "The nucleocapsid protein of the SARS coronavirus is capable of self-association through a C-terminal 209 amino acid interaction domain." Biochem Biophys Res Commun **317**(4): 1030-1036.

Suzuki, A., M. Okamoto, et al. (2005). "Detection of human coronavirus-NL63 in children in Japan." Pediatr Infect Dis J **24**(7): 645-646.

Tan, T. H., T. Barkham, et al. (2005). "Genetic lesions within the 3a gene of SARS-CoV." Virol J **2**: 51.

Tan, Y. W., S. Fang, et al. (2006). "Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells." Nucleic Acids Res **34**(17): 4816-4825.

Tang, X. C., J. X. Zhang, et al. (2006). "Prevalence and genetic diversity of coronaviruses in bats from China." J Virol **80**(15): 7481-7490.

Timani, K. A., Q. Liao, et al. (2005). "Nuclear/nucleolar localization properties of C-terminal nucleocapsid protein of SARS coronavirus." Virus Res **114**(1-2): 23-34.

Tresnan, D. B., R. Levis, et al. (1996). "Feline aminopeptidase N serves as a receptor for feline, canine, porcine, and human coronaviruses in serogroup I." J Virol **70**(12): 8669-8674.

Tyrrell, D. A. and M. L. Bynoe (1965). "Cultivation of a Novel Type of Common-Cold Virus in Organ Cultures." Br Med J **1**(5448): 1467-1470.

Tyrrell, D. A. and M. L. Bynoe (1966). "Cultivation of viruses from a high proportion of patients with colds." Lancet **1**(7428): 76-77.

Vabret, A., J. Dina, et al. (2008). "Human (non-severe acute respiratory syndrome) coronavirus infections in hospitalised children in France." J Paediatr Child Health **44**(4): 176-181.

van der Hoek, L., K. Pyrc, et al. (2006). "Human coronavirus NL63, a new respiratory virus." FEMS Microbiol Rev **30**(5): 760-773.

van der Hoek, L., K. Pyrc, et al. (2004). "Identification of a new human coronavirus." Nat Med **10**(4): 368-373.

van der Hoek, L., K. Sure, et al. (2005). "Croup is associated with the novel coronavirus NL63." PLoS Med **2**(8): e240.

Vennema, H., G. J. Godeke, et al. (1996). "Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes." EMBO J **15**(8): 2020-2028.

Verheije, M. H., M. C. Hagemeijer, et al. (2010). "The coronavirus nucleocapsid protein is dynamically associated with the replication-transcription complexes." J Virol **84**(21): 11575-11579.

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

Weiss, S. R. and S. Navas-Martin (2005). "Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus." Microbiol Mol Biol Rev **69**(4): 635-664.

Wesseling, J. G., G. J. Godeke, et al. (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.

Woo, P. C., Y. Huang, et al. (2010). "Coronavirus genomics and bioinformatics analysis." Viruses **2**(8): 1804-1820.

Woo, P. C., S. K. Lau, et al. (2009). "Coronavirus diversity, phylogeny and interspecies jumping." Exp Biol Med (Maywood) **234**(10): 1117-1127.

Woods, R. D. (2001). "Efficacy of a transmissible gastroenteritis coronavirus with an altered ORF-3 gene." Can J Vet Res **65**(1): 28-32.

Wurm, T., H. Chen, et al. (2001). "Localization to the nucleolus is a common feature of coronavirus nucleoproteins, and the protein may disrupt host cell division." J Virol **75**(19): 9345-9356.

Yanagawa, H., Y. Nakamura, et al. (1986). "Nationwide epidemic of Kawasaki disease in Japan during winter of 1985-86." Lancet **2**(8516): 1138-1139.

Yang, Y., Z. Xiong, et al. (2005). "Bcl-xL inhibits T-cell apoptosis induced by expression of SARS coronavirus E protein in the absence of growth factors." Biochem J **392**(Pt 1): 135-143.

Yeager, C. L., R. A. Ashmun, et al. (1992). "Human aminopeptidase N is a receptor for human coronavirus 229E." Nature **357**(6377): 420-422.

You, J., B. K. Dove, et al. (2005). "Subcellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein." J Gen Virol **86**(Pt 12): 3303-3310.

Yu, I. M., C. L. Gustafson, et al. (2005). "Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain." J Biol Chem **280**(24): 23280-23286.

Yu, X., W. Bi, et al. (1994). "Mouse hepatitis virus gene 5b protein is a new virion envelope protein." Virology **202**(2): 1018-1023.

Yuan, X., J. Li, et al. (2005). "Subcellular localization and membrane association of SARS-CoV 3a protein." Virus Res **109**(2): 191-202.

Zakhartchouk, A. N., S. Viswanathan, et al. (2005). "Severe acute respiratory syndrome coronavirus nucleocapsid protein expressed by an adenovirus vector is phosphorylated and immunogenic in mice." J Gen Virol **86**(Pt 1): 211-215.

Zhang, X., K. Wu, et al. (2007). "Nucleocapsid protein of SARS-CoV activates interleukin-6 expression through cellular transcription factor NF-kappaB." Virology **365**(2): 324-335.

Chapter II

Cloning and expression of HCoV-NL63 structural

proteins in a bacterial system
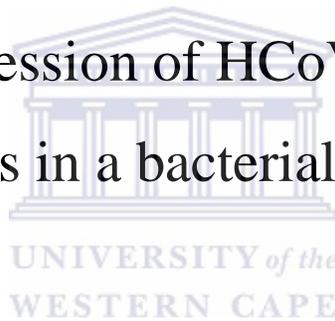
2.1. ABSTRACT

The M, E, ORF3 and N structural proteins of HCoV-NL63 level of toxicity remains a mystery. However based on the cytotoxicity exhibited by other coronaviruses, it is necessary for the molecular and proteomic characterisation of these proteins to be determined. In this study we detail the cloning and expression of the structural proteins of HCoV-NL63.

Molecular characterisation was initiated by the reverse transcription of NL63 RNA to generate cDNA. Upon amplification, amplified genes were ligated into a pGEM shuttle vector before transformation into *E. coli* (JM109) cells. After confirmation through sequencing, these pGEM - gene constructs where then extracted from the cell via alkaline lysis and restriction enzyme digested. The restriction enzyme digest used two rare enzymes; *Sgf1* and *Pme1* to produce "sticky ends", which then allowed the genes to be further ligated into a bacterial expression vector-pFLEXI. KRX *E. coli* served as the protein expression model for these gene-vector constructs. Thereafter, protein expression was analysed via Western blot against the GST fusion proteins. Optimisation of expression using various parameters such as temperature control, auto-induction, and solubilisation resulted in inadequate expression of the HCoV-NL63 M, E, and ORF3 proteins, with only N protein being expressed successfully. These results indicate that the low level of toxic protein expression in the lag and growth phase may be due to an unstable promoter sequence seen on the pFLEXI vector system.

2.2. INTRODUCTION

Since the advent of recombinant DNA technology, there has been a great expansion in the study of the many different proteins found in living organisms. As a result, the expression and purification of proteins has become a fundamental part of many aspects of research. However, the success of the expression of recombinant proteins often remains frustratingly elusive because of the complex structure of proteins in a host system.

There are numerous systems available for protein production, but the use of the Gram-negative *Escherichia coli* bacterium is preferred for its ability to grow rapidly, the low cost of culturing and the availability of a large number of cloning plasmids (Baneyx 1999). Bacterial expression has also been shown to be beneficial in the expression of a heteromultimeric protein like the α and β subunits of human haemoglobin to form a fully functional tetrameric structure (Hoffman, Looker et al. 1990) Although a relatively easy-to-use system, the over-expression of the heterologous proteins in *E. coli* often yields inadequate soluble proteins (Baneyx 1999; Edwards, Arrowsmith et al. 2000; Hammarstrom, Hellgren et al. 2002). To overcome this limitation, the optimisation of expression conditions such as temperature, growth media, induction parameters, promoter and *E. coli* expression strains is necessary (Braun, Hu et al. 2002) Once proteins have been expressed, the next step is to purify those proteins. This is achieved via the use of an affinity tag. Common affinity tags include GST, HIS and MBP, which enhance protein solubility. Although the addition of a fusion tag onto a protein can improve expression, the tag has also been shown to interfere with protein structure and function (Kapust and Waugh 1999; Wu and Filutowicz 1999; Bucher, Evdokimov et al. 2002). It is thus advisable to remove the tag after purification.

This chapter reports the cloning and expression of the HCoV-NL63 structural proteins: E, M, ORF3, and N proteins. Deletion mutants of the hydrophilic regions of the M and ORF3 have also been generated to increase the solubility of the fusion proteins, which would allow for subsequent purification. The N protein and truncated clones were previously cloned and sequenced and included in expression studies to allow for further immunological studies in the proceeding chapter.

The expression of all the above-mentioned clones occurred downstream from the glutathione-S-transfer (GST) fusion protein, which was used to allow for detection and purification. The GST tag can be used for protein expression as it allows for by the use of a T7 RNA polymerase promoter found in the pFLEXI vector system. To carry out expression, the use of a suitable bacterial strain that expresses the T7 RNA polymerase is required. The bacterial strain used here was K12 *E. coli* derivative KRX. KRX has high transformation efficiency because of the absence of proteases and nucleases. Good protein expression is possible as the vector contains a chromosomal copy of the T7 RNA polymerase gene that has a rhamnose promoter, which helps control the expression of recombinant proteins. This promoter is activated via the addition of rhamnose to the medium, thus providing precise control of T7 RNA polymerase levels that, in turn, control the recombinant protein expression. Rhamnose induces the activation of RhaR, which in turn will induce the production of active RhaS, which will bind to the rhamnose activating transcription from RhaBAD. The isomerase (RhaA), kinase (RhaB) and aldolase (RhaD) are deleted and replaced with the T7 RNA polymerase in KRX. Rhamnose is not metabolised by the cell and is not consumed during growth, but is merely required for efficient expression (Hartnett, Gracyalny *et al.* 2006).

In this study, the expression of HCoV-NL63 E, M, ORF3, MΔN and ORFΔN could not be detected in either the soluble or insoluble fractions, even though the GST tag is known to increase protein solubility and expression. Several factors could possibly have contributed to this occurrence, viz cloning errors, vector toxicity and lethality of the Barnase gene, and protein degradation by bacterial proteases. Plasmids can be lost at low frequency during cell division. This loss can, however, be offset if the ligated genes it carried are toxic to the cell (Baneyx 1999). Sequencing prior to expression and the insertion of the AmpR gene for ampicillin resistance prevented the growth of plasmid-free cells, and in so doing ruling out the possibility of cloning errors.

2.3. MATERIALS AND METHODS

2.3.1 Reagents

HCoV-NL63 RNA was a kind gift from Prof L. van der Hoek, of The Netherlands. A clinical sample was obtained and the RNA was extracted from the fifth passage of an Amsterdam 1 viral strain. The kits used for SDS gel and Western blot preparation were obtained from Sigma Aldrich. Glycerol stocks of GST-tagged nucleocapsid constructs, cloned into pFlexi vector, were kindly provided by Mr M. Berry (Department of Medical Biosciences, University of the Western Cape). Recombinant constructs included SARS-CoV N, HCoV-NL63-N, and N and C terminus deletion mutants of HCoV-NL63 N. For the purpose of this and the following chapter, the constructs were named SARS-N, N1, N2 and N3, respectively (Table 2.1).

Table 2.1: Constructs used for expression in this chapter

Constructs Name	Description of construct	a.a. position
SARS-N	Full-length SARS-CoV N gene	1-423
N1	HCoV-NL63 full length	1-378
N2	N-terminal half of the HCoV-NL63 N	1-189
N3	C-terminal half of the HCoV-NL63 N	191-378

2.3.2 Bacterial strains and plasmids

Bacterial strains used in this study were all obtained from Promega. These include the KRX, BL21 and JM109 *E. coli*-competent cells. KRX is a K12 derivative that can be used with a pFLEXI expression vector. On the other hand, a λ (DE3) pLysS strain was used with a pGEX-4T-2 expression vector to express the E, M and N HCoV-NL63 structural proteins for further studies (Appendix III). Both the pFLEXI and pGEX-4T-2 expression vectors are under the control of a T7 promoter. The pGEM-T-easy vector system was included to allow for the confirmation of the gene of interest because of its convenience for cloning PCR products. Initial transformation with the pGEM vector was carried out using the JM109 competent cells. KRX competent cells were used for optimisation and the final expression of the recombinant proteins. The pFLEXI vector was used due to its unique ability to express fusion and native proteins to allow for further studies into protein function and structure.

2.3.3 Generation of recombinant plasmid vectors

2.3.3.1 Reverse Transcription

1st strand cDNA was generated from HCoV-NL63 RNA, using the Reverse Transcription System (Promega) according to the manufacturer's instruction. cDNA was transcribed from HCoV-NL63 RNA using 25 mM MgCl$_2$, Reverse Transcription 10x buffer, 10 mM dNTP mix, RNAsin (Promega), 25 U/µl Reverse transcriptase, Oligo dT Primer (0.5 µg/µl), test RNA and nuclease-free water to make up a final volume of 20.0 µl in a sterile microcentrifuge. The control was a water-blank with all the reagents mentioned above except RNA, which was replaced by nuclease-free water, in order to ascertain the purity of the reagents. The mixture was made while carefully working on ice. The transcription conditions were set at 42ºC for 60 minutes, 95ºC for 5 minutes and 0ºC for 5 minutes. Thereafter the product was stored at -4ºC until required.

2.3.3.2 Primer design
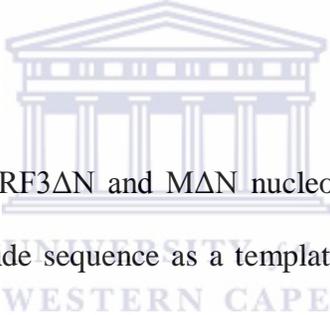
The HCoV-NL63 E, M, ORF3, ORF3ΔN and MΔN nucleotide sequences was retrieved from NCBI. Using the obtained nucleotide sequence as a template, the forward and reverse primers were designed using Promega primer design tools. For ligation into the Flexi™ protein expression vector, a *SgfI* (5´-GCGATCGC-3´) restriction-site was incorporated into the forward primer and a *PmeI* (5´-GTTTAAAC-3´) restriction site was incorporated into the reverse primer. An additional C residue was appended upstream of the ATG in the forward sequence to prevent a frameshift in the reading frame. These primers were used as 100 ρ.mole/µl stock solutions and their nucleotide sequences are shown in table 2.2

Table 2.2: Sequences specific forward and reverse primers for each gene used in PCR amplification

Primer Name	Sequence
ORF3-Flexi. Fw	*GC***GAATTC**A*ATG*CCTTTTGGTGGCCTATTC
ORF3-Flexi. Rw	*GC***TCTAGA**TCAA*TTA*ATCGAAGGAACATGC
M-Flexi. Fw	TGCG**GCGATCGCC***ATG*ACTTTGGCGCCGTGTTAAA
M-Flexi Rw	TGTG**GTTTAAAC***TTA*GATTAAATGAAGCAACTTCTCA
E-Flexi. Fw	GCGC**GCGATCGCC***ATG*TTCCTTCGATTAATTG
E-Flexi. Rw	GCGC**GTTTAAAC***TTA*GACATTTAGTACTTCAGC
ORF3ΔN. Fw	AGCA**GCGATCGCC***ATG*TACTCATTCGTTTCGGA
ORF3ΔN. Rw	TTGT**GTTAAAC***TTA*GACCAGAAGATCAGGAA
MΔN. Fw	AGCA**GCGATCGCC***ATG*TACTCATTCGTTTCGGA
MΔN. Rw	TTGT**GTTAAAC***TTA*GACCAGAAGATCAGGAA

Bold: - Sgf1 cut site (Fw); Pme1 cut site (Rw), underlined and italics:-stop/start codon

2.3.3.3 Gene amplification-PCR

The PCR reaction was set-up in a final volume of 25μl with 0.2mM dNTP mix, 1.5 mM $MgCl_2$, 4μM forward and reverse primers, 5μl 5x green GoTaq Flexi buffer, 0.5 μl 5u/μl GoTaq polymerase and 2μl cDNA template.

Of the 10 primers the lowest Tm was 64.68°C for the NL63 N2 reverse primer and, therefore, an annealing temperature of 60°C was chosen to ensure high specificity of binding of base pairs to the template strand. PCR amplification was therefore run under the following conditions: 95°C

for 3 min followed by 30 cycles of 95°C for 1 min, 60°C for 1 min and 72°C for 2min. The final elongation was run for 15 min at 72°C.

2.3.3.4 Agarose gel electrophoresis

The PCR product was diluted in a ratio of 1:5 with 6 x Blue/Orange loading dye (0.4% (v/v) orange G, 0.03% (v/v) bromophenol blue, 0.03% (v/v) xylene cyanol FF, 15% (v/v) Ficoll® 400, 10 mM Tris-HCl (pH 7.5) and 50 mM EDTA (pH 8.0). The amplification reactions were run on a 1% (w/v) agarose gel, containing 0.001% (v/v) ethidium bromide, by electrophoresis in TBE buffer (89 mM tris (hydroxymethyl) aminomethane, 0.089 mM boric acid, 2 mM EDTA (pH 8)). The gel was run at 90 volts, 2 amperes for a period of approximately 1 hour. The size of the DNA fragment of interest was determined according to the DNA migration pattern in an agarose gel compared to that of a pre-determined DNA molecular weight marker. When the PCR fragment was visualised under the high frequency ultraviolet (UV) light, exposure to UV light was minimised to avoid damage to the DNA.

2.3.3.5 Purification of PCR products

The PCR product to be ligated was gel purified by means of the Wizard® PCR Preps DNA purification systems (Promega) according to the manufacturer's specifications. Briefly, the bands to be cloned were excised from the agarose gel by means of a clean scalpel and razor blade. A weighed gel slice was completely dissolved by incubation at 50°C for 10 min, with occasional vortexing in the presence of a membrane binding solution (4.5 M guanidine isothiocyanate and 0.5 M potassium acetate) at a ratio of 1:1. The gel mixture was then transferred into an SV minicolumn, where it was washed with membrane wash solution (10 mM potassium acetate,

80% (v/v) ethanol and 16.7 μM EDTA). The DNA was finally eluted with nuclease-free water into a sterile micro-centrifuge tube.

2.3.4 Cloning and verification of amplified products

2.3.4.1 Ligation into pGEM and JM109transformation

A ligation reaction was then performed using the Promega pGEM T-easy vector system and the purified PCR product from above. The reaction mixture was set up using 2 x ligation buffer, pGEM vector (50ng/μl), T4 DNA ligase (3U/μl) and 3.5 μl of PCR product. The positive control used here had the control insert DNA (Promega) in place of PCR product, and the negative control had nuclease-free water replacing the DNA. Ligation was done overnight at 4ºC.

Transformation into E. coli (JM109) competent cells was performed using the pGEM T-easy system as per manufacturer's instructions (Promega). 4 μl of ligation product, including the controls, was dispensed into sterile 1.5 ml microcentrifuge tubes and placed on ice. Competent cells were thawed on ice for 5 minutes and 50 μl dispensed into each reaction tube. After gently flicking the tube to mix the contents, the tube was placed on ice for 20 minutes. The cells were then heat-shocked by placing the tubes in a dry heating block at 42°C for 45 seconds and immediately returned to ice for a further 2 minutes. 950 μl sterile LB broth (1% bacto-tryptone, 0.5% bacto -yeast extract and 0.5% NaCl) was preheated to 37°C and then added to each tube and incubated at 37°C for 90 minutes on a shaker at approximately 150rpm.

LB agar plates were prepared with 1% Bacto – Tryptone, 0.5% Bacto – Yeast extract, 0.5% NaCl and 1.5% agar. This agar mixture was autoclaved and upon cooling Ampicillin was added at 1 μl per ml before pouring the plates and allowing them to set on a flat surface. Working aseptically to avoid contamination, and with the aid of a hockey stick, 100 μl of transformed sample was

spread onto duplicate LB agar plates containing 100 µl IPTG (100 mM), 20 µl Xgal (50mg/M) and 10 µl ampicillin (100mg/M). The positive and negative controls from the transformation reactions were used, together with a plate for competent cells only, to verify the effectiveness of the ampicillin. Plates were then incubated overnight at 37°C. Thereafter plates were sealed with parafilm and placed in the fridge for approximately 8 hours to intensify the colour of the blue colonies to distinguish between blue (untransformed) and white (transformed) colonies. Eight white colonies were inoculated into 8 tubes, each containing 10 ml of LB broth and 10 µl of ampicillin, and the culture was grown overnight at 37ºC with agitation.

2.3.4.2 Colony PCR

In addition to the above-mentioned grow-up, a colony PCR was carried out to verify that the colony picked for culturing does in fact contain the inserts (transformed). The colony PCR mixture was made up using 5 µl dNTP mix (10mM); 1.5 µl MgCl$_2$, 2.0µl 10X buffer, 1.0 µl forward and reverse primers (100 µM), 0.5µl Taq Polymerase and 3.5µl nuclease-free water to make up a final volume of 10µl of sample. The cDNA was obtained by briefly inserting a sterile pipette that was used to pick the colony into the PCR mixture. The samples were then placed in a thermocycler under conditions specific for each gene. The reaction was run for 30 cycles with a final extension temperature was set at 72 ºC for 15 minutes, and thereafter a soak temperature of 4ºC. After which the samples were run on a 1% agarose gel to verify which of the colonies contain the correct inserts of gene. Upon verification via electrophoresis, the overnight cultures that contained the inserts were then mini-prepped to extract the DNA for further use.

2.3.4.3 Extraction of plasmid DNA

Plasmid DNA was isolated from the transformed cells using a Wizard® plus SV minipreps DNA purification system. Briefly, a single bacterial colony that was confirmed to contain the amplicon by colony PCR was inoculated into 10 ml of LB medium containing antibiotics, and incubated overnight at 37ºC. The culture was pelleted for 5 minutes and then immediately resuspended in 250 µl of cell re-suspension solution (50 mM Tris-HCl, 10 mM EDTA). This was mixed with another 250 µl of cell lysis solution (1% (w/v) SDS, 0.2 M NaOH) and then mixed by inverting. To this mixture, 10 µl of alkaline protease solution was added, mixed and then incubated for 5 minutes at room temperature. Following incubation, the sample was mixed with 350 µl of neutralizing solution mM potassium acetate, 8.3 mM Tris-HCl, 40 µM EDTA), inverted and then centrifuged at top speed (13 000 rpm) at room temperature for 10 minutes. Once the cleared lysate was obtained, the lysate was decanted into a spin column and then centrifuged for 1 minute at top speed. The flow-through was discarded and the column reinserted into the collection tube. 750 µl of wash solution was added and the sample was centrifuged at top speed for 1 minute. The spin column was then washed with 250 µl wash solution and centrifuged for 2 minutes. The plasmid DNA was eluted by transferring the spin column into a sterile centrifuge tube, followed by addition of 30 µl of nuclease free water to the spin column. The latter was centrifuged at top speed for 1 minute and the DNA stored at -20°C.

2.3.4.4 Sequencing

Since PCR was used to create the original insert, the sequence of the protein coding region in the initial clone was verified by automated DNA sequencing. Cycle sequencing was performed at Inqaba Biotech. The clone corresponding to the initial nucleotide sequence (NCBI) was selected for further experimentation. Following *in silico* analysis, the construct was digested with pFlexi

restriction enzymes (*Sgf1* and *Pme1*) for ligation into a pFlexi vector that has been digested with the same restriction enzymes.

2.3.4.5 Restriction endonuclease digest

4µl of DNA from the mini preps mentioned above was cut with 1µl of Sgf1 and Pme1 restriction enzymes. 4 µl of 5X flexi digest buffer and 11 µl of nuclease-free water were micro-centrifuged, and then placed in a dry heating block for 3 hours at 37°C. Thereafter the temperature was changed to 65°C to inactivate the restriction enzymes. The entire product of each sample was then loaded onto a 1% agarose gel to allow for the purification of the product of interest. The gel was purified using the Wizard® SV gel and PCR cleanup system (Promega), as per protocol.

2.3.4.6 KRX transformation and ligation into pFLEXI vector system

After cutting the DNA with the Sgf1 and Pme1 enzymes, 'sticky ends' are produced which can then be ligated into compatible sites of pFLEXI (pFN2a-GST) bacterial expression vectors using the pFLEXI vector systems as per protocol (Promega). Ligation into the pFLEXI vector was done using 11 µl DNA, 5µl Flexi vectors (GST), 2µl of 10 x ligation buffer and 2µl T4 DNA ligase (3 U/µl) in a final volume of 20 µl. Ligation was done overnight at 4ºC.

Transformation into KRX competent cells $>10^6$ cfu/ µg (Promega) was done per protocol using 50µl aliquots of competent cells. The KRX competent cells were thawed on ice for 5 minutes. 5µl of ligation product was dispensed into the 50µl aliquot of competent cells, and after gently flicking the tubes were returned to ice for 20 minutes. The cells were then heat shocked on a dry heating block at exactly 42°C for 15 seconds and, without shaking, immediately returned to ice for another 2 minutes. 450 µl sterile LB broth (1% bacto-tryptone, 0.5% bacto -yeast extract and 0.5% NaCl) at room temperature was added to each tube and incubated at 37°C for 90 minutes

on a shaker at approximately 150rpm. 100 µl of transformed samples containing the GST tagged

construct were spread-plated onto duplicate LB agar plates containing 10µl of ampicillin

(100mg/M).

Negative controls were set up by plating the viral constructs on kanamycin LB agar plates. Plates

were then incubated overnight at 37ºC, together with untransformed *E.coli* cells, cells

transformed with control DNA and ligation controls to check efficiency of ligation and reagents.

Ten colonies were inoculated into 10 greiner tubes, each containing 10 ml of LB broth and 10 µl

ampicillin for the GST tagged proteins. The cultures were grown overnight at 37ºC with agitation

by placing them on a shaker.

2.3.5 Expression studies

2.3.5.1 Autoinduction

The autoinduction protocol as described by Schagat *et al.* (2008) utilises glucose and rhamnose

to tightly control the T7 RNA polymerase gene via a rhamnose promoter (rhaP$_{BAD}$), which has

been shown to be inhibited by the metabolites of glucose via a cyclic AMP (cAMP)

pathway(Holcroft and Egan 2000). Glucose may therefore be used to inhibit pre-induction

protein expression, where rhamnose will induce protein expression, but only once the glucose is

completely metabolised from the culture medium.

We utilised an early autoinduction protocol where glycerol stocks of HCoV-NL63 M, ORF3, E,

ORF3ΔN and MΔN in the pFLEXI vector where inoculated into a starter culture of 10ml LB

medium and incubated at 37°C for 14 hours with shaking at 180rpm. Starter cultures were

diluted 1:100 into a LB medium supplemented with 0.05% (w/v) glucose and 0.1% (w/v)

rhamnose, and incubated at 37°C with shaking at 180rpm. Samples of 50ml were taken at 8 and 24 hours and harvested by centrifugation at 6000rpm for 15 minutes at 4°C to prevent protein degradation. Dry cell pellet was frozen at -80°C until further use. We were unable to detect any protein expression and therefore chose to modify the expression protocols as follows.

2.3.5.2 Induction at specific OD_{600} reading with fermentation at 37°C

Glycerol stocks were inoculated into 10ml of LB medium and incubated for 14 hours at 37°C with shaking at 180rpm. Starter cultures were diluted 1:100 into 100ml LB medium and incubated for approximately 4 hours, at which time OD_{600} readings were taken. Once the culture reached an OD_{600} of 0.3-0.4, expression was induced by the addition of 0.1% (w/v) rhamnose. Cultures were incubated as previously described for 24 hours when cells were harvested by centrifugation at 6000rpm for 15 minutes. Expression was undetectable, and so we chose to further modify the expression protocol by reducing the fermentation temperature to 25°C.

2.3.5.3 Induction at specific OD_{600} reading with fermentation at 25°C

The previous method was maintained until after induction of expression where the fermentation temperature was reduced to 25°C. Low yields of expressed proteins may indicate toxicity to the cell by the heterologous protein, and decreasing the fermentation temperature will as a result slow the initial expression of the protein, allowing the cell to acclimatise and in so doing increase final yields. We, however, were not able to detect any protein expression and therefore returned all fermentation temperatures to 37°C.

2.3.5.4 Time course expression

Expression protocols described above were maintained where starter cultures were inoculated into 1500ml LB medium. Samples of 200ml were then taken, at 4, 8, 12, 16, 20 and 24 hours post-induction to determine at which point expression is at its highest. Protein catabolism was not detected at these various time points, and thus large-scale expression was adopted.

2.3.5.5 Large-scale expression

Due to the low levels of protein expression and a reduction in cell density, a large-scale culture was deemed necessary to obtain an ideal concentration for protein purification. The expression protocols described previously was maintained where starter cultures were inoculated in 1500ml LB medium with the addition of 1mM IPTG on induction. A small level of protein expression was detected for the M protein (Appendix II); however, upon repetition of this protocol expression was no longer detected.

UNIVERSITY of the
WESTERN CAPE
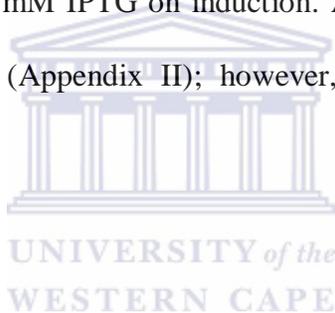
2.3.6. Protein Analysis

2.3.6.1 Protein extraction and Preparation of Cell Lysates

Following centrifugation, cell pellets were resuspended by vortexing in 2ml/g lysis buffer containing 1% (v/v) TritonX, 150mM NaCl, 10mM Tris(hydroxymethyl)-amino-methane (Tris), 5mM EDTA and a protease inhibitor cocktail. Table 2.3 describes the function of all constituents in the lysis buffer. Cells were subsequently lysed by sonication in an ice bath to prevent protein degradation. All subsequent work was done on ice or at 4°C. Following sonication, insoluble cell fraction was separated from soluble fraction by centrifugation at 14000rpm for 15 minutes. The

soluble fraction was removed and stored at -20°C. The insoluble fraction was resuspended in 1ml lysis buffer by sonication as vortexing posed insufficient.

Table 2.3: Function of all components of chosen cell lysis buffer

Reagent	Function
TritonX	Non-ionic detergent improves solubility of GST fusion proteins and prevents aggregation of lysed cells.
NaCl	Provides an osmotic shock to cells
Tris	Interacts with lipopolysaccharides in the outer membrane of the cell and thereby increases permeability.
EDTA	Inhibits divalent cation-dependant proteases

Protein expression was not detectable in either the soluble or insoluble fraction of the cell and therefore several lysozymes were introduced to the lysis buffer in a final concentration of 0.1mg/ml. The insoluble fraction was incubated for 3 hours at room temperature on a rocker in the buffer containing lysozymes. Subsequently samples were centrifuged at 14000rpm for 15 minutes and the supernatant was removed and stored at -20°C. The pellet was further resuspended as previously described to determine whether any expressed proteins remained in the insoluble fraction.

2.3.6.2 Polyacrylamide gel electrophoresis SDS

A 3% stacking and a 12.5% separation SDS-PAGE gel was used to separate proteins in a bio-rad system. Electrophoresis was run at 15mA/gel for 80 minutes in a Tris-glycine SDS running

buffer (25 mM Tris, 192 mM Glycine, 0.1% (w/v) SDS (pH 8.3)). The SDS-PAGE gel was then either stained in Coomassie Brilliant Blue stain (40% (v/v) methanol, 10% (v/v) acetic acid and 0.025% (w/v) Coomassie Brilliant Blue (R250) overnight, to identify all proteins expressed in the cell, or transferred to a Western blot to identify the GST tag with the appended N-protein. Coomassie stain was removed using a coomassie destain solution (50% (v/v) methanol and 10% (v/v) acetic acid) and preserved in cellophane sheeting after treatment with gel drying solution (50% (v/v) methanol, 10% acetic acid and 10% (v/v) glycerol).

2.3.6.3 Western transfer

For Western blots, the proteins, separated by SDS-PAGE, were transferred to a nitrocellulose or PVDF membrane, where neither membrane had any noticeable preferential properties. Prior to transfer, the nitrocellulose membrane was equilibrated in 20% (v/v) methanol and the PVDF in 100% (v/v) methanol. Proteins were transferred in transfer buffer (27mM tris(hydroxymethyl)-amino-methane, 191mM glycine and 20% (v/v) methanol) in a submersion system at 100V for 90 minutes, after which, the membrane was blocked with a 5% (w/v) milk and 0.05% (v/v) Tween- 20 in PBS solution for 30 minutes on a rocker. The membrane was then incubated at 4°C overnight on a roller in 3% (w/v) milk and 0.05% (v/v) tween 20 in PBS solution with the primary antibody, rabbitαGST, in a dilution of 1:5000. The membrane was then washed in a wash solution (0.05% (v/v) tween 20 in PBS) for 1 hour. The secondary antibody, HRP-conjugated Goatαrabbit, was added in a dilution of 1:5000 in fresh solution previously described and incubated at room temperature for 1 hour on a roller, following which the membrane was washed in wash solution with subsequent addition of the peroxidase substrate. The presence of the GST-fusion protein was determined by colorimetric analysis.

2.4. RESULTS AND DISCUSSION

E. coli is genetically easy to manipulate and inexpensive to culture, thus making it the preferred host for the expression of heterologous proteins. One of the disadvantages in using *E.coli.*, however, is that it is incapable of undergoing any post-translational modifications, such as glycosylation, which is imperative for the production of active, folded proteins (Singh and Panda 2005; Rinas, Hoffmann et al. 2007)

2.4.1 PCR

The HCoV-NL63 M, ORF3, E, ORF3ΔN and MΔN genes were amplified from cDNA generated by reverse transcription of the viral RNA. Primers used in amplification were designed to append the *Sgf1* and *Pme1* restriction enzyme cut sites to the amplicon. The *Sgf1* cut site is located upstream of the start codon within the protein coding region. This will allow for the expression of untagged proteins as well as the N-terminal tagged region. The *Pme1* cut site contains stop codon for the coding region of the protein and will add a Valine residue to the carboxyl terminus. As a result of these two enzymes cut sites, the reading frame and orientation of the insert is maintained. The gene sizes, as determined by the accession numbers from NCBI, were as follows: M: 700bp; 0RF3: 678bp; E: 240bp; ORF3ΔN: 340bp and MΔN: 390bp. Following PCR, the products were run on a 1% (w/v) agarose gel and viewed under UV light (Figures 2.1 to 2.5). The M, ORF3, E, ORF3ΔN and MΔN genes were identified as the correct size and gel excised for purification with a PCR gel Clean-Up kit (Promega). The viral genes were subsequently ligated into a pGEM® vector and used to transform JM109 strain competent *E. coli.*

Figure 2.1: PCR amplification of HCoV-NL63 M viral gene using primers described in table 2.1. Lane M: 100BP marker; Lane 1-6: HCoV-NL63 M-gene (±700bp); Lane 7: Blank; Lane 8: Negative control using all primers described in table 2.1.

Figure 2.2: **PCR amplification of HCoV-NL63 MΔN viral gene using primers described in table 2.1.** Lane M: 100BP marker; Lane 1&2: HCoV-NL63 MΔN (±390bp); Lane 3-6: Blank.

Figure 2.3: PCR amplification of HCoV-NL63 ORF3 viral gene using primers described in table 2.1. Lane M: 100BP marker; Lane 1-4: HCoV-NL63 ORF3 (±678bp).
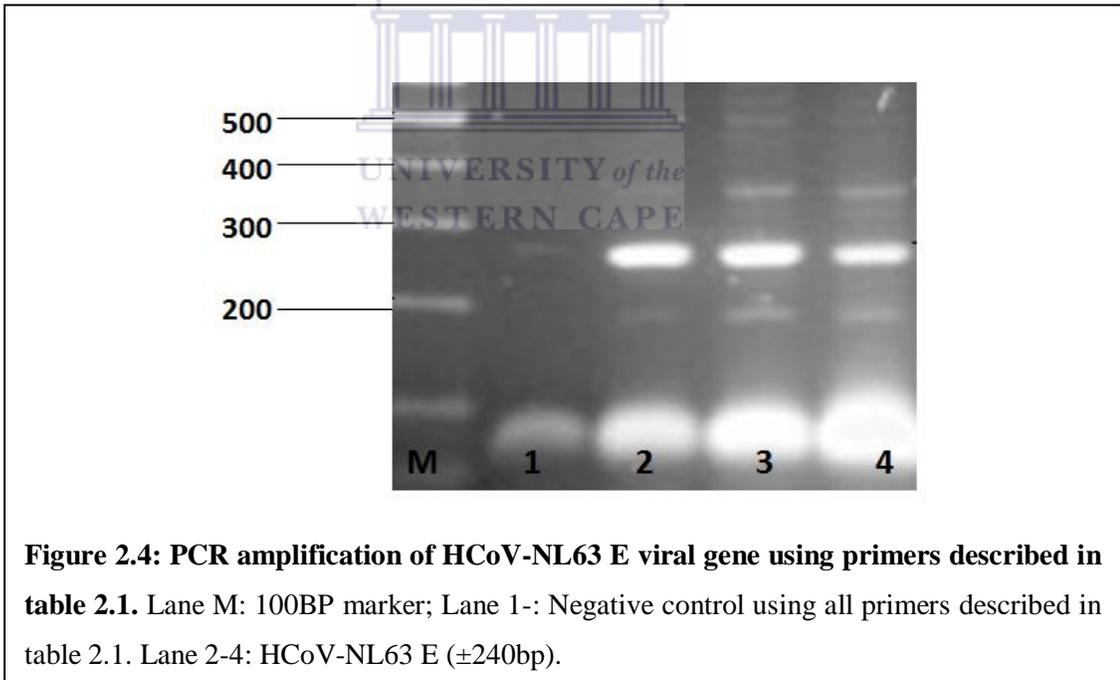
Figure 2.4: PCR amplification of HCoV-NL63 E viral gene using primers described in table 2.1. Lane M: 100BP marker; Lane 1-: Negative control using all primers described in table 2.1. Lane 2-4: HCoV-NL63 E (±240bp).

Figure 2.5: PCR amplification of HCoV-NL63 ORF3ΔN viral gene using primers described in table 2.1. Lane M: 100BP marker; Lane 1-: Blank; Lane 2: Negative control using primers described in table 2.1; Lane 3: HCoV-NL63 ORF3ΔN.
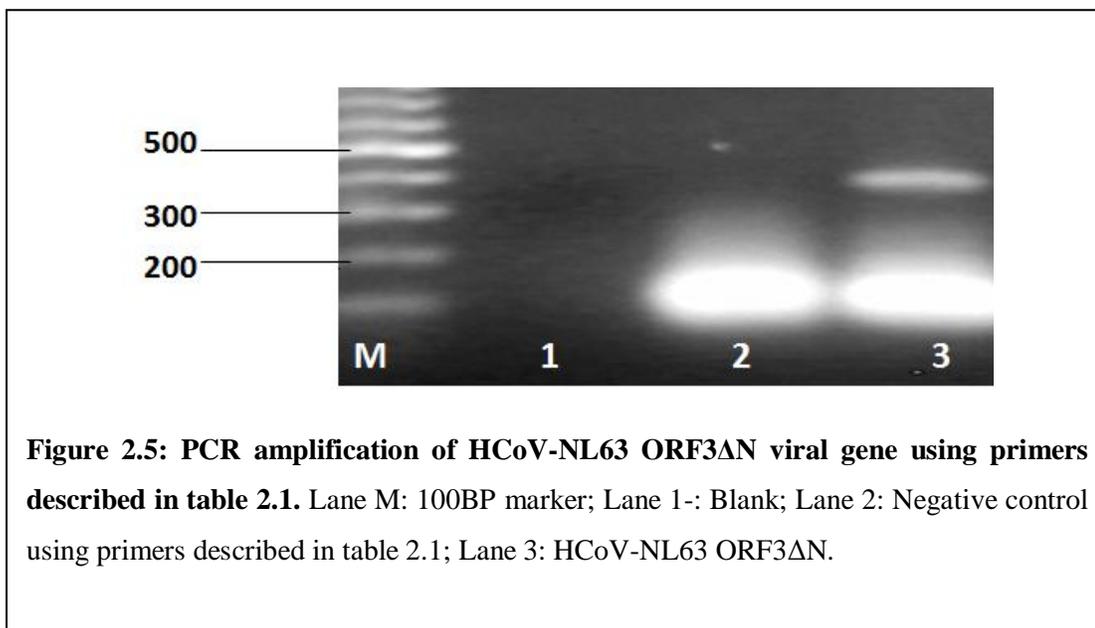
2.4.2. Colony PCR of JM109 Competent *E. coli*

Confirmation of the amplified genes of interest, based purely on the size observed following agarose gel electrophoresis, is not completely accurate. For this reason, amplicons were purified and subsequently ligated into the pGEM® T-easy vector for sequencing. This would also allow for identification of any base pair mutations. Recombinant pGEM® T-easy vectors were then transformed into JM109 competent *E. coli*. Following overnight incubation, colony forming units (CFUs) were selected and screened for insert using colony PCR (Figures 2.6 to 2.10). CFUs confirmed to contain the correct insert size were cultured overnight. From these, plasmids were extracted and confirmed by sequencing; all reported sequences were confirmed by determining the sequence for both strands. Forward and reverse sequences were aligned using the multiple sequence alignment tool ClustalX (Version 2.0.12) and viewed in GeneDoc (Version 2.7.000) (Appendix I).
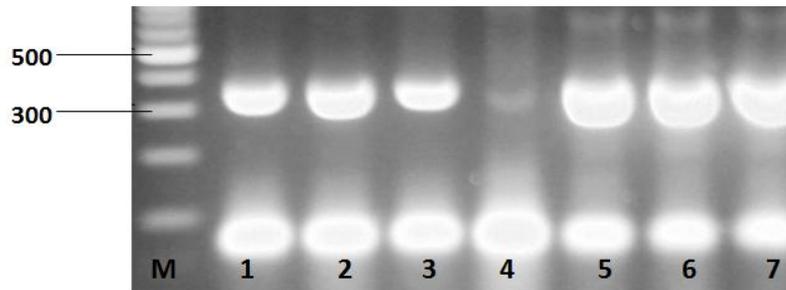
Figure 2.6: Detection of ORF3ΔN viral gene in JM109 competent *E. coli transformed* with *pGEM® vector previously ligated with viral genes. Genes were detected using PCR* with *primers described in Table 2.1.* Lane M: 100BP marker; Lane 1-7: Amplification of the ORF3ΔN gene (±360bp). All positive amplifications indicate successful ligation into the pGEM vector with subsequent transformation of JM109 competent *E. coli* with the ligation product
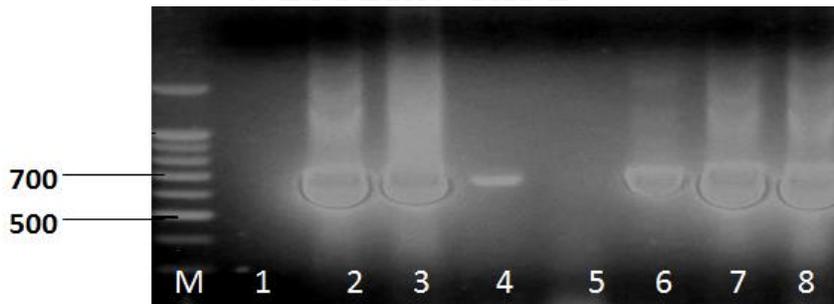

Figure 2.7: Detection of ORF3 viral gene in JM109 competent *E. coli transformed with* *pGEM® vector previously ligated with viral genes. Genes were detected using PCR with* *primers described in Table 2.1.* Lane M: 100BP marker; Lane 1-: Blank; Lane 2-4&6-8: Amplification of ORF3 gene (±678bp); Lane 5: Negative control using the primers described in Table 2.1. All positive amplifications indicate successful ligation into the pGEM vector with subsequent transformation of JM109 competent *E. coli* with the ligation product
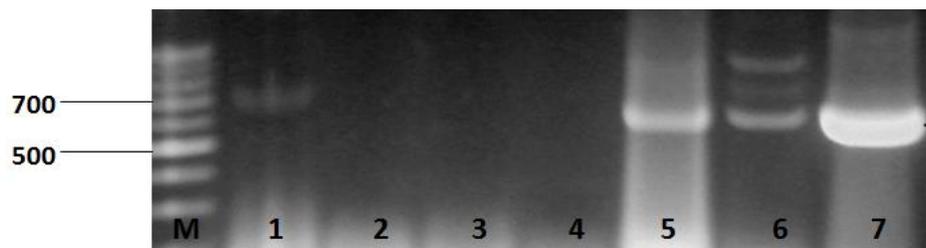
Figure 2.8: Detection of M viral gene in JM109 competent *E. coli transformed with pGEM® vector previously ligated with viral genes. Genes were detected using PCR with primers described in Table 2.1.* Lane M: 100BP marker; Lane 1-3: Amplification of M gene (±700bp). Lane 2-4: Unsuccessful amplification of the N1 gene; Lanes 5-7: Amplification of the M gene. All positive amplifications indicate successful ligation into the pGEM vector with subsequent transformation of JM109 competent *E. coli* with the ligation product
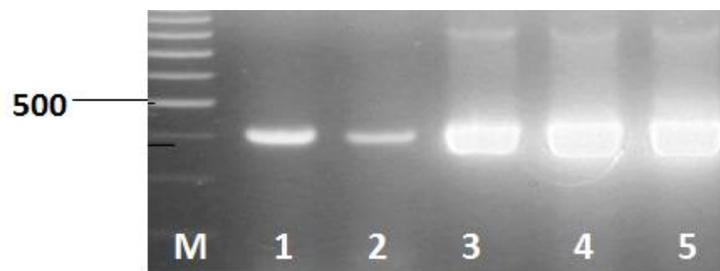

Figure 2.9: Detection of MΔN viral gene in JM109 competent *E. coli transformed with pGEM® vector previously ligated with viral genes. Genes were detected using PCR with primers described in Table 2.1.* Lane M: 100BP marker; Lane 1-5: Amplification of MΔN gene (±390bp). All positive amplifications indicate successful ligation into the pGEM vector with subsequent transformation of JM109 competent *E. coli* with the ligation product
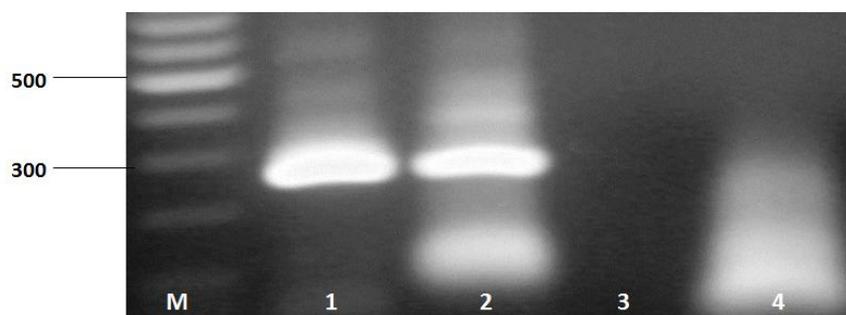
Figure 2.10: Detection of Eviral gene in JM109 competent *E. coli transformed with pGEM® vector previously ligated with viral genes. Genes were detected using PCR with primers described in Table 2.1.* Lane M: 100BP marker; Lane 1&2: Amplification of E gene (±; Lanes 3: Blank; Lane 4: Negative control using all primers described in table 2.1. All positive amplifications indicate successful ligation into the pGEM vector with subsequent transformation of JM109 competent *E. coli* with the ligation product

2.4.3. *SgfI* and *PmeI* Restriction

Sequenced-confirmed viral gene inserts were excised from recombinant pGEM constructs and subsequently cloned into compatible restriction enzyme sites (*SgfI* and *PmeI*) of the pFlexi™ vector. First, to confirm the release of the insert from the recombinant pGEM constructs, digests were viewed on a 1% (w/v) agarose gel under UV light (Figure 2.11 to 2.13) and viral genes were identified by their sizes. Bands were then excised from the gel, purified and ligated into the Flexi™ protein expression vector, which was then transformed into KRX competent cells. Restriction enzyme digests of the recombinant pFLEXI-N constructs from a previous study (refer to Section 2.3.1) were also done at this stage to confirm presence of the inserts (Figure 2.14).
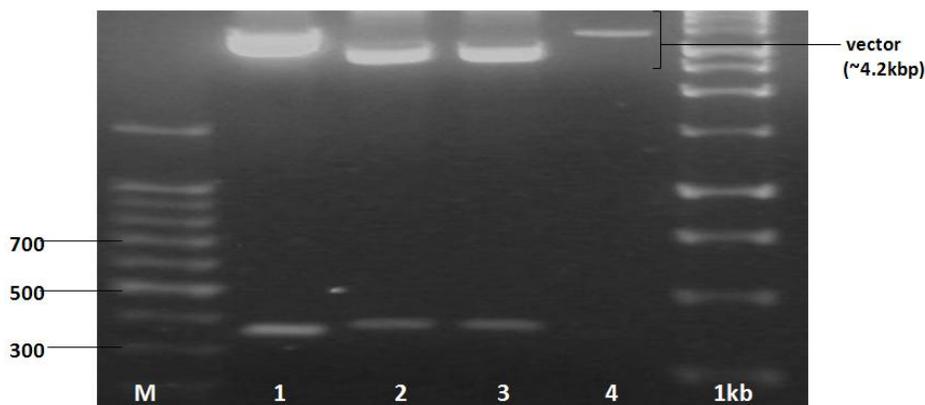
Figure 2.11: Restriction of the pGEM® vector, with *SgfI* and *PmeI*, isolated from JM109 competent *E. coli* to remove ORF3ΔN and MΔN viral genes. Lane M: 100BP marker; Lane 1: HCoV-NL63 ORF3ΔN (±360bp); Lane 2&3: HCoV-NL63 MΔN (±360bp); Lane 4: Negative control. All viral genes have an appended 5' SgfI and a 3' PmeI restriction site to allow for digestion with respective enzymes and subsequent ligation into the Flexi™ vector.
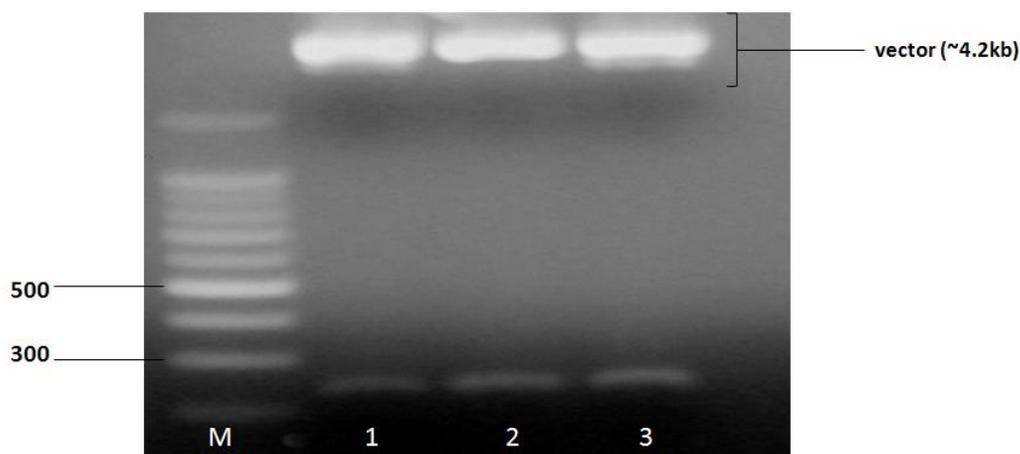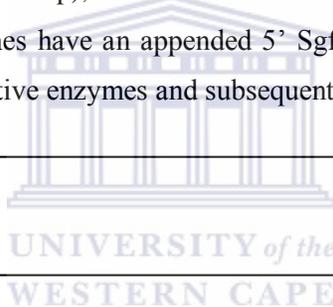

Figure 2.12: Restriction of the pGEM® vector, with *SgfI* and *PmeI*, isolated from JM109 competent *E. coli* to remove E viral genes. Lane M: 100BP marker; Lane 1-3: HCoV-NL63 E. These viral genes have an appended 5' SgfI and a 3' PmeI restriction site to allow for digestion with respective enzymes subsequent ligation into the Flexi™ vector.

Figure 2.13: Restriction of the pGEM® vector, with *SgfI* and *PmeI*, isolated from JM109 competent *E. coli* to remove M and ORF3 viral genes. Lane M: 100BP marker; Lane 1-3: cut pFLEXI vector (~4.2kbp); Lane 4: Negative control; Lane 5 HCoV-NL63 ORF3 (±678bp); Lane 6: HCoV-NL63 M (±700bp); All viral genes have an appended 5' SgfI and a 3' PmeI restriction site to allow for digestion with respective enzymes and subsequent ligation into the Flexi™ vector.

Figure 2.14: Restriction enzyme digest of recombinant N vector constructs: Lane M: 100BP marker. Lane 1: SARS N-gene (±1200BP); Lane 2: NL63 full length N-gene (±1100BP) (N1); Lane 3: 5' half of the NL63 N-gene (±550BP) coding for the N-terminal fragment of the N-protein (N2); Lane 4: 3' half of the NL63 N-gene (±550BP) coding for the C-terminal fragment of the N-protein (N3)

## 2.4.4. Colony PCR of KRX Competent *E. coli*

To confirm ligation of the various inserts into the Flexi™ vectors, CFUs were inoculated into a culture and corresponding PCR reaction and viral genes were amplified as previously described. Results were again used to select successfully transformed colonies (Figure 2.15 to 2.19).
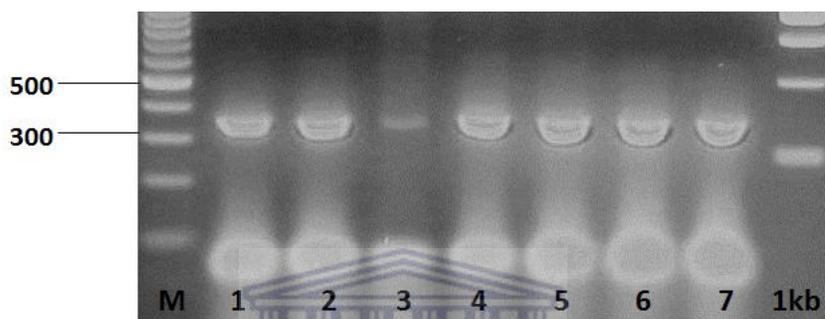

Figure 2.15: Colony PCR of KRX competent *E. coli* previously transformed with Flexi™ vector ligated with ORF3ΔN-genes. Colony PCR was performed with primers described in Table 2.1 and used as confirmation of successful transformation and ligation Lane M: 100BP marker; Lane 1-7: amplification of the ORF3ΔN-gene (±360bp). All successful amplifications of viral genes represent colonies, picked from agar plates, which were successfully transformed with the respective Flexi™-viral gene construct.
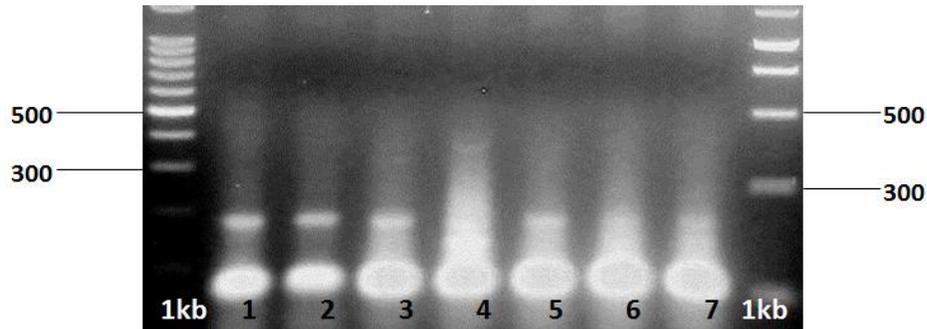
Figure 2.16: Colony PCR of KRX competent *E. coli previously transformed with Flexi™ vector ligated with* **E-***genes. Colony PCR was performed with primers described in Table 2.1 and used as confirmation of successful transformation and ligation* Lane M: 100BP marker; Lane 1-7: amplification of the E-gene (±240bp). All successful amplifications of viral genes represent colonies, picked from agar plates, which were successfully transformed with the respective Flexi™-viral gene construct.
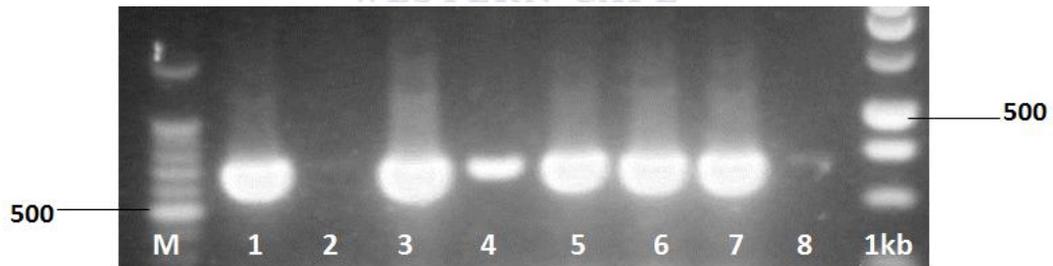
Figure 2.17: Colony PCR of KRX competent *E. coli previously transformed with Flexi™ vector ligated with* **M-***genes. Colony PCR was performed with primers described in Table 2.1 and used as confirmation of successful transformation and ligation* Lane M: 100BP marker; Lane 1&3-7: Amplification of the M- gene(±700bp); Lane 2&8: Blank. All successful amplifications of viral genes represent colonies, picked from agar plates, which were successfully transformed with the respective Flexi™-viral gene construct.
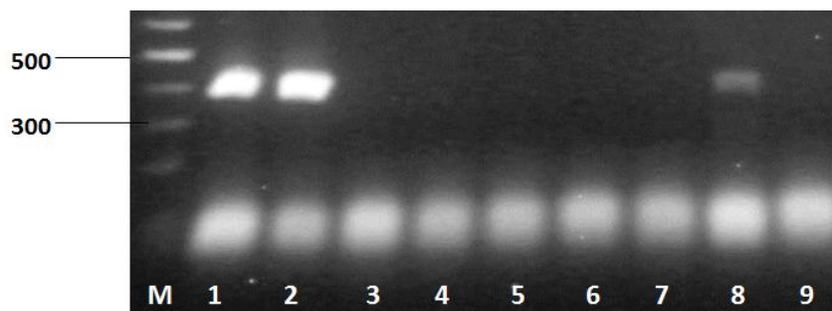
**Figure 2.18: Colony PCR of KRX competent *E. coli* previously transformed with Flexi™ vector ligated with MΔN-*genes. Colony PCR was performed with primers described in Table 2.1 and used as confirmation of successful transformation and ligation* Lane M: 100BP marker; Lane 1, 2&8: Amplification of the MΔN- gene (±390bp); Lane 3-7: Unsuccesful amplification of MΔN- gene; Lane 9: Negative control. All successful amplifications of viral genes represent colonies, picked from agar plates, which were successfully transformed with the respective Flexi™-viral gene construct
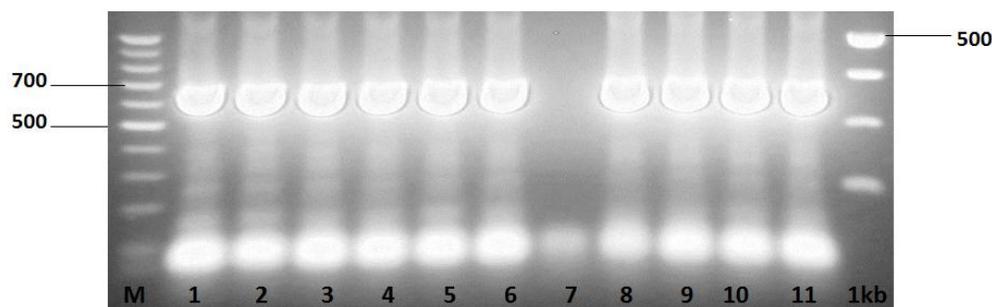
**Figure 2.19: Colony PCR of KRX competent *E. coli* previously transformed with Flexi™ vector ligated with ORF3-*genes. Colony PCR was performed with primers described in Table 2.1 and used as confirmation of successful transformation and ligation* Lane M: 100BP marker; Lane 1-6: Amplification of the ORF3-gene (±678bp); Lane 7: Negative control; Lane 8-11: Amplification of the ORF3-gene. All successful amplifications of viral genes represent colonies, picked from agar plates, which were successfully transformed with the respective Flexi™-viral gene construct.

In addition to restriction endonuclease digests (see 2.4.3), a colony PCR was also initially done to confirm insert size. Figure 2.20 A and B indicates the successful transformation of nucleocapsid genes onto KRX cells From these figures it is clear that the nucleocapsid constructs were successfully cloned and transformed into KRX competent cells.
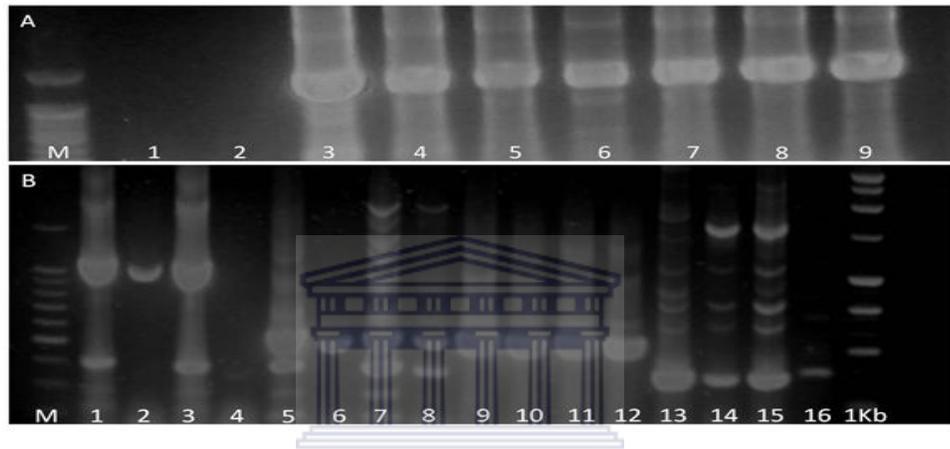
Figure 2.20 Colony PCR confirming the presence of construct that have been transformed into KRX cells (A) Lane M: 100BP marker; Lane 1: Blank; Lane 2: Negative control using all primers described in table 1 and 2; Lanes 3-9: Amplification of SARS N. (B) Lane M: 100BP marker; Lane 1-3: Amplification of N1 gene. Lane 4: Unsuccessful amplification of the N1 gene; Lanes 5-8: Amplification of the N2 gene; Lanes 9-12: Amplification of the N3 gene; Lanes 13-16.

2.4.5 Bacterial Expression of Recombinant Viral Proteins

The Flexi™ vector system contains a chromosomal copy of the T7 RNA polymerase gene that has a rhamnose promoter to control recombinant protein expression. Although the T7 polymerase system like this facilitates higher level of expression, they are prone to leaky expression. The presence of an unstable promoter sequence in an expression vector has been

previously seen to result in low levels of toxic protein expression in both the initial and lag phases, as was observed in our study. (Weickert, Doherty et al. 1996; Giacalone, Gentile et al. 2006).

In this study, the expression protocol used included sub-culturing of starter cultures into an expression culture at a dilution of 1:100. Cultures were then grown to an OD_{600} of 0.3-0.4 (Table 2.4), at which point expression was induced by the addition of 1mM IPTG and 0.1% rhamnose. Cells were harvested 5 hours post-induction and lysed using protocols as previously described.

Table 2.4: OD_{600} (mg/ml) of cultures after time specified

|	Starter Culture	5 hours post-induction
Uninduced	0.63	0.73
E	0.57	0.40
M	0.62	0.44
ORF3	0.57	0.57
ORF3ΔN	0.58	0.63
ORF3ΔN	0.62	0.57

As shown in Table 2.4, the expression of the viral-GST fusion proteins decreased the OD_{600} readings in comparison to the un-induced control. This suggests possible toxicity of the proteins to the bacterial cell, which would explain the difficulty in modifying protocols for successful expression of the viral proteins (the Western blots for the modified expression protocols can be found in Appendix II). In a biological system like *E. coli*, protein expression will occur at a

subcellular location during a defined period of time and quantity (Peti and Page 2007). This expression of a foreign protein occurs at a significantly higher level than normal, and as a result of this over-expression creates a metabolic burden on the host cell (Bianchi and Baneyx 1999). This often results in a low cell density, as seen in our OD readings obtained pre- and post-rhamnose induction. In addition, the presence of background expression at a high level could indicate a further reduction in bacterial growth, delaying the onset of a bacterial equilibrium phase (Glick 1995; Georgiou and Valax 1996). Induction usually occurs at the final bacterial growth phase, when all the glucose in the media is metabolised and the cells now begin to use the L-Rhamnose that allows expression to occur.

M and E glycoproteins failed to express. Since both are membrane-bound proteins, the former could be due to the interaction with cellular membranes. A previous study by Lia et al has also shown that the E protein is able to slow *E. coli* growth by increasing membrane permeability and subsequently leading to cell lysis. (Liao, Lescar et al. 2004). Also, other membrane proteins simply fail to express in *E. coli* and, if they do express, they do so *via* the formation of an insoluble product called an inclusion body (Shimamoto, Kasciukovich et al. 1998; Singh and Panda 2005). An inclusion body is a dense aggregate that is formed between an expression fusion protein and RNA, decreasing the yield of purified protein (de Groot, Espargaro et al. 2008; Alibolandi and Mirzahoseini 2011; Garcia-Fruitos, Vazquez et al. 2012). As a result, the need to solubilise these proteins becomes a crucial step in obtaining a high concentration of the protein of interest. There are various solubilising techniques available; however the efficacy of each method is dependent on the protein being expressed.

In an attempt to express the previously cloned SARS-N, HCoV-NL63 N1, N2 and N3 proteins, we found that there was a need to solubilise said proteins in order to obtain a workable

concentration for further studies. To do this, the insoluble fraction was treated with lysozymes as this technique does not have an impact on the expressed proteins obtained. Lysozymes digest the peptidoglycan layer of the bacterial cell wall, and are a relatively inexpensive means to solubilise insoluble proteins. Figure 2.21 shows the expression results obtained after the treatment of the cell pellet with lysozymes in the same buffer used for lysis. Expression of the proteins of interest indicates that heterologously expressed viral-fusion proteins interact with bacterial cellular membranes. In this study a variety of growth factors were investigated to provide a greater yield of non-degraded soluble proteins. These included lower growth temperature between 20°C and 30°C, inducing for a shorter period and increasing aeration, all of which proved unsuccessfully in detecting the E, M, ORF3 proteins in the soluble or insoluble fractions.
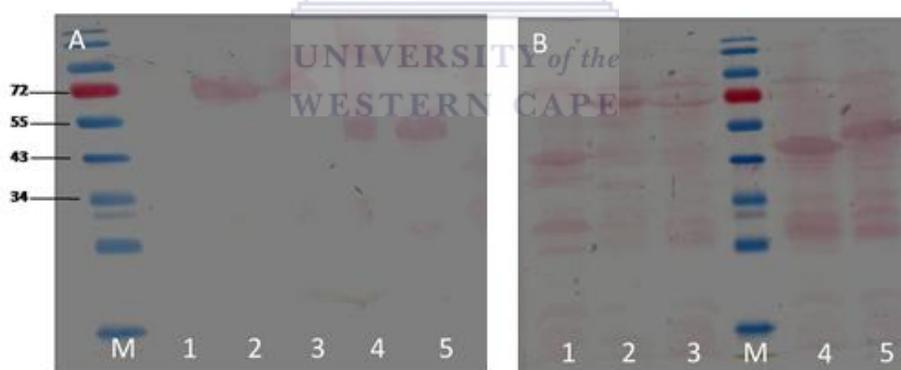
Figure 2.21: Digestion of cellular membranes using lysozymes to release membrane-bound, insoluble proteins. (A) Western blots representing proteins which remained in the insoluble fraction following digestion with lysozymes. (B) Proteins made soluble following digestion with lysozymes.(A) Lane M: Prestained molecular weight protein marker; Lane 1: Uninduced control; Lane 2: GST-SARS N; Lane 3: GST-N1; Lane 4: GST-N2; Lane 5: GST-N3. Lane order is again maintained for B.

In a further attempt to maximise expression of the E, M and ORF 3 fusion proteins the KRX strain proved to be the most suitable *E.coli* strain as it is deficient of any cytoplasmic protease genes, causing lower protein degradation. However, the rhamnose-inducible system is not a widely used system when compared to the common IPTG-inducible system that boasts extensive troubleshooting techniques. Therefore the E, M and N proteins were also cloned in a pGEX -4T-2 system to determine if the proteins are indeed toxic (Appendix III; Cloning was done as part of this Chapter, but expression of these proteins will form part of another study). Even so, the Flexi™ vector still proved to be a suitable system for the cloning and expression of the HCoV-NL63 N and SARS-CoV N proteins.

Table 2.5:-Expression of the HCoV-NL63 structural proteins

Construct	Structural Protein	Expression in this study	Source of the construct
E-pFLEXI	E-GST	-	This study
M-pFLEXI	M-GST	NR	This study
ORF3-pFLEXI	ORF3-GST	-	This study
MΔN-pFLEXI	MΔN-GST	-	This study
ORF3ΔN-pFLEXI	ORF3ΔN-GST	-	This study
SARS-N-pFLEXI	SARS-N-GST	+	Berry et al, 2012
N1-pFLEXI	N1-GST	+	Berry et al 2012
N2-pFLEXI	N2-GST	+	Berry et al 2012
N3-pFLEXI	N3-GST	+	Berry et al 2012

-: No expression; **NR:** Not repeatable; **+:** Expression

Our study concluded that the E, M and ORF3 proteins are toxic to the KRX *E. coli* strain, but further improvements could still be investigated before completely discounting this system for use in the expression of these proteins (Table 2.5). One such improvement could be the inclusion of a Kozak sequence located upstream on the mRNA. This consensus sequence has been shown to play a major role in the initiation of translation, as well as increasing protein expression without altering the nucleotide sequence of the amplified gene (Kozak 1984; Kozak 1986; Seeber 1997). A bacterial system is easy to use and high expression levels are often attained, but one major drawback is the frequent inability to synthesise native eukaryotic proteins. For this reason an expression system like the vaccinia virus vector system is useful. The vaccinia vector uses components from the bacteria in a eukaryotic model, to offer a highly efficient and specific solution for eukaryotic protein synthesis (Moss and Earl 2002). This expression system synthesises the bacterial T7 RNA polymerase in the cytoplasm as opposed to the nucleus of infected cells. As a vector, vaccinia virus has a number of useful characteristics, including a capacity to clone large fragments of foreign DNA (20+ kbp) with retention of infectivity, a wide host range, a relatively high level of protein synthesis (Moss and Earl 2001)

2.6 REFERENCES

Alibolandi, M. and H. Mirzahoseini (2011). "Chemical assistance in refolding of bacterial inclusion bodies." Biochem Res Int **2011**: 631607.

Baneyx, F. (1999). "Recombinant protein expression in Escherichia coli." Curr Opin Biotechnol **10**(5): 411-421.

Bianchi, A. A. and F. Baneyx (1999). "Hyperosmotic shock induces the sigma32 and sigmaE stress regulons of Escherichia coli." Mol Microbiol **34**(5): 1029-1038.

Braun, P., Y. Hu, et al. (2002). "Proteome-scale purification of human proteins from bacteria." Proc Natl Acad Sci U S A **99**(5): 2654-2659.

Bucher, M. H., A. G. Evdokimov, et al. (2002). "Differential effects of short affinity tags on the crystallization of Pyrococcus furiosus maltodextrin-binding protein." Acta Crystallogr D Biol Crystallogr **58**(Pt 3): 392-397.

de Groot, N. S., A. Espargaro, et al. (2008). "Studies on bacterial inclusion bodies." Future Microbiol **3**(4): 423-435.

Edwards, A. M., C. H. Arrowsmith, et al. (2000). "Protein production: feeding the crystallographers and NMR spectroscopists." Nat Struct Biol **7 Suppl**: 970-972.

Garcia-Fruitos, E., E. Vazquez, et al. (2012). "Bacterial inclusion bodies: making gold from waste." Trends Biotechnol **30**(2): 65-70.

Georgiou, G. and P. Valax (1996). "Expression of correctly folded proteins in Escherichia coli." Curr Opin Biotechnol **7**(2): 190-197.

Giacalone, M. J., A. M. Gentile, et al. (2006). "Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system." BioTechniques **40**(3): 355-364.

Glick, B. R. (1995). "Metabolic load and heterologous gene expression." Biotechnol Adv **13**(2): 247-261.

Hammarstrom, M., N. Hellgren, et al. (2002). "Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli." Protein Sci **11**(2): 313-321.

Hoffman, S. J., D. L. Looker, et al. (1990). "Expression of fully functional tetrameric human hemoglobin in Escherichia coli." Proc Natl Acad Sci U S A **87**(21): 8521-8525.

Holcroft, C. C. and S. M. Egan (2000). "Roles of cyclic AMP receptor protein and the carboxyl-terminal domain of the alpha subunit in transcription activation of the Escherichia coli rhaBAD operon." J Bacteriol **182**(12): 3529-3535.

Kapust, R. B. and D. S. Waugh (1999). "Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused." Protein Sci **8**(8): 1668-1674.

Kozak, M. (1984). "Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin in vivo." Nature **308**(5956): 241-246.

Kozak, M. (1986). "Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes." Cell **44**(2): 283-292.

Liao, Y., J. Lescar, et al. (2004). "Expression of SARS-coronavirus envelope protein in Escherichia coli cells alters membrane permeability." Biochem Biophys Res Commun **325**(1): 374-380.

Moss, B. and P. L. Earl (2001). "Overview of the vaccinia virus expression system." Curr Protoc Protein Sci **Chapter 5**: Unit5 11.

Moss, B. and P. L. Earl (2002). "Overview of the vaccinia virus expression system." Curr Protoc Mol Biol **Chapter 16**: Unit16 15.

Peti, W. and R. Page (2007). "Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost." Protein Expr Purif **51**(1): 1-10.

Rinas, U., F. Hoffmann, et al. (2007). "Inclusion body anatomy and functioning of chaperone-mediated in vivo inclusion body disassembly during high-level recombinant protein production in Escherichia coli." J Biotechnol **127**(2): 244-257.

Seeber, F. (1997). "Consensus sequence of translational initiation sites from Toxoplasma gondii genes." Parasitol Res **83**(3): 309-311.

Shimamoto, N., T. Kasciukovich, et al. (1998). "Efficient solubilization of proteins overproduced as inclusion bodies by use of an extreme concentration of glycerol." Technical Tips Online **3**(1): 141-143.

Singh, S. M. and A. K. Panda (2005). "Solubilization and refolding of bacterial inclusion body proteins." J Biosci Bioeng **99**(4): 303-310.

Weickert, M. J., D. H. Doherty, et al. (1996). "Optimization of heterologous protein production in Escherichia coli." Curr Opin Biotechnol **7**(5): 494-499.

Wu, J. and M. Filutowicz (1999). "Hexahistidine (His6)-tag dependent protein dimerization: a cautionary tale." Acta Biochim Pol **46**(3): 591-599.

UNIVERSITY *of the*
WESTERN CAPE
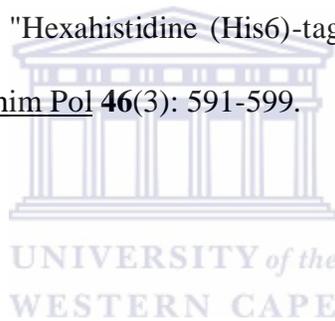
Chapter III

HCoV-NL63 nucleocapsid stimulates an immune response in whole blood cultures: A pilot study
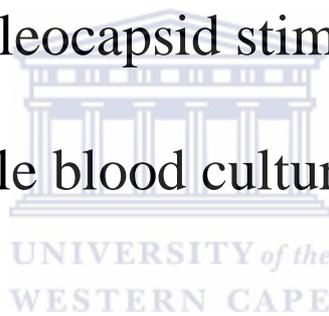
3.1 ABSTRACT

This chapter is prepared for submission as a Short Communication to *Archives of Medical Sciences*. For this reason, the format of this Chapter is different to that of Chapter II.

The Human coronavirus-HCOV-NL63 (HCoV-NL63) was first isolated in 2003 from the nasal aspirate of an infant. This virus is closely related to the severe acute respiratory coronavirus (SARS-CoV), but unlike SARS-CoV, HCoV-NL63 circulates continuously in the human population. Recent studies have focused on the effect of coronavirus infections on the host immune response, in particular on the cytokines that are produced as a response to these virus infections. The coronavirus nucleocapsid (N) protein is a multifunction protein and plays an important role in viral assembly and pathogenesis. In this study, the effects of the HCoV-NL63 N protein on the cytokines that regulate immune responses in lymphocytes were determined. Full-length as well as deletion-mutants of HCoV-NL63 N protein were expressed in *Escherichia coli* KRK cells and column-purified. Cellular responses induced by these recombinant N proteins were evaluated in whole blood culture. A double antibody sandwich ELISA assay was used to measure the immune responses; interferon-gamma (IFN-γ), interleukin-10 (IL-10) and interleukin-6 (IL-6) responses against the N proteins were observed. While lymphocyte activation resulted in high expression levels of IL-6, IFN-γ and IL-10, cytokines were secreted at much lower levels. This study showed that the HCoV-NL63 N protein elicits a broad-based cellular immune response.

3.2. INTRODUCTION

Coronaviruses are among the most common causes of respiratory infections, accounting for 10-30% of respiratory infections (Bastien, Robinson et al. 2005). Coronaviruses are a family of 27-32kb sized RNA viruses that have a positive sense single-stranded genome with helical genetic symmetry that infect mammalian and avian hosts (van der Hoek, Sure et al. 2005).

It was initially believed that there were only two human coronaviruses (HCoV) in circulation: HCoV-OC43 and HCoV-229E (van der Hoek, Pyrc et al. 2006) With the exception of SARS-CoV, it was commonly believed that human coronaviruses caused only mild respiratory tract infections (Monto and Lim 1974) As a result, the presence of HCoV within the population was not monitored. In the 1960s (McIntosh 2005), six samples with essentially different viruses causing URTs were isolated from patients, four of which shown to be serologically distinct (Tyrrell and Bynoe 1965; Hamre and Procknow 1966; Almeida and Tyrrell 1967). These four clinical samples were unfortunately lost prior to characterisation and consequently focus was directed at the two remaining samples. This resulted in HCoV-229E and HCoV-OC43 being well characterised. Additionally, the detection of HCoV-NL63 in samples collected in 1981 (Talbot, Shepherd et al. 2009) and 1988 (Fouchier, Hartwig et al. 2004) indicates that the virus has been circulating in the human population for a considerable amount of time, and it is unknown whether these represent the same strain as the previously lost HCoV's. (Pyrc, Dijkman et al. 2006). The detection of SARS-CoV, however, showed that coronaviruses are able to undergo considerable recombination that can result in a more lethal form of the virus. SARS-CoV is characterised by symptoms such as fever, myalgia, a non-productive cough and dyspnea, and has lead to more than 8000 infections in children and 800 deaths in 26 countries (Cheng, Lau et al. 2007).

This increased interest and research into CoVs has led to the discovery of two new human coronaviruses, HKU1 (Woo, Lau et al. 2005) and HCoV-NL63 (Pyrc, Berkhout et al. 2007). Infection with HCoV-NL63 presents with common cold-like symptoms such as fever, sore throat, and rhinitis (Pyrc, Jebbink et al. 2004; Vabret, Mourez et al. 2005; Han, Chung et al. 2007; Kuypers, Martin et al. 2007; Wu, Chang et al. 2008; Fielding 2011). HCoV-NL63 is also associated with acute respiratory illness and croup in young children, the elderly and immune-compromised patients. (Pyrc, Jebbink et al. 2004). Studies have found that between 1.0 and 9.3% of respiratory tract infections in children have been positive for HCoV-HCOV-NL63 (Fielding 2011) with 75% of children aged 2.5-3.5 years being sero-positive for HCoV-NL63 (Dijkman, Jebbink et al. 2008).

To date, research groups in Canada, Brazil and Denmark have all reported fatalities following infections by HCoV-NL63 (Bastien, Anderson et al. 2005; Oosterhof, Christensen et al. 2010; Cabeca and Bellei 2012). This indicates that HCoV-NL63 could play a crucial role in more severe respiratory diseases. In the Denmark report, a patient that underwent an allo-haematopoietic cell transplant presented with influenza-like symptoms five months after surgery. Clinical samples of the tracheal and nasopharyngeal fluid tested negative for various respiratory viruses such as the RSV, influenza A and B, para-influenza virus, meta-pneumovirus, adenovirus and rhinovirus. PCR of the bronchlear lavage fluid tested positive for HCoV-NL63. The patient's condition continued to deteriorate and, five weeks after admission into hospital, the patient died as a result of progressive respiratory failure. This isolated case, along with the previously mentioned reports in Canada and Brazil, all presented with a suppressed or compromised immune systems, indicating that routine screenings for coronaviruses in particular HCoV-NL63 should be conducted. (Oosterhof, Christensen et al. 2010).

Phylogenetic analysis using complete genomes have divided coronaviruses into three distinct groups (Liu and Sun 2008). However, the International Committee for Taxonomy of Viruses (ICTV) recently proposed that these three groups be replaced by three new genera, namely *Alpha-*, *Beta-, Gamma-*and *Delta-coronavirus*(Woo, Huang et al. 2010).HCoV-NL63 belongs to the group I coronaviruses, according to phylogenetic analyses. The highest similarity is observed with HCoV-229E and porcine epidemic diarrhoea virus (PEDV), 65% and 61%, respectively (van der Hoek, Pyrc et al. 2004) HCoV-NL63 also presents with symptoms similar to those caused by other coronaviruses. This makes the identification of the causative coronavirus based solely on the observation of symptoms in patients difficult. (Koetz, Nilsson et al. 2006).

Coronaviruses have a positive sense single-stranded RNA genome, approximately 25-30kbs in size. There are four major structural proteins that are encoded by the genome; these are the spike (S), membrane (M), envelope (E), and the nucleocapsid (N) proteins. Of these structural proteins, the coronavirus nucleocapsid protein is the most abundantly expressed protein with an infected cell (Surjit, Liu et al. 2004; Hsieh, Chang et al. 2005; Zuniga, Sola et al. 2007) The N protein is responsible for several functions, one of the most vital being the binding to viral RNA to allow for the formation of the ribonuclocapsid core(Hsieh, Chang et al. 2005). The N protein is also believed to play a role in the replication, transcription and translation of the virus (Sawicki, Sawicki et al. 2007).

Currently, the N proteins plays a role in several functions, including the inhibition of interferon production (Kopecky-Bromberg, Martinez-Sobrido et al. 2007), the up-regulation of the AP1 signal transduction pathway (He, Leeson et al. 2003), interaction with various cellular proteins including cyclophilin A (Luo, Chen et al. 2005) as well as the deregulation in the host cell cycle (Surjit and Lal 2008).

Several studies have shown that the N protein of the order *Nidovirales* localises to the nucleus or the nucleolus (Hiscox, Wurm et al. 2001; Timani, Liao et al. 2005). On the other hand, some studies have also shown that the N protein may in fact localise to the cytoplasm (Rowland, Chauhan et al. 2005; You, Dove et al. 2005) . These two opposing findings indicate that the localisation of the N protein is not a conserved process in all the nidoviruses. The N protein is proposed to be an important target protein in the development of effective coronavirus diagnostic tools. This is due to that fact that the N protein has been shown to be highly antigenic and is thus able to induce an effective antibody response in the host cell (Leung, Tam et al. 2004).

In addition, several reports have shown that the SARS-CoV N protein, as is with other coronaviruses, induces a specific T cell response and also interferes with various cellular pathways, indicating that the N protein can play a crucial regulatory role (Hogue 1995; Narayanan, Maeda et al. 2000; Narayanan, Chen et al. 2003). The importance of this in relation to host immune response is unknown and warrants further investigation.

The host immune response to HCoV-NL63 N is still not well understood. Therefore, this chapter focuses on characterization of the host immune response to HCoV-NL63 N in human whole blood culture. Recombinant HCoV-NL63 full-length N, as well as two deletion mutants, expressed in a bacterial cell system, were used to treat whole blood cultures. An ELISA was used to measure pro- and anti-inflammatory activity after 18 hours. A statistically significant IL-6 response was detected, indicating that all HCoV-NL63 recombinant N proteins stimulated an inflammatory response.

3.3. MATERIALS AND METHODS

3.3.1 Protein expression and preparation of cell lysate

Selected plasmids from the glycerol stocks mentioned in 3.1 were spread-plated and grown overnight at 37°C. Thereafter a single colony was selected from the plates and grown for 14 hours at 37°C in 10 ml LB (10 g pancreatic digest of casein, 5 g yeast extract powder, and 5 g NaCl, per 1L water) broth containing ampicillin (1.0 mmol/L) as an antibiotic. These were diluted 1:100 with LB containing ampicillin, and thereafter they were grown at 37°C with shaking. Protein expression was induced with 0.05% (w/v) glucose and 0.1% (w/v) Rhamnose. At 8 hours, post-induction cells were harvested by centrifugation at 6000rpm for 15 minutes at 4°C to prevent protein degradation. Dry cell pellet was frozen at -80°C until further use. Every 2 grams of cells were resuspended in 2ml of lysis buffer (1% (v/v) TritonX, 150mM NaCl, 10mM Tris (hydroxymethyl)-amino-methane (Tris), 5mM EDTA and a protease inhibitor. Cells were lysed by sonication for 3 minutes at 30-second intervals in an ice bath to prevent protein degradation. Following sonication, insoluble cell fractions were separated from soluble fractions by centrifugation at 14000rpm for 15 minutes at 4°C. The insoluble fraction was discarded and the soluble supernatant was stored at -20°C until further use.

3.3.2 Protein analysis

3.3.2.1 SDS Page

For SDS-PAGE analysis, sterile gel plates were assembled and verified for any leakages. 15% of SDS gel was prepared using these plates according to the Fluka specifications. For quality purpose, APS (10%w/v) was prepared every week. Prior to sample loading into gel wells, a 1.5:1 mix ratio of protein sample with loading dye was performed. Approximately 25

µl and 10 µl of total protein and protein marker, respectively, were loaded into protein wells. These were separated according to their electrophoretic ability, applying constant current at 20 amperes per gel. Following this reaction the gel was either submerged in Commassie stain (45 % methanol, 10 % glacial acetic acid, 45 % water, and 3 g/L Coomasie brilliant blue R250) overnight or subjected to Western blotting for further experiments.

3.3.2.2 Western blot

The separated proteins were transferred to a nitrocelullose membrane for Western blot studies. The SDS gel was transferred to a Western blot transfer system covered with nitrocelullose membrane that has been sandwiched with Whatmann paper and sponges immersed in pre-cooled transfer buffer. The experiment was carried out at 4°C with the transfer system running at 100 V and constant amperes for a period of 120 min.

On completion of the transfer, the nitrocellulose membrane was immersed in Ponceau stain with shaking for 3 minutes and then blocked with 3% (w/v) milk for approximately 30 minutes. The membrane was then incubated in 3% (w/v) milk and 0.05% (v/v) tween 20 PBS solution with a rabbit anti-GST, primary antibody in a ratio of 1:5000 overnight at 4°C. This was followed by a washing step with PBS containing 0.05% Tween-20 for 45 minutes, and then covered with a secondary antibody, peroxidase-labeled goat anti-rabbit for 60 minutes in a 1:2000 ratio. The membrane was again washed twice with PBS/0.05% Tween-20 at 30 minute intervals, after which 1000 µl of 1 Component TMB membrane peroxidase substrate was added on the membrane to detect the presence of recombinant protein from the total bacterial proteins. The presence of the protein of interest was indicated by the appearance of a purple colour on the nitrocellulose membrane.

3.3.2.3 Quantification of protein concentrations

For quantification of proteins, a Qubit-iTTM Assay was used. Also, 199 x n µl of Quant-iTTM buffer (where n represents the number of standards plus number of samples) was mixed up with 1 x n µl of Quant-iTTM reagent to a make up a Quant-iTTM working solution. 190 µl of the latter was then mixed with 10 µl of standards, whereas 180-199 µl of working solution was mixed with user samples to make up a final volume of 200 µl. Following equilibration of the QubitTM fluorometer with known standards, fluorometer readings were taken and the concentrations (mg/ml) were calculated using the formula C = QF x 200/10 (where QF = Qubit fluorometer reading, C = concentration).

3.3.2.4 Purification

A MagneGST™ purification system was then utilised to bind GST fusion proteins. The MagneGST™ particles were equilibrated by washing 3 times in the MagneGST™ binding/wash buffer (4.2mM Na_2HPO4, 2mM K_2HPO4, 140mM NaCl, 10mM KCl). Particles were separated from the binding/wash buffer by placing tubes in the magnetic stand. Glutathione particles were then resuspended in 250µl binding/wash buffer and 25µl lysozyme treated pellet, and 25µl soluble fraction was added and incubated at room temperature on an orbital shaker for 30 minutes. Glutathione particles (now bound to GST fusion proteins) are isolated in the magnetic stand and washed 3 times in the binding/wash buffer. The particles were then resuspended in 100µl elution buffer (50mM Glutathione, 50mM Tris-HCl) and incubated at room temperature for 30 minutes on an orbital shaker. Particles are again separated in the magnetic stand and the supernatant (containing the purified GST fusion proteins) was removed and stored at -20°C.

3.3.3 Whole blood cytokine assays

3.3.3.1. Blood sample collection

Blood was obtained from healthy, male volunteers and stored at ambient room temperature. Consent was obtained from all participants. Blood samples were collected in 10 ml citrate-containing vacuum tubes by veni-puncture and used within 8 hours of collection. All procedures were performed under sterile conditions. Whole blood was diluted 1:10 with Roswell Park Memorial Institute (RPMI)-1640 (Sigma-Aldrich, Germany) before being stimulated with the full-length SARS-COV and HCOV-NL63 N and Deletion Mutant proteins.

3.3.3.2. ELISA test to measure cytokine production

Cytokine production was measured in the supernatants, following incubation with the recombinant viral proteins for 18 hours. In this study, human blood was treated with the B-cell mitogen, lipopolysaccharide (LPS) and the T cell mitogen, phytohemagglutinin, (PHA) which served as controls. Whole blood culture supernatants were screened for IL-6 as a biomarker of inflammation, IFN-γ as a biomarker of cell mediated immunity and IL-10 as a biomarker for humoral immunity. Cytokine production by whole blood cultures were measured using ELISA kits (eBioscience, USA). The kits contained all reagents required for the assay. Briefly, 96-well plates (Nunc-Immuno plate, Serving Life Science, Denmark) were coated with 100 µl pre-titrated purified capturing antibody per well diluted in coating buffer and incubated overnight at 37°C. After 5 washes with wash buffer (autoclaved phosphate buffered saline containing 0.05 % Tween-20), non-specific binding sites were blocked with assay diluent for 1 hour at room temperature. 50 µl supernatants were then added to their respective wells. Recombinant human cytokine standards were also included on each plate to generate standard curves and calibrate samples. The plate was sealed and incubated at room temperature for 2 hours. After 5 washes, 100 µl of detection antibody (Biotin-conjugated

anti-human cytokine) was added to each well. The plate was incubated for 1 hour at room temperature. The plate was washed again 5 times and the biotineylated sandwich was detected by adding 100 µl of the Avidin-horseradish peroxidase conjugate (HRP) to all wells. The plate was then incubated for 30 minutes at room temperature. After 7 washes the bound peroxidase was monitored by adding 100 µl of the substrate solution (Tetramethlybenzidine solution) to every well. The plate was incubated for approximately 15 minutes at room temperature after which the reaction was stopped by the addition of 50 µl stop solution to all wells. Absorbance was read at 450 nm on an ELISA plate reader.

3.3.3.3 Statistical Analysis

All data was statistically analysed via one-way ANOVA using SigmaStat software (Systat Software Inc., USA). All cytokine assays were conducted in triplicates and repeated 4 times to avoid statistical errors. Results were expressed in terms of the mean ± standard deviation or the mean ± standard error of the mean (SEM).
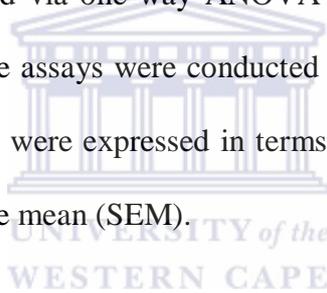
3.4 RESULTS

3.4.1 Expression of the recombinant viral proteins

Proteins were successfully expressed and analysed using SDS-PAGE, and Western blot.

Figure 3.1 represents total proteins that were expressed in a bacterial system.

Figure 3.1 Western blot analyses if recombinant proteins expressed in bacterial cell in the soluble fraction. On the figure, M represents the pre-stained protein marker. Lane 1-4 represents detected proteins in the following order: SARS-CoV N, N1, N2, and N3.
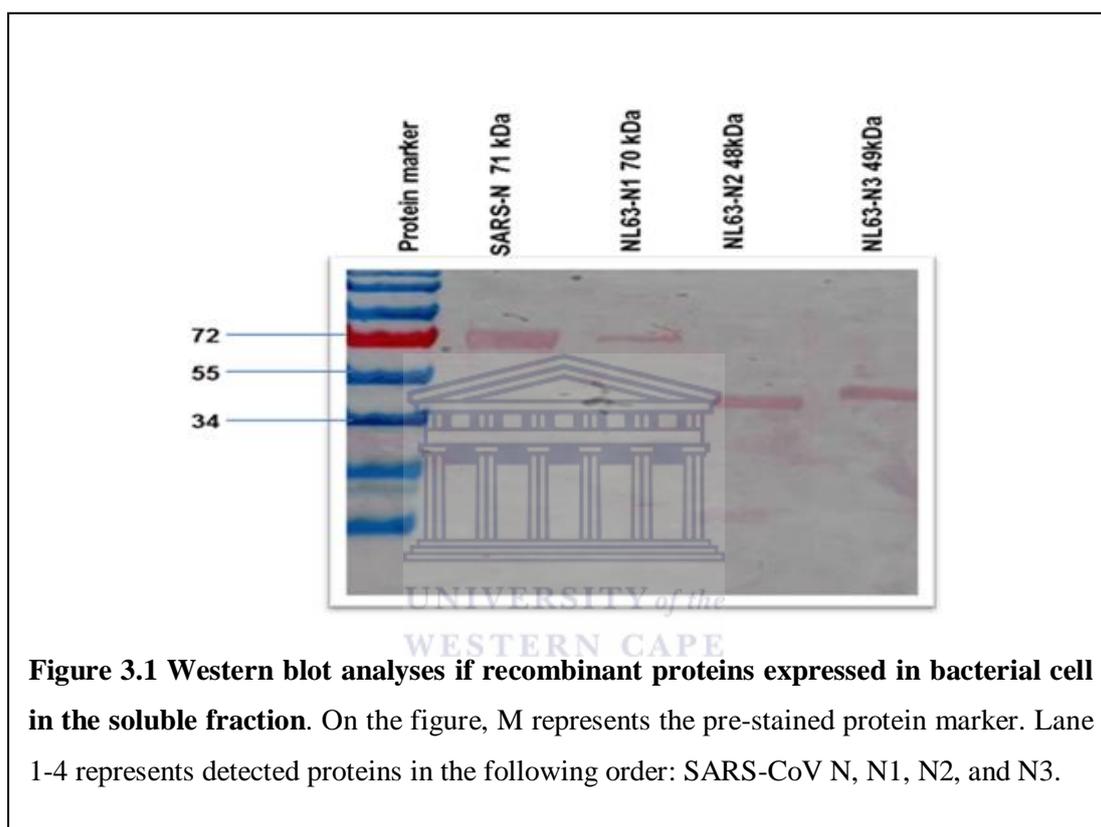
3.4.2 Purification of the recombinant proteins by MagneGST™

The purification of recombinant protein was conducted using the MagneGST protein purification system. Following the optimisation of protein expression, the presence of the nucleocapsid proteins was confirmed by small-scale purification; this is clearly shown in Figure 3.2 below.
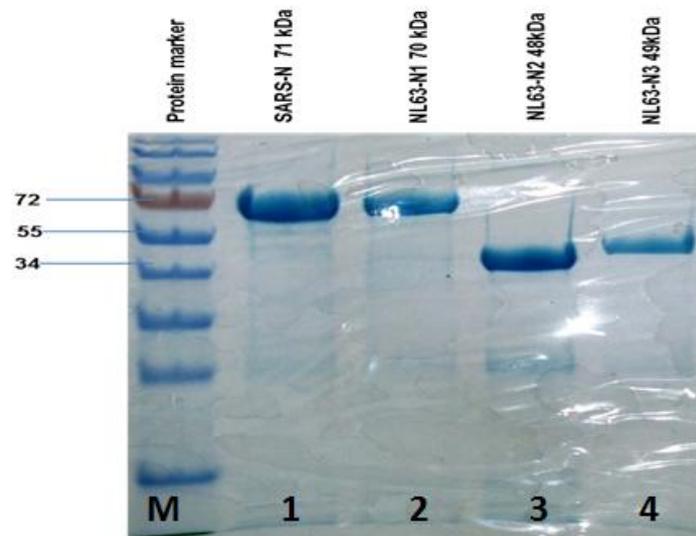
Figure 3.2: Coomassie stain of GST fusion viral proteins, purified using the MagneGST™ purification system. "M" indicates the protein marker. Lane 1-4 represent purified proteins in the following order: SARS-CoV-N, N1, N2, and N3

3.4.3 Quantification of protein concentrations

Table 3.1.represents the purified protein concentrations obtained using the Qubit-iTTM system.

Construct	Concentration (mg/ml)
SARS-COV-N	0.45
N1	0.11
N2	0.34
N3	0.18

3.4.4 The effect of Nucleocapsid viral proteins on unstimulated whole blood cultures

IL-6 synthesis was used as a biomarker to detect an inflammatory response in LPS unstimulated whole blood cultures. Figure 4.5 illustrates a typical standard curve used to determine IL-6 secretion in WBCs incubated with various viral proteins. A polynomial relationship between OD and IL-6 synthesis (pg/ml), with a correlation coefficient of 0.997, was observed. Upon PHA stimulation of WBCs, the T cell derived cytokines IFN-γ and IL-10 are produced. IFN-γ was used as a biomarker for cell mediated immunity and IL-10 as a biomarker to determine the humoral immune response. Figure 4.3 also illustrates the typical standard curves constructed to determine IFN-γ and IL-10 synthesis in WBCs exposed to viral proteins. A linear progression for both IL-10 and IFN-γ standard curves, with correlation coefficients of 0.985 and 0.998, respectively, were observed.
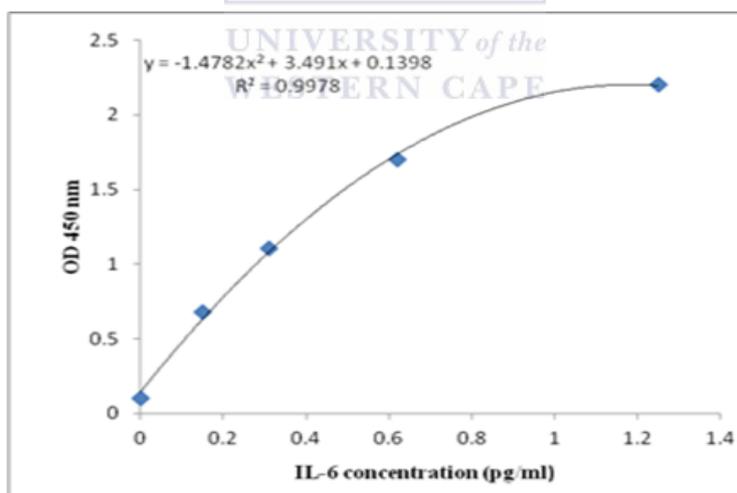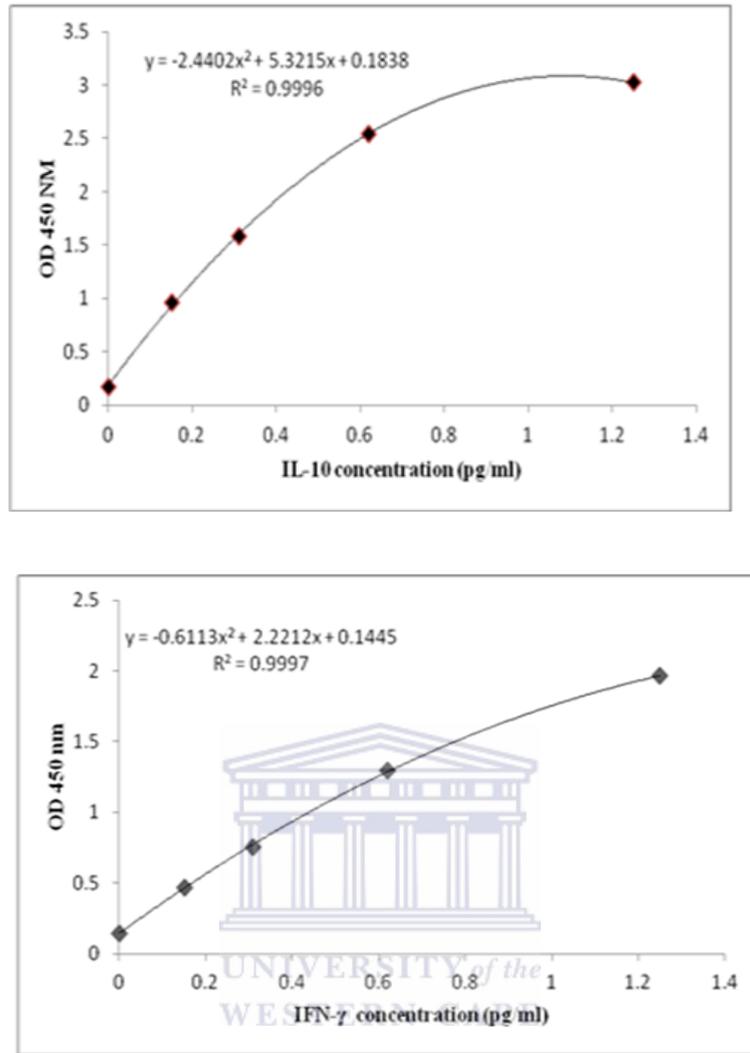
Figure 3.3 Standard curves for IL-6, IFN-γ and IL-10 synthesis (pg/ml) versus OD (450 nm). A polynomial relationship for the IL-6 standard curve with a correlation coefficient of 0.993 was generated. Standard curves obtained for IL-10 and IFN-γ showed a linear progression with correlation coefficients of 0.985 and 0.998, respectively.

3.4.5 The inflammatory activity of the recombinant viral proteins

Results show that all the recombinant viral proteins are significantly different from the positive control LPS. But HCoV-NL63 full length N protein (N1) and the truncated clones N2 and N3 released the higher levels of IL-6 when compared to SARS-CoV-N.

Figure 3.4 presents the average IL-6 concentrations obtained for all three donors on incubation with the various proteins.

Figure 3.4 Induction of IL-6 (pg/ml) of whole blood cultures in the presence of 10μg/ml of viral proteins various and distilled water (control). Each point represents the mean and standard deviation of three replicates. An asterisk (*) designates significant difference to the control (P<0.05).
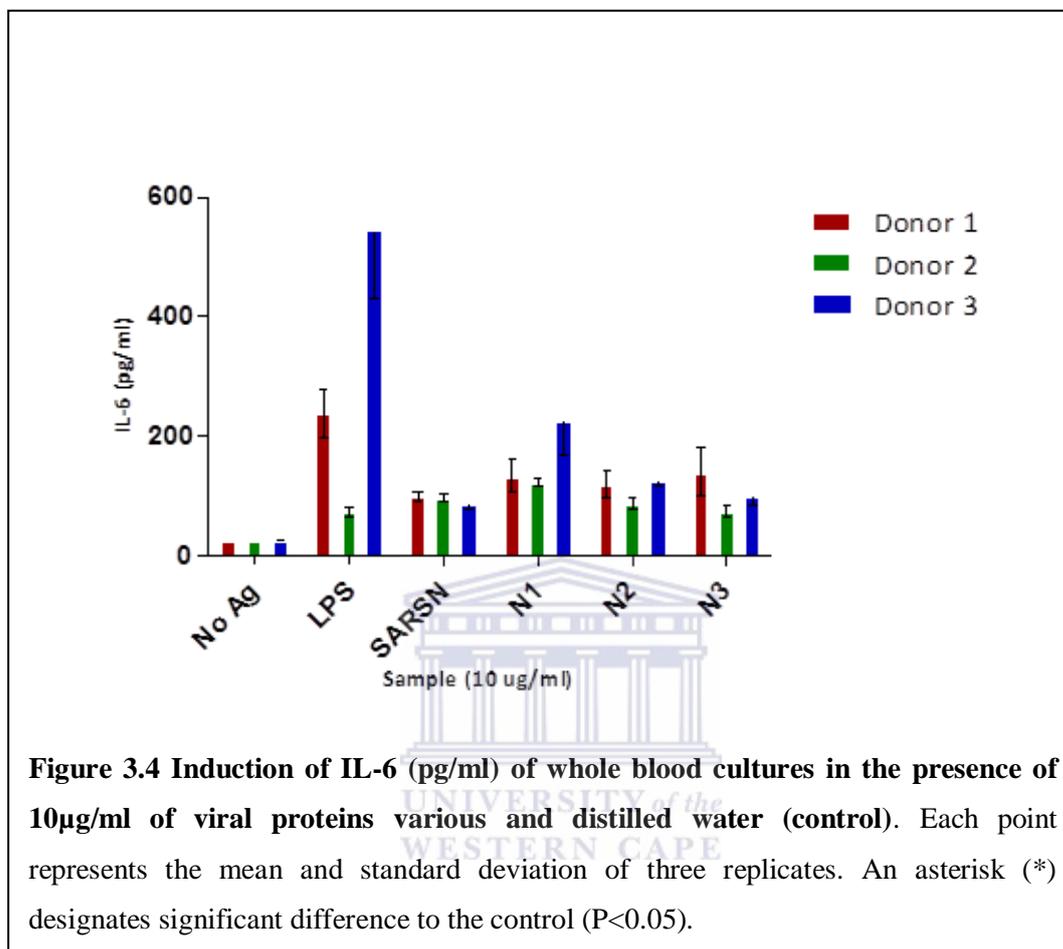
3.4.6 Effect of recombinant viral proteins on T-cell differentiation

Figures 3.5 and 3.6 represents the average IL-10 and IFN-γ levels obtained for all donors treated with the viral proteins. Results for IL-10 show that all viral proteins reacted equally when compared to the positive control, PHA. These results reflect that the N proteins of both SARS-COV and HCOV-NL63 may not have an effect on the humoral immune response. IFN-γ has been shown to enhance the activation of macrophages as well as antibody dependent cellular cytoxicity. This form of immune response serves as a defence against intracellular pathogens ((Langezaal, Coecke et al. 2001) and the comparison of the various

viral proteins to the control, PHA, showed that none of the proteins had an effect on IFN-γ synthesis and, subsequently, cell-mediated immunity.

Figure 3.5 Induction of IL-10 (pg/ml) of whole blood cultures, in the presence of 10µg/ml of the recombinant viral proteins and distilled water (control). Each point represents the mean and standard deviation of three replicates (n=12). An asterisk (*) designates significant difference to the control (P>0.05).
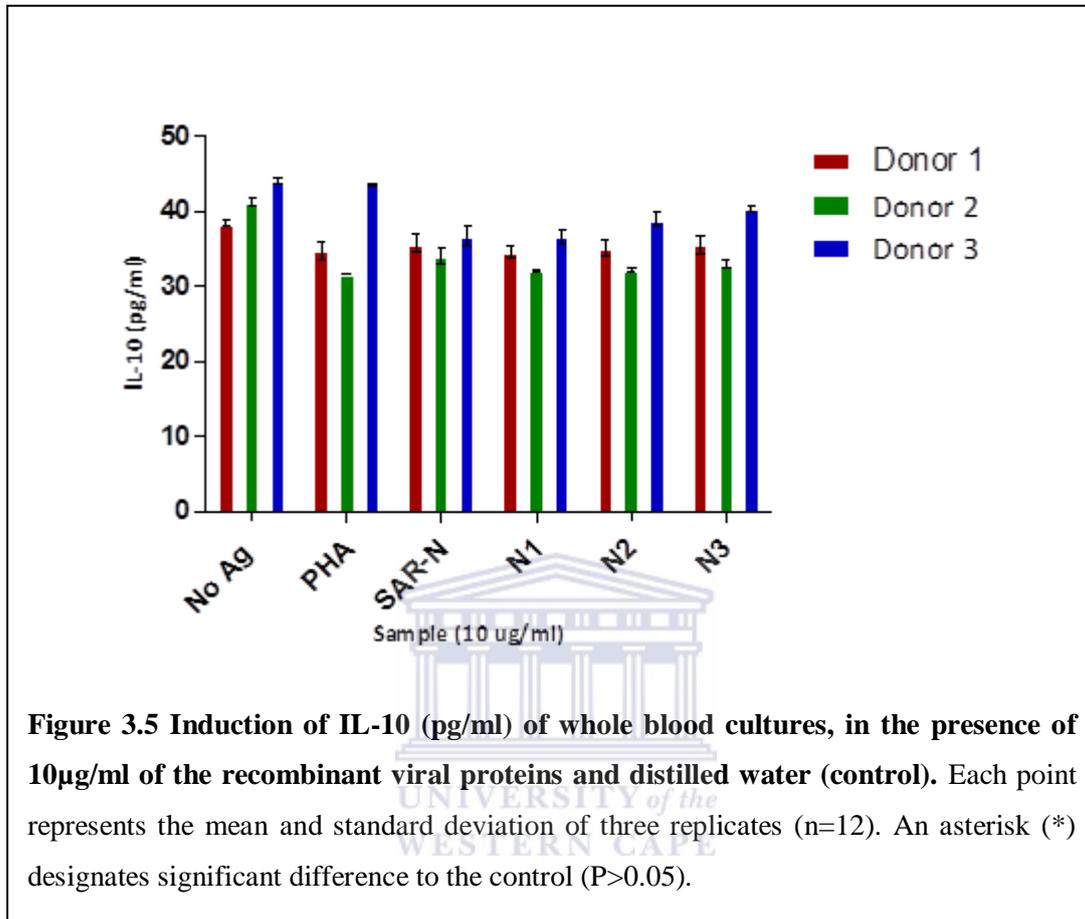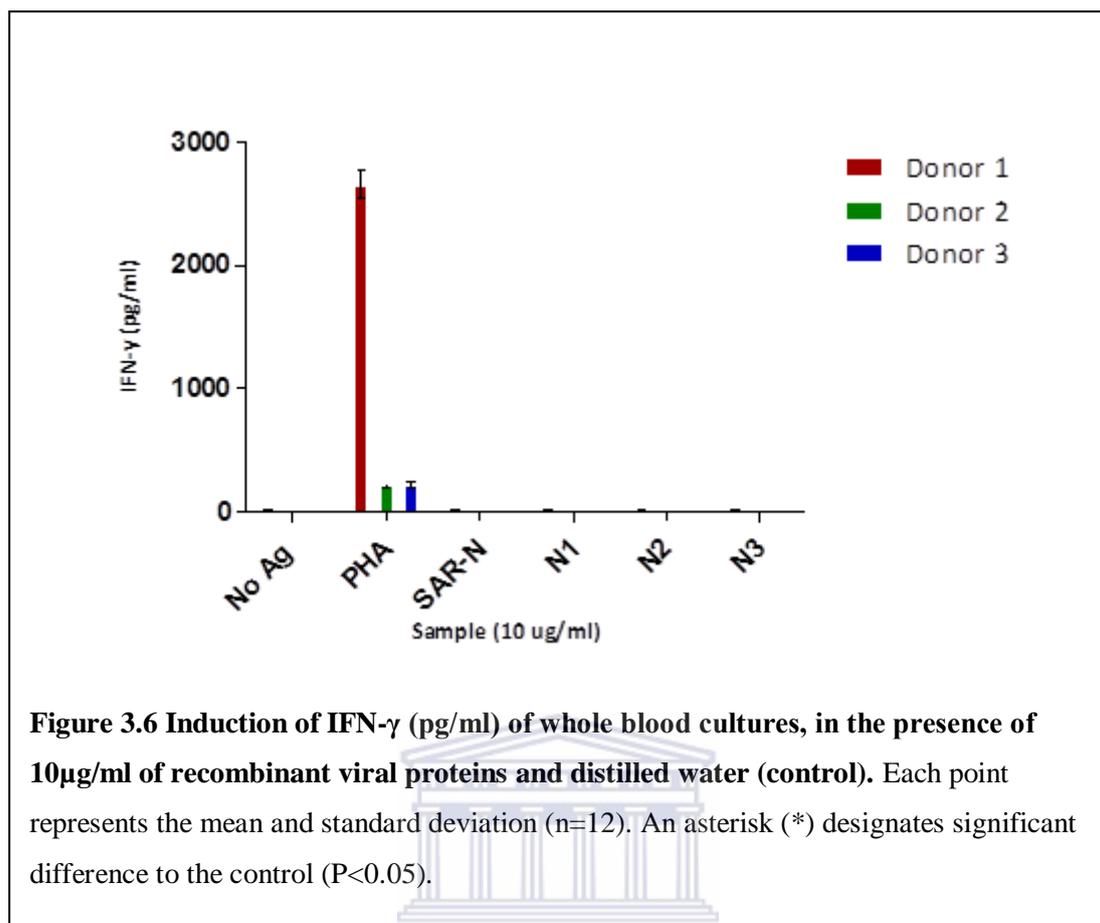
Figure 3.6 Induction of IFN-γ (pg/ml) of whole blood cultures, in the presence of 10μg/ml of recombinant viral proteins and distilled water (control). Each point represents the mean and standard deviation (n=12). An asterisk (*) designates significant difference to the control (P<0.05).

4.5. DISCUSSION AND CONCLUSION

Stimulation of the innate immune response is important for the control and elimination of invading pathogens. As a result, an interaction between a virus and innate immune mediators will allow for the limitation of viral spread within the host system during the early phases of disease (Katze, He et al. 2002). Studies have also shown that many viruses use a specific genes to encode products that are able to affect both the innate and acquired immune system responses (Symons, Alcami et al. 1995; Gale, Korth et al. 1997; Basler, Wang et al. 2000; Parisien, Lau et al. 2001; Park, Shaw et al. 2003; Andrejeva, Childs et al. 2004; Meylan, Curran et al. 2005; Ye, Hauns et al. 2007). It has been shown that the epithelial cells of the airways are a primary site for HCoV-NL63 infection, allowing for synthesis of many

cytokines. Therefore, an investigation into how these specific viruses and their viral proteins interact with the host innate immune system is essential to understand the way in which virulence and disease outcome is regulated. This present pilot study is the first to explore the cytokines that are produced when the HCoV-NL63 viral protein N interact with human whole blood cultures.

In this study, human blood was treated with the B-cell mitogen - lipopolysaccharide (LPS) - and the T cell mitogen - phytohemagglutinin (PHA) -, which served as controls in our study. LPS stimulates an inflammatory response and initiates the release of IL-6 from monocytes and B-lymphocytes. On the other hand, PHA stimulates the cell mediated and humoral immune pathways, initiating the release of IFN-γ and IL-10 from T lymphocytes (Hussain, Kaleem et al. 2002).
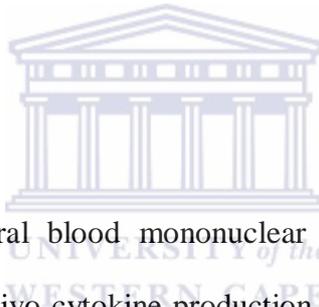
The isolation of peripheral blood mononuclear cells (PBMC) remains the most popular methods to measure ex vivo cytokine production as it is simple and less expensive. However PBMC cultures have been shown to have an inconsistent monocyte to lymphocyte ratio when compared to whole blood cultures (Popa, Barrera et al. 2009). Whole blood cultures on the other hand contain the efficient physiological proportions of blood bourn factors such as moncytes, lymphocytes and platelets that are able to maintain the cells their natural environment in vivo (Muller, Herner et al. 1998; Hussain, Kaleem et al. 2002) and as such can be used to accurately determine the immune response to various antigens and. Our study used whole blood cultures as we conclude that it is a well suited low cost method to measure monocytic cytokine production. Also, whole blood cultures can be carried out immediately after sampling and only requires the use of small volumes of blood to be obtained.

Our results showed that HCoV-NL63 N1, as well as N2 and N3, promoted the synthesis of the cytokine IL-6 (Figure 3.4). This is similar to the results of a previous study that showed that the SARS-CoV N protein activates IL-6 expression (Zhang, Wu et al. 2007). IL-6 is a pleiotrophic inflammatory cytokine that is produced by various cells including mononuclear phagocytes, fibroblasts and endothelial cells (Akira, Taga et al. 1993; Muller, Herner et al. 1998; Papanicolaou, Wilder et al. 1998). It also plays a crucial role in cellular humoral responses, as well as in antibody production In addition, to identify the antigenic epitopes, Zhang and colleagues also used different regions corresponding to the N- and C-termini of the N protein in their study. They concluded that both the N- (86-96 amino acid regions) and the C-terminus (341-422 amino acid region) stimulate an increase in IL-6 activity. This agrees with our findings, as our study also showed that both N2 and N3 elicited similarly strong IL-6 responses; however, in our study this was donor dependent. The increased release of the IL-6 cytokine has also been reported for other coronaviruses, *viz* HCoV-OC43 and the Feline Infectious Peritonitis virus (Edwards, Denis et al. 2000; Foley, Rand et al. 2003). Various studies of SARS-CoV tissue pathology also show the presence of high levels of the IL-6 inflammatory cytokine (Zhang, Li et al. 2004; Huang, Su et al. 2005; Yu, Hu et al. 2005; Roberts, Deming et al. 2007). This up-regulation of IL-6 is probably due to a secondary response to the initial activation of the immune system (Thiel and Weber 2008). Taken together, our and previous IL-6 data, suggested that the recombinant HCoV-NL63 N proteins have the potential to induce an inflammatory action in a healthy immune system.

The innate immune system response to SARS-CoV infection is related to the pathological changes that occur during disease progression (Huang, Su et al. 2005; Theron, Huang et al. 2005; Okabayashi, Kariwa et al. 2006). The synthesis of Type 1 interferon in

response to an intracellular pathogen, like a virus, is important for the effective functioning of the innate immune system during infection (Stark, Kerr et al. 1998; Takaoka and Yanai 2006). In fact, SARS-CoV N suppresses IFN expression by targeting cellular recognition receptors on the host cells in the innate immune pathway (Frieman, Heise et al. 2008; Thiel and Weber 2008). The production of innate immune system cytokines during HCoV-NL63 infection has not yet been studied in detail. Therefore, in this chapter, we looked at the response of two cytokines, namely IFN-γ and IL-10, to the treatment of whole blood with HCoV-NL63 N. Results for N1, N2 and N3 showed that HCoV-NL63 N did not stimulate IFN-γ secretion (Figure 3.6). This could indicate that the HCoV-NL63 N protein may have a low-level impact on cell mediated immunity, although there are no previous studies to support this. IL-10 is an anti-inflammatory cytokine produced by Th2 cells, which stimulates humoral immunity i.e. B cell activation and maturation resulting in antibody production (Storni, Kundig et al. 2005; Viveros-Paredes, Puebla-Perez et al. 2006). Results showed that none of the recombinant viral proteins had an impact on IL-10 synthesis. This could indicate that the HCoV-NL63 N1, N2 and N3 proteins do not have an immune-suppressive effect on humoral immune response system (Figure 3.5); however optimization and repetition of this experiment is required to substantiate this viewpoint.

Although donors showed similar IL-6 responses, there were slight differences in the secretion levels of this cytokine (Figure 3.4.). This individual variability could be attributed to the fact that patient blood samples can rarely be standardised, as responses are dependent on the functionality of an individual's existing immune system. Nonetheless, data obtained provided a vital understanding into the immune response elicited by the HCoV-NL63 N protein. Even though the results for this pilot study showed that HCoV-NL63 N stimulated a

pro-inflammatory response, further investigation to identify immuno-stimulatory epitopes for N is required.

3.6 REFERENCES

Akira, S., T. Taga, et al. (1993). "Interleukin-6 in biology and medicine." <u>Adv Immunol</u> **54**: 1-78.

Almeida, J. D. and D. A. Tyrrell (1967). "The morphology of three previously uncharacterized human respiratory viruses that grow in organ culture." <u>J Gen Virol</u> **1**(2): 175-178.

Andrejeva, J., K. S. Childs, et al. (2004). "The V proteins of paramyxoviruses bind the IFN-inducible RNA helicase, mda-5, and inhibit its activation of the IFN-beta promoter." <u>Proc Natl Acad Sci U S A</u> **101**(49): 17264-17269.

Basler, C. F., X. Wang, et al. (2000). "The Ebola virus VP35 protein functions as a type I IFN antagonist." <u>Proc Natl Acad Sci U S A</u> **97**(22): 12289-12294.

Bastien, N., K. Anderson, et al. (2005). "Human coronavirus NL63 infection in Canada." <u>J Infect Dis</u> **191**(4): 503-506.

Bastien, N., J. L. Robinson, et al. (2005). "Human coronavirus NL-63 infections in children: a 1-year study." <u>J Clin Microbiol</u> **43**(9): 4567-4573.

Cabeca, T. K. and N. Bellei (2012). "Human coronavirus NL-63 infection in a Brazilian patient suspected of H1N1 2009 influenza infection: description of a fatal case." <u>J Clin Virol</u> **53**(1): 82-84.

Cheng, V. C., S. K. Lau, et al. (2007). "Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection." <u>Clin Microbiol Rev</u> **20**(4): 660-694.

Dijkman, R., M. F. Jebbink, et al. (2008). "Human coronavirus NL63 and 229E seroconversion in children." J Clin Microbiol **46**(7): 2368-2373.

Edwards, J. A., F. Denis, et al. (2000). "Activation of glial cells by human coronavirus OC43 infection." J Neuroimmunol **108**(1-2): 73-81.

Fielding, B. C. (2011). "Human coronavirus NL63: a clinically important virus?" Future microbiology **6**(2): 153-159.

Foley, J. E., C. Rand, et al. (2003). "Inflammation and changes in cytokine levels in neurological feline infectious peritonitis." J Feline Med Surg **5**(6): 313-322.

Fouchier, R. A., N. G. Hartwig, et al. (2004). "A previously undescribed coronavirus associated with respiratory disease in humans." Proceedings of the National Academy of Sciences of the United States of America **101**(16): 6212-6216.

Frieman, M., M. Heise, et al. (2008). "SARS coronavirus and innate immunity." Virus Res **133**(1): 101-112.

Gale, M. J., Jr., M. J. Korth, et al. (1997). "Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein." Virology **230**(2): 217-227.

Hamre, D. and J. J. Procknow (1966). "A new virus isolated from the human respiratory tract." Proc Soc Exp Biol Med **121**(1): 190-193.

Han, T. H., J. Y. Chung, et al. (2007). "Human Coronavirus-NL63 infections in Korean children, 2004-2006." Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology **38**(1): 27-31.

He, R., A. Leeson, et al. (2003). "Activation of AP-1 signal transduction pathway by SARS coronavirus nucleocapsid protein." Biochem Biophys Res Commun **311**(4): 870-876.

Hiscox, J. A., T. Wurm, et al. (2001). "The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus." J Virol **75**(1): 506-512.

Hogue, B. G. (1995). "Bovine coronavirus nucleocapsid protein processing and assembly." Adv Exp Med Biol **380**: 259-263.

Hsieh, P. K., S. C. Chang, et al. (2005). "Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent." J Virol **79**(22): 13848-13855.

Huang, K. J., I. J. Su, et al. (2005). "An interferon-gamma-related cytokine storm in SARS patients." J Med Virol **75**(2): 185-194.

Hussain, R., A. Kaleem, et al. (2002). "Cytokine profiles using whole-blood assays can discriminate between tuberculosis patients and healthy endemic controls in a BCG-vaccinated population." J Immunol Methods **264**(1-2): 95-108.

Katze, M. G., Y. He, et al. (2002). "Viruses and interferon: a fight for supremacy." Nat Rev Immunol **2**(9): 675-687.

Koetz, A., P. Nilsson, et al. (2006). "Detection of human coronavirus NL63, human metapneumovirus and respiratory syncytial virus in children with respiratory tract infections in south-west Sweden." Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases **12**(11): 1089-1096.

Kopecky-Bromberg, S. A., L. Martinez-Sobrido, et al. (2007). "Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists." J Virol **81**(2): 548-557.

Kuypers, J., E. T. Martin, et al. (2007). "Clinical disease in children associated with newly described coronavirus subtypes." Pediatrics **119**(1): e70-76.

Langezaal, I., S. Coecke, et al. (2001). "Whole blood cytokine response as a measure of immunotoxicity." Toxicol In Vitro **15**(4-5): 313-318.

Leung, D. T., F. C. Tam, et al. (2004). "Antibody response of patients with severe acute respiratory syndrome (SARS) targets the viral nucleocapsid." J Infect Dis **190**(2): 379-386.

Liu, Z. H. and X. Sun (2008). "Coronavirus phylogeny based on base-base correlation." International journal of bioinformatics research and applications **4**(2): 211-220.

Luo, H., Q. Chen, et al. (2005). "The nucleocapsid protein of SARS coronavirus has a high binding affinity to the human cellular heterogeneous nuclear ribonucleoprotein A1." FEBS Lett **579**(12): 2623-2628.

McIntosh, K. (2005). "Coronaviruses in the limelight." J Infect Dis **191**(4): 489-491.

Meylan, E., J. Curran, et al. (2005). "Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus." Nature **437**(7062): 1167-1172.

Monto, A. S. and S. K. Lim (1974). "The Tecumseh study of respiratory illness. VI. Frequency of and relationship between outbreaks of coronavirus infection." J Infect Dis **129**(3): 271-276.

Muller, K., E. B. Herner, et al. (1998). "Inflammatory cytokines and cytokine antagonists in whole blood cultures of patients with systemic juvenile chronic arthritis." Br J Rheumatol **37**(5): 562-569.

Narayanan, K., C. J. Chen, et al. (2003). "Nucleocapsid-independent specific viral RNA packaging via viral envelope protein and viral RNA signal." J Virol **77**(5): 2922-2927.

Narayanan, K., A. Maeda, et al. (2000). "Characterization of the coronavirus M protein and nucleocapsid interaction in infected cells." J Virol **74**(17): 8127-8134.

Okabayashi, T., H. Kariwa, et al. (2006). "Cytokine regulation in SARS coronavirus infection compared to other respiratory virus infections." J Med Virol **78**(4): 417-424.

Oosterhof, L., C. B. Christensen, et al. (2010). "Fatal lower respiratory tract disease with human corona virus NL63 in an adult haematopoietic cell transplant recipient." <u>Bone Marrow Transplant</u> **45**(6): 1115-1116.

Papanicolaou, D. A., R. L. Wilder, et al. (1998). "The pathophysiologic roles of interleukin-6 in human disease." <u>Ann Intern Med</u> **128**(2): 127-137.

Parisien, J. P., J. F. Lau, et al. (2001). "The V protein of human parainfluenza virus 2 antagonizes type I interferon responses by destabilizing signal transducer and activator of transcription 2." <u>Virology</u> **283**(2): 230-239.

Park, M. S., M. L. Shaw, et al. (2003). "Newcastle disease virus (NDV)-based assay demonstrates interferon-antagonist activity for the NDV V protein and the Nipah virus V, W, and C proteins." <u>J Virol</u> **77**(2): 1501-1511.

Popa, C., P. Barrera, et al. (2009). "Cytokine production from stimulated whole blood cultures in rheumatoid arthritis patients treated with various TNF blocking agents." <u>Eur Cytokine Netw</u> **20**(2): 88-93.

Pyrc, K., B. Berkhout, et al. (2007). "The novel human coronaviruses NL63 and HKU1." <u>J Virol</u> **81**(7): 3051-3057.

Pyrc, K., R. Dijkman, et al. (2006). "Mosaic structure of human coronavirus NL63, one thousand years of evolution." <u>J Mol Biol</u> **364**(5): 964-973.

Pyrc, K., M. F. Jebbink, et al. (2004). "Genome structure and transcriptional regulation of human coronavirus NL63." <u>Virol J</u> **1**: 7.

Roberts, A., D. Deming, et al. (2007). "A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice." <u>PLoS Pathog</u> **3**(1): e5.

Rowland, R. R., V. Chauhan, et al. (2005). "Intracellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein: absence of nucleolar

accumulation during infection and after expression as a recombinant protein in vero cells." J Virol **79**(17): 11507-11512.

Sawicki, S. G., D. L. Sawicki, et al. (2007). "A contemporary view of coronavirus transcription." J Virol **81**(1): 20-29.

Stark, G. R., I. M. Kerr, et al. (1998). "How cells respond to interferons." Annu Rev Biochem **67**: 227-264.

Storni, T., T. M. Kundig, et al. (2005). "Immunity in response to particulate antigen-delivery systems." Adv Drug Deliv Rev **57**(3): 333-355.

Surjit, M. and S. K. Lal (2008). "The SARS-CoV nucleocapsid protein: a protein with multifarious activities." Infect Genet Evol **8**(4): 397-405.

Surjit, M., B. Liu, et al. (2004). "The nucleocapsid protein of the SARS coronavirus is capable of self-association through a C-terminal 209 amino acid interaction domain." Biochem Biophys Res Commun **317**(4): 1030-1036.

Symons, J. A., A. Alcami, et al. (1995). "Vaccinia virus encodes a soluble type I interferon receptor of novel structure and broad species specificity." Cell **81**(4): 551-560.

Takaoka, A. and H. Yanai (2006). "Interferon signalling network in innate defence." Cell Microbiol **8**(6): 907-922.

Talbot, H. K., B. E. Shepherd, et al. (2009). "The pediatric burden of human coronaviruses evaluated for twenty years." Pediatr Infect Dis J **28**(8): 682-687.

Theron, M., K. J. Huang, et al. (2005). "A probable role for IFN-gamma in the development of a lung immunopathology in SARS." Cytokine **32**(1): 30-38.

Thiel, V. and F. Weber (2008). "Interferon and cytokine responses to SARS-coronavirus infection." Cytokine Growth Factor Rev **19**(2): 121-132.

Timani, K. A., Q. Liao, et al. (2005). "Nuclear/nucleolar localization properties of C-terminal nucleocapsid protein of SARS coronavirus." Virus Res **114**(1-2): 23-34.

Tyrrell, D. A. and M. L. Bynoe (1965). "Cultivation of a Novel Type of Common-Cold Virus in Organ Cultures." Br Med J **1**(5448): 1467-1470.

Vabret, A., T. Mourez, et al. (2005). "Human coronavirus NL63, France." Emerg Infect Dis **11**(8): 1225-1229.

van der Hoek, L., K. Pyrc, et al. (2006). "Human coronavirus NL63, a new respiratory virus." FEMS Microbiol Rev **30**(5): 760-773.

van der Hoek, L., K. Pyrc, et al. (2004). "Identification of a new human coronavirus." Nat Med **10**(4): 368-373.

van der Hoek, L., K. Sure, et al. (2005). "Croup is associated with the novel coronavirus NL63." PLoS Med **2**(8): e240.

Viveros-Paredes, J. M., A. M. Puebla-Perez, et al. (2006). "Dysregulation of the Th1/Th2 cytokine profile is associated with immunosuppression induced by hypothalamic-pituitary-adrenal axis activation in mice." Int Immunopharmacol **6**(5): 774-781.

Woo, P. C., Y. Huang, et al. (2010). "Coronavirus genomics and bioinformatics analysis." Viruses **2**(8): 1804-1820.

Woo, P. C., S. K. Lau, et al. (2005). "Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia." J Virol **79**(2): 884-895.

Wu, P. S., L. Y. Chang, et al. (2008). "Clinical manifestations of human coronavirus NL63 infection in children in Taiwan." Eur J Pediatr **167**(1): 75-80.

Ye, Y., K. Hauns, et al. (2007). "Mouse hepatitis coronavirus A59 nucleocapsid protein is a type I interferon antagonist." J Virol **81**(6): 2554-2563.

You, J., B. K. Dove, et al. (2005). "Subcellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein." J Gen Virol **86**(Pt 12): 3303-3310.

Yu, S. Y., Y. W. Hu, et al. (2005). "Gene expression profiles in peripheral blood mononuclear cells of SARS patients." World J Gastroenterol **11**(32): 5037-5043.

Zhang, X., K. Wu, et al. (2007). "Nucleocapsid protein of SARS-CoV activates interleukin-6 expression through cellular transcription factor NF-kappaB." Virology **365**(2): 324-335.

Zhang, Y., J. Li, et al. (2004). "Analysis of serum cytokines in patients with severe acute respiratory syndrome." Infect Immun **72**(8): 4410-4415.

Zuniga, S., I. Sola, et al. (2007). "Coronavirus nucleocapsid protein is an RNA chaperone." Virology **357**(2): 215-227.

UNIVERSITY *of the*
WESTERN CAPE

Chapter IV

Thesis Synopsis

The human coronavirus NL63 was discovered in 2004 and shown to be a clinically important virus. HCoV-NL63 affects mainly children, the immunocompromised and the elderly. Infection with HCoV-NL63 may result in upper respiratory infections and can also present with complications like asthma exacerbation, febrile seizures and high fever.

The virus is assembled through the interaction of four main structural proteins, the spike (S), membrane (M), envelope (E) and the nucleocapsid (N). Of these structural proteins, the nucleocapsid protein is of vital importance in the formation of the ribonucleocapsid core via the interaction with viral RNA. Assembly occurs when the N protein undergoes N-N homotypic interactions and thereafter interacts with the viral RNA. The N protein has also been shown to interact with the M protein on the virus surface during the assembly process. In this study, molecular techniques were used to characterize the properties of the various HCoV-NL63 structural proteins.

Using viral RNA as a template, RT-PCR was used to generate cDNA of the HCoV-NL63 genome. The E, M, ORF3 as well as the hydrophilic regions of the ORF3 (ORF3ΔN) and M (MΔN) proteins were PCR amplified and ligated into the pGEM® T-easy vector for sequencing. Once sequences were confirmed, using M13 primers, the genes were removed from the vector by *SgfI* and *PmeI* enzyme restriction; these enzymes allow for cloning into the pFLEXI bacterial expression vector. The pFLEXI vector was used to transform KRX strain of competent *E. coli* cells which served as our expression model. SARS-CoV N, HCoV-NL63 N, as well as well as truncated mutants corresponding to the N- (N2) and C-terminal (N3) of the HCoV-NL63 N protein, were included for expression.

Transformed cells were grown in LB culture and protein expression was induced with the addition of rhamnose. Cultures were subsequently harvested by centrifugation and lysed by sonication. Protein detection was carried out by means of a Western Blot against the GST

affinity tag. E, M, ORF3, ORF3ΔN and MΔN HCoV-NL63 protein expression could not be optimally expressed as a result of several drawbacks encountered. A major one being high levels of background expression, which put the cells under unnecessary stress in the logarithmic phase of growth and thereby proving toxic to the cell. Therefore, in future studies, the expression conditions need to be further optimized to allow expression. Recommendations to improve protein expression include (1) the use of another expression vector system that may not prove toxic to *E. coli*; (2) change the GST tag to the newly used Halo-tag; studies have shown the latter to be more efficient in facilitating detection and solubility; (3) the use of Kozak sequences to initiate an effective protein translation process. However, the pFLEXI/KRX system proved to a good model in the expression of the HCoV-NL63 N proteins and its truncated clones. The use of lysozymes in the lysis buffer proved effective in increasing protein solubility from the membrane bound cell fraction allowing the HCoV-NL63 N1, N2 and N3 proteins to be detected.

These HCoV-NL63 N proteins were subsequently purified using a MagneGST™ purification system. Using whole blood cultures from three patients, IL-6, IL10 and IfN-γ cytokine levels stimulated by these recombinant N viral proteins was indentified. We found that the HCoV-NL63 N1, N2 and N3 proteins caused significant IL-6 levels in culture indicating that an inflammatory response is activated by these proteins. Therefore, the hypothesis that the HCoV-NL63 N protein is able to elicit an effective inflammatory response in a whole blood culture system was proven. The N1, N2 and N3 proteins did not stimulate a detectable IL-10 and IfN-γ response and thus, definitive results concerning these cytokines would require further study. Both the N- and C-terminal of the N protein, however, equally stimulated an IL-6 response, indicating the presence of several immunogenic peptides dispersed along the entire length of the HCoV-NL63 N-protein. It could also be indicative of the division of one immunogenic peptide between the N- and C-terminal. Inclusion of the

middle disordered region would therefore aid in characterizing the immunogenicity of the N-protein. Future studies will include identifying the regions involved in dimerization and RNA binding, production of antibodies in a murine model and developing an antibody capture ELISA.

Of particular interest, since the conclusion of this thesis, a new coronavirus was discovered in the Middle East earlier in September 2012. Scientist first identified the new virus HCoV-EMC/2012 in a 60-year-old man suffering from acute pneumonia. The virus has since resulted in three fatalities, even though it has been shown to be closely related to viruses in Asian bats. This served as reminder to health authorities around the world of the threat posed by coronaviruses. Since the SARS epidemic in 2002, it has become evident that highly pathogenic coronavirus strains can still evolve and thus ongoing research is vital in understanding coronaviruses.

Appendix

APPENDIX I: Vectors and Sequences

Figure 1: This diagram represents the vectors that were used to clone the genes of interest. The figure marked "A" indicates PGEM-T Easy cloning vector.

pGEM®-T easy vector was chosen because of its convenience for cloning PCR products acting as a shuttle vector to allow for confirmation via sequencing. The

figure marked "B", on the other hand, represents the GST tagged bacterial expression vector that is appended by *Sgf1* and *Pmel* restriction sites. The T7

promoter upstream controls the expression of the DNA insert.

Figure 2: Sequence verification of the cloned HCoV-NL63-E gene in pGEM. Cloned genes were sequenced with M13 forward and reverse primers. M13 primer binding sites are located up- and down-stream of the multiple cloning site allowing for easy sequencing (This data applies to figures 2-5). E1&2_F_rev4C represents the gene strand sequenced with M13 forward primers, where the E1&2_R represents the gene strand sequenced with M13 reverse primers. The label original is the sequence of the HCoV-NL E gene obtained from NCBI with accession number AY697421

Figure 3: Sequence verification of the truncated clone of HCoV-NL63 ORF3ΔN-gene in pGEM. The first HALF_ORF3_ represents the gene strand sequenced with M13 reverse primers, whereas the second HALF_ORF3_R below represents the gene strand sequenced with M13 forward primers. NL63-ORF3 is the sequence of the HCoV-NL63 ORF 3gene obtained from NCBI with accession number AY697419.

Figure 4: Sequence verification of the truncated clone of HCoV-NL63 MΔN-gene in pGEM. The first HALF_M_For represents the gene strand sequenced with M13 forward primers, whereas the second HALF_M_Rev represents the gene strand sequenced with M13 reverse primers. NL63-M_AY6 is the sequence of the HCoV-NL63 M-gene obtained from NCBI with accession number AY697422

Figure 5: Sequence verification of the cloned full length HCoV-NL63 ORF3-gene in pGEM. The ORF3-5Cons represents the gene strand sequenced with M13 forward primers, whereas the ORF3_1Cons represents the gene strand sequenced with M13 reverse primers. The middle sequence labelled original is the sequence of the HCoV-NL63 ORF3-gene obtained from NCBI with accession number AY567487.2

APPENDIX II: Results of protein optimisation protocols using the pFLEXI vector system

Induction at specific OD600 reading with fermentation at 37°C

As mentioned in section 2.3.3.2 of Chapter II glycerol stocks were inoculated into 10ml of LB medium and incubated for 14 hours at 37°C with shaking at 180rpm. Starter cultures were diluted 1:100 into 100ml LB medium and incubated for approximately 4 hours when OD600 readings were taken. Once the culture reached an OD600 of 0.3-0.4, expression was induced by the addition of 0.1% (w/v) rhamnose. Cultures were incubated as previously described for 24 hours when cells were harvested by centrifugation at 6000rpm for 15min. Detection was carried out using SDS-page gel and Western blot. The observed molecular weights of the recombinant proteins appended with an N-terminal GST tag were predicted to be MΔN: ~42kDa; ORF3ΔN: ~42kDa; M: ~72kDa; ORF3 ~72kDa and E: ~36kDa. Figure six A and B below represents the expression results obtained for this protocol. Expression was undetectable and thus we further modified the expression protocol.

Figure 6: (A) Coomassie stain, (B) Western Blot detection of GST fusion viral proteins. Lane 1: E; Lane 2: M; Lane 3: MΔN, Lane 4: ORF 3Lane 5: ORF3ΔN; Lane 6: Prestained molecular weight protein marker; Lane 7: uninduced control. Lane order has been maintained for Figure 6B.
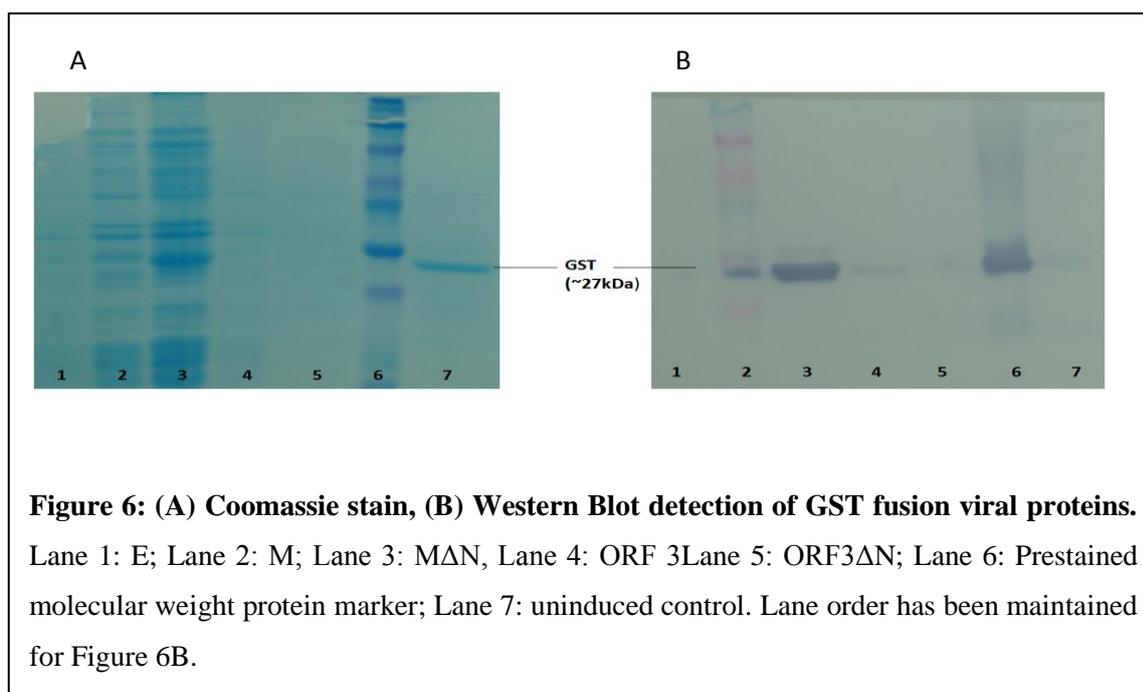
Large scale expression

The inability to detect protein expression and a reduced cell density a large scale culture was necessary to obtain an ideal concentration for protein purification. The expression protocols described previously was maintained where starter cultures were inoculated in 1500ml LB medium with the addition of 1mM IPTG on induction. A small level of protein expression was seen detected for the M protein (Figure 7) however; upon repetition of this protocol, expression was no longer detected.


Figure 7: HCoV-NL63 proteins by Western Blot with peroxidase labelled antibodies against the GST fusion protein. Lane 1: E; Lane 2: M; Lane 3: MΔN, Lane 4: ORF 3, Lane 5: Prestained molecular weight protein marker; Lane 6: ORF3ΔN; Lane 7: uninduced control.

APPENDIX III: Cloning into the pGEX-4T-2 expression vector for further studies using the IPTG inducible system.

*Cloning of the HCoV-NL63 E, M and N genes into the pGEX-4T-2 expression vector*

The pGEX-4T-2 vector was used to enable the expression of the HCoV-NL63 E, M and N proteins outside the scope of this thesis. Protein expression from a pGEXplasmid is under the control of the *tac* promoter, which is induced using the lactose analog isopropyl b-D-thiogalactoside (IPTG).

The HCoV-NL63 E, M and N genes were PCR amplified (Figures 8, 9 and 10) using pCMV primers (not shown). The forward and reverse primers contained an EcoR1 and BamH1 respectively to enable ligation into the "cut"pGEX-4T-2 vector.

Figure 8: PCR amplification of HCoV-NL63 E viral gene using pCMV primers. Lane M: 100BP marker; Lane 1-4: HCoV-NL63 E; Lane 5: Blank; Lane6: Negative control using all pCMV primers.

Figure 9: PCR amplification of HCoV-NL63 M viral gene using pCMV primers.
Lane M: 100BP marker; Lane 1: HCoV-NL63 M-gene; Lane 2: Blank; Lane 3&4:
Negative control using all pCMV primers.

Figure 10: PCR amplification of HCoV-NL63 N viral gene using pCMV primers.
Lane M: 100BP marker; Lane 1: Blank; Lane 2, 3&7: Negative control using all pCMV
primers; lanes 5&6 HCoV-NL63 N-gene.

The pGEX-4T-2 vector was linearised by digestion with EcoR1 and BamH1 (result not
shown) Following restriction digest, 20ul of the linearised plasmid was treated with calf

intestinal phosphatase (CIAP) to remove the 5' phosphates groups from the DNA. This will prevent he pGEX-4T-2 vector from self re-ligation. The HCoV-NL63 E, M and N genes were subsequently ligated BamH1 and EcoR1 digested and visualized on a 1% agarose gel to confirm the presence of the gene of interest (Figure 11). Thereafter 50ng of insert was ligated into the cut and CIAP treated pGEX-4T-2 plasmid. 10 ul of the ligation reaction was transformed into competent BL21 *E.coli*. Transformants were selected on LB/agar/ampicillin plates from which colonies were picked and grown in 10ml of LB cultures. The colonies forming units (CFU's) were PCR screened (Figures 12, 13 and 14) for the presence of the correct HCoV-NL63 E, M and N inserts using the previously mentioned pCMV primers. Glycerol stocks were prepared if these bacterial clones for further use. To verify the cloned inserts were indeed present, plasmid DNA isolated was purified and sent for sequencing (results not shown).
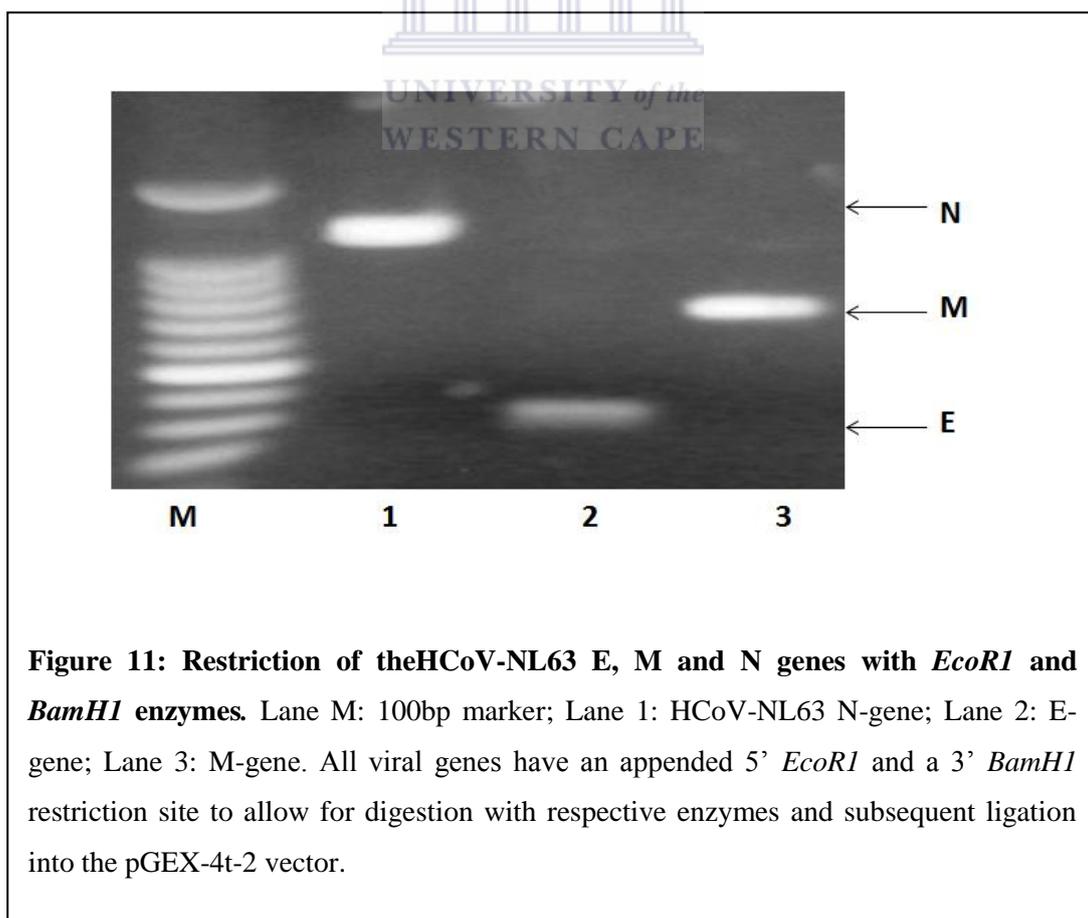

Figure 11: Restriction of theHCoV-NL63 E, M and N genes with *EcoR1* and *BamH1* enzymes. Lane M: 100bp marker; Lane 1: HCoV-NL63 N-gene; Lane 2: E-gene; Lane 3: M-gene. All viral genes have an appended 5' *EcoR1* and a 3' *BamH1* restriction site to allow for digestion with respective enzymes and subsequent ligation into the pGEX-4t-2 vector.
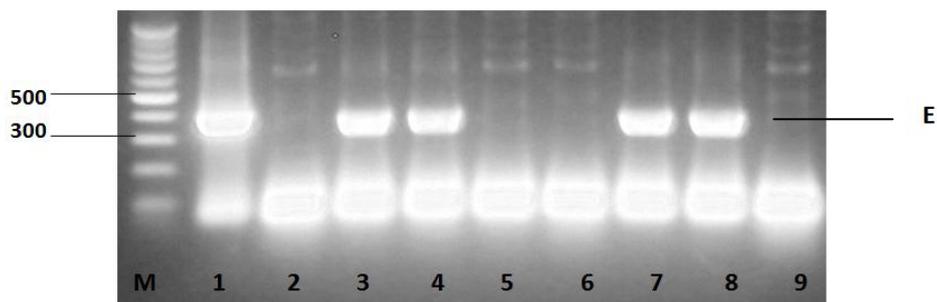
Figure 12: Colony PCR of BL21 competent *E. coli* transformed with pGEX-4T-2 vector ligated with HCoV-NL63 E viral genes. Colony PCR was performed with pCMV primers and used as confirmation of successful transformation and ligation. Lane M: 100bp marker; Lane 1-9: E gene.
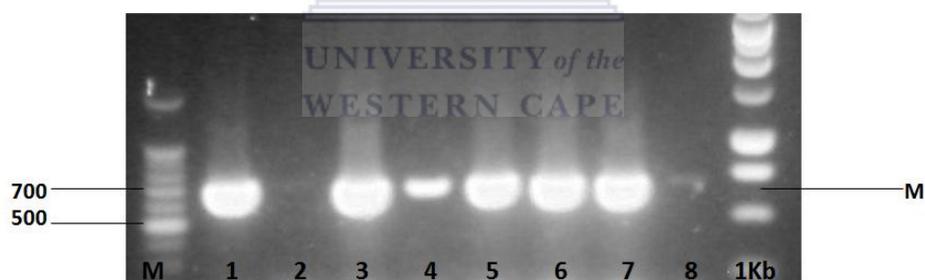
Figure 13: Colony PCR of BL21 competent *E. coli* transformed with pGEX-4T-2 vector ligated with HCoV-NL63 M viral genes. Colony PCR was performed with pCMV primers and used as confirmation of successful transformation and ligation. Lane M: 100bp marker; Lane 1,3-8: M gene; Lane 2: Blank.
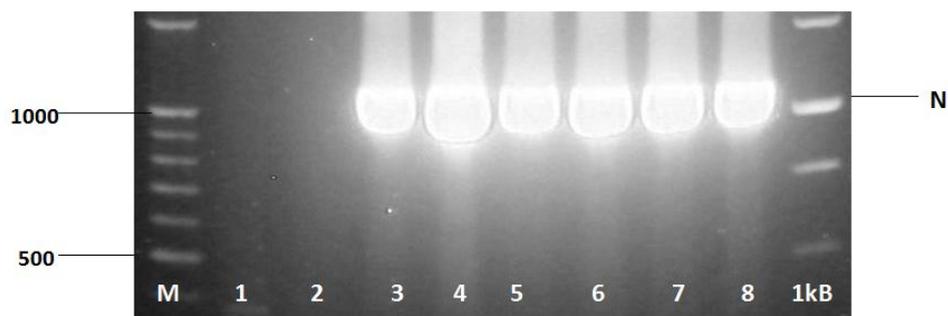
Figure 14: Colony PCR of BL21 competent *E. coli* transformed with pGEX-4T-2 vector ligated with HCoV-NL63 N viral genes. Colony PCR was performed with pCMV primers and used as confirmation of successful transformation and ligation. Lane M: 100bp marker; Lane 1&2:- Blank;3-8: N gene.