# The Generation of Antibodies against Human Coronavirus NL63 (HCoV-NL63) Nucleocapsid Protein

by

**Marjorie Smith**

**Thesis submitted in partial fulfilment for the degree**

**MSc**

**Department of 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

**15 March 2017**

# DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis, "*The generation of antibodies against Human Coronavirus NL63 nucleocapsid protein"*, is the result of my own work and has not been previously submitted, in part or in its entirety, to any other university or institution. All the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

Full name: _____

Date: _____



Signature: _____

# DEDICATION

*This thesis work is dedicated to my husband, Shane, who has given me constant encouragement during the challenges of life as a science student and wife.*

*I am truly thankful for having you in my life.*

*This work is also dedicated to my parents, Kobus and Elma van Zyl, whose good examples have taught me to work hard for the things that I aspire to achieve.*

# ACKNOWLEDGEMENTS


- ✧ I would first like to express my gratitude to my supervisor, Professor Burtram C. Fielding, for his insight, constant motivation when the going got tough and his patience and guidance to get the job done.

- ✧ I would also like to thank Professor Edmund Pool for his collaboration on research conducted.

- ✧ I sincerely thank the members of the Molecular Biology Virology Laboratory, with special recognition to the senior students, namely Aasiyah, Michael, Yanga, Tiza and Taryn for their mentoring and support. To the remaining students for your help rendered, sharing of equipment and reagents and the discussions about my work.

- ✧ Thank you to my funders, the National Research Foundation (NRF), the Poliomyelitis Research Fund (PRF) and the Ernst and Ethel Eriksen Trust (EEET) for personal and research funds which made this MSc possible.

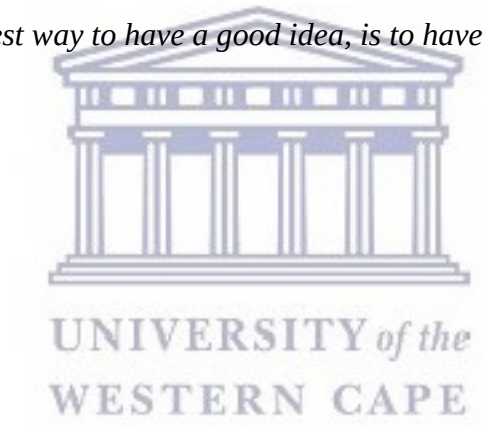
- ✧ A heart-felt thank you to my parents, brother, family and my dearest friends (Janske Nel, Aasiyah Chafekar and Beynon Abrahams) for your constant encouragement, support and love; getting me through late nights in the lab and never stop believing in me.

- ✧ To my loving husband, Shane Smith, I can never thank you enough for your understanding, patience and taking such good care of me!

- ✧ So many years so many people had a hand in the final product. I acknowledge and thank you.

*The best way to have a good idea, is to have lots of ideas!*

~Linus Pauling



# ABSTRACT

Coronaviruses (CoVs) are enveloped viruses composed of single-stranded, positive sense RNA viruses with the largest viral genomes among RNA viruses (27-33 kb). In 2004, human coronavirus NL63 (HCoV-NL63) was discovered by researchers from The Netherlands.

Since its discovery, it has been found in countries all over the world and affects mainly young children between ages 0-5 years old, immunocompromised and elderly people. HCoV-NL63 has been indicated to cause both mild upper and more serious, but less common, lower respiratory tract infections and are commonly associated with other respiratory viruses in co-infections that can increase the severity of HCoV-NL63 infection. The nucleocapsid (N) protein of HCoV-NL63 is a multifunctional phosphoprotein that modulates the assembly of the ribonucleocapsid core of mature virions. During infection, the N protein is expressed at high levels within an infected cell and elicits a strong immune response in infected patients. For this reason, N protein was used to generate antibodies against an immune response to the presence of HCoV-NL63 N protein in mouse blood serum.

In this study, recombinant plasmid constructs of SARS-CoV N and HCoV-NL63 N protein and its truncates, were used. The presence of the gene insert in the Flexi vector was confirmed by restriction endonuclease digest. The recombinant proteins were expressed using the KRX *E. coli* bacterial expression system. The successful expression of these recombinant proteins was confirmed by Western Blot analysis against the GST tag which also facilitated purification of the fusion proteins.

Peptides were designed and synthesized from two regions in the HCoV-NL63 N gene sequence that were identified as antigenic determinants. Polyclonal antibodies were

successfully produced and validated using an indirect ELISA. Splenic B-lymphocytes were fused with Sp2 myeloma cells to form hybridomas. Hybridoma colonies were selected, cloned and screened for the production of monoclonal antibodies by indirect ELISA. The purified GST-tagged viral fusion N proteins of SARS-CoV and HCoV-NL63 were subsequently used to characterize the immune activity of the produced antibodies by Western Blotting.

Results showed that the synthesized peptides could elicit an immune response and produce polyclonal antibodies to HCoV-NL63 N-specific peptides. Polyclonal serum anti-N antibodies detected SARS-CoV N and HCoV-NL63 N protein and it's truncates, but cross-reactivity between SARS-N and HCoV-NL63 N was observed. Due to time constraints, no monoclonal antibodies were produced successfully that would have been able to specifically detect only HCoV-NL63 N protein. Further studies will need to be conducted to gain better understanding of cross-reactivity between HCoV-NL63 and other human coronaviruses.

# KEY WORDS

Human Coronavirus NL63, Nucleocapsid Protein, Protein Expression, GST Purification, Miniprep, Midiprep, Polyclonal Antibodies, ELISA, Peptides

# TABLE OF CONTENTS

# LIST OF ABBREVIATIONS

| | |
|---|---|
| Δ | Delta |
| ® | Registered |
| ™ | Trademark |
| ~ | Approximately |
| < | Less than |
| > | More than |
| °C | Degrees Celsius |
| µg | Microgram |
| µl | Microliter |
| µM | Micromolar |
| 3’ | 3 prime |
| 5’ | 5 prime |
| aa | Amino acid |
| ACE-2 | Angiotensin converting enzyme 2 |
| AdV | adenovirus |
| APS | Ammoniumpersulfate |
| ARTIs | Acute respiratory tract infections |
| BM | Biomass |
| bp | Basepair |
| BR | Broad range |
| BSA | Bovine serum albumin |
| C | Concentration |
| cAMP | Cyclic adenosine monophosphate |

| | |
|---|---|
| cDNA-AFLP | Complementary DNA amplified fragment-length polymorphism |
| CFA | Complete Freund's adjuvant |
| $CO_2$ | Carbon dioxide |
| CoV(s) | Coronavirus(es) |
| CTD | Carboxy terminal domain |
| C-terminal | Carboxy terminal |
| DMEM | Dulbecco's Modified Eagle's Medium |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxyribose nucleoside triphosphate |
| dsDNA | Double-stranded DNA |
| DTT | dithiothreitol |
| E | Envelope |
| *E. coli* | *Escherichia coli* |
| EDTA | Ethylenediaminetetraacetic acid |
| EIA | Enzyme immunoassay |
| ELISA | Enzyme-linked immunosorbent assay |
| ER | Endoplasmic reticulum |
| ERGIC | Endoplasmic reticulum Golgi-intermediate compartment |
| FBS | Fetal bovine serum |
| flu | Influenza virus |
| g | Gram |
| GSH | Glutathione |
| GST | Glutathione-S-Transferase |
| $H_2SO_4$ | Sulfuric acid |

| | |
|---|---|
| HAT | Hypoxanthine aminopterim thymidine |
| hBoV | Human bocavirus |
| HCl | Hydrochloric acid |
| HCoV | Human coronavirus |
| HCoV-229E | Human coronavirus 229E |
| HCoV-EMC | Erasmus Medical College human coronavirus |
| HCoV-HKU1 | Human coronavirus HKU1 |
| HCoV-NH | New Haven coronavirus |
| HCoV-NL63 | Human coronavirus NL63 |
| HCoV-NL63 | Human coronavirus NL63 |
| HCoV-OC43 | Human coronavirus OC43 |
| HI-FBS | Heat-inactivated foetal bovine serum |
| hMPV | Human metapneumovirus |
| HRP | Horseradish peroxidase |
| hRV | Human rhinovirus |
| HT | Hypoxanthine thymidine |
| IBV | Infectious bronchitis virus |
| ICTV | International Committee of Taxonomy of Viruses |
| IFA | Incomplete Freund's adjuvant |
| In vitro | Outside a living organism |
| In vivo | In a living organism |
| KAc | Potassium acetate |
| Kb | Kilobase |
| KCl | Potassium chloride |
| KD | Kawasaki disease |

| | |
|---|---|
| kDa | Kilodalton |
| KLH | Keyhole limpit hemocyanin |
| L | Litre |
| LB | Luria Bertani |
| M | Molar |
| M | Membrane |
| mA | Milliamps |
| MAP | Multiple antigenic peptides |
| MBS | *m*-maleimidobenzoyl-*N*-hydroxysuccinimide ester |
| MERS-CoV | Middle eastern respiratory syndrome coronavirus |
| mg | Milligram |
| $MgCl_2$ | Magnesium chloride |
| MHV | Mouse hepatitis virus |
| Min(s) | Minute(s) |
| ml | millilitre |
| mM | Millimolar |
| mRNA | Messenger RNA |
| *n* | Any number |
| N | Nucleocapsid |
| N/A | Not applicable |
| $Na_2HPO_4$ | Disodium phosphate |
| NaCl | Sodium chloride |
| NaOH | Sodium hydroxide |
| NCBI | National Bioinformatics Institute |
| NLS | Nuclear localization signal |

| | |
|---|---|
| NTD | amino-terminal domain |
| N-terminal | Amino terminal |
| OA | Ovalbumin |
| OD | Optical density |
| ORFs | Open reading frames |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| PEG | Polyethylene glycol |
| pFN2A | Flexi vector |
| pH | Measure of acidity |
| PIV | Para-influenza virus |
| PVDF | Polyvinylidine difluoride |
| QF | Qubit fluorometer |
| RFS | Ribosomal frameshift |
| RNA | ribonucleic acid |
| rpm | Revolutions per minute |
| RPMI | Roswell Park Memorial Institute medium |
| RSV | Respiratory syncytial virus |
| S | Spike |
| SARS-CoV | Severe acute respiratory syndrome coronavirus |
| SDS | Sodium dodecyl sulphide |
| SDS-PAGE | Sodium dodecyl sulphide polyacrylamide gel electrophoresis |
| sg mRNA | Sub-genomic messenger RNA |
| SR domains | serine and arginine amino acid residues |

| | |
|---|---|
| ssDNA | Single-stranded DNA |
| ssRNA | Single-stranded RNA |
| TBE | Tris-Borate-EDTA |
| TE | Tris-EDTA |
| TEMED | Tetramethylethelenediamine |
| TEV | Tobacco etch Virus |
| TGEV | Transmissible gastroenteritis coronavirus |
| TMB | Tetramethylbenzidine |
| TRS | Transcription regulatory sequence |
| u | Units |
| URTIs | Upper respiratory tract infections |
| UTR | Untranslated region |
| UV | Ultraviolet |
| V | Volts |
| v/v | Volume-to-volume |
| VIDISCA | Virus-discovery-cDNA-AFLP |
| VLPs | Virion-like particles |
| w/v | Weight-to-volume |
| WHO | World health organization |

# LIST OF PUBLICATIONS

McBride, R., van Zyl, M., et al. (2014). "The coronavirus nucleocapsid is a multi-functional protein." Viruses 6(8): 2991-3018.

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF APPENDIXES

# CHAPTER 1

# INTRODUCTION


Department of Medical Biosciences, Faculty of Natural Sciences, University of the Western Cape, South Africa

# CHAPTER 1 INTRODUCTION

## 1.1 Discovery of Human Coronaviruses

Coronaviruses (CoVs) are enveloped viruses belonging to the *Coronaviridae* family in the order *Nidovirales,* which are composed of single-stranded, positive sense RNA viruses with the largest viral genome among RNA viruses (27-33 kb) (van der Hoek, L., Pyrc, K., et al. 2004). Previously, CoVs were phylogenetically divided into three distinct groups: Groups I and II viruses mainly have mammalian hosts whereas group III viruses are found exclusively in birds (van der Hoek, L., Pyrc, K., et al. 2006; Pyrc, K., Berkhout, B., et al. 2007b). Recently, the International Committee for Taxonomy of Viruses (ICTV) have replaced these groups by four genera, namely *Alpha-*, *Beta-*, *Gamma- and Deltacoronaviruses* (ICTV 2011).

Acute respiratory tract illnesses (ARTIs) are of the most common causes of disease in humans and rank among the top three causes of death in children under five years of age (Abdul-Rasool, S. and Fielding, B.C. 2010; Jevsnik, M., Ursic, T., et al. 2012). ARTIs have a great impact and economic burden on our health care system and effects people of all ages (Cabeca, T.K. and Bellei, N. 2012). Approximately 80% of these ARTIs are viral. Some of the viruses responsible for respiratory infections include respiratory syncytial virus (RSV), parainfluenza virus (PIV), influenza virus (flu), adenoviruses (AdV), human rhinoviruses (hRV) and enteroviruses. Viruses that are less commonly responsible for ARTIs are human metapneumovirus (hMPV), human bocavirus (hBoV) and human coronavirus (HCoV) (Jevsnik, M., Ursic, T., et al. 2012; Zhang, G., Hu, Y., et al. 2012). In humans, CoVs are commonly associated with upper

respiratory tract illnesses (URTIs) which are known to manifest as common cold symptoms.

The first two HCoVs, HCoV-229E and OC43, were discovered in the early 1960's and were thought to be relatively harmless pathogens resulting in the common cold (Tyrrell, D.A. and Bynoe, M.L. 1965; Almeida, J.D. and Tyrrell, D.A. 1967; McIntosh, K., Dees, J.H., et al. 1967). The outbreak of atypical pneumonia late in 2002 in China lead to a high mortality rate in infected people and was termed severe acute respiratory syndrome resulting from infection with the causative agent severe acute respiratory syndrome coronavirus (SARS-CoV) (Drosten, C., Gunther, S., et al. 2003; Fouchier, R.A., Kuiken, T., et al. 2003; Ksiazek, T.G., Erdman, D., et al. 2003). The outbreak of SARS-CoV sparked new interest in the detection and characterization of CoVs after subsequent assessments determined that the SARS-CoV crossed to human hosts most likely from zoonotic reservoirs. This cross-species transmission raised the question of the likelihood of other pathogenic HCoVs (Graham, R.L. and Baric, R.S. 2010). In 2004, HCoV-NL63 was identified as a new group I CoV (now *Alphacoronavirus*), closely related to HCoV-229E based on analyses of the complete genome sequence (van der Hoek, L., Pyrc, K., et al. 2004). In January 2005, a new group II virus was identified in patients from Hong Kong presenting with pneumonia, namely HCoV-HKU1 (Woo, P.C., Lau, S.K., et al. 2005a). Then, in 2012 a new novel HCoV, Middle East Respiratory Syndrome coronavirus (MERS-CoV) (previously termed human coronavirus Erasmus Medical College (HCoV-EMC)), was identified. This novel CoV was sequenced and found to be closely related to a cluster of bat viruses in the *betacoronavirus* genera together with SARS-CoV (Corman, V.M., Eckerle, I., et al. 2012; Zaki, A.M., van Boheemen, S., et al. 2012). Latest figures

reported by the World Health Organisation (WHO) indicate that as of June 2015, there were 1338 laboratory-confirmed cases of MERS-CoV infections reported, including at least 475 deaths (WHO, 2015).

## 1.2 Human Coronavirus NL63

HCoV-NL63 was first isolated from the nasopharyngeal aspirate from a 7-month-old child with bronchiolitis in January 2003. Diagnostic tests for several respiratory viruses yielded negative results. A modified cDNA amplified fragment length polymorphism (cDNA-AFLP) technique (Virus-Discovery-cDNA-AFLP or VIDISCA) was then developed to clone and amplify unknown viral genomes. Results showed similarities to the complete genome of other CoVs. Sequence divergence from known CoVs allowed for a novel CoV to be identified by Lia van der Hoek *et al*. (2004) in Amsterdam, who named the CoV human coronavirus NL63 (HCoV-NL63) (van der Hoek, L., Pyrc, K., et al. 2004). During the same period, Fouchier *et al* (2005) isolated a previously unrecognized group I CoV from an 8-month-old boy with pneumonia and found the virus to be similar to HCoV-NL63, naming it HCoV-NL (Fouchier, R.A., Hartwig, N.G., et al. 2004). Also in 2005, Esper *et al* (2005) identified the genomic sequences of another novel HCoV in New Haven, Connecticut and named it "New Haven Coronavirus" (HCoV-NH). Subsequent sequence comparisons with HCoV-NL and HCoV-NL63 revealed these viruses to represent the same species (Esper, F., Weibel, C., et al. 2005). The virus is known today only as HCoV-NL63.

## 1.3 Pathogenesis and Clinical Features of HCoV-NL63

As CoVs evolve over time, so too will its pathogenicity. The high prevalence and high incidence of co-infection among HCoV-NL63, makes it more susceptible for

recombination (Pyrc, K., Dijkman, R., et al. 2006). In a recent study done in Colorado, USA, three distinct genotypes for HCoV-NL63 were identified: Genotype A predominantly circulated in the year of 2005 and genotypes B and C were predominant during 2009 and 2010. This may suggest yearly antigenic variation of HCoV-NL63 infection (Dominguez, S.R., Sims, G.E., et al. 2012). The virus can remain infective at room temperature and survive for up to seven days in aqueous solution and respiratory secretions, making direct person-to-person transmission still the major route of HCoV-NL63 spread (Hofmann, H., Pyrc, K., et al. 2005; Muller, A., Tillmann, R.L., et al. 2008).

### 1.3.1 Co-infections

HCoV-NL63 is commonly associated with other respiratory viruses in co-infections such as with influenza A virus H3N2 (Chiu, S.S., Chan, K.H., et al. 2005), RSV-A (Arden, K.E., Nissen, M.D., et al. 2005; Esper, F., Weibel, C., et al. 2005; van der Hoek, L., Sure, K., et al. 2005), parainfluenza-3 (Arden, K.E., Nissen, M.D., et al. 2005; van der Hoek, L., Sure, K., et al. 2005) or hMPV (Arden, K.E., Nissen, M.D., et al. 2005; Esper, F., Shapiro, E.D., et al. 2005). Co-infected patients are more likely to be hospitalized due to the severity that arises from this kind of super-infection (Abdul-Rasool, S. and Fielding, B.C. 2010). In a study conducted in Germany, the most common co-infection detected in children younger than three years of age was RSV-A and HCoV-NL63 (van der Hoek, L., Sure, K., et al. 2005). In South Africa, co-infection of HCoV-NL63 with bocavirus has also been reported (Smuts, H. and Hardie, D. 2006). Infection with HCoV-NL63 is usually more severe where multiple viruses are involved (Mackay, I.M., Arden, K.E., et al. 2012).

Patients with a co-infection have a lower viral load of HCoV-NL63 possibly due to (i) the weakened immune system that allow for secondary infection to occur; (ii) competition for the same receptor; (iii) activation of an innate immune response to inhibit HCoV-NL63 replication or (iv) prolonged persistence of low levels HCoV-NL63 (van der Hoek, L., Sure, K., et al. 2005; Abdul-Rasool, S. and Fielding, B.C. 2010; Fielding, B.C. 2011).

Even though co-infection can increase the severity of HCoV-NL63 infection, some studies have reported that this kind of patient rarely had underlying conditions that could compromise their immune systems and make them more susceptible to a co-infection with another virus as one would believe (Mackay, I.M., Arden, K.E., et al. 2012; Xin, C., Yong, Z.Z., et al. 2012). The ability of CoV to establish infection is not affected by the CoV load in which co-pathogens were detected and shouldn't be interpreted as an incidental infection that doesn't contribute to disease (Gaunt, E.R., Hardie, A., et al. 2010) and further study would enable better assessment of the association between HCoV infection and severe illness (Dare, R.K., Fry, A.M., et al. 2007).



### 1.3.2 Clinical Features

It has been shown by several clinical studies on HCoV-NL63 that the virus is found in both upper and lower respiratory tract infections, not differing much from symptoms associated with other "old" CoVs such as HCoV-OC43 and HCoV-229E. The subject of HCoV-NL63 being associated with other body systems is still very much controversial (Vabret, A., Dina, J., et al. 2009). Symptoms associated with HCoV-NL63 infections can include mild symptoms including sore throat, rhinitis and cough,

to more severe asthma exacerbation, febrile seizures, high fever (Chiu, S.S., Chan, K.H., et al. 2005), pneumonia, bronchiolitis and croup (Arden, K.E., Nissen, M.D., et al. 2005; van der Hoek, L., Sure, K., et al. 2005). HCoV-NL63 has since been identified in clinical specimens from infants younger than six years of age, adults of frail age and immunocompromised individuals (Arden, K.E., Nissen, M.D., et al. 2005; van der Hoek, L., Sure, K., et al. 2005; van der Hoek, L., Pyrc, K., et al. 2006).

One should keep in mind that reports of symptoms in young children is primarily based on parental observations and may lead to other symptoms being missed. Most of the clinical and epidemiological studies done on HCoV-NL63 have several shortcomings, such as small sample size and short periods of assessment which can lead to underestimation of the true effect of the virus on the population (Fielding, B.C. 2011). More comprehensive studies will most probably reveal associations of HCoV-NL63 with other body systems that is most often over-looked due to co-infections or misdiagnosis of HCoV-NL63 infection.



Three fatal cases have been reported where HCoV-NL63 was identified as the aetiological agent: a 42-year old male transplant patient (Oosterhof, L., Christensen, C.B., et al. 2010), a 46-year old female with diabetes (Cabeca, T.K. and Bellei, N. 2012) and a 92-year old male (Bastien, N., Anderson, K., et al. 2005). These fatal outcomes were reported in adults and not children as depicted in majority of publications where HCoV-NL63 infection have been identified (Cabeca, T.K. and Bellei, N. 2012). The occurrence of HCoV-NL63 infection among adults has not been fully documented.

### 1.3.3 Association with Other Diseases

Croup, also known as laryngotracheobronchitis, is the swelling around the vocal cords and is most commonly diagnosed in young children (Kaneshiro, N.K. and Zieve, D. 2012). Children infected with HCoV-NL63 have a 6.6% higher chance of developing croup than children not infected with the virus. A study conducted in Germany, reported that 17.4% of patients younger than three years who were diagnosed with croup, were infected with HCoV-NL63, thus establishing a clear link between HCoV-NL63 and croup (van der Hoek, L., Sure, K., et al. 2005). The etiological agent has been believed to be mainly PIV or RSV, but is it now abundantly clear that HCoV-NL63 also plays a major role in this disease (Kaneshiro, N.K. and Zieve, D. 2012).



There are compelling evidence from one study that suggests an association between HCoV-NL63 and Kawasaki disease (KD), a form of early childhood systemic vasculitis that presents as prolonged fever, polymorphic exanthema, oropharyngeal erythema and bilateral conjunctivitis (Esper, F., Shapiro, E.D., et al. 2005). KD was first described by Kawasaki in 1967 as an acute systemic febrile illness of unknown etiology that predominantly affects children of approximately five years of age (Chang, L.Y., Chiang, B.L., et al. 2006). The problem with the association with KD is that no subsequent significant statistical link could be established when this association was tested in other countries including Taiwan (Belay, E.D., Erdman, D.D., et al. 2005), Japan (Ebihara, T., Endo, R., et al. 2005) and California, and therefore controversy exists as to whether there is an association between HCoV-NL63 infection and KD (McIntosh, K. 2005; Shimizu, C., Shike, H., et al. 2005;

Chang, L.Y., Chiang, B.L., et al. 2006; Dominguez, S.R., Anderson, M.S., et al. 2006; Kahn, J.S. 2006; Lehmann, C., Klar, R., et al. 2009).

In several studies it states that there is a link between HCoV-NL63 and gastrointestinal findings (Moes, E., Vijgen, L., et al. 2005; Vabret, A., Mourez, T., et al. 2005; Leung, T.F., Li, C.Y., et al. 2009), which appear to be a direct consequence of invasion of the intestinal mucosa by the virus (Principi, N., Bosis, S., et al. 2010). Even though HCoV-NL63 could be detected in four of 878 stool samples from children with acute gastric illness in one study, it was also reported to be absent in another study where stool samples have been collected from 479 patients with gastroenteritis (Esper, F., Ou, Z., et al. 2010; Risku, M., Lappalainen, S., et al. 2010). However, of the patients that tested positive for HCoVs, also tested positive for other gastroenteritis viruses (Risku, M., Lappalainen, S., et al. 2010), but this doesn't mean HCoV-NL63 should be excluded as the causative agent. Identifying the exact role of HCoV-NL63 in gastrointestinal illnesses is still not clear since many HCoV-NL63 patients can present with co-infections making determination of the etiological agent a difficult task (Fielding, B.C. 2011).



## 1.4 Distribution and Prevalence

A worldwide distribution of this virus, after its isolation in The Netherlands (van der Hoek, L., Pyrc, K., et al. 2004), was indicated by its subsequent identification in various countries including China (Chiu, S.S., Chan, K.H., et al. 2005), Japan (Ebihara, T., Endo, R., et al. 2005), Thailand (Dare, R.K., Fry, A.M., et al. 2007), Taiwan (Wu, P.S., Chang, L.Y., et al. 2008), Australia (Arden, K.E., Nissen, M.D., et al. 2005), Canada (Bastien, N., Anderson, K., et al. 2005), Belgium (Moes, E., Vijgen,

L., et al. 2005), France (Vabret, A., Dina, J., et al. 2006), Italy (Canducci, F., Debiaggi, M., et al. 2008), Germany (Forster, J., Ihorst, G., et al. 2004), Sweden (Koetz, A., Nilsson, P., et al. 2006), USA (Esper, F., Weibel, C., et al. 2005) and South Africa (Smuts, H., Workman, L., et al. 2008), to name a few, which indicated a global spread of this virus (Figure 1.1). The detection of HCoV-NL63 was primarily observed in winter months, but several reports indicate seasonal variations (Fielding, B.C. 2011). In China, HCoV-NL63 infection was primarily observed to peak in spring and summer (Chiu, S.S., Chan, K.H., et al. 2005) and in Taiwan, infections peaked during autumn (Wu, P.S., Chang, L.Y., et al. 2008). In Thailand, however, HCoV-NL63 infection showed no predilection for a particular season and could it be detected throughout the year (Dare, R.K., Fry, A.M., et al. 2007). Recently, a study was done in Colorado, USA, where respiratory specimens of paediatric patients with respiratory disease was used to characterize the epidemiological and clinical characteristics associated with circulating HCoV-NL63 viruses over a period of three years. It was found, in correlation with previous studies, that even here there is clear variation in prevalence for a particular season where HCoV-NL63 infection was detected in both winter (2010 to 2011) and spring (2004 to 2005) (Dominguez, S.R., Sims, G.E., et al. 2012). A significant number of hospitalizations for children under eighteen years of age, elderly people and immunocompromised individuals have been reported for HCoVs. In a one-year study done in Hong Kong (China) it was shown that for 587 children hospitalized with respiratory tract infection, CoVs accounts for 4.4% of all submission for acute respiratory infections of which HCoV-NL63 was the most common with an incidence of 2.6% (Chan, K.C., Tang, N.L., et al. 2005). Moreover, HCoV-NL63 was detected in five (1.2%) of 419 specimens in a study done in Japan (Suzuki, A., Okamoto, M.,

et al. 2005) and in another Japanese report it was indicated that three (2.5%) of 118 nasopharyngeal swab samples from hospitalized children younger than two years old, tested positive for HCoV-NL63 (Ebihara, T., Endo, R., et al. 2005).

High prevalence for HCoV-NL63 infection among children was also noted in countries around Europe. A study conducted in Italy on 322 infants suffering from acute respiratory disease has shown that 21.4% of the cases were caused by HCoV-NL63 (Canducci, F., Debiaggi, M., et al. 2008). In France, 9.3% of respiratory specimens were positive for HCoV-NL63 (Vabret, A., Mourez, T., et al. 2005). Of the 664 specimens that were sampled from 592 children under the age of six years old that were hospitalized in Slovenia, 15% of them were infected with HCoV-NL63 (Jevsnik, M., Ursic, T., et al. 2012). A recent study done in Australia, HCoV-NL63 was detected in 12.3% of patients with respiratory symptoms. Noteworthy of this study was that HCoV-NL63 was only found in the male population and that this gender distribution of HCoV-NL63 was unexpected and warrants further investigation when conducting clinical or epidemiological studies (Mackay, I.M., Arden, K.E., et al. 2012). Another study done in Saudi Arabia on children younger than sixteen years of age hospitalized with acute respiratory tract infections, have shown that four of 144 (2.8%) of children tested positive for HCoV-NL63 and all had underlying medical conditions and the highest incidence was found among children between the ages of one to four years old (Al Hajjar, S., Al Thawadi, S., et al. 2011).

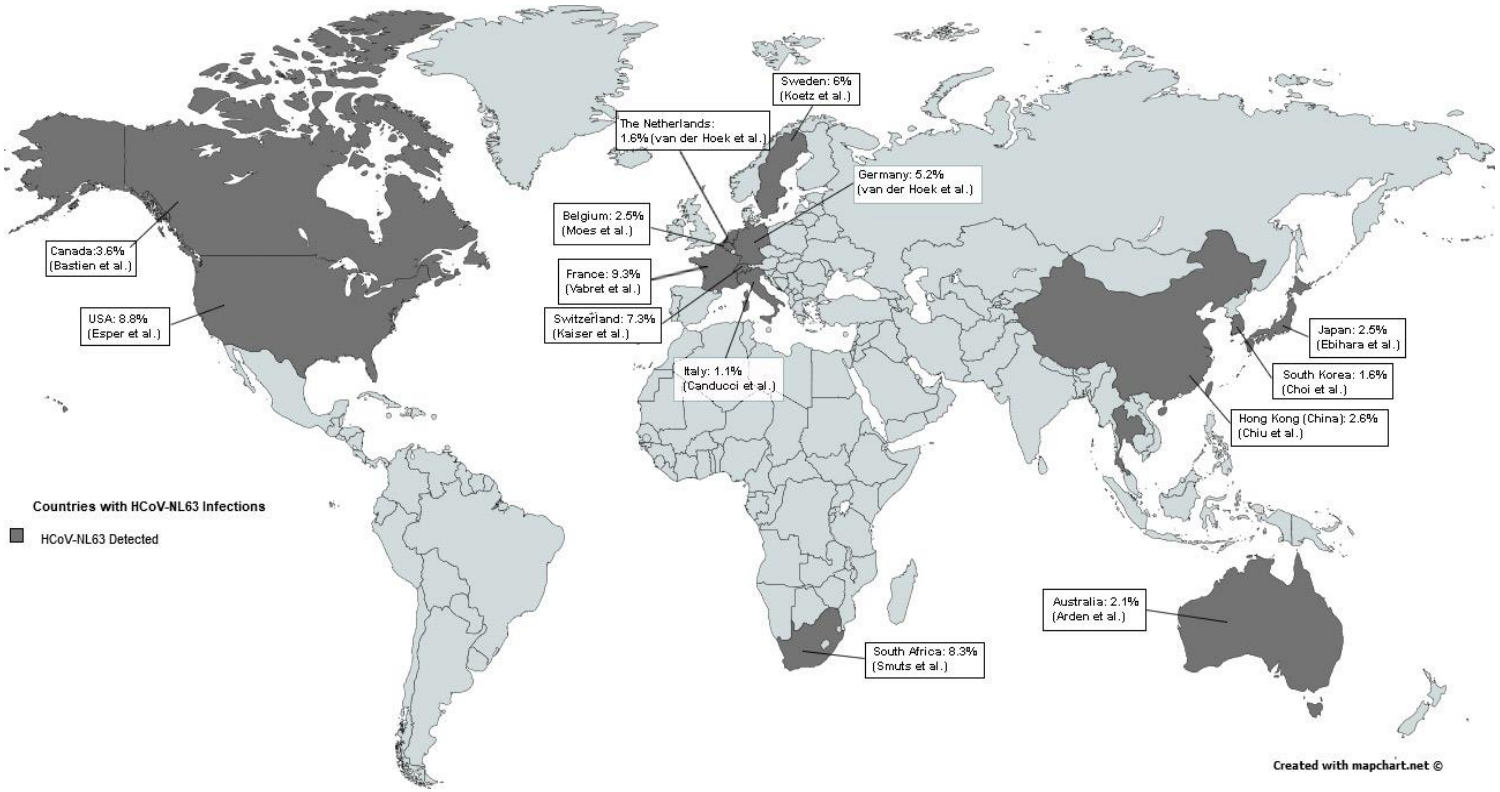



**Figure 1.1 Worldwide distribution of HCoV-NL63.** Countries with documented HCoV-NL63 infections are indicated in grey. The boxes indicate the frequency of HCoV-NL63 infections for each respective country (Adapted from Pyrc, 2005).

Since a great majority of these studies only report on children that was hospitalized, the epidemiological and clinical data is limited for adults. Thus, researchers recently started focusing on the number of adults that become infected with HCoV-NL63. A study done in China showed that 1.1% of adults were infected with HCoV-NL63 and in another study in Beijing detected in less than one percent of HCoV-NL63 among infected adults (Lu, R., Yu, X., et al. 2012; Yu, X., Lu, R., et al. 2012). These numbers are still relatively low in comparison with that found in earlier studies done on young children.

The percentage of detection and its accuracy is hampered by two major problems: the suitability of the samples examined such as those samples collected by nose/throat swabs and by nasopharyngeal aspirate that can differ and can influence the detection and identification of respiratory pathogens (Kleines, M., Scheithauer, S., et al. 2007); and secondly, diagnostic tests for HCoVs aren't normally used in routine testing for viruses, resulting in underestimation of percentage of HCoV infections (van Elden, L.J., van Loon, A.M., et al. 2004).



## 1.5 Genome Organization and Replication Cycle

HCoV-NL63 has a single-stranded, positive sense RNA genome of 27 553 nucleotides (van der Hoek, L., Pyrc, K., et al. 2004) in the order: 5'-1a-1b-S-ORF3-E-M-N-3' that is capped[1] at the 5' (five prime) end and polyadenylated [2] at the 3' (three prime) end (Pyrc, K., Berkhout, B., et al. 2007b). Two thirds of the HCoV-NL63 genome, from the 5' end, contains large open reading frames (ORFs) that encode for non-structural,

[1] The 5' cap is a specially altered nucleotide on the 5' end of mRNA that facilitates stabilization of mRNA.
[2] A 3' poly-A tail is a long chain of adenine nucleotides added to the 3' end of mRNA that functions in the stabilization and export of mRNA from the nucleus to the cytoplasm for translation.

regulatory replicase polyproteins, 1a and 1b. The remaining one third on the 3' end of the genome contains genes that encode for structural proteins: Spike (S), Envelope (E), Membrane (M) and Nucleocapsid (N) proteins; as well as an additional accessory gene, ORF3, located between the S and E structural proteins that encode for non-structural accessory proteins (Pyrc, K., Berkhout, B., et al. 2007b). Accessory proteins, such as ORF3, have been shown to be modulators of pathogenicity in the natural host (Haijema, B.J., Volders, H., et al. 2004; Casais, R., Davies, M., et al. 2005). The genome of HCoV-NL63 is commonly known to present that of a "mosaic-structured genome" due to frequent recombination (Pyrc, K., Dijkman, R., et al. 2006). Figure 1.2 shows the genomic organization of HCoV-NL63 and other CoVs causing infection in humans and their different open reading frames and accessory proteins (Figure 1.2).

### 1.5.1 Transcriptional Regulation and Replication

The transcription and replication of CoVs are very complex processes and the exact mechanism is not fully understood. The HCoV-NL63 genome contains seven ORFs in total that are produced from six distinct messenger RNAs (mRNAs): the full-length genomic mRNA (g mRNA) and a set of five subgenomic (sg) mRNAs (Pyrc, K., Jebbink, M.F., et al. 2004) (Figure 1.3). The mRNAs of CoVs are believed to be generated in the membrane-associated replication centres (Brockway, S.M., Clay, C.T., et al. 2003; Woo, P.C., Lau, S.K., et al. 2005b). The g mRNA encodes for polyproteins 1a and 1b; the structural and accessory proteins S, ORF3, E, M and N are encoded by the five sg mRNAs.

Untranslated regions (UTRs) are found at the 5' and 3' ends of the genome. A leader sequence of 72 nucleotides is found within the UTR on the 5' end of the genome and are located upstream of all the ORFs. The sg mRNAs are generated during minus

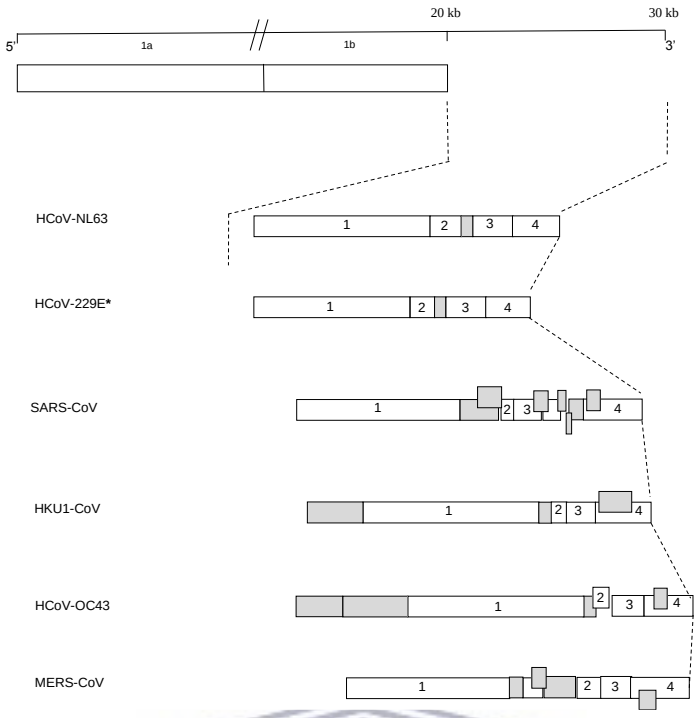



**Figure 1.2 Schematic organization of coronavirus genomes causing infections in humans.** Genomic maps shown are based on the complete genome sequences (NCBI accession numbers): HCoV-NL63: Human coronavirus NL63 (NC_005831); HCoV-229E: Human coronavirus 229E (NC_002645); SARS-CoV: Severe acute respiratory syndrome coronavirus (NC_004718); HKU1-CoV: human coronavirus HKU1 (NC_006577); HCoV-OC43: Human coronavirus OC43 (NC_005147); ORFs S (1), E (2), M (3) and N (4) are shown and open reading frames encoding for accessory genes are shaded in grey.



strand (-) synthesis by means of a discontinuous replication strategy which are then copied into plus strand (+) mRNAs (Pyrc, K., Jebbink, M.F., et al. 2004; Pyrc, K., Berkhout, B., et al. 2007a; Pyrc, K., Berkhout, B., et al. 2007b). A common transcription regulatory sequence (TRS) with a core sequence of AACUAAA contained within the UTR upstream of each ORF, is crucial for sg mRNA generation

(Pyrc, K., Berkhout, B., et al. 2007a; Pyrc, K., Berkhout, B., et al. 2007b). In a recent study, it was found that this core sequence was found in duplicate preceding M and N genes which could possibly affect production (increased production) of sg mRNAs because these two proteins are essential for virus budding and packaging of viral RNA respectively (Nguyen, V.P. and Hogue, B.G. 1997; de Haan, C.A., Kuo, L., et al. 1998; Dominguez, S.R., Sims, G.E., et al. 2012).

During translation one large polyprotein is encoded by the 1a gene and with a (-1) ribosomal frameshift, it generates the 1ab polyprotein. This frameshift is facilitated by a pseudoknot located within the polyprotein that consists of a hairpin loop that acts as a slippery sequence (Pyrc, K., Berkhout, B., et al. 2005). The 1ab polyprotein functions mainly in RNA replication (van der Hoek, L., Pyrc, K., et al. 2006).

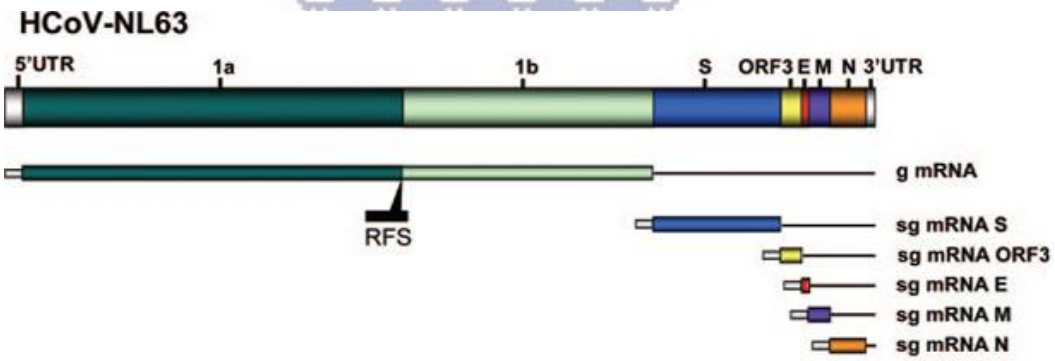



**Figure 1.3 Schematic organization of HCoV-NL63 genome.** Below the scheme the corresponding subgenomic mRNAs (sg mRNA) generated during discontinuous transcription process. Untranslated regions (UTR) on the 5' and 3' ends of the genome is marked. All ORFs contain UTRs upstream with a common transcription regulation sequence (TRS). The ribosomal frameshift (RFS) element is indicated between 1a and 1b genes (Pyrc, K., Berkhout, B., et al. 2007b).

### 1.5.2 Cellular Entry

HCoV-NL63 and SARS-CoV use the same cellular receptor, angiotensin converting enzyme (ACE)-2, for entry into the host cell with overlapping binding regions (Figure 1.4) (Donoghue, M., Hsieh, F., et al. 2000; Hofmann, H., Pyrc, K., et al. 2005). Even though they may use the same receptor, their method of binding to the receptor and the events that follow are very different (Wu, K., Chen, L., et al. 2011). One of the differences between HCoV-NL63 and SARS-CoV is that the S protein of HCoV-NL63 binds to ACE-2 with a lower affinity than SARS-CoV (Glowacka, I., Bertram, S., et al. 2010). This can be due to the structural differences between S proteins of SARS-CoV and HCoV-NL63: the core structure of HCoV-NL63 consists of two layer beta-sandwich where SARS-CoV only has one; the receptor-binding motif of HCoV-NL63 has three discontinuous short loops whereas SARS-CoV has one long continuous sub-domain (Wu, K., Chen, L., et al. 2011). It has also been found that ACE-2 cellular protein levels in HCoV-NL63 are reduced three days post-infection, which coincides with the increase of HCoV-NL63 production in infected cells during early phase replication (Dijkman, R., Jebbink, M.F., et al. 2012). Although not proven, it is speculated that this reduction in ACE-2 could be due to the binding of S protein to the host receptor (Glowacka, I., Bertram, S., et al. 2010).

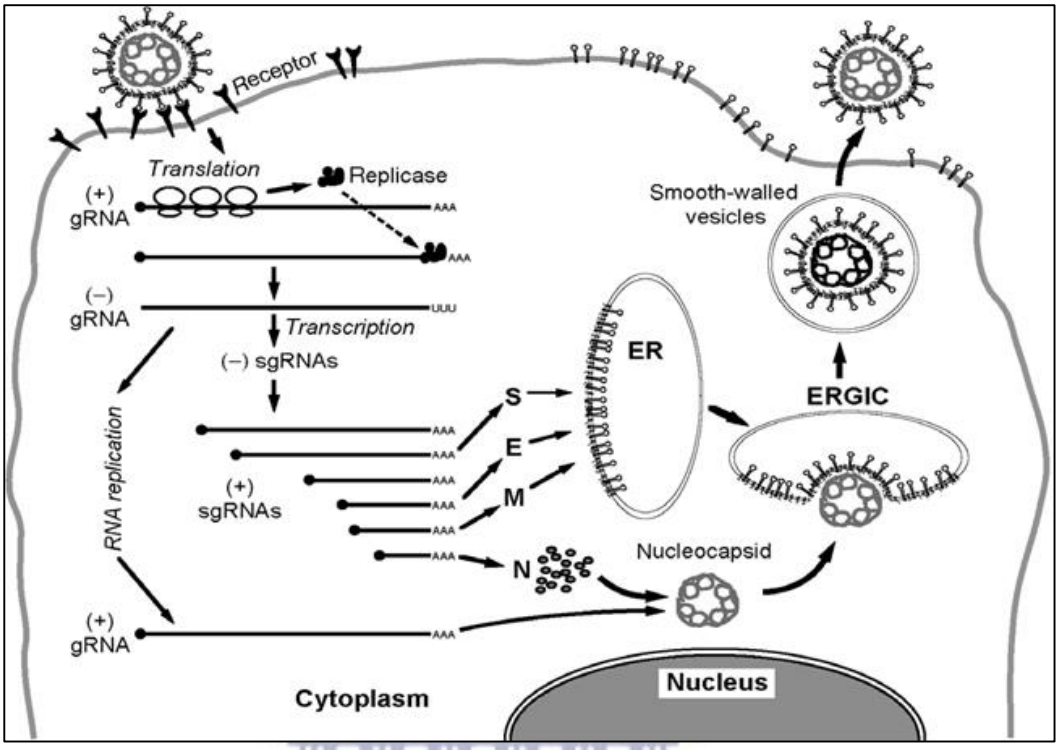



**Figure 1.4 Viral replication and virion assembly.** The virus particle fuses with the host cell through interactions between S protein of the virus particle and ACE-2 receptor on the host cell membranes. Positive-strand genomic RNA (gRNA) is released into the cytoplasm used for translation of ORF1a and b to replicase polyprotein 1ab. Subgenomic RNAs (sg RNAs) translate to form structural and accessory proteins S, ORF3, E and M that inserts into the endoplasmic reticulum (ER) and transits to the ER-Golgi intermediate compartment (ERGIC). N protein associates with replicated gRNA to form nucleocapsids that bud into the ERGIC to form smooth-walled vesicles before being transported and released from the cell. Adapted from *The Molecular Biology of Coronaviruses* by P.S. Masters (Masters, P.S. 2006).

## 1.6 Structural Proteins of HCoV-NL63

### 1.6.1 Structural (S) Protein

The S glycoprotein on the membrane surface provide a typical crown-like structure (crown = *corona*) to the virion (van der Hoek, L., Pyrc, K., et al. 2006). This spike protein facilitates entry into host cells through recognition, attachment and fusion of S with the host membrane or endocytosis, which will then trigger the release of the viral RNA genome into the host cytoplasm (Pyrc, K., Berkhout, B., et al. 2007b).

### 1.6.2 Membrane (M) Protein

The most abundant glycoprotein in the virus particle in infected cells is the membrane (M) protein characterized by three hydrophobic domains spanning the membrane bilayer with three potential N-glycosylation sites: a short N (amino)-terminal ectodomain towards the lumen, short hydrophilic triple-spanning transmembrane regions and a long C (carboxy)-terminal endodomain towards the cytoplasm of the virion (de Haan, C.A., Kuo, L., et al. 1998; de Haan, C.A., Smeets, M., et al. 1999; Narayanan, K., Maeda, A., et al. 2000; Fang, X., Ye, L., et al. 2005; Voss, D., Pfefferle, S., et al. 2009; Arndt, A.L., Larson, B.J., et al. 2010; Muller, M.A., van der Hoek, L., et al. 2010). The basic overall structures of M are shared among all CoVs (Rottier, P., Brandenburg, D., et al. 1984; Baudoux, P., Carrat, C., et al. 1998; Arndt, A.L., Larson, B.J., et al. 2010). This triple spanning M protein contained within the viral envelope plays an important role in envelope formation and virus entry into host cells (Kuo, L. and Masters, P.S. 2003). CoVs acquire their lipid envelope by budding of the nucleocapsid through the endoplasmic reticulum (ER)-Golgi intermediate compartment (ERGIC) membranes (Figure 1.4) (de Haan, C.A., Kuo, L., et al. 1998;

de Haan, C.A., Smeets, M., et al. 1999). In the case where M protein is expressed without the presence of the other viral proteins, N and E, it localizes in the late-Golgi complex (Klumperman, J., Locker, J.K., et al. 1994).

The C-terminal and the amphipathic domains were found to be structurally essential requirements of crucial importance for virus and viral envelope assembly by interacting with the nucleocapsid to which M has a high affinity, as well as the interaction with E and S proteins (de Haan, C.A., Kuo, L., et al. 1998; de Haan, C.A., Smeets, M., et al. 1999; Siu, Y.L., Teoh, K.T., et al. 2008; Arndt, A.L., Larson, B.J., et al. 2010). M protein contains highly conserved glycosylated sequences due to probable interaction between the virus and host cells (de Haan, C.A., Kuo, L., et al. 1998). Besides mediating virus particle formation, M protein also recruits other viral structural components to the ERGIC for virus assembly and budding (Voss, D., Pfefferle, S., et al. 2009). The restriction of M protein to the budding compartment is believed to be driven by interaction with the viral nucleocapsid, which in turn facilitates the envelopment of the nucleocapsid at the budding compartment (Narayanan, K., Maeda, A., et al. 2000). This binding of M protein to the nucleocapsid appears to function in stabilization of the core (He, R., Leeson, A., et al. 2004).



### 1.6.3 Envelope (E) Protein

Although relatively small in size (76 amino acids) and expressed in low quantities *in vivo,* the E protein have been revealed to play a significant multifunctional role in the CoV virion life cycle (Vennema, H., Godeke, G.J., et al. 1996). Apart from being involved in viral envelope formation, it also plays an important role in viral replication, such as with transmissible gastroenteritis CoV (TGEV) and mouse hepatitis virus

(MHV) (Kuo, L. and Masters, P.S. 2003). The E protein has also been linked to apoptosis of the E-protein-expressing cells in MHV (An, S., Chen, C.J., et al. 1999).

E protein is a poorly characterized small, minor structural protein that plays a part in the extracellular release of virion-like particles (VLPs) and act together with M to form the envelope (Fischer, F., Stegen, C.F., et al. 1998; de Haan, C.A., Smeets, M., et al. 1999). In several studies done on these VLPs, the formation of CoVs is mediated by only the M and E proteins and not together with S and/or N proteins (Baudoux, P., Carrat, C., et al. 1998; Corse, E. and Machamer, C.E. 2000; de Haan, C.A., Vennema, H., et al. 2000; Hurst, K.R., Kuo, L., et al. 2005; Yang, Y., Xiong, Z., et al. 2005; DeDiego, M.L., Alvarez, E., et al. 2007). It is not clear whether the E protein acts independently of M during the budding process (Corse, E. and Machamer, C.E. 2000) as opposed to some evidence that suggested an interaction between E and M (Baudoux, P., Carrat, C., et al. 1998; Corse, E. and Machamer, C.E. 2000; Hurst, K.R., Kuo, L., et al. 2005). In contrast to these findings, it has been suggested that in the formation of SARS-CoV-like particles, M and N proteins are the key role players instead of the E protein (Huang, Y., Yang, Z.Y., et al. 2004).



However, when a study was done on whether E gene plays a critical role in viral assembly, the E gene was deleted and it was found that even though the cell remained viable, it had very low infectivity and replicated poorly; in another this disruption was found to be lethal to the cell. This suggested that E protein is both a vital protein, and it is required for the production of an infectious virus (Curtis, K.M., Yount, B., et al. 2002; Kuo, L. and Masters, P.S. 2003; Hurst, K.R., Kuo, L., et al. 2005).

### 1.6.4 Nucleocapsid (N) Protein

The N protein, one of the most abundant viral proteins and a multifunctional phosphoprotein (McBride, R., van Zyl, M., et al. 2014), is highly basic, exists as a disulfide-linked multimer and is the only RNA-binding protein of mature virions to form a ribonucleocapsid structure. The N protein of MHV, a member of the CoV family, is 60 kD long and has an amino acid sequence of 455 amino acids long (Baric, R.S., Nelson, G.W., et al. 1988; You, J., Dove, B.K., et al. 2005). The CoV N protein was identified to contain three conserved regions, separated by two spacers, that serve as RNA-binding sites and is involved in cell signalling (Parker, M.M. and Masters, P.S. 1990; You, J., Dove, B.K., et al. 2005). The N molecule contains two independently folding domains: N-terminal domain (NTD) and C-terminal domain (CTD). The NTD is separated from CTD by SR-domains of splicing factors resembling a serine- and arginine-rich region (Hurst, K.R., Koetzner, C.A., et al. 2009).



The primary function of CoV N protein is to modulate assembly of the ribonucleocapsid core. For proper assembly of virions, specific protein-protein as well as protein-nucleic acid interactions must exist (Nguyen, V.P. and Hogue, B.G. 1997). It is proposed that there's an interaction between N and M glycoproteins at the site of budding (Nguyen, V.P. and Hogue, B.G. 1997). It is generally accepted that CoVs form a helical nucleocapsid consisting of the viral genome RNA and N protein (Figure 1.5) (Sturman, L.S., Holmes, K.V., et al. 1980). The envelopment of this nucleocapsid is speculated to be driven by the interactions between N and M or by M and the viral mRNA at the pre-Golgi compartment at the ER membrane. The binding of M to N through a packaging signal lead to the binding of M to the viral genomic RNA and

due to the association of N with mRNA, the conformation of N protein may be altered to specifically bind to M protein (Cologna, R. and Hogue, B.G. 1998; Narayanan, K., Maeda, A., et al. 2000; Chen, C.Y., Chang, C.K., et al. 2007). Subsequent interaction with E protein follows interaction with M protein that leads to extracellular release of virion particles (Fischer, F., Stegen, C.F., et al. 1998; Hurst, K.R., Koetzner, C.A., et al. 2009). After synthesis, N is phosphorylated and is associated with cellular membranes through the association with E (Baric, R.S., Nelson, G.W., et al. 1988). Phosphorylation of N occurs by kinases such as cyclin-dependent kinase, glycogen synthase kinase 3, casein kinase II, and mitogen-activated protein kinase (Surjit, M., Kumar, R., et al. 2005).

SARS-CoV N protein can regulate cellular processes (Surjit, M., Liu, B., et al. 2004; Luo, H., Chen, Q., et al. 2005). It was also postulated that N protein is involved in viral transcription and replication (Surjit, M., Kumar, R., et al. 2005) and possibly in translation (Ye, Y., Hauns, K., et al. 2007), as well as RNA synthesis (Baric, R.S., Nelson, G.W., et al. 1988; Almazan, F., Galan, C., et al. 2004; Surjit, M. and Lal, S.K. 2008; Verheije, M.H., Hagemeijer, M.C., et al. 2010).



Other reported functions of CoV N protein may include its self-association to form dimers (Surjit, M., Kumar, R., et al. 2005), acting as an RNA chaperone (Zuniga, S., Sola, I., et al. 2007; Chang, C.K., Hsu, Y.L., et al. 2009; Verheije, M.H., Hagemeijer, M.C., et al. 2010), contributing to the perturbation of cellular processes in several host cells (Surjit, M., Liu, B., et al. 2004; Verheije, M.H., Hagemeijer, M.C., et al. 2010), inducing apoptosis and actin reorganization in mammalian cells (Surjit, M., Kumar, R., et al. 2005) and acting as an interferon antagonist (Ye, Y., Hauns, K., et al. 2007). N protein can also be used as a diagnostic tool in the early detection of SARS-CoV

infection in patients due to its high immunogenicity (Lau, S.K., Woo, P.C., et al. 2004).

HCoV N protein is highly antigenic and abundantly expressed during viral infection (Hiscox, J.A., Cavanagh, D., et al. 1995; Timani, K.A., Ye, L., et al. 2004; Zhao, J., Wang, W., et al. 2007) and detectable as early as from the first day after infection (Che, X.Y., Qiu, L.W., et al. 2004). Several studies have used N as an antigen for the detection of antibodies to HCoV infection (Hajimorad, M.R. and Francki, R.I. 1991; Woo, P.C., Lau, S.K., et al. 2004a; Dijkman, R., Jebbink, M.F., et al. 2008; Lehmann, C., Wolf, H., et al. 2008; Severance, E.G., Bossis, I., et al. 2008). Antibodies have become essential for diagnostic and therapeutic purposes due to their high specificity, high binding affinity, long half-lives and low toxicity, and are invaluable reagents for antigen detection in immunoblotting (Harlow, E. and Lane, D. 1988; Hancock, D.C. and O'Reilly, N.J. 2005; Hayes, A.J., Hughes, C.E., et al. 2008). The N protein therefore makes a great target to be used in enzyme-linked immunosorbent assay (ELISA) assays for the direct detection of viral antigens and the production of suitable antibodies that can be used to detect the native viral protein.

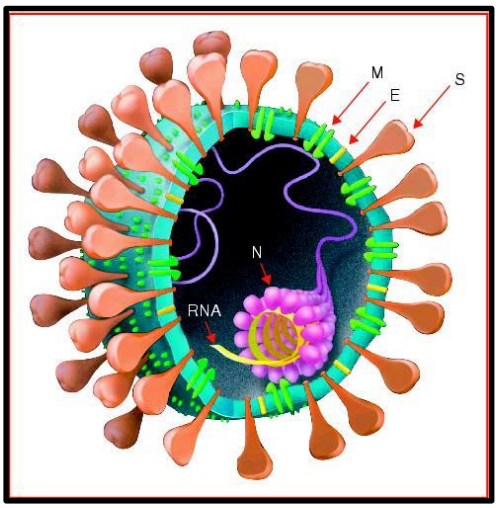



**Figure 1.5 Coronavirus structure**. CoV structure with structural proteins: nucleocapsid protein (N), encapsulating the RNA and forming the nucleocapsid. Membrane protein (M), envelope protein (E) and spike protein (S) found on the surface of the protein (Holmes, K.V. and Enjuanes, L. 2003).



## 1.7 Aims of this thesis

HCoVs are commonly associated with both mild URTIs known to manifest as the common cold. As CoVs evolve over time, so too will its pathogenicity. HCoV-NL63 is commonly associated with other respiratory viruses in co-infections that may lead to increased severity of symptoms and most likely more hospitalizations. Apart from SARS-CoV, HCoV-NL63 and four more known HCoVs, HCoV-229E, HCoV-OC43, HCoV-HKU1 and HCoV-MERS, are continually circulating the human population and are not well studied. Constant development of molecular tools can facilitate accurate detection of CoVs in clinical samples, which in turn would increase our understanding of the extent to which CoVs affect human health.

HCoV N protein has been shown to be highly antigenic. The N protein is abundantly expressed during infection and able to induce an effective antibody response in the host cell. Thus, making the N protein an important target protein when used in the development of effective CoV diagnostic tools. The N protein can, therefore, be utilized as an antigen for detection of antibodies to HCoV infections.

This dissertation focuses on the production of antibodies against HCoV-NL63 N protein in an animal model. Major outcomes will include:

- Design and synthesis of peptides from the HCoV-NL63 N protein amino acid sequence.
- Immunization of mouse animal model with synthetic peptides to produce polyclonal anti-N antibodies.
- Cloning and production of monoclonal anti-N antibodies.
- Characterization of N-specific antibodies for detection of expressed purified HCoV-NL63 N protein.



Sub-aims include:

- Expression and purification of recombinant SARS-CoV and HCoV-NL63 N proteins and its truncated clones.

# CHAPTER 2

# METHODOLOGY


Department of Medical Biosciences, Faculty of Natural Sciences, University of the Western Cape, South Africa

# CHAPTER 2 METHODOLOGY

## 2.1 Bacterial Strains and Plasmids

KRX (Promega) strain of competent *Escherichia coli* (*E. coli)* is a K12 derivative that was used for the expression of all constructs utilizing a Flexi® (pFN2A) Vector System (Promega). The Flexi® vector (Appendix 1) was used as a bacterial expression vector due to its unique ability to express fusion or native proteins for study into the structure and function of proteins. The Flexi® vector possess two unique enzyme sites, *SgfI* and *PmeI*, that allow for easy insertion and transfer of the protein coding region of interest and a lethal barnase gene, located between these cloning sites, that is replaced upon successful ligation and prevents growth of unsuccessfully ligated cells. This vector also contains an ampicillin-resistance gene for plasmid selection, a T7 promoter for bacterial or *in vitro* protein expression and a glutathione-S-transferase (GST) coding region that facilitate detection and purification of expressed proteins (Promega).



## 2.2 Recombinant Constructs

In this thesis, the DNA constructs used for expression of proteins, were provided as glycerol stocks of transformed *E. coli* cultures and were kind gifts from Mr. M. Berry (Department of Medical Biosciences, University of the Western Cape). These recombinant constructs are clones of full-length SARS-N and HCoV-NL63 N genes, as well as deletion mutants of the N- and C-terminus of HCoV-NL63 N. For the purpose of this thesis the constructs were named SARS-N, N1, N2, N3 respectively (Figure 2.1). All constructs were expressed downstream from a GST fusion protein and maintained and propagated in KRX *E. coli* competent cells. Before proceeding to

the expression studies, recombinant plasmids were verified for correct insert sizes. SARS-CoV N has been highly characterized and therefore included as a positive control.

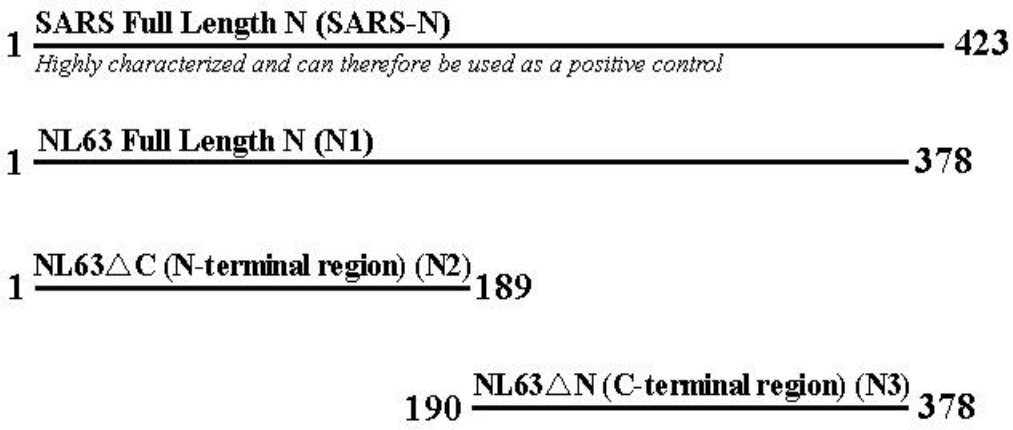



**Figure 2.1 Schematic description of SARS-CoV N (SARS-N), HCoV-NL63 N protein (N1) and recombinant mutants, N2 and N3.** Numbers indicate amino acid positions**.**



## 2.3 Cultivation and Verification of Recombinant Constructs

### 2.3.1 Preparation of Culture Medium

Luria Bertani (LB): LB medium (10 g pancreatic digest of casein or tryptone powder, 5 g yeast extract powder and 5 g NaCl per 1 L, pH 7.2) is one of the most common liquid media used in bacterial cultivation such as *E. coli*. The media was autoclaved to sterilize it after which, upon cooling, 1 µl/ml ampicillin was added.

LB agar: LB agar was prepared in the same way as LB medium with the addition of 15 g agar bacteriologica per 1 L. Agar was thoroughly dissolved by convection heat prior to autoclaving.

### 2.3.2 Preparation of Starter Culture

Bacterial glycerol stocks of the respective recombinant constructs were streaked onto solidified LB agar plates and incubated overnight at 37°C. Single colonies were picked and inoculated into a starter culture of 10 ml LB medium which was incubated for 14 hours at 37°C with shaking at 150 rpm. Own glycerol stocks were prepared from this starter culture with 1.5 ml culture media and 500 µl glycerol and stored at -80°C. The remainder of the starter culture was used for plasmid DNA extraction using the peqGOLD Plasmid Miniprep I kit (Peqlab).

### 2.3.3 Small-scale Plasmid DNA Extraction (Miniprep)

Plasmid DNA was isolated from bacterial overnight culture using the peqGOLD Plasmid Miniprep Kit I (Peqlab). The principle of this procedure is based on the consistency of alkaline-SDS lysis of bacterial cells combined with the PerfectBind technology which is a column-based isolation system that include selective and reversible DNA binding on a silica membrane to simplify the extraction and purification of plasmid DNA from contaminants and enzyme inhibitors. Purified plasmid DNA is then eluted from the column to deliver high-quality plasmid DNA.



The manufacturer's protocol advises using 5 ml overnight culture for plasmid isolation. It was found however that using a biomass (BM) index was more relevant to ensure the highest yield of plasmid DNA. A BM of approximately 4 mg was used for all minipreps. The BM was calculated using the following equation:

**BM (mg) = Optical Density of 600 ($OD_{600}$) (mg/ml) x Volume (ml)**

Cells were harvested by centrifugation at 14,000 rpm for 5 minutes and supernatant discarded. The cell pellet was then immediately resuspended in 250 µl Solution I (containing 100 µg/ml RNase A) and vortexed thoroughly to obtain high yields. Cells

were then lysed by adding 250 µl Solution II and the tube gently inverted 6-10 times. The cell suspension was incubated for 2 minutes at room temperature to denature all DNA to its single stranded form in the presence of the high pH levels provided by this solution and to obtain a clear lysate. The lysate was then neutralized with the addition of 350 µl Solution III and gently inverted 6-10 times until a flocculent white precipitate formed which allows the normally circular plasmid DNA to realign and the chromosomal DNA to bind the cellular debris in its still single-stranded form. This was followed by centrifugation at 14,000 rpm for 10 minutes at room temperature whereby chromosomal DNA, bound to cellular debris, was separated as a pellet from the lysate containing the plasmid DNA. The clear supernatant was then transferred to the PerfectBind DNA column, assembled in a 2 ml collection tube (supplied by Peqlab), and centrifuged at 14,000 rpm for 1 minute. The lysate passed through the silica membrane, allowing the plasmid DNA to bind to the membrane due to the generated electrostatic interactions in the presence of chaotropic salts. The flow-through was discarded and column reassembled. The membrane was washed with 500 µl PW plasmid buffer for the complete removal of protein contamination and centrifuged at 14,000 rpm for 1 minute at room temperature. The flow-through was discarded and the column reassembled after which it was washed with 750 µl DNA Wash Buffer that was diluted with absolute ethanol before use. This was followed by centrifugation at 14,000 rpm for 2 minutes to ensure complete evaporation of residual ethanol in the membrane. The plasmid DNA was released from the membrane with the addition of 30 µl nuclease-free water / elution buffer which neutralizes the electrostatic interactions between the DNA and silica membrane. The eluted plasmid DNA was then subjected to a final centrifugation at 14,000 rpm for 1 minute in a clean

1.5 ml microcentrifuge tube and stored at -20°C until further use. This was followed by quantitation of the plasmid DNA by Qubit® fluorometry.

### 2.3.4 Quantification of Plasmid DNA (fluorometry)

The miniprep-purified plasmid DNA was quantitated using the Quant-iT™ dsDNA Broad Range (BR) Assay Kit with the Qubit® 1.0 Fluorometer (Invitrogen) to record the DNA concentrations of each plasmid easily and accurately. The Qubit® quantitation platform allows for highly sensitive fluorescence-based quantitation. Briefly, prepare the Quant-iT™ dsDNA BR working solution by diluting 1 x *n* µl of Quant-iT™ dsDNA BR reagent in 199 µl x *n* µl of Quant-iT™ buffer (where *n* represent the number of samples plus the number of standards). Assay tubes were prepared to a final volume of 200 µl by mixing 190 µl of working solution with 10 µl of each standard as well as mixing 180-199 µl of working solution with 1-20 µl of each sample. All the assay tubes were vortexed for 2-3 seconds and incubated at room temperature for 2 minutes. Following calibration of the Qubit® 1.0 Fluorometer with known standards, fluorescence was measured of the isolated plasmid DNA samples and the final concentrations calculated using the following formula where C = concentration of your samples, QF = the Qubit fluorometer and $x$ = number of µl of sample added to assay tube:

$$\mathbf{C\ (\mu g/ml) = QF\ value\ (\mu g/ml) \times \left(\frac{200}{x}\right)}$$

### 2.3.5 Polymerase Chain Reaction

Plasmid DNA was amplified by Polymerase Chain Reaction (PCR) using the GoTaq® Flexi DNA Polymerase kit (Promega). PCR amplification is a process that allows the

duplication of the target DNA sequence of interest and generating millions of copies of this specific DNA sequence (Saiki, R.K., Scharf, S., et al. 1985).

PCR reaction was set up in a final volume of 25 µl with 5µl 5X Green GoTaq® Flexi buffer, 2 µl $MgCl_2$, 0.5 µl PCR nucleotide mix, 1 µl forward primer, 1 µl reverse primer, 0.5 µl GoTaq® Flexi polymerase, 2 µl plasmid DNA template and 13 µl nuclease-free water. Table 2.1 illustrates the concentrations of each reagent and its ranges used in a typical PCR reaction (depending on the final volume) and the order in which components should be added to the PCR tube. The PCR reaction tubes are then placed in a thermal cycler for amplification that was run under the following conditions: initial denaturation at 95°C for 3 minutes followed by 30 cycles of denaturation at 95°C for 1 minute, annealing at 60°C for 1 minute, elongation at 72°C for 2 minutes and a final elongation at 72°C for 15 minutes. The amplified product was then loaded onto an agarose gel and separated by electrophoresis to monitor migration progress.



### 2.3.6 Agarose Gel Electrophoresis

For the verification of the plasmid DNA, the PCR product was diluted with 6X 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)] in a 1:6 ratio and separated on 1% (w/v) agarose gel, containing 0.001% (v/v) ethidium bromide, by electrophoresis in a TBE [89 mM tris (hydroxymethyl) aminomethane, 0.089 mM boric acid, 2 mM EDTA (pH 8.0)] buffer at 90 V for 35 minutes. Results were recorded by viewing the agarose gel under ultraviolet (UV) light and taking a picture thereof to verify the presence of amplified genes of interest.

**Table 2.1 Typical ranges used in a PCR reaction.**

| Components & Order | Final Concentrations |
|---|---|
| Nuclease-free water to final volume | - |
| 5X Green GoTaq Flexi Buffer | 1X |
| PCR nucleotide mix (10 mM each) | 0.2 mM each dNTP |
| GoTaq DNA polymerase (5 u/µl) | 1.25 u |
| $MgCl_2$ solution (25 mM) | 1-4 mM |
| Forward primer | 0.1–1.0 µM |
| Backward primer | 0.1–1.0 µM |
| Template DNA | <0.5 µg/50 µl |



### 2.3.7 Restriction Endonuclease Digestion of Miniprep

A restriction endonuclease digestion was performed following the isolation of plasmid DNA, to verify further the presence of the insert. Double stranded DNA (dsDNA) are cleaved by what is more commonly known as restriction enzymes, which recognizes short, specific (often palindromic) DNA sequences. The reaction was set up according to the manufacturer's protocol (Promega).

**Table 2.2 Restriction digest with Flexi enzymes *SgfI* and *PmeI*.**

| Reagent | Digest Reaction (µl) |
|---|---|
| 5X Flexi Digest | 10 |
| DNA constructs (0.5 µg/µl) | 10 |
| *SgfI* and *PmeI* enzyme blend (10 u/µl) | 1 |
| Nuclease-free water | 29 |

Briefly, the reaction to remove the gene insert from the Flexi® vector was set up (as shown in Table 2.2) with 10 µl of 5X Flexi® Digest Buffer [50 mM Tris-HCl (pH 7.9), 250 mM NaCl, 150 mM KCl, 50 mM $MgCl_2$, 5 mM DTT, 0.5 mg/ml acetylated bovine serum albumin (BSA)], 1 µl Flexi® Enzyme Blend (*SgfI* and *PmeI*), 10 µl of plasmid DNA product (calculated for 0.5 µg DNA per loading sample) and nuclease-free water to a final volume of 50 µl. The reaction tubes were incubated in a thermocycler at 37°C for 2 hours followed by the inactivation of enzymes at 65°C for 20 minutes. The cut inserts were characterized electrophoretically on 1% agarose gel (as described in 2.3.6). Once the presence of the insert was confirmed, large-scale cultivation could be performed for midiprep.



## 2.4 Large-scale Plasmid Extraction (Midiprep)

The NucleoBond® PC100 Plasmid DNA Purification Kit (Macherey-Nagel) was used for the midiprep procedure to isolate plasmid DNA from bacterial cells. The principle of this procedure is based on that of the miniprep procedure (as described in 2.3.3),

but implemented due to the use of BM index for higher yield volumes of plasmid DNA. A BM of approximately 4 mg was used.

Briefly, bacterial glycerol stocks of the respective recombinant constructs were streaked onto solidified LB agar plates and incubated overnight at 37°C. Single colonies were picked and inoculated into a starter culture of 10 ml LB medium which was incubated for 8 hours at 37°C with shaking at 150 rpm. The starter cultures were diluted 1:100 into 400 ml LB and incubated at 37°C for 6 hours with agitation at 150 rpm. Harvesting of the cells were done by centrifugation at 3,000 rpm for 10 minutes at 4°C. After culturing cells, pellets were resuspended in 8 ml resuspension Buffer S1 (50 mM Tris-HCl, 10 mM EDTA, 100 µg/ml RNase A at pH 8.0) with RNase A. This was followed by lysis of cells in 8 ml lysis Buffer S2 (200 mM NaOH, 1% SDS). The cells were then neutralized in 8 ml neutralization Buffer S3 [2.8 M KAc (pH 5.1)], inverted 6-8 times and incubated at room temperature for 5 minutes. Lysate was then equilibrated in 2.5 ml equilibration Buffer N2 [100 mM Tris, 15% ethanol, 900 mM KCl, 0.15% Triton X-100 (pH 6.3)] on the NucleoBond AX100 column after which the suspension was centrifuged at 14,000 rpm for 25 minutes at 4°C to separate chromosomal DNA from cellular debris. The lysate was then allowed to bind onto the column by means of gravitational flow-through and was washed with 12 ml wash Buffer N3 [100 mM Tris, 15% ethanol, 1.15 M KCl (pH 6.3)]. The plasmid DNA was then eluted from the column by the addition of 5 ml elution Buffer N5 [100 mM Tris, 15% ethanol, 1 M KCl (pH 8.5)] after which the eluted plasmid DNA was precipitated in 3.5 ml isopropanol and centrifuged at 14,000 rpm for 30 minutes at 4°C to harvest plasmid DNA. The pellet was washed and dried with 1 ml 70% ethanol and centrifuged at 14,000 rpm for 10 minutes at room temperature. The plasmid DNA was

reconstituted by dissolving the DNA pellet in buffer TE or sterile distilled water under constant spinning for 1 hour. The plasmid DNA from midiprep was quantitated using the Quant-iT™ dsDNA Broad Range (BR) Assay Kit with the Qubit® 1.0 Fluorometer (Invitrogen) as described in 2.3.4.

### 2.4.1 Restriction Endonuclease Digestion of Midiprep

Restriction endonuclease digestion was performed as a second confirmatory step on the eluted plasmid DNA from midiprep to verify the presence of the plasmid DNA insert on a larger scale in comparison to miniprep. The same protocol was followed as described previously in 2.3.7.

## 2.5 Expression of Full-Length N and deletion mutants in KRX *E. coli* cells through Autoinduction.

The autoinduction protocol as described by Schagat *et al*. (2008) employs the tight control of the T7 RNA polymerase gene, contained in the *E. coli* KRX strain, by glucose and rhamnose via a rhamnose promoter, (rhaPBAD). The promoter is repressed by metabolism of glucose via a cyclic adenosine monophosphate (cAMP) pathway and only activated in the presence of rhamnose once the glucose is consumed from the culture medium. This process thus controls the production of recombinant protein of interest (Holcroft, C.C. and Egan, S.M. 2000; Schagat, T., Ohana, R., et al. 2008).

For autoinduction, previously prepared glycerol stocks of SARS-N, N1, N2 and N3 were inoculated into 10 ml LB medium and incubated at 37°C with agitation at 150 rpm for 14 hours. The starter cultures were diluted 1:100 into 400 ml LB expression cultures that was autoinduced with 0.1% (w/v) rhamnose and 0.05% (w/v) glucose

immediately after dilution occurred and then incubated at 37°C for 6 hours with agitation at 150 rpm. Harvesting of the cells were done by centrifugation at 3,000 rpm for 10 minutes at 4°C after which the pellets were stored at -80°C until further use.

## 2.6 Protein Analysis

### 2.6.1 Protein Extraction

For proteins to be extracted it is necessary to lyse the cells that contain the protein by bringing it into solution that can break down the cell and release the protein from cell membranes. The chosen lysis buffer contains a non-ionic detergent that solubilizes the proteins so they can migrate individually through a separating gel.

Following centrifugation, each gram of cell pellet was resuspended in 2 ml lysis buffer (1% (v/v) TritonX, 150 mM NaCl, 10 mM Tris (hydroxymethyl)-amino-methane (Tris), 5 mM EDTA) and lysed in an ice bath by probe sonication. Table 2.3 describes the function of all constituents in the lysis buffer. Probe sonication was performed in five short bursts of 45 second intervals at medium speed with 30 second rest periods between each interval to prevent protein degradation. Subsequent centrifugation at 14,000 rpm for 10 minutes at 4°C separated the soluble from insoluble cell fractions after which the soluble fraction was removed and stored at -20°C until further use.

**Table 2.3 Function of all constituents in lysis buffer.**

| Compound | Function |
|---|---|
| Triton-X | 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-dependent proteases. |

## 2.6.2 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE describes a technique widely used to separate proteins according to their electrophoretic mobility. SDS (sodium dodecyl sulfate) is an anionic detergent that linearize proteins and impart a negative charge to the proteins. The reaction between acrylamide and bis-acrylamide (N,N-methylenebisacrylamide) result in a cross-linking polyacrylamide gel matrix. This process is initiated by the addition of ammoniumpersulfate (APS) along with TEMED (tetramethylethylenediamine). By applying an electrical current to the gel, the negatively charged proteins will move through the pores of the acrylamide gel and migrate towards the positive electrode and will subsequently separate according to the size of the polypeptide chain and its charge.

### *2.6.2.1 Preparation of Cell Lysate for SDS-PAGE*

For proteins to be accurately separated by size, it is necessary to unfold, or denature, the proteins to overcome the globular form of protein secondary and tertiary structures. A loading buffer, like Laemmli buffer that contains the anionic detergent, SDS, are normally used to denature proteins together with heating of the sample. Therefore, the extracted protein lysate was mixed with 2X Laemmli loading buffer (4% SDS, 10% 2-mercaptoethanol, 20% glycerol, 0.004% bromophenol blue, 0.125 M Tris HCl) in a 1:1 ratio, vortexed and heated on a heating block at 92°C for 3 minutes to linearize the protein. Centrifuge at 14,000 rpm for 5 minutes to remove undissolved particles. Sample preparation should follow shortly after gels have been prepared and should be loaded onto the gel for electrophoresis immediately after centrifugation.

### *2.6.2.2 Preparation of SDS Gels and Electrophoresis*

For SDS-PAGE analysis, self-cast gel plates were assembled with spacers. The Sigma-Aldrich Gel Preparation kit (Fluka) was used to prepare the gel. A 12.5% separation gel and 3% stacking gel was prepared according to the manufacturer's protocol (Fluka) using the Hoefer SE245 Dual Gel Caster (Hoefer). The gel was allowed to set for 20-30 minutes at room temperature. Once the gel has polymerized, the glass sandwich was transferred to a Hoefer SE250 Mighty Small II Mini-Vertical gel electrophoresis unit (Hoefer) and filled with 1X SDS running buffer [25 mM Tris-HCl, 200 mM Glycine, 0.1% SDS (w/v), (pH 8.3)]. Approximately 30 µl of total standard volume of linearized protein samples and 5 µl of peqGOLD Protein-Marker IV (prestained) (Peqlab) respectively, were loaded into the wells. Proteins were separated according to their electrophoretic ability by applying constant current at 20 mA per gel for 70 minutes or until dye has reached the bottom of the gel. Table 2.4 describes the function

of all constituents needed for SDS-PAGE. Following electrophoresis, gels can either be stained by Coomassie brilliant blue or subjected to Western Blotting for detection of proteins.

**Table 2.4 Function of all constituents used in SDS-PAGE**

| Reagent | Function |
|---|---|
| • Sodium dodecyl sulfate (SDS) | Gives proteins negative charge, denatures proteins |
| • ß-mercaptoethanol or dithiothreitol (DTT). | Reduce disulphide bridges in proteins |
| • Glycerol in loading buffer | Increase density |
| • Bromophenol Blue | Dye to monitor separation of sample through gel |
| • Acrylamide | Cross-linking agent for gels |
| • N-N-methylenebisacrylamide (Bis) | Cross-linking agent for gels |
| • Tetramethylethylenediamine (TEMED) | Free-radical stabilizer – Catalyze polymerization of gels |
| • Ammoniumpersulfate (APS) | Oxidizing agent - Catalyze polymerization of gels |



### *2.6.2.3 Coomassie Staining*

Coomassie Brilliant Blue R-250 staining solution (40% (v/v) methanol, 10% (v/v) acetic acid and 0.025% (w/v) Coomassie Brilliant Blue R-250) was used to fix proteins to the gel to prevent them from diffusing into aqueous solution as well as stain them to visualize and examine the proteins of interest. The gel was left to stain on a shaker

overnight at room temperature in the Coomassie staining solution. The gel was then put in Coomassie destain solution (50% (v/v) methanol and 10% (v/v) acetic acid), placed on a shaker at room temperature and washed several times to remove all excess dye. The dye doesn't bind to the acrylamide gel, but does fix to the proteins instead leaving a clear gel with deep blue protein bands. The gel was then preserved in cellophane sheeting after treatment with gel drying solution [40% methanol (v/v), 10% glycerol (v/v), 7.5% acetic acid (v/v)] and left to dry overnight.

### 2.6.3 Western Blotting of Proteins

The principle of Western blotting was first described by Towbin *et al*. in 1979 and describes the transfer of proteins resolved on SDS-PAGE onto a membrane made of either nitrocellulose or polyvinylidine difluoride (PVDF) using electric current that "blots" the proteins from the gel (Towbin, H., Staehelin, T., et al. 1979). Detection of the proteins of interest are facilitated by incubation of the blot with primary antibody and enzyme-linked secondary antibody that bind to specific amino acid sequences of the protein known as epitopes. To view the antibody-protein complex, a chromogenic substance that reacts with the enzyme-linked secondary antibody is added to cause a colour reaction.

Before the proteins, that were separated by SDS-PAGE, can be transferred to the nitrocellulose membrane, the membrane first needed to be equilibrated in 20% (v/v) methanol for nitrocellulose or 100% (v/v) methanol if using PVDF membrane. Proteins were then transferred by electrophoresis in a Hoefer T22 Mighty Small Transfer Tank (Hoefer) in transfer buffer (27 mM tris (hydroxymethyl)-amino-methane, 191 mM glycine, 20% (v/v) methanol) at 90 V for 90 minutes. The membrane was then stained with Ponceau stain [0.1% (w/v) Ponceau S, 5% acetic acid

(v/v)] for approximately 3 minutes and rinsed with sterile distilled water to confirm positive transfer. Blocking of the unoccupied sites was done by incubating the membrane in blocking solution (5% (w/v) milk, 0.05% tween 20 (v/v) in 1X PBS) for 30 minutes on a rocker at room temperature. This was followed by incubation of the membrane on a roller overnight at 4°C in a blocking solution (5% non-fat dry milk (w/v), 0.05% Tween-20 in 1X PBS) containing the primary antibody, rabbit anti-GST, at a dilution of 1:1000. The membrane was then washed with wash solution for 30 minutes on a rocker at room temperature. The membrane was then incubated again in 5% milk solution containing the goat anti-rabbit secondary antibody, labelled with horseradish peroxidase (HRP), in a dilution of 1:1000 at room temperature on a roller for 1 hour in the same manner as described previously. After the membrane was washed with wash solution, membrane peroxidase substrate, 3,3',5,5'-tetramethylbenzidine (TMB) (Sigma-Aldrich), was added to the membrane to determine the presence of GST fusion proteins by the production of coloured bands.



## 2.7 Purification of Recombinant Fusion Proteins

The MagneGST™ Protein Purification system (Promega) allows for the rapid purification of GST fusion proteins by utilizing paramagnetic particles that immobilizes the reduced glutathione and thus isolating the GST fusion protein from cell lysate leaving the unbound proteins to be washed away.

The MagneGST™ particles were equilibrated by washing it 3 times in 250 µl MagneGST™ binding/wash buffer (4.2 mM $Na_2HPO_4$, 2 mM $K_2HPO_4$, 140 mM NaCl, 10 mM KCl). By placing tubes on a magnetic stand, the particles were separated from the binding/wash buffer. The particles were then resuspended in 100 µl binding/wash buffer and 200 µl cell lysate and incubated on an orbital shaker for 1 hour at room

temperature. The GSH-bound GST fusion proteins were isolated in the magnetic stand, resuspended in 250 µl binding/wash buffer and incubated for 5 minutes at room temperature. The particles were then again captured in the magnetic stand and washed 3 times with 250 µl binding/wash buffer, after which the particles were resuspended in 200 µl elution buffer (50 mM Glutathione [pH 7.0-8.0), 50 mM Tris-HCl (pH 8.1)] and incubated at room temperature for 30 minutes on an orbital shaker. After capturing the particles, a final time in the magnetic stand, the supernatant containing the eluted GST fusion protein was subject to quantification using Quant-iT™ protein assay kit with the Qubit™ 1.0 Fluorometer (Invitrogen) to record the concentration of proteins in lysate.

### 2.7.1 Quantification of Purified Proteins

Protein quantification was performed in the same manner as previously described in 2.3.4 with the exception of incubating the assay tubes for 15 minutes before reading of the samples.



### 2.7.2 SDS-PAGE and Coomassie Staining of Purified Proteins

A 12.5% separation and 3% stacking gel was prepared for SDS-PAGE according to manufacturer's protocol as previously described in 2.6.2.2. The gel was then stained overnight in Coomassie Brilliant Blue for the identification of all proteins expressed in the cell as described in 2.6.2.3.

### 2.7.3 Western Blotting of Purified Proteins

Western blotting of purified proteins was performed following the same protocols previously described in 2.6.3. The membrane was incubated in 5% milk-PBS/Tween

solution containing 1:1000 primary, rabbit anti-GST antibody, followed by subsequent incubation with 1:5000 secondary HRP-conjugated goat anti-rabbit antibody.

## 2.8 Polyclonal Antibody Production

### 2.8.1 Peptide selection and synthesis

The sequence of the HCoV-NL63 N gene was obtained from the National Centre for Biotechnology Information (NCBI, GenBank) with accession number ABE97141.1 (Appendix 2). The sequence was used for the design and synthesis of peptides needed for antibody production. The sequences were subjected to several tests and algorithms to identify those sequences that are most antigenic and could serve as peptides for the detection of the putative protein. First, the sequence was used to identify antigenic determinants (Immunomedicine Group). In this study, an antigen profiler tool (Thermo Fischer Scientific) was used to identify which of the determinants are potentially excellent antigens. Following the identification of excellent antigens, the physiochemical properties of each selected sequence were determined (Pepcalc). From these results, only those sequences that were identified to have good water solubility were selected for further analysis. The average antigenic propensity was determined for each of the selected sequences and compared to the average antigenic propensity of that of the full-length N protein. These sequences (RKKFPPPSFYMPLLVSSD, aa residues 13-30 and PLEPKFSIALPPELSVVE, aa residues 122-139) were chosen as highly antigenic sites and labelled peptide 1 and peptide 2 respectively.

These sequences (peptide 1 and 2) were synthesized using conventional solid-phase chemistry and purified by Bio-Synthesis, Inc. The synthetic peptides were covalently

conjugated to a Lysine backbone to form a multiple antigen peptide (MAP)-8 (MAP, Sigma).

### 2.8.2 Animal Models

Specific pathogen free Balb/c female mice, purchased from the University of Cape Town's Animal Unit (Cape Town, South Africa), were used as animal models after obtaining animal ethical clearance from the University of the Western Cape. All experiments were conducted in accordance with the guidelines of the institutional animal ethics committee. The mice were kept in a well-ventilated, temperature controlled environment under a twelve-hour day/night cycle and given standard mouse feed and drinking water *ad libitum* (Medical Research Council, Cape Town, South Africa).

### 2.8.3 Immunization

Immunization was performed both *in vivo* and *in vitro*.



#### *2.8.3.1 In vivo Immunization*

Initial immunisation was performed *in vivo* via intraperitoneal injection of two female Balb/c strain mice with a standard dose mixture of 250 µg (5 mg/ml) of each peptide (Biosynthesis Inc.) combined with 150 µl 1X PBS and emulsified in 100 µl Complete Freund's Adjuvant (CFA) per mouse. Each mouse was immunised with 200 µl of the antigen and CFA emulsion. A total of two booster immunizations was administered with the same amount of peptide emulsified in Incomplete Freund's Adjuvant (IFA) every two weeks.

### *2.8.3.2 Antibody titre*

The antibody titre was monitored throughout the immunization process. A total volume of 20 µl of blood was collected from the tail vein and resuspended in 180 µl 1X PBS on the day of the first immunization and then with each subsequent immunization. Blood plasma and red blood cells were separated by centrifugation at 12,000 rpm for 5 minutes. The supernatant was collected and stored at -20°C for further use. Thereafter, all the samples were screened for antibody titre with an indirect ELISA. The animal with the highest antibody titre were used for polyclonal antibody production. The animal was asphyxiated with carbon dioxide prior to the procedure and sacrificed by cervical dislocation. No anaesthetics could be used due to its effect on the spleen cells which was required for monoclonal antibody production. The spleen was harvested and macerated after which it was prepared in medium, centrifuged and erythrocytes removed by centrifugation at 14,000 rpm.



## 2.8.4 Indirect Enzyme-Linked Immunosorbent Assay (ELISA) for Screening of Mouse Serum for Production of Polyclonal Antibodies (Antibody Titre).

Each well of a Nunc™ Maxisorp™ Clear Flat-Bottom Immuno Nonsterile 96-well plate (Thermo Fisher Scientific) was coated with 50 µl of each peptide (1 µg/ml in 1X PBS) and incubated for two hours at room temperature or at 4°C overnight. Microtiter plates were washed once with washing buffer (0.1% Tween-20 (v/v) in 1X PBS), followed by blocking of unoccupied sites with 200 µl of 5% milk solution on a shaker for 1 hour at room temperature. The plates were decanted and washed five times. Blood serum samples were diluted in 1X antibody dilution buffer (0.1% of 5% non-fat dry milk in

1X PBS) to reduce non-specific binding, after which 100 µl of diluted serum was serially diluted in a 96-well plate and incubated on a shaker for 1 hour at room temperature. This was followed by washing of the plates five times with wash buffer. The secondary antibody, goat anti-mouse conjugated to HRP (Sigma Aldrich), was diluted 1:2000 in antibody dilution buffer of which 50 µl were added per well and incubated for 1 hour at room temperature on a shaker. The plates were washed seven times with wash buffer after which 50 µl SureBlue™ TMB Microwell Peroxidase Substrate (KPL, Whitehead Scientific), was added to each well. The plates were incubated for 15 minutes in a dark place at room temperature. The enzyme reaction was stopped by the addition of 50 µl stop solution [2 M sulphuric acid ($H_2SO_4$)] to enable accurate measurement of intensity at 450 nm using the plate reader [Ascent Software version 2.6 (Original Multiskan EX, Type 355, Thermo Electron Corporation, China)].

### 2.8.5 *In vitro* Immunization



*In vitro* immunisation was performed on one half of the spleen which received a booster immunization. The other half of the spleen was used for fusion with the Sp2/0-Ag14 murine myeloma cell line (Sigma Aldrich). The one half of the spleen was homogenized and made up in medium consisting of RPMI-1640 (contain no pyruvate) or DMEM (contain no asparagine), 1X Glutamax, 1X antibiotic/antimycotic, 50 µg/ml Gentamicin, Sodium Pyruvate, and 1% mercaptoethanol. Peptides (1 µg/ml) was added to the suspension and incubated for 30 minutes at 37°C, 5% carbon dioxide, and 95% humidity. After incubation, the spleen culture was made up to 50 ml medium (as above) supplemented with heat inactivated foetal bovine serum (HI-FBS) and

incubated for three days under the conditions described above to allow for the antigens to bind to B-lymphocytes. Fusion was performed after the three day incubation.

## 2.9 Monoclonal Antibody Production

### 2.9.1 Fusion of Myeloma Cells with Splenic B-lymphocytes

The Sp2/0-Ag14 murine myelomas is a non-producer cell line used as fusion partner for splenic B-lymphocytes. Sp2 myelomas were cultured in full medium (Ex-Cell 610-HSF (SAFC Biosciences) supplemented with 1X antibiotic/antimycotic (PAA), 1X Glutamax (Sigma Aldrich), 50 µg/ml Gentamicin (Sigma Aldrich), 10% HI-FBS and sodium pyruvate. Incubation was at 37°C, 5% carbon dioxide, and 95% humidity until 70-80% confluency has been reached.

The half of the spleen being used for fusion was macerated, centrifuged at 1,500 rpm for 10 minutes and resuspended in wash buffer [1X Dulbecco's PBS (dPBS) (Whitehead Scientific) supplemented with 1X antibiotic/antimycotic, and 1X Glutamax)]. A cell count of splenic B-lymphocytes and Sp2 myelomas were performed using a haemocytometer (Marienfeld, Germany). The Sp2 myelomas were added to the suspension of spleen cells in a 1:2 ratio and made up to 50 ml with wash buffer. The suspension was centrifuged at 1,000 rpm for 10 minutes and the wash buffer decanted to retain the cell pellet. The pellet was resuspended over 1 minute in 1 ml of Polyethylene Glycol (PEG) (Sigma Aldrich) and incubated for 1 minute at 37°C. The addition of PEG will result in higher fusion frequency and greater reproducibility (Carter 1996). Two mililitres of serum free medium (RPMI-1640 (Sigma Aldrich) containing 50 µg/ml Gentamicin, 1X antibiotic/antimycotic, and 1X Glutamax) was added drop-wise over 2 minutes. This was followed by the drop-wise

addition of 2 ml serum free medium over 1 minute. The next 2 ml of serum-free medium was added over 30 seconds. Serum free medium was then added to a final volume of 50 ml. The suspension was centrifuged at 1,000 rpm for 10 minutes and the supernatant was removed to retain the cell pellet. The pellet was resuspended in 50 ml of full medium supplemented with Hypoxanthine-Aminopterin-Thymidine (HAT) (Sigma Aldrich), 1% mercaptoethanol (Sigma Aldrich) and sodium pyruvate (Sigma Aldrich) shortly after fusion and plated in four 96-well tissue culture treated microtiter plates (Nunc, Denmark). The plates were incubated for 72 hours. After 72 hours, HAT supplemented full medium was replaced by full medium supplemented with Hypoxanthine-Thymidine (HT) (prepared same as HAT media) (Sigma Aldrich) and the hybridomas incubated for another 72 hours. The use of the selective media, HAT will allow the growth of hybridomas and kill any unfused myeloma cells. This process was repeated until there was sufficient growth and colonies were macroscopic. Once macroscopic, hybridoma colonies were selected for screening of monoclonal antibody production by indirect ELISA (as previously described 2.8.4).



The data was compared and linear regressions were produced using Microsoft Excel™ computer software and correlations determined by calculating the Spearman's Correlation Coefficient.

### 2.9.2 Screening of Hybridoma Colonies

The culture supernatant of hybridomas were screened for antigen affinity (as described in the ELISA procedure in 2.8.4) to identify potential colonies to be used for cloning. Pre-immune mouse plasma was used as positive control and post-immune mouse plasma was used for negative control at the same concentrations. All recorded

absorbance readings higher than the positive control, were selected for further cloning to produce monoclonal hybridoma colonies.

### 2.9.3 Cloning

After screening, cloning of hybridomas was performed on the wells which generated the highest antibody titre in a plate. The objective of cloning was to obtain a single hybridoma cell per well in a 96-well microtiter plate. Cells were resuspended in medium. A cell count was done using a haemocytometer. The cells were diluted to give one cell per 100 µl. From the dilution, the cells were resuspended in HT supplemented full medium. A 96-well microtiter plate was then seeded with 100 µl per well HT supplemented full medium containing cells. After 96 hours, another 50 µl of HT supplemented full medium was added to each well. Another 50 µl of HT supplemented medium was added to each well every 72 hours until colonies were once again macroscopic and ready for screening by ELISA (as previously described in 2.8.4).



The same method was used as previously described in 2.9.2 with selection of hybridoma colonies, where all values higher than the positive control would be cloned further to produce monoclonal antibodies.

## 2.10 Characterization of Generated Polyclonal Antibodies

### 2.10.1 Western Blot analysis using Generated Polyclonal Antibodies against Synthetic Peptides.

Western blotting of purified proteins (SARS-N, N1, N2 and N3) with mouse blood plasma containing polyclonal antibodies generated against HCoV-NL63 N peptides,

was performed following the same protocol previously described in 2.6.3. The membrane was incubated in 5% milk-PBS/Tween solution containing 1:1000 primary antibody, mouse polyclonal antiserum (containing anti-peptide 1 and anti-peptide 2 antibodies), followed by subsequent incubation with secondary HRP-conjugated goat anti-mouse antibody at a dilution of 1:3000.

# CHAPTER 3

# RESULTS


Department of Medical Biosciences, Faculty of Natural Sciences, University of the Western Cape, South Africa

# CHAPTER 3 RESULTS

## 3.1 Polymerase Chain Reaction (PCR) of Recombinant Constructs

Plasmid DNA of SARS-CoV and HCoV-NL63 N recombinant gene constructs (SARS-N, N1, N2 and N3) that were isolated and purified by miniprep were quantified by Qubit® fluorometry to determine plasmid concentrations (Table 3.1). This was followed by the amplification of the purified plasmid DNA by PCR. The gene sizes, as determined by the accession number from NCBI, were as follows: SARS-CoV N: 1269 bp; N1: 1134 bp; N2: 567 bp and N3: 566 bp. Following PCR, the products were run on a 1% (w/v) agarose gel and viewed under UV light (Figure 3.1). The N genes were identified as the correct size (Table 3.2) and subsequent confirmation could be performed.

**Table 3.1 Concentrations of purified plasmid DNA (miniprep).**

| Construct | Concentration (µg/ml) |
|---|---|
| **SARS-N** | 0.05 |
| **N1** | 0.04 |
| **N2** | 0.06 |
| **N3** | 0.05 |

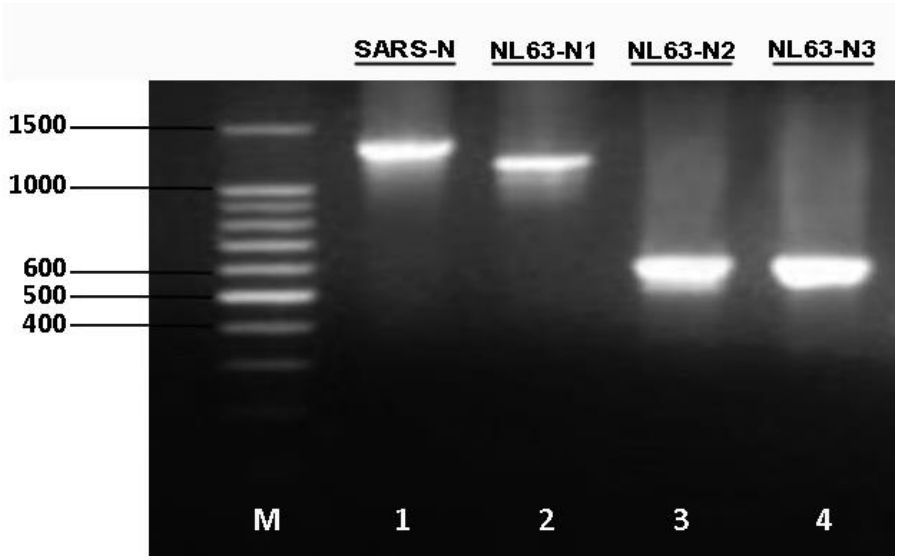



**Figure 3.1 PCR amplification of purified plasmid DNA.** Lane M: 100 bp DNA marker. Lane 1: SARS-CoV N gene (SARS-N) (± 1200 bp); Lane 2: HCoV-NL63 full-length N gene (N1) (± 1100 bp); Lane 3: 5' half of HCoV-NL63 N gene (± 550 bp) coding for the N-terminal fragment of the N gene (N2); Lane 4: 3' half of the HCoV-NL63 N gene (±550 bp) coding for the C-terminal fragment of the N gene (N3).



**Table 3.2 Description of heterologously expressed proteins.**

| Gene Name | Nucleotide Size (bp) | Protein Size (aa) | Molecular Weight (kDa) | Fusion Protein (kDa) |
|---|---|---|---|---|
| **SARS-N** | 1269 | 422 | 46.05 | 72.05 |
| **N1** | 1134 | 377 | 42.60 | 68.60 |
| **N2** | 567 | 189 | 21.51 | 47.51 |
| **N3** | 566 | 187 | 20.90 | 46.90 |

## 3.2 Restriction endonuclease digestion (Miniprep) with Flexi™ Enzymes (*SgfI and PmeI*).

The presence of the genes of interest in the recombinant constructs were verified by a single restriction endonuclease digestion of the purified plasmid DNA sample, obtained from miniprep, to separate HCoV-NL63 N and SARS-CoV N gene fragments from the Flexi™ vector using specific Flexi™ restriction enzymes, *SgfI* and *PmeI*. The size fractionation of the digest was resolved on a 1% agarose gel electrophoresis and viewed under UV light (Appendix 3). The results confirmed the presence of SARS-N (~1200 bp), N1 (~1100 bp), N2 (~ 600 bp) and N3 (~ 600 bp) genes on the pFN2A GST Flexi™ vector and were identified by their correct sizes (Table 3.2). The vector (4 kb) is presented on top of the figure in each lane of the agarose gel (Appendix 1).

## 3.3 Restriction endonuclease digestion (Midiprep) with Flexi™ Enzymes (*SgfI* and *PmeI*).



Following the small-scale isolation and verification of the HCoV-NL63 and SARS-CoV N genes from the Flexi™ vector, large-scale cultivation could be performed and the isolation of plasmid DNA with midiprep to produce higher yield per BM index. The presence of the genes of interest in the recombinant constructs were verified by a single restriction endonuclease digestion of the purified plasmid DNA sample, obtained from midiprep with specific Flexi™ restriction enzymes, *SgfI* and *PmeI*. The size fractionation of the digest was resolved on a 1% agarose gel electrophoresis and viewed under UV light (Figure 3.2). The results confirmed the presence of SARS-N (~1200 bp), N1 (~1100 bp), N2 (~ 600 bp) and N3 (~ 600 bp) genes on the pFN2A

GST Flexi™ vector and were identified by their correct sizes (Table 3.2). The vector (4 kb) is presented on top of the figure in each lane of the agarose gel (Appendix 1).

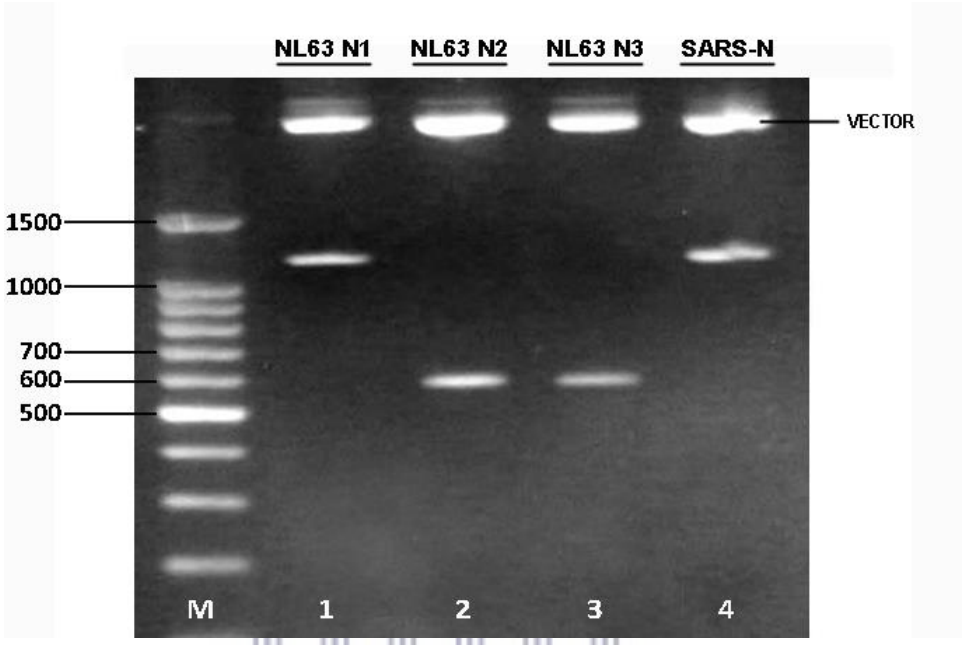



**Figure 3.2 Restriction enzyme digestion of SARS-CoV and HCoV-NL63 N plasmid DNA from midiprep lysate.** Lane M: 100 bp DNA ladder. Lane 1: HCoV-NL63 full-length N gene (N1) (~1100 bp); Lane 2: N-terminal fragment of HCoV-NL63 N gene (N2) (~550 bp); Lane 3: C-terminal fragment of HCoV-NL63 N gene (N3) (~550 bp); Lane 4: SARS-CoV full-length N gene (SARS-N) (~1200 bp).



## 3.4 Analysis of Recombinant Proteins Expressed through Autoinduction.

The expression protocol included sub-culturing of overnight starter cultures into an expression culture in a dilution of 1:100. Cultures were autoinduced with 0.1% (w/v) rhamnose and 0.05% (w/v) glucose immediately after dilution and incubated for

approximately 6 hours. After incubation, cells were lysed and total proteins were extracted from *E.coli* cells using protocols previously described in 2.6.1. Heterologous proteins were separated on 15% SDS-PAGE and subsequently transferred to a nitrocellulose membrane for Western Blotting (Figure 3.3). Expression of total proteins were detected by incubating the blot with specific rabbit anti-GST primary antibody and goat anti-rabbit secondary antibody. As depicted in Table 3.2, total proteins were confirmed to be of correct estimated sizes: SARS-N (~72 kDa), N1 (~69 kDa), N2 (~48 kDa) and N3 (~ 47 kDa). From the results, it is evident that the technique is highly specific to the GST tag that was fused to the proteins of interest.

## 3.5 Purification of Expressed Recombinant Proteins

The GST affinity tag used in this is study is a naturally occurring 26 kDa protein found in the parasite *Schistosoma japonicum* and aid in the purification of recombinant proteins. Protein purification was done using the MagneGST™ Protein Purification system that allows GST-tagged proteins to be eluted from crude cell lysate by addition of glutathione. The supernatant containing the eluted GST fusion proteins were quantified with the Qubit® fluorometer system to record the concentrations of proteins in the purified cell lysate. The results are depicted in Table 3.3.

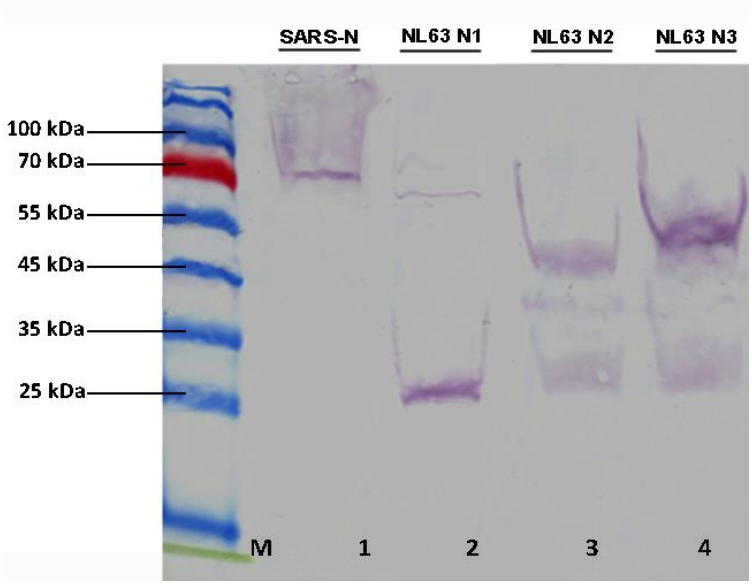



**Figure 3.3 Detection of GST-tagged SARS-CoV N and HCoV-NL63 N proteins and its truncates by Western Blot.** Extracted proteins were separated on SDS-PAGE and subjected to Western blotting for the detection of expressed GST-tagged fusion proteins with specific GST primary antibody and appropriate secondary antibody. Lane M: Pre-stained molecular weight protein marker. Lane 1: Expression of unpurified GST-SARS-N (~72 kDa); Lane 2: Expression of unpurified GST-NL63-N1 (~69 kDa); Lane 3: Expression of GST-NL63-N2 (~48 kDa); Lane 4: Expression of unpurified GST-NL63-N2 (~47 kDa).

Following quantification, the purified recombinant proteins were resolved on 15% SDS-PAGE gel and stained with Coomassie brilliant blue to visualize and confirm expression of purified recombinant proteins. The expression of purified SARS-N and HCoVNL63 N1, N2 and N3 could further be verified by subsequent Western blotting and incubation with anti-GST primary antibody and appropriate secondary antibody. The correct sizes for the purified fusion proteins SARS-N (~72 kDa), N1 (~69 kDa), N2 (~48 kDa) and N3 (~47 kDa) could be confirmed in Figure 3.4.

**Table 3.3 Concentrations of purified total proteins in mg/ml**

| Construct | Concentration (mg/ml) |
|---|---|
| **SARS-CoV N-GST** | 0.26 |
| **N1-GST** | 0.11 |
| **N2-GST** | 0.09 |
| **N3-GST** | 0.17 |

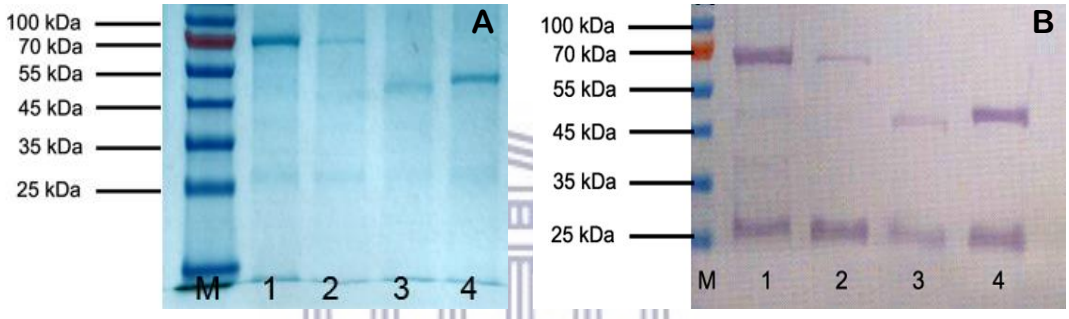



**Figure 3.4 Detection of purified recombinant proteins using (A) Coomassie stained SDS-PAGE gel and (B) Western Blot.** Proteins were purified using the MagneGST™ purification system and subjected to SDS-PAGE (Coomassie stained) and Western Blot. Lane M: Pre-stained molecular weight protein marker; Lane 1: Purified GST-SARS-N; Lane 2: Purified GST-N1; Lane 3: Purified GST-N2; Lane 4: Purified GST-N3. Lane order is maintained for B.

## 3.6 Polyclonal Antibody Production against HCoV-NL63 Peptides

Peptides for HCoV-NL63 N protein were synthesized and used to immunize mice for the production of antibodies to viral proteins. Immunization of mice was performed using a mixture of the peptides. ELISA was performed to evaluate the antibody titers in mouse plasma. Spleen cells were harvested following a booster injection. Spleen and myeloma cells were fused and screened for antibody-producing hybridoma colonies. Colonies were selected and cloned.

### 3.6.1 Peptide Design & Synthesis

#### *3.6.1.1 Antigenic Determinants*

The full sequence of HCoV-NL63 N protein (Appendix 2) was submitted to an Antigenic Peptide Prediction Tool (Immunomedicine Group) that follows the Kolaskar and Tongaonkar (1990) method (Kolaskar, A.S. and Tongaonkar, P.C. 1990). The full sequence of HCoV-NL63 N protein is 377 bases long with an average antigenic propensity of 1.0123 (Figure 3.5). A total of thirteen residues were identified to be potentially antigenic residues as depicted in Table 3.4.



#### *3.6.1.2 Antigen Profile*

All thirteen potentially antigenic residues were scored for their antigenicity profile with the Thermo Fischer Antigen Profiler Peptide Tool (Thermo Fischer Scientific). The antigenicity of each sequence is rated on a scale between <1.0 and 5.0. From the thirteen residues profiled, only six residues were rated as potentially excellent peptides. These were highlighted in Table 3.4.

### *3.6.1.3 Water Solubility Profile*

These six residues were submitted on the Pepcalc Peptide property calculator that makes calculations and estimations on physiochemical properties of a given sequence (results not shown). The most important factor of interest is that of water solubility. It also provides a peptide hydropathy plot (Figure 3.6) based on the index used by Hopp & Woods (1981) (Hopp, T.P. and Woods, K.R. 1981). From the six residues tested, four were identified to have good water solubility (Table 3.4).

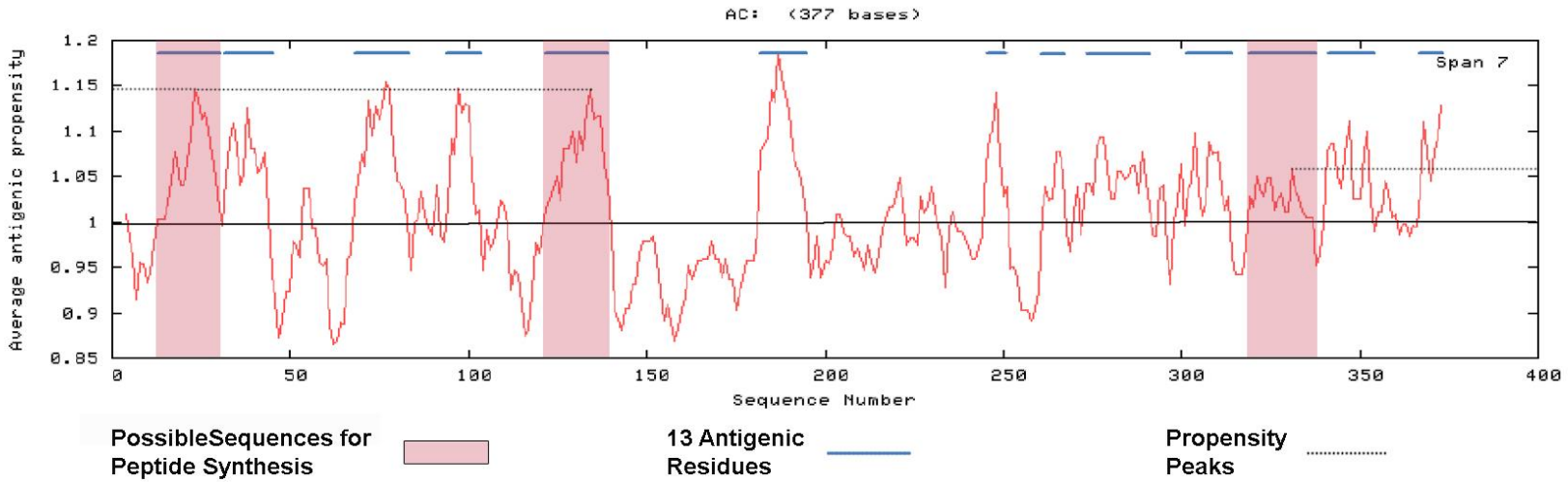



**Figure 3.5 Antigenic propensity plot for HCoV-NL63 N protein.** Sequence is 377 bases long. Thirteen residues were identified to be possibly antigenic. Plot indicates peaks higher than 1 which indicates propensity towards the surface. The average antigenic propensity for HCoV-NL63 N is 1.0123. The highlighted regions indicate sequences that could possibly be selected for peptide synthesis. The first two highlighted sections were chosen due to their peaks in antigenicity and their propensity was the highest at 1.15 respectively compared to the third selection at 1.05.

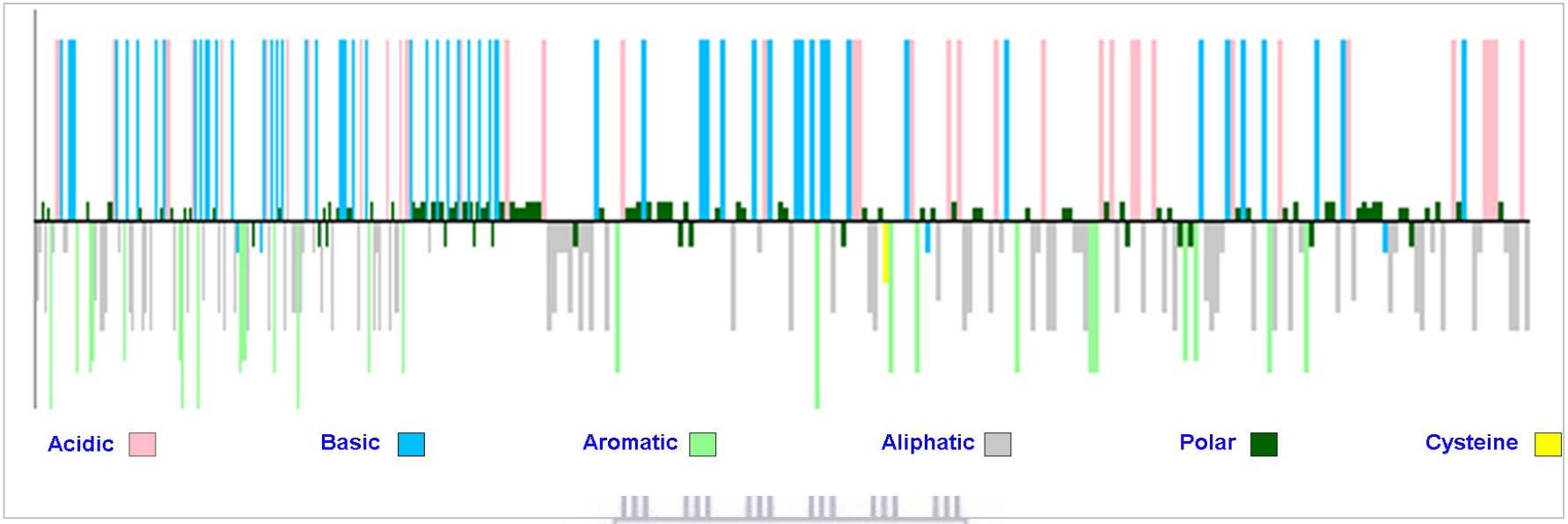



**Figure 3.6 Hydropathy plot of HCoV-NL63 N protein sequence.** Top bars represent hydrophilic residues and bottom bars represent hydrophobic residues in the complete sequence of HCoV-NL63 N protein. Plot generated using algorithm adapted from Hopp and Woods (Hopp, T.P. and Woods, K.R. 1981).

**Table 3.4 Summary of antigenic properties of HCoV-NL63 N protein residues.**

| | *n* | Sequence | Position of amino acids | Number of Residues | Antigen Profile[3] Score | Water Solubility[4] | Average Antigenic Propensity | Antigenic Propensity Peaks |
|---|---|---|---|---|---|---|---|---|
| ** | 1 | RKKFPPPSFYMPLLVSSD | 13 – 30 | 17 | 3.8 | Good | 1.0524 | 1.15 |
| | 2 | APYRVIPRNLVPI | 32 – 44 | 12 | 2.5 | - | - | - |
| * | 3 | QRVDLPPKVHFYYLGT | 68 – 83 | 15 | 3.8 | Poor | - | - |
| * | 4 | RLDGVVWVAK | 94 – 103 | 9 | 3.0 | Good | 1.1077 | - |
| ** | 5 | PLEPKFSIALPPELSVVE | 122 – 139 | 17 | 3.6 | Good | 1.1021 | 1.15 |
| | 6 | SSDLVAAVTLALKN | 182 – 195 | 13 | 1.5 | - | - | - |
| | 7 | NVIQCFG | 245 – 251 | 6 | 1.4 | - | - | - |
| | 8 | DSDLVQN | 261 – 267 | 6 | 2.1 | - | - | - |
| * | 9 | GFPQLAELIPNQAALFFDS | 273 – 291 | 18 | 3.0 | Poor | - | - |
| | 10 | VQITYTYKMLVA | 302 – 313 | 11 | 1.8 | - | - | - |
| * | 11 | LPKFIEQISAFTKPSSVKE | 319 – 337 | 16 | 3.5 | Good | 1.0402 | 1.05 |
| | 12 | QSSHVAQNTVLNA | 341 - 353 | 12 | 1.0 | - | - | - |
| | 13 | AIIEIVN | 367 - 373 | 6 | 1.4 | - | - | - |

[3] Antigen profiler peptide tool uses an algorithm to determine antigenicity of your sequence and scores it on a scale from poor (<1) to excellent (5.0). An excellent antigen ranges from 2.7 – 5.0.

[4] Profile of water solubility is only an estimate. Further algorithm programs should be used to determine water solubility with more accuracy.

* Six sequences with the highest antigen profile scores making them potentially excellent peptides.

** Two sequences selected from the six sequences identified to be excellent peptides to be used for antibody production due to their good water solubility and high peaks of antigenic propensity.

#### *3.6.1.4 Average Antigenic Propensity*

The average antigenic propensity for each of the selected sequences that can function as potentially excellent peptides with high water solubility, was determined and compared to the average antigenic propensity of the full-length N protein.

Three possible sequences were identified: RKKFPPPSFYMPLLVSSD (aa 13-30), PLEPKFSIALPPELSVVE (aa 122-139) and LPKFIEQISAFTKPSSVKE (aa 319-337) (Table 3.4). All three had relatively good antigenic propensity values of 1.0524 (Appendix 4), 1.1021 (Appendix 5) and 1.0402 (Appendix 6) respectively, but it was the peaks of the first two sequences at ~1.15 each that resulted in those two residues being chosen for synthesis of peptides that will be used to immunize mice and screen for antigenic immune responses.

### 3.6.2 ELISA Screening for Polyclonal Antibody Titre.

Indirect ELISA was performed for an initial screening to determine the antibody titre produced by the peptides (peptides 1 and 2) that was injected into two mice. Absorbance readings from mouse serum taken pre-immunization and with booster immunizations, post-immune 1 and post-immune 2, was recorded on a graph using Microsoft Excel computer software.

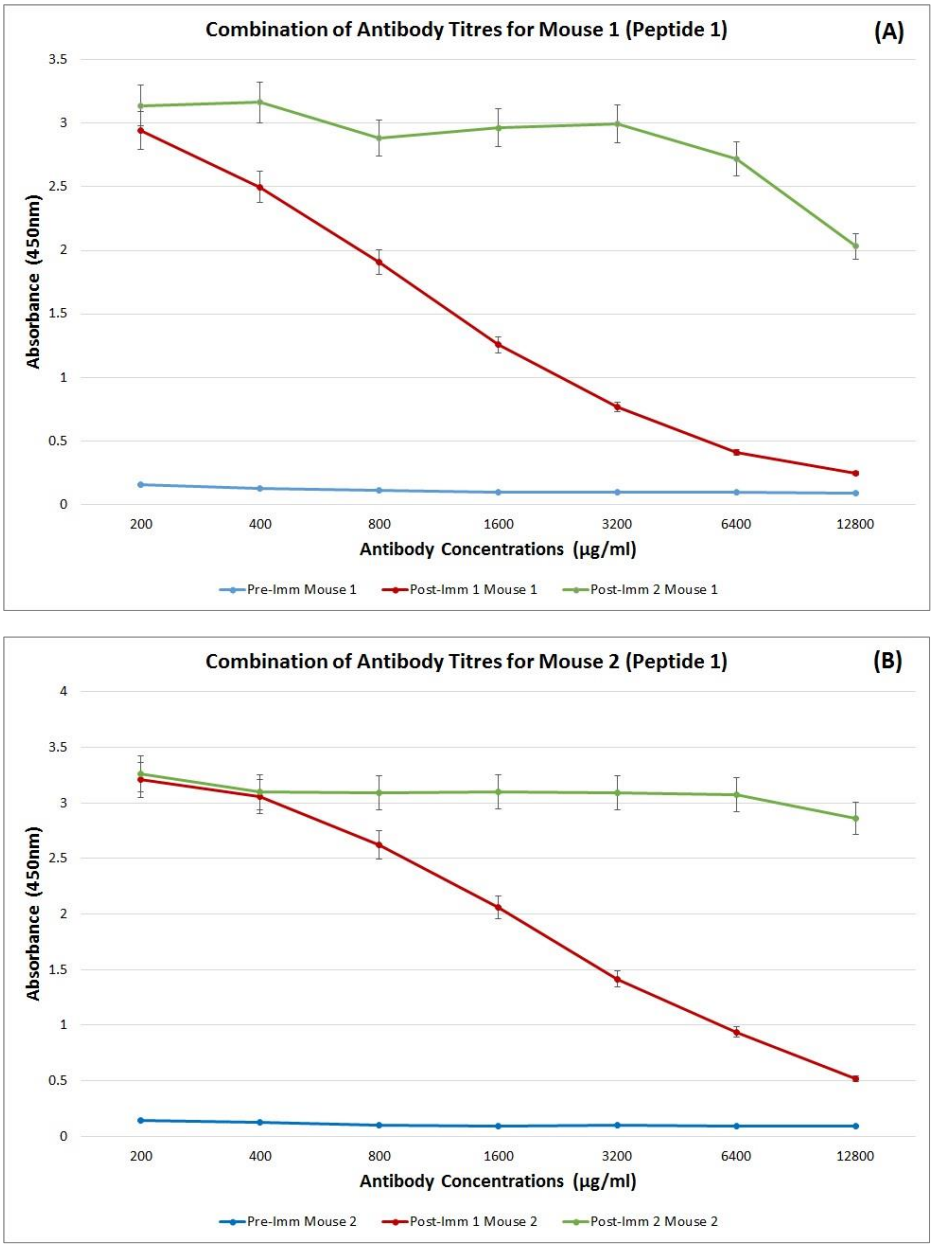
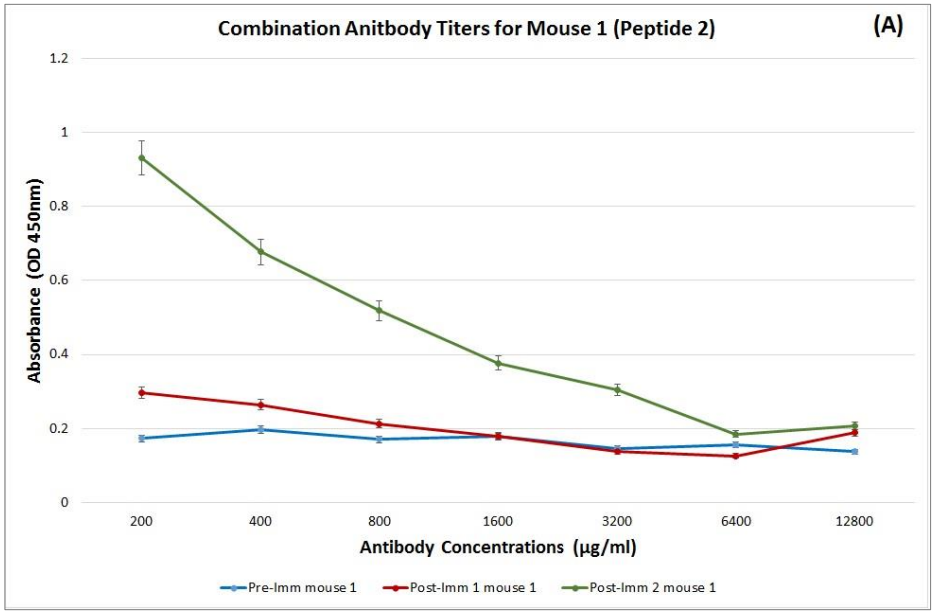



**Figure 3.7 Combined dilution curves of antibody titres with peptide 1.** Dilution curves represent plasma antibody titres of pre-immunization (blue), 1st post-immunization (red) and 2nd post-immunization (green) absorbance readings from both mouse 1 and 2 immunized with peptide 1. The *x*-axis represents the concentration of serum polyclonal antibodies at a 1/200 dilution. The *y*-axis represents absorbance values at an OD of 450nm. Error bars represent 5% error.

Figure 3.7 represent the reaction of mouse 1 and 2 to peptide 1. Both mice responded well after the first and second immunizations. The mice showed relatively the same response after the first immunization, but after the second immunization, mouse 2 had a high and prolonged antibody titre in response to peptide 1 (Figure 3.7B), whereas the antibody titre of mouse 1 gradually reduced (Figure 3.7A). Figure 3.8 represent antibody titres of both mice in response to peptide 2. Both mice responded equally to peptide 2 with a sharp decrease in response after the second post-immunization.

The first and second post-immunization responses for both mouse 1 and 2 against peptide 1 (Figure 3.9A) and peptide 2 (Figure 3.9B) is shown in combination. The post-immune 1 polynomial graphs for mouse 1 and 2 correlate with co-efficient $R^2$-values of ~0.98 respectively. The post-immune 2 polynomial graphs for mouse 1 and 2 show correlation co-efficient $R^2$-values of ~0.76 for mouse 1 and ~0.95 for mouse 2. In response to peptide 2 (Figure 3.9B), the post-immune polynomial graphs for mouse 1 show correlation co-efficient $R^2$-values of ~0.86 for post-immune 1 and ~0.93 for post-immune 2. The post-immune polynomial graphs for mouse 2 show correlation co-efficient $R^2$-values of ~0.85 for post-immune 1 and ~0.96 for post-immune 2.



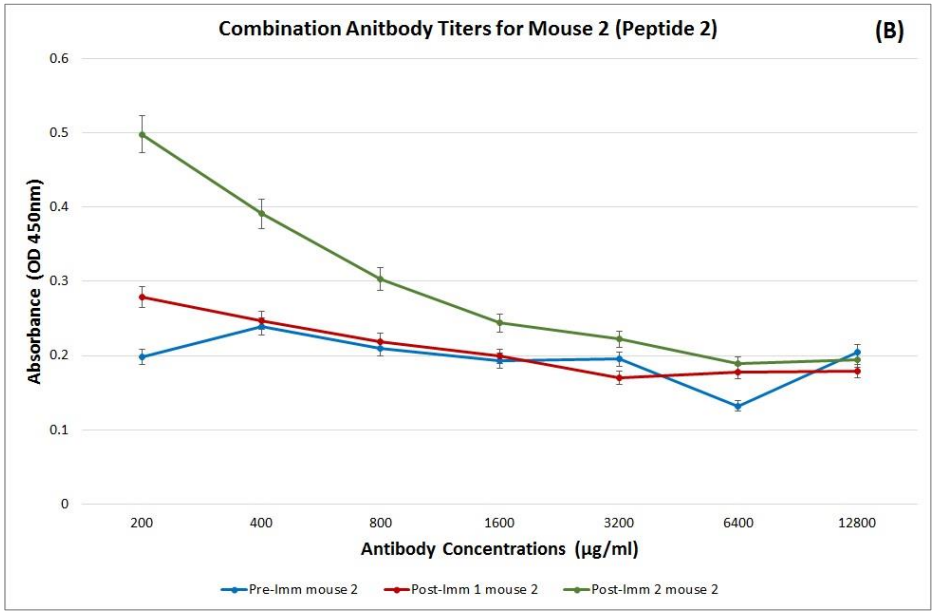
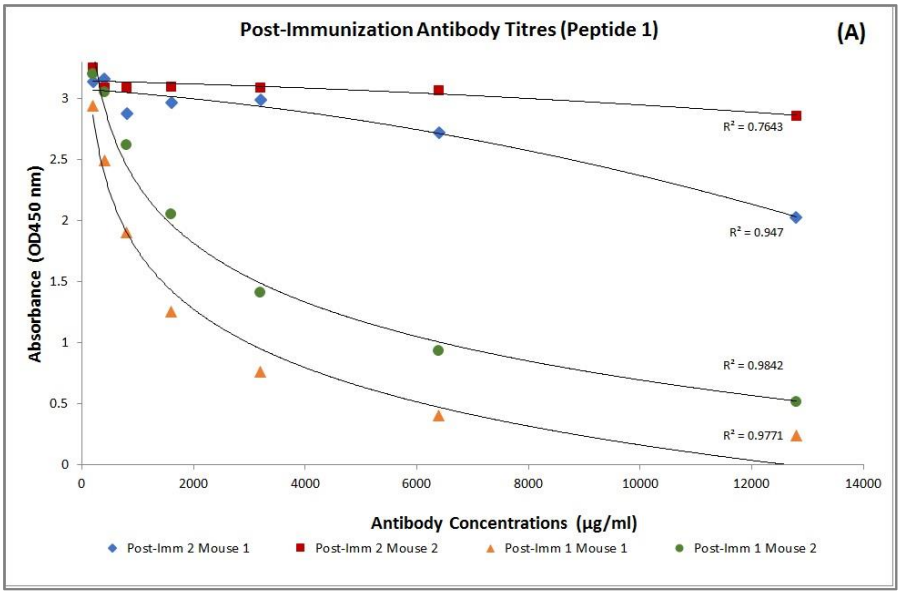




**Figure 3.8 Combination dilution curves of antibody titres to peptide 2.** Dilution curves represent plasma antibody titres of pre-immunization (blue), 1st post-immunization (red) and 2nd post-immunization (green) absorbance readings from both mouse 1 and 2 immunized with peptide 1. The *x*-axis represents the concentration of serum polyclonal antibodies at a 1/200 dilution. The *y*-axis represents absorbance values at an OD of 450nm. Error bars represent 5% error.



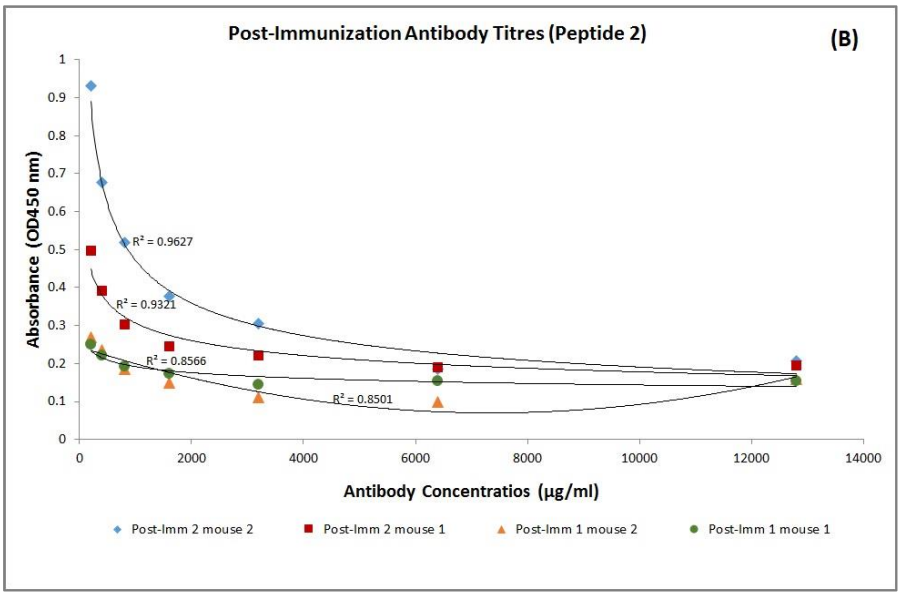




**Figure 3.9 Dilution curves of post-immunization titres with peptide 1 and 2.** Dilution curves represent 1st post-immunization of mouse 1 plasma (orange triangles), 1st post-immunization of mouse 2 plasma (green circles), 2nd post-immunization of mouse 1 plasma (red squares) and 2nd post-immunization of mouse 2 plasma (blue diamonds) absorbance readings with peptide 1. The *x*-axis represents the concentration of serum polyclonal antibodies at a 1/200 dilution of 100µl. The *y*-axis represents absorbance values at an OD of 450 nm. The $r^2$-values represent the correlation co-efficient of each of the dilution curves.

### 3.6.3 Hybridoma Screening

Myeloma cells were fused with splenic B-lymphocytes extracted from the spleen of mouse 1 (producing highest antibody titre). The supernatant with hybridomas producing antibodies against peptide 1 (Figure 3.10) and peptide 2 (Figure 3.11) were screened for antigen affinity. Absorbance readings were measured using the corrected values from the microtiter plate reader. Pre-immune serum was used as the negative control and post-immune 2 serum was used for the positive control. Measured absorbance readings were recorded on an Excel Spreadsheet (Microsoft Corporation). All values that exceeded that of the positive control was assumed indicative of a positive response in producing antibodies and were thus chosen as hybridoma colonies for further cloning.

**Table 3.5 Summary of ELISA screening of hybridoma culture supernatants.**

| Hybridomas (PAb) | Typical OD 450 (nm) | | Specificity to Antigens |
|---|---|---|---|
| | Peptide 1 | Peptide 2 | |
| **1D6** | | 1.787 | Specific |
| **2A1** | 1.516 | 0.301 | Non-specific |
| **2E4** | | 0.762 | Specific |
| **3A6** | | 0.779 | Specific |
| **3G4** | 1.473 | | Specific |
| **3G7** | 1.619 | | Specific |
| **4F7** | | 0.767 | Specific |
| **4F8** | 2.053 | | Specific |
| **4G4** | 2.131 | | Specific |
| **4G8** | 2.277 | 0.586 | Non-specific |

A total of 22 wells contained colonies that produced antibody titres higher than that of the positive control against peptide 1 (Figure 3.10). Only five colonies were producing antibodies above that of the positive control in response to peptide 2 (Figure 3.11). In accordance to the above calculations, all results were recorded at OD450 nm and are depicted in Table 3.5. According to the ELISA results obtained, ten wells were selected for cloning to contain hybridomas secreting antibodies against plasma from antigen-treated mice (in response to peptide 1 and 2). Specificity of antibodies to the respective peptides are also depicted in the table and indicates a cross-reactivity between peptides which in turn is indicative of non-specific antibody production to either peptide. This cross-reactivity was observed in two of the wells (hybrids 2A1 and 4G8).



**Table 3.6 Summary of second ELISA screening of hybridoma culture supernatant after cloning.**

| Hybridomas (PAb) | Typical OD450 (nm) | |
|---|---|---|
| | **Peptide 1** | **Peptide 2** |
| **1D6** | N/A | N/A |
| **2A1** | 0.74 | 0.12 |
| **2E4** | 0.09 | 0.14 |
| **3A6** | 0.07 | 0.17 |
| **3G4** | 2.75 | 0.14 |
| **3G7** | 0.09 | 0.05 |
| **4F7** | 0.14 | 0.08 |
| **4F8** | 2.22 | 0.05 |
| **4G4** | 1.04 | 0.05 |
| **4G8** | 1.76 | 0.13 |

The chosen ten hybridoma colonies were cloned and screened a second time by ELISA (Table 3.6) for colonies possibly producing monoclonal antibodies. In Figure 3.12 (B), only five colonies were producing antibody titres higher than the positive control in response to peptide 2 (2E4, 3A6, 3B9, 3G4 and 4G8). No colonies were producing antibody titres to peptide 1 exceeding that of the positive control (Figure 3.12A).

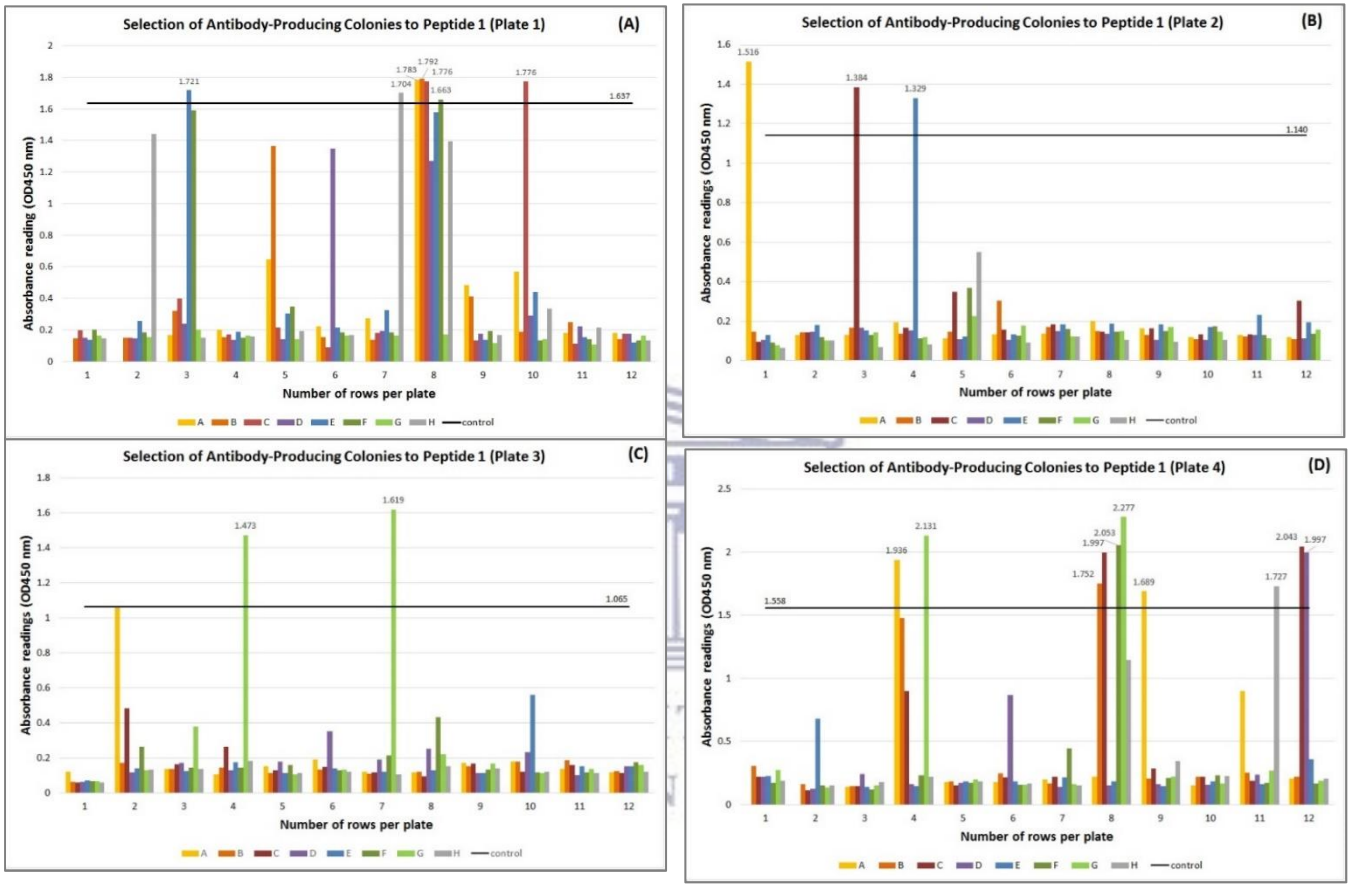



**Figure 3.10 Selection of hybridoma producing colonies against peptide 1.** Supernatant of fused myeloma cells were screened with ELISA for hybridomas producing antibodies against peptide 1 and the absorbance readings recorded at OD450nm. The data is presented as percentage and all values higher than 100% represent that of antibody-producing hybridoma colonies (labelled bars). Coloured bars represent columns for each plate.

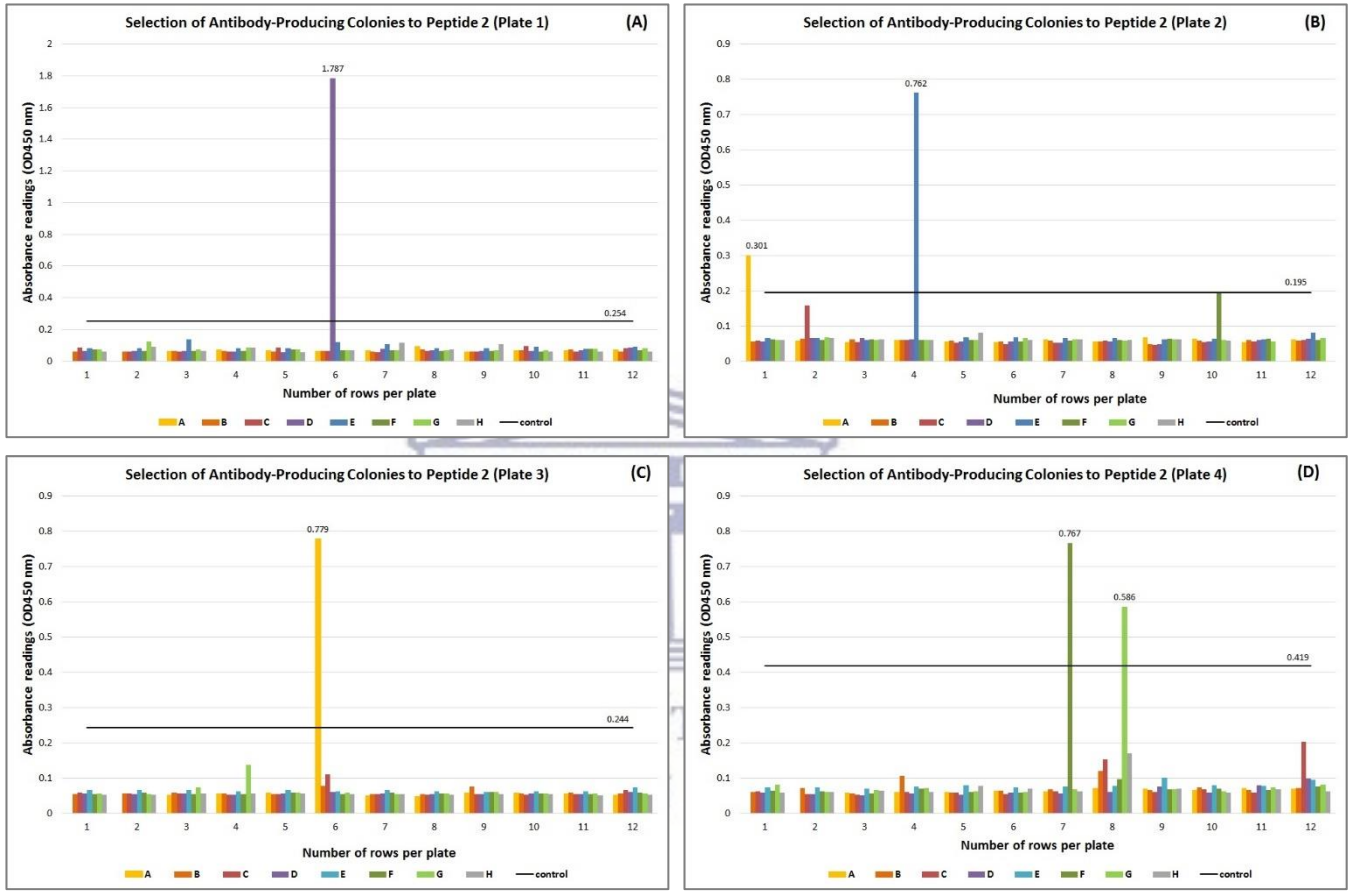
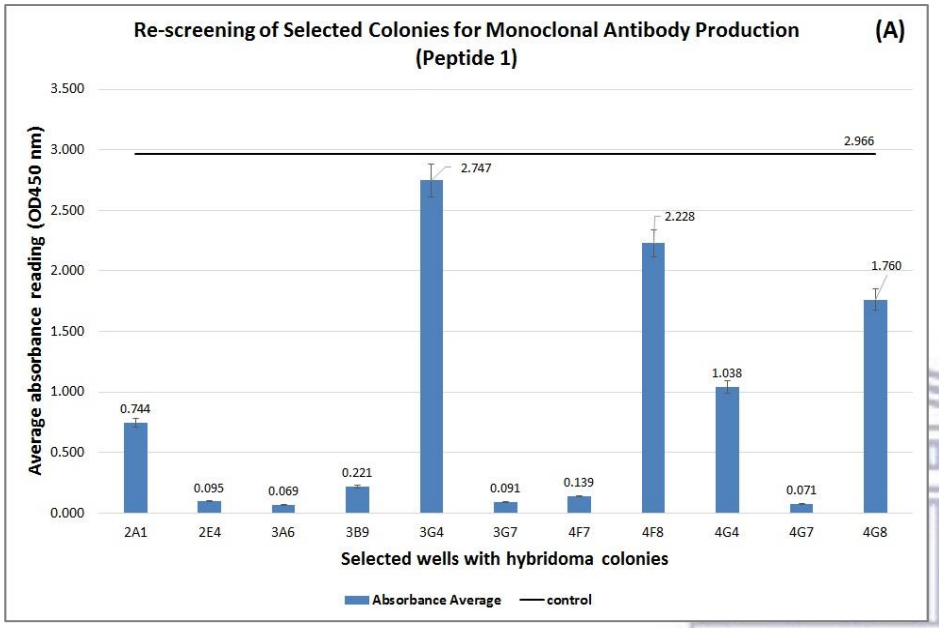



**Figure 3.11 Selection of hybridoma producing colonies against peptide 2.** Supernatant of fused myeloma cells were screened with ELISA for hybridomas producing antibodies against peptide 2 and the absorbance readings recorded at OD450nm. The data is presented as percentage and all values higher than 100% represent that of antibody-producing hybridoma colonies (labelled bars). Coloured bars represent columns for each plate**.**



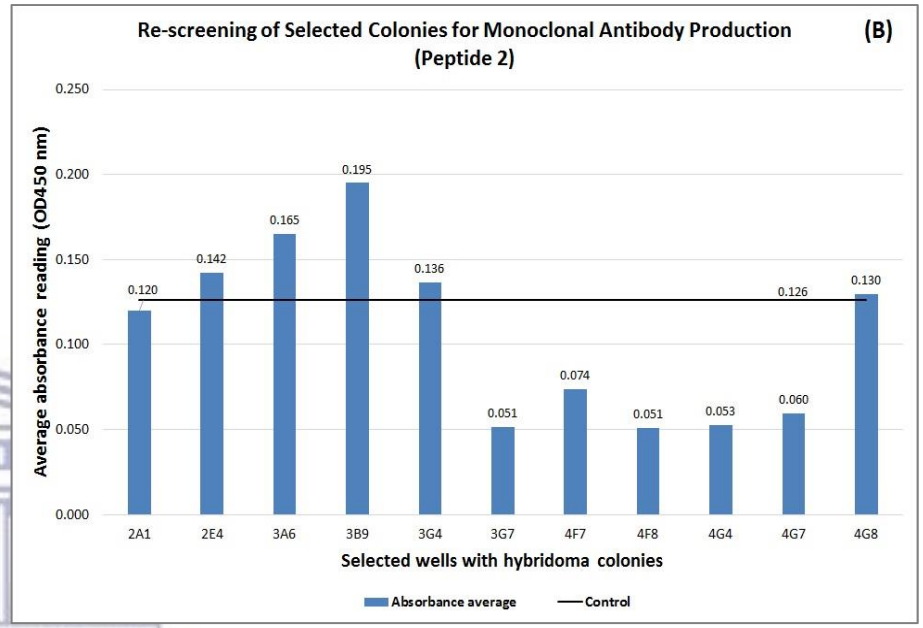




**Figure 3.12 Re-screen of selected colonies for cloning.** Supernatant of wells containing hybridoma colonies producing polyclonal antibodies was screened with ELISA for further cloning. All wells that produce readings higher than the 100% baseline was chosen for cloning in production of monoclonal antibodies. Myeloma cells that was treated with peptide 1 did not produce readings acceptable for cloning. Myeloma cells that was treated with peptide 2 reacted favourably with three wells containing hybridoma colonies that produced well above the 100% baseline to continue cloning.

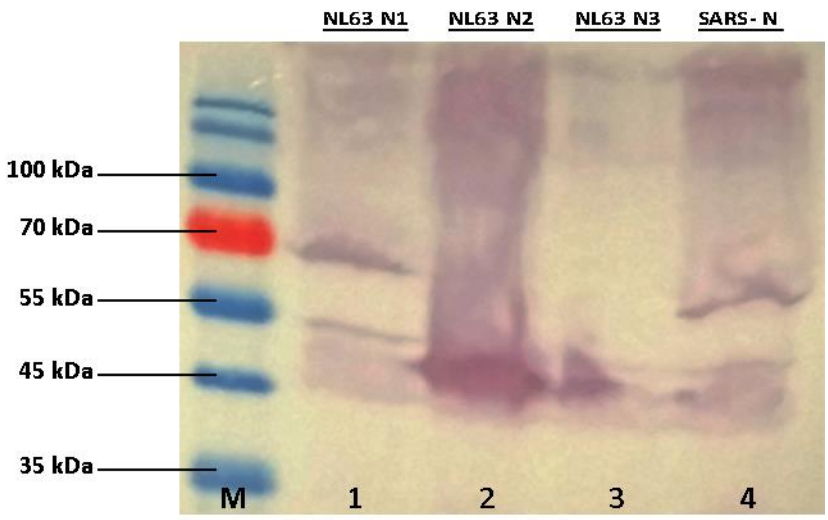



**Figure 3.13 Western blot of polyclonal mouse antisera used for detection of recombinant proteins.** Mouse plasma containing polyclonal antibodies generated against HCoV-NL63 N- and C- termini was used as primary antibody for the detection of recombinant proteins SARS-N, N1, N2 and N3. Lane M: Pre-stained protein molecular marker; Lane 1: GST-N1 (~69 kDa); Lane 2: GST-N2 (~48 kDa); Lane 3: GST-N3 (~47 kDa); Lane 4: SARS-GST N (~72 kDa).

### 3.6.4 Characterization of Polyclonal Antibodies

Mouse blood plasma containing polyclonal antibodies generated against HCoV-NL63 N peptides (peptides 1 and 2) were used as primary antibodies in western blot for the detection of purified recombinant proteins (confirmed in 2.7). The membrane was incubated in 5% milk-PBS/Tween solution containing 1:1000 primary antibody, mouse polyclonal antiserum (a combination of polyclonal antibodies, anti-peptide 1 and anti-peptide 2), followed by subsequent incubation with secondary HRP-conjugated goat anti-mouse antibody at a dilution of 1:3000. Even though the quality of the picture taken of the western blot is poor, the correct sizes for the purified fusion proteins, SARS-N (~72 kDa), N1 (~69 kDa), N2 (~48 kDa) and N3 (~47 kDa), could be identified in Figure 3.13.

# CHAPTER 4

# DISCUSSION


Department of Medical Biosciences, Faculty of Natural Sciences, University of the Western Cape, South Africa

# CHAPTER 4 DISCUSSION

## 4.1 Expression of Recombinant HCoV-NL63 Nucleocapsid Protein

Expression in a bacterial system remains a popular method for producing heterologous, recombinant proteins. The most widely used hosts for the production of recombinant proteins are *E. coli* due to its ability to grow rapidly and at high-density on inexpensive substrates as well as the availability of an increasingly large number of cloning vectors and mutant host strains (Baneyx, F. 1999). A variety of plasmid vectors for positive control have been described that rely on inactivation of a lethal gene or lethal site, a dominant function presenting cell sensitivity, or a repressor of antibiotic-resistance function (Yazynin, S., Lange, H., et al. 1999). A diverse selection of *E. coli* expression vectors is available with fusion tags under the control of different promoters (Peti, W. and Page, R. 2007).

For the expression of all the constructs used in this experiment, KRX strain of competent *E. coli* was used through utilization of a Flexi™ (pFN2A) vector system (Promega 2009b). The Flexi™ vector is of particular value due to its several properties contributing towards control during protein expression. The Flexi™ vector contains a T7 promoter for bacterial or *in vitro* protein expression and it appends a GST coding region to the protein of interest that facilitate detection and purification of the expressed viral proteins (Promega 2009b). This vector contains two unique restriction sites, *SgfI* and *PmeI*, either side of the cloning region, that allows for the insertion or transfer of the sequence of interest and ensures high fidelity and unidirectional cloning (Promega 2009b). The vector also contains an ampicillin-resistance gene for selection of the plasmid and a lethal barnase gene located between cloning sites, or *SgfI* and

*PmeI* restriction sites, for positive selection of the recombinant clones (Yazynin, S., Lange, H., et al. 1999; Promega 2009b). Barnase is a major extra-cellular ribonuclease isolated from *Bacillus amyloliquefaciens* and its expression in *E. coli* is lethal to the cells (Paddon, C.J., Vasantha, N., et al. 1989; Qin, Q., Liu, Y.L., et al. 2005). Upon successful ligation, the barnase gene is replaced by the cloned gene and growth of cells will persist.

The recombinant constructs used in this study are clones of full-length SARS-CoV N and HCoV-NL63 N genes. Specific truncated mutants were synthesized to resemble the N- and C-termini of HCoV-NL63 N gene (Figure 2.1). Before expression studies could be done, it was important to first confirm the presence of plasmid DNA in the glycerol stock samples. A PCR amplification was performed and the gene inserts identified by gel electrophoresis (Figure 3.1) to be present at the correct sizes as determined by the NCBI (Table 3.2). SARS-CoV N is slightly larger in size (1269 bp) than HCoV-NL63 N (1134 bp) due to the additional accessory proteins in its genome. The truncated clones are approximately half the size of HCoV-NL63 N at ~560 bp. Restriction endonuclease digestion was performed utilizing the Flexi enzymes, *SgfI* and *PmeI,* to cut plasmid DNA from the Flexi vector. The results obtained from the digests serve as verification that plasmid DNA was completely cut from the vector (4137 bp) (Appendix 1) in both small (Appendix 3) and large scale (Figure 3.2) preparations and can therefore be utilized for expression of proteins.



Autoinduction used in expression of recombinant proteins depends on the metabolic mechanisms of bacteria in regulating use of energy sources available in growth medium (Peti, W. and Page, R. 2007). KRX *E. coli* utilizes a T7 RNA polymerase

gene to tightly control recombinant protein expression via a T7 promoter provided by the Flexi™ vector. This gene is controlled by the rhamnose promoter (*rha*PBAD). This promoter is subject to repression by glucose metabolites and therefore inhibits pre-induction protein expression, whereas with the addition of rhamnose to the medium it will induce protein expression, but only after all glucose have been completely metabolized from the culture media (Holcroft, C.C. and Egan, S.M. 2000; Schagat, T., Ohana, R., et al. 2008; Promega 2009c). Typical expression of recombinant proteins in bacteria is a time-consuming process where cells are grown overnight to an optimal density and then induced (usually by inducers such as IPTG) (Sambrook, J. and Russell, D.W. 2001; Samuelson, J. 2011). The single-step autoinduction protocol simplifies the process with addition of glucose and rhamnose to the growth media which eliminates the need to monitor culture density (Schagat, T., Ohana, R., et al. 2008).

The recombinant clones in this experiment were all expressed downstream from a GST coding region whereby it attaches the GST moiety to the protein of interest. This makes it very useful for the expression, detection and purification of fusion proteins. For purification of the fusion protein, the MagneGST™ purification system was used due to its ability to incorporate glutathione to bind GST fusion proteins. It provides a simple, rapid and reliable method for purification. Glutathione is immobilized on paramagnetic particles and is used to isolate GST fusion proteins directly from crude cell lysates under mild, non-denaturing conditions and preserving protein antigenicity and function (Amersham 2001; Promega 2009a).

Detection of the purified fusion proteins is done using colorimetric or immunological methods such as SDS-PAGE and Western Blot analysis (Amersham 2001).

Following expression of the recombinant viral proteins and analysing the western blots of unpurified (Figure 3.3) and purified proteins (Figure 3.4B) it was detected that HCoV-NL63 full-length N protein (N1) and its truncates (N2 and N3) started showing partial degradation early in the experiment. Bands were detected in the region of 25 kDa which represent the GST tag. In Figure 3.4, the GST tag was also separately detected from SARS-CoV N protein.

At 26 kDa, GST is considerably larger than many other fusion protein affinity tags. It is not clearly understood why the structure of the GST fusion tag often degrades upon denaturation and reduction for protein gel electrophoresis (e.g., SDS-PAGE). As a result, electrophoresed samples of GST fusion proteins often appear as a ladder of lower molecular weight bands below the full-sized fusion protein (Thermo Fischer).

The degradation of the recombinant fusion proteins results in the GST tag being cleaved from the target protein. Normally, protein degradation is caused by either incorrect folding of a protein or subsequent protein solubility, which in turn is determined by protein folding (ExpTec 2009). It has been previously reported that SARS-CoV (Wang, Y., Wu, X., et al. 2004), 229E (Tang, T.K., Wu, M.P., et al. 2005) and infectious brochitis virus (IBV) (Fan, H., Ooi, A., et al. 2005) N proteins does harbour some instability at its CTD and NTD. Together with the reported instability, it has also been reported that bacterial proteases are the leading cause of intracellular degradation. Autolytic degradation has also been reported to occur at 4°C where purified IBV N protein degraded rapidly (Fan, H., Ooi, A., et al. 2005). Inhibition of these proteases is crucial for efficient expression of heterologous proteins as *E. coli* employs cytoplasmic degradation to conserve cellular resources (Baneyx, F. 1999).

A rate of protein turnover is maintained in all living cells by continuous degradation (or breakdown) of proteins by proteases (Ryan, B.J. and Henehan, G.T. 2013). In order for the cell to conserve cellular resources, it will degrade improperly folded proteins into their constituent amino acids for recycling (Baneyx, F. 1999). However, correct folding and solubility of proteins are usually increased by the addition of a fusion tag, such as GST. It has been assumed that fusion tags increase correct folding of their downstream partner as they reach a native conformation, thereby promoting correct folding of the entire protein (Baneyx, F. 1999). Since the target proteins could still be detected and seemed to be intact, this possibility could be discarded. Other factors to consider if proteolysis is expected, could be the presence of proteases or protein aggregation due to excessive sonication (Amersham 2001; Peti, W. and Page, R. 2007; Promega 2009a). Protease is an enzyme that could cleave the GST tag from the protein or even degrade the protein itself. Excessive forces delivered by sonication could cause alteration in the GST moiety conformation and prevent fusion proteins from binding to glutathione agarose beads during purification step and could lead to ineffective binding and therefore purification of the target protein (Ryan, B.J. and Henehan, G.T. 2013).



Factors that could lead to possible degradation of GST fusion proteins can be reduced or prevented by introducing protease inhibitor cocktails to the lysis buffer that will inhibit the liberation of proteases from the subcellular compartments by blocking the action of proteases (Ryan, B.J. and Henehan, G.T. 2013). Also by avoiding excessive sonication and frothing during sonication, by purification at 4°C, by reducing the purification steps and keeping samples and buffers on ice to reduce the activity of proteases or by lowering the cultivation and expression temperatures if degradation

was expected during expression of proteins (Georgiou, G. and Valax, P. 1996; Baneyx, F. 1999; Promega 2009a; Ryan, B.J. and Henehan, G.T. 2013). In this experiment, no protease inhibitor was added and could be the cause of the degradation seen on the western blots. Generally, proteolysis of proteins does not carry over into purification steps of recombinant proteins if the proper steps to inactivate proteases have been taken, but on occasion this may be observed during storage and can be addressed by further purification or reapplying specific protease inhibitors.

Fusion tags, such as GST, are often utilized for rapid purification and detection of recombinant proteins, but aren't vital for these particular functions and not considered to be permanent fixtures on recombinant proteins. The fusion protein can easily be cleaved using Tobacco Etch Virus (TEV) protease after purification (Amersham 2001; Promega 2009a). For further analytical studies of recombinant protein structure and function, it would be required to remove or cleave the GST tag from the protein with TEV or other proteases (Sigma-Aldrich).



In cleavage of a GST fusion protein with TEV, one should observe three lanes of bands on an SDS-PAGE gel: the native full-length N protein plus the tag (~72 kDa) (used as positive control), the native full-length N protein without the tag (~42 kDa) and the GST tag (~26 kDa). If the GST tag was indeed cleaved somehow from the native protein (due to whatever factor as discussed above), bands in the regions of ~42 kDa and ~21 kDa (protein without the GST tag) should have been visible on the coomassie-stained SDS-PAGE gel in Figure 3.4A, but was not. It should be noted that expression of unpurified recombinant N proteins was only confirmed with Western Blot analysis (Figures 3.3) without coomassie stains of the SDS-PAGE gels to validate that it is

indeed partial degradation of the GST tag. The secondary anti-GST antibody used in the western blot analysis should be able to bind to GST proteins (whether fused to the native protein or not).

Since cleavage of the GST tag from the recombinant proteins was not necessary in this study for further downstream applications, no cleavage of GST tag was performed. The possibility of degradation, as seen from these results, does therefore not have any relevant impact on the current study.

## 4.2 Production of Polyclonal Antibodies against HCoV-NL63 nucleocapsid protein.

Antibodies are produced by the immunization of animals like mice and rabbits using an adjuvant to initiate an immune response and boost the production of antibodies (Harlow, E. and Lane, D. 1988). One can either immunize with highly purified antigens to produce monospecific polyclonal antibodies or immunize with a mixture of antigens to produce polyspecific polyclonal antibodies. For the production of monoclonal antibodies, further cloning and screening strategies are implemented (Schlom, J., Wunderlich, D., et al. 1980). It is also possible to use synthetic peptides to produce antibodies if the native antigen is absent and the amino acid sequence of the desired antigen is known (Saravanan, P., Kataria, J., et al. 2003).

Purified GST fusion proteins are extensively used as immunogens for the production of appropriate antibodies that can be used as probes for specific proteins. In this study, however, synthetic peptides were used to raise polyclonal antibodies. Peptides were chosen for the high frequency at which anti-peptide antibodies cross-react with the corresponding native protein with the additional advantage of an already well defined epitope that will be recognized by the antibody (Hearn, M.T.H., Bishop, C.A., et al.

1979; Niman, H.L., Houghten, R.A., et al. 1983; Atassi, M.Z. 1984). With the aid of peptides, antibodies can therefore be raised against sequence-specific sites of interest like in this case the C- and N- termini of the NL63 nucleocapsid protein (Niman, H.L., Houghten, R.A., et al. 1983). Sera can rapidly be screened for anti-peptide activity using ELISA due to the readily available peptide immunogen against which the antibody was raised (Engvall, E. and Perlmann, P. 1971; Voller, A., Bertlett, A., et al. 1978). Diagnostic tools used for analysis in identifying native proteins include Western blotting, immunoimaging, ELISA and immunoprecipitation (Schulz, S., Rocken, C., et al. 2006; Ohara, K., Horibe, T., et al. 2011; Zarei, O., Irajian, G.R., et al. 2011).



The complete sequence for HCoV-NL63 nucleocapsid protein, isolated in Amsterdam (Appendix 2), was used to design and synthesize specific peptides used in this study. The sequence was submitted to a series of different algorithms, each determining different aspects of the sequence. The complete HCoV-NL63 N sequence was used to produce an antigenic propensity plot (Figure 3.5) that aids in determining potentially antigenic residues within a given sequence (Immunomedicine Group). In total, 13 residues could be identified to be antigenic residues from the NL63 N protein sequence of 377 residues and is listed in Table 3.4. To narrow down which of the 13 potentially antigenic residues are excellent antigens to consider for peptide synthesis, each of the 13 residues were subjected to being profiled by a bioinformatics protein sequence tool (Thermo Fischer Scientific). The profiler tool uses an algorithm to determine the antigenicity of each sequence according to a scale between <1.0 and 5.0. Excellent peptides score between 2.7 – 5.0. Six of the 13 residues were identified to have high

antigen profile scores ranging between 3.0 and 3.8 (indicated by an asterisk (*) in Table 3.4), giving them an excellent rating as potential peptides.

In order to cross-react, the synthetic peptide structure must have access to the external surface of the protein. Predictive algorithms are used to provide data on the physiochemical properties of a peptide (Pepcalc results not shown) (Chou, P.Y. and Fasman, G.D. 1978; Garnier, J., Osguthorpe, D.J., et al. 1978). These properties include antigenicity, hydrophilicity, flexibility, surface probability and charge distribution of a given amino acid sequence (Shinnick, T.M., Sutcliffe, J.G., et al. 1983; Thornton, J.M., Edwards, M.S., et al. 1986; Rubinstein, N.D., Mayrose, I., et al. 2008). The water solubility of a peptide is probably the most important physiochemical factor because there can be many antigenic residues, but if they aren't soluble, the antibody won't be able to detect it in solution. A hydropathy plot using an algorithm adapted from Hopp and Woods show HCoV-NL63 N protein to be mainly hydrophilic (Figure 3.6) (Hopp, T.P. and Woods, K.R. 1981). The six residues that scored excellent antigenicity scores were tested for their water solubility and only four resulted in having good water solubility (Table 3.4).



Accessibility of a peptide antibody to recognize the native protein is an important factor to consider and may exposed regions, such as the N- and C-termini and secondary structural features like loops and turns, make good choices for peptides (Shinnick, T.M., Sutcliffe, J.G., et al. 1983; Hancock, D.C. and O'Reilly, N.J. 2005; Rubinstein, N.D., Mayrose, I., et al. 2008).The average antigenic propensity gives an indication of the surface accessibility or probability of the peptide of being detected. Propensity is measured at values >1. The higher the value, the greater chance of an antibody recognizing the epitope which will then effectively not only result in an

immune response but also in a desired antigenic response. Individual antigenic propensity plots (Appendix 4 to 6) were graphed for the four sequences with good water solubility scores. Three possible sequences were identified (highlighted in Figure 3.5): RKKFPPPSFYMPLLVSSD (aa 13-30), PLEPKFSIALPPELSVVE (aa 122-139) and LPKFIEQISAFTKPSSVKE (aa 319-337) (Table 3.4). All three these residues proved to be excellent choices for the synthesis of peptides that could possibly elicit antigenic responses for the production of antibodies. All three had relatively good antigenic propensity values of 1.0524 (Appendix 4), 1.1021 (Appendix 5) and 1.0402 (Appendix 6) respectively, but it was the peaks of the first two sequences at ~1.15 each that resulted in those two residues (indicated with double asterisk (**) in Table 3.4) being chosen for synthesis of peptides that will be used to immunize mice and screen for antigenic immune response.

N- and C- termini have specific sequences that are exposed and facilitate accessibility of a peptide antibody to recognize the native protein (Niman, H.L., Houghten, R.A., et al. 1983; Houen, G., Jakobsen, M.H., et al. 1997; Hancock, D.C. and O'Reilly, N.J. 2005; Rubinstein, N.D., Mayrose, I., et al. 2008). Specially designed peptides for this study were to consist of sequences found at or near the N- and C- termini of HCoV-NL63 N sequence that will facilitate detection of those regions of HCoV-NL63 on a western blot for example. With the results obtained (Table 3.4), sequence $n$ = 1, is found at the N-terminal end of the HCoV-NL63 N sequence and, for the purpose of this thesis, named peptide 1 (aa 13-30). Sequence $n$ = 5 was named peptide 2 (aa 122-139). This peptide sequence, unfortunately, is not allocated at the C-terminal end of the sequence of HCoV-NL63 N protein. Sequence $n$ = 11 (aa 319 -337) (Table 3.4) would have included a sequence from the C-terminal end, but due to its lower antigenic

propensity peak value (~1.05) in comparison to the other two (~1.15), it was not selected as a possible peptide in this study if chances of it being accessible is low.

For immunization, the peptide sequence should be between 8 and 20 amino acids in size. Peptides smaller than eight amino acids won't function as an epitope and may elicit antibodies that do not recognize the native protein. Peptides larger than twenty amino acids may adopt conformations that are not specific to and resemble that of the native protein (Hancock, D.C. and O'Reilly, N.J. 2005). Peptide 1 and 2 are both 17 residues in length (Table 3.4). Many peptides are antigenic, but not all are immunogenic because of their small size (Hermanson, G.T. 2008). For that reason, peptides are usually conjugated to a carrier protein such as keyhole limpet hemocyanin (KLH), BSA or ovalbumin (OA) (Strynadka, N.C., Redmond, M.J., et al. 1988). KLH are immunogenic carrier proteins commonly used for polyclonal anti-peptide antibodies. BSA should rather be avoided due to its interference in subsequent studies where the antibody is used in tissue culture with BSA (Schaaper, W.M., Lankhof, H., et al. 1989; Houen, G., Olsen, D.T., et al. 2003; Hancock, D.C. and O'Reilly, N.J. 2005). MAPs are used as an alternative to the traditional peptide-carrier conjugates (Tam, J.P. 1988). It makes use of Lysine core matrix that functions as a scaffold (Posnett, D.N., McGrath, H., et al. 1988) that binds four to eight copies of the antigenic peptide that can induce a potent immune response (Kowalczyk, W., Monso, M., et al. 2011; Wang, G.Z., Tang, X.D., et al. 2011). The selected peptides were coupled MAP8 conjugates to make the peptide more antigenic and induce a potent immune response compared to other conjugate proteins (Posnett, D.N., McGrath, H., et al. 1988; Tam, J.P. 1988; Kowalczyk, W., Monso, M., et al. 2011).

These conjugated peptides (especially if water-soluble) are mixed with an adjuvant as a water-in-oil emulsion that gets injected into an animal during immunization (Goding, J.W. 1996). Adjuvants are commonly used to enhance the immune response that efficiently leads to high-titred antisera. CFA has been the most utilized and effective adjuvant since its original description (Freund, J., Stern, E.R., et al. 1947) and its immunostimulatory capabilities could not be surpassed by any other adjuvant and is considered the gold standard among immunologists (Altman, A. and Dixon, F.J. 1989; Stills, H.F., Jr. 2005). CFA is composed of a light mineral oil, mannide monooleate and avirulent *M. tuberculosis* H37Ra or *Mycobacterium butyricum* (Stewart-Tull, D.E.S. 1995). The mycobacterial components produce a stronger delayed-typed hypersensitivity (Lindblad, E.B., Elhay, M.J., et al. 1997). IFA lacks the killed mycobacterial cells and are used in booster immunizations to counteract any hypersensitivity to CFA (Goding, J.W. 1996). The use of CFA has been associated with a variety of lesions including localized injection site granulomas and a review written by Stills 2005 describes studies on other replacement adjuvants for CFA commercially available that could potentially be equally effective with fewer side effects (Stills, H.F., Jr. 2005).



For successful peptide antibody production, the method of choice for immunization and which animal to use are dependent on the intended application and the type of antibody and antigen that is required (Harlow, E. and Lane, D. 1988; Burns, R. 2005). One must keep in mind that different animals within the same group, can respond differently to the same immunogen. A given anti-peptide antibody may work in one assay but not in another (e.g. ELISA vs Western Blot) (Hancock, D.C. and O'Reilly, N.J. 2005). To produce polyclonal peptide antibodies, the New Zealand White rabbit

is the most commonly used host animal for its ability to respond to broad classes of antigens and provide good yields in a short period of time. Balb/c mice are the ideal hosts suited for monoclonal peptide antibody production (Lee, B.S., Huang, J.S., et al. 2010; Zarei, O., Irajian, G.R., et al. 2011).

Two balb/c mice were immunized with a mixture of the peptides in CFA water-in-oil emulsion. Initial immunization was followed by two to three booster immunizations with IFA that counters the effects of CFA on the animal host but maintain the elicited immune response. After each booster immunization, tail blood was drawn that was used to test for polyclonal antibody titre. The purpose of *in vivo* immunization was to obtain pre- and post-immunization polyclonal antibodies to be used as negative and positive controls for screening by indirect ELISA. ELISA, also known as enzyme immunoassay (EIA), is a plate based assay technique used to detect and quantify antibodies, peptides, proteins, and hormones. Indirect ELISA describes a process whereby antigen is immobilized on the surface of a microtiter well. A primary antibody specific to the antigen is added to form an antibody-antigen complex. This complex is then detected by the addition of enzyme conjugated secondary antibody that recognizes the primary antibody. Specific enzymatic substrate is added that result in visible signal, in most cases colour change, that indicates the quantity of antigen in the sample. The amount of coloured product is measured spectrophotometrically to determine antigen concentrations.



Antibody titres were determined for each mouse to peptide 1 and 2 respectively. In Figure 3.7B post-immune 2 graph for mouse 2 had a higher antibody titre and prolonged response to peptide 1 when compared to mouse 1 (Figure 3.7A). In Figure

3.8 no clear deductions can be made as to which mouse had a higher or better antibody titre in response to peptide 2. There was no real difference between either the pre-immune and post-immune 1 responses in either mouse and an overlap was noticed at low antibody concentrations. The antibody titre, after the second post-immunization, was not maintained and showed a sudden decrease in response. In comparison to the results obtained from Figure 3.7, it was noted that the highest reported OD reading for antibody titre response after the second post-immunization in Figure 3.8 was lower than the pre-immune values obtained with peptide 1. This could indicate that neither mouse had a good and sustainable response to the peptide 2 antigen or that peptide 2 was unable to elicit a strong enough immune response to produce polyclonal antibodies.

Figures 3.9 show a combination of post-immune 1 and post-immune 2 antibody titres for both mice to peptide 1 (A) and peptide 2 (B) respectively. The data presented in Figure 3.9 produced polynomial trend lines with high correlation coefficients ($R^2$) seen with both immunizations in both mice, except for post-immune 2 in mouse 2 where the $R^2$ = ~0.76. Correlation coefficients indicate how well the regression line (in this case the polynomial trend line) approximates the real data points and should deliver values close to one for a perfect fit. The trend line for mouse 2 post-immune 2 (Figure 3.9A) corresponds with what was seen in Figure 3.7B. The dilution curves presented in Figure 3.9B for post-immunizations against peptide 2 for both mice have correlation coefficients that are low and far from the desired value of one and indicates what was seen with the antibody titres in Figure 3.8. From the $R^2$-values presented here, these dilution curves do not correlate well and are the results undesirable in response to peptide 2.

ELISA results confirmed the production of polyclonal antibodies with high titres observed with peptide 1 (Figure 3.10). *In vitro* immunization was performed on the spleen cells of the mouse that exhibited the highest antibody titre. According to our results, mouse 1 responded better to each peptide and was therefore sacrificed to harvest its spleen that contain B-lymphocytes used to fuse with myeloma cells to form hybridomas. Splenocytes were fused with myeloma cells and subjected to ELISA screening (Figure 3.10 & 3.11) that identified ten hybridoma colonies (Table 3.5) that could be used for cloning. The chosen colonies are indicative of hybridomas that can produce monoclonal antibodies to the antigens. Two of the chosen colonies showed non-specific affinity for peptides 1 and 2. This cross-reactivity was observed in two of the wells (hybrids 2A1 and 4G8). The hybridoma colonies were cloned and screened a second time (Figure 3.12). In Figure 3.12A, no colonies produced antibody titres higher than that of the positive control in response to peptide 1. Even though it looks like colonies produced antibody titres higher than that the positive control in response to peptide 2 (Figure 3.12B), it became clear upon further investigation that the positive control value itself is too low to be accepted as values that indicates the positive production of monoclonal antibodies and was the results obtained with peptide 2 discarded as negative production of monoclonal antibodies. After the second cloning, no further cloning could be performed. Apart from having trouble with constant contamination in attempts to repeat the experiment, no monoclonal antibodies could be produced in this study, due to the time constraints permitted for this masters and limiting resources available.

Further characterization of HCoV-NL63 N protein was performed with mouse 1 blood plasma from the second post-immunization containing polyclonal antibodies generated against HCoV-NL63 N peptides (peptide 1 and 2). The serum was used as primary antibody in western blot for the detection of purified recombinant N proteins. Even though the quality of the picture used in figure 3.13 is poor due to electrophoresis error resulting in skewed lanes (too high temperature or running too fast), the recombinant proteins of HCoV-NL63 N could be detected using the blood serum containing polyspecific polyclonal antibodies to HCoV-NL63 N protein (Figure 3.13) with some background detected in lanes 1 and 4. The synthetic peptides were able to produce polyclonal antibodies that could detect N1, N2 (N-terminus) and N3 (C-terminus), but also SARS-N. The results from the western blot in Figure 3.13 are quite interesting for two reasons; 1) SARS-N was detected by the polyclonal antibodies generated against a mixture of peptides synthesized from the HCoV-NL63 N sequence, and 2) N3 was detected with polyclonal antibodies produced from a mixture of peptides synthesized from sequences located primarily at the N-terminal fragment of HCoV-NL63 N protein (Figure 2.1). In Figure 4.1, a schematic diagram illustrates the proposed reactions expected to be seen on the western blot of the polyconal antibodies binding to the epitopes at each terminal.



Studies have suggested sequence homology between N protein of SARS-CoV and HCoV-NL63 and even though this western blot wasn't carried out using monoclonal antibodies, the sequences of the synthetic peptides used to raise polyclonal antibodies could detect sequences within SARS-N that were similar to that of NL63 N (Figure 4.1). SARS-CoV N was included as a control that would allow us to identify any cross-

reactivity between CoV species which is important for determining sensitivity and specificity. It has been found that SARS-CoV infections stimulate cross-reactive antibody responses to other CoVs (Chan, K.H., Cheng, V.C.C., et al. 2005). Some studies have indicated cross-reactivity between SARS-CoV and 229E, NL63 and OC43 N proteins (Che, X.Y., Qiu, L.W., et al. 2004; Woo, P.C., Lau, S.K., et al. 2004b; Chan, K.H., Cheng, V.C.C., et al. 2005; Shao, X., Guo, X., et al. 2007; Dijkman, R., Jebbink, M.F., et al. 2008). Other studies showed polyclonal antibodies from group I CoVs (NL63, 229E, TGEV) are capable of reacting with Vero cells infected with SARS-CoV (Sun, Z.F. and Meng, X.J. 2004). Cross-reaction of SARS-CoV N protein was found to be localized between amino acids 120-208 and for HCoV-NL63 N protein it was found to be between amino acids 1-39 of the N-terminal (Vlasova, A.N., Zhang, X., et al. 2007). It has been proposed that cross-reactivity either occurs at antigenic level with cross-reactivity of epitopes or at the antibody level with the production of polyclonal antibodies (Chan, K.H., Cheng, V.C.C., et al. 2005).

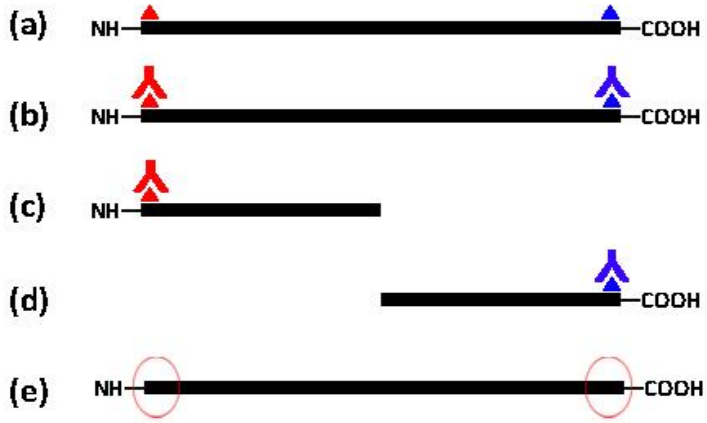



**Figure 4.1 Schematic diagram of HCoV-NL63 sequence with epitopes on the N- and C-terminal for detection of polyclonal antibodies.** (a) Genome representing that of full-length HCoV-NL63 N with epitopes that recognize peptide-antibodies on the N-terminal and C-terminal ends. (b) Full-length HCoV-NL63 N protein with epitopes on N- and C-terminal ends to which both or either polyclonal antibodies can bind. (c) Epitope on N-terminal fragment of HCoV-NL63 N protein. Antibody only able to recognize this epitope. (d) Epitope on C-terminal fragment of HCoV-NL63 N protein. Antibody only able to recognize this epitope. (e) Genome of SARS-CoV N protein. Even though this CoV shares homology between N proteins with NL63, we are not sure if antibodies would be able to detect either N- or C-terminal epitopes or even both.



The use of synthetic peptides opposed to the full length protein, may show a reduction in cross-reactivity (Sun, Z.F. and Meng, X.J. 2004; Woo, P.C., Lau, S.K., et al. 2004a; Sastre, P., Dijkman, R., et al. 2011). Specificity can be increased by identifying and directing detection against immunogenic epitopes. Monoclonal antibodies against N

protein are highly specific and sensitive for the detection of NL63, 229 (Sastre, P., Dijkman, R., et al. 2011) and SARS-CoV (Lau, S.K., Woo, P.C., et al. 2004). In typical tests where the native protein needs to be identified, polyclonal antibodies make a more robust reagent due to its broader specificity to antiserum containing many types of antibodies with different epitopes, making the antibodies therefor less sensitive to changes in the antigen than monoclonal antibodies. Monoclonal antibodies on the other hand are specific to a single epitope on the antigen. This highly specific nature allows for the development of assays that can differentiate between closely related substances (Schlom, J., Wunderlich, D., et al. 1980; Kohler, G. and Milstein, C. 2005).

Where polyclonal peptide antibodies are not sufficient to detect the native protein, monoclonal peptide antibodies may be used for more quantitative and expanded analyses. Studies showed where peptides were used for the direct detection against immunogenic epitopes, it lead to an increased immune response in SARS (Woo, P.C., Lau, S.K., et al. 2004a) NL63 and 229E (Sastre, P., Dijkman, R., et al. 2011). Chosen peptides were administered as impure antigens (as a mixture) that will effectively produce polyspecific polyclonal antibodies. To optimize your chance of producing useful antibodies to a protein it is advisable to use peptides from different regions of the protein sequence and to immunize more than one animal (or species of animal) with each individual peptide (Hancock, D.C. and O'Reilly, N.J. 2005) to produce antibodies that are more specific in the detection of native protein.



HCoV-NL63 N recombinant protein and its truncated clones were expressed in a bacterial system and validated with Western Blot analysis with the use of anti-GST

antibody that can bind to the GST tag appended to the recombinant proteins and thus detect the proteins on a blot. Polyclonal anti-N antibodies were successfully generated by immunizing mice with a mixture of synthetic peptides specifically designed to elicit production of polyspecific polyclonal antibodies. A mixture of peptides was used to immunize mice which resulted in polyclonal antibodies not specific enough in the detection of HCoV-NL63 N recombinant proteins. Further studies should be conducted on cross-reactivity with SARS-CoV N protein. The production of monoclonal antibodies will increase specificity in the detection of recombinant HCoV-NL63 N protein and its truncated clones.

# CHAPTER 5

# CONCLUSION


Department of Medical Biosciences, Faculty of Natural Sciences, University of the Western Cape, South Africa

# CHAPTER 5 CONCLUSION

Coronaviruses are enveloped viruses belonging to the *Coronaviridae* family in the order *Nidovirales,* which are composed of single-stranded, positive sense RNA viruses with the largest viral genome among RNA viruses (27-33 kb). In humans, CoVs are commonly associated with URTIs which are known to manifest as common cold symptoms. The outbreak of SARS-CoV sparked new interest in the detection and characterization of CoVs. HCoV-NL63 was first isolated from the nasopharyngeal aspirate from a 7-month-old child with bronchiolitis in January 2003. HCoV-NL63 has since been identified in clinical specimens from infants younger than six years of age, adults of frail age and immunocompromised individuals. Three fatal cases have been reported where HCoV-NL63 was identified as the aetiological agent. HCoV-NL63 is commonly associated with other respiratory viruses. Infection with HCoV-NL63 is usually more severe where multiple viruses are involved. HCoV-NL63 have been associated with other diseases such as croup, KD and even gastrointestinal illnesses. A worldwide distribution of this virus, after its isolation in The Netherlands, was indicated by its subsequent identification in various countries which indicated a global spread of this virus.



These recombinant constructs are clones of full-length SARS-N and HCoV-NL63 N genes, as well as deletion mutants of the N- and C-terminus of HCoV-NL63 N. For the purpose of this thesis, the constructs were named SARS-N, N1, N2 and N3 respectively. The N protein, one of the most abundant viral proteins and a multifunctional phosphoprotein, is highly basic, exists as a disulfide-linked

multimer and is the only RNA-binding protein of mature virions to form a ribonucleocapsid structure. N protein can also be used as an early diagnostic tool in the early detection of SARS-CoV infection in patients due to its high immunogenicity. HCoV N protein is highly antigenic and abundantly expressed during viral infection and detectable as early as from the first day after infection. Antibodies have become essential for diagnostic and therapeutic purposes due to their high specificity, high binding affinity, long half-lives and low toxicity, and are invaluable reagents for antigen detection in immunoblotting

In this study, recombinant plasmid constructs of SARS-CoV N and HCoV-NL63 N protein and its truncates, were used. The presence of the gene insert on the Flexi vector was confirmed by restriction endonuclease digest. The recombinant proteins were expressed using the KRX *E. coli* bacterial expression system. All constructs were expressed downstream from a GST fusion protein. Restriction enzyme digests were performed on mini- and midiprep samples of the recombinant plasmids to isolate the plasmid DNA and to confirm the insert of the genes of interest. Autoinduction was performed under the tight control of a rhamnose promotor by the addition of glucose. Expression is only activated in the presence of rhamnose once glucose is consumed from the culture medium. Protein samples were lysed by means of probe sonication that at high speeds tear the bacterial cells walls to release the proteins into the lysate. The lysate was separated electrophoretically on SDS-PAGE gel and successful expression of the recombinant proteins were confirmed by Western Blot analysis against the GST tag. GST fusion proteins were purified with the MagneGST™ paramagnetic particles that immobilize reduced glutathione

thus isolating fusion proteins from the cell lysate. Purified recombinant proteins was detected with anti-GST antibody and validated on Western Blot.

HCoV-NL63 N sequence was subjected to a series of algorithms for the prediction of antigenic sequences that can be used to synthesize peptides for detection of the putative protein. Peptides were designed and synthesized from two regions in the HCoV-NL63 N gene sequence that were identified as antigenic determinants that would most likely elicit an immune response that would lead to the production of polyclonal antibodies against N protein. Ideally the peptides were to be synthesized from regions within the sequence of NL63-N that present N- and C-terminal fragments. Peptides 1 and 2 were however synthesized from regions within the N-terminal region of HCoV-NL63 N sequence that proved to be most antigenic. Balb/c mice were immunized with a mixture of peptides 1 and 2 resulting in the production of polyspecific polyclonal antibodies. Polyclonal antibodies were successfully produced and validated using an indirect ELISA. Data from the ELISA screening showed peptide 1 to produce the best immune response. Splenic B-lymphocytes were fused with Sp2 myeloma cells to form hybridomas. Supernatant of hybridoma colonies screened by indirect ELISA and ten colonies selected for cloning of monoclonal antibodies which proved unsuccessful due to contamination and time constraints of this study. The purified GST-tagged viral fusion N proteins of SARS-CoV and HCoV-NL63 were subsequently used to characterize the immune activity of the produced polyclonal antibodies by Western Blotting. Polyclonal serum anti-N antibodies detected SARS-CoV N and HCoV-NL63 N protein and it's truncates, where cross-reactivity between SARS-N and HCoV-

NL63 N was observed. Further studies will need to be conducted to gain better understanding of cross-reactivity between HCoV-NL63 and other HCoVs.

It was possible to detect expression of HCoV-NL63 N protein and its truncated clones with synthetic anti-GST antibodies that can be purchased, as well as with blood serum of mice that were immunized with peptides that were designed and synthesized from the HCoV-NL63 N protein sequence to produce polyclonal antibodies.

In future studies, with the production of monoclonal antibodies against HCoV-NL63 N protein, it will be possible to detect the expression of NL63 N protein with more specificity and less background. Instead of using peptides, purified N protein may also be used as immunogen to elicit an antibody response to HCoV-NL63 N protein. For the design of peptides, antigenic properties of HCoV-NL63 N protein sequence should be determined using other bioinformatics tools, such as DNAStar, that can produce results about antigenic residues/epitopes with more accuracy. Instead of using a mixture of peptides when immunizing an animal host, the individual peptides should rather be administered by immunizing each animal with a single peptide. A better understanding of cross-reactivity among CoVs warrant for additional studies to improve serological responses to CoVs for purposes of laboratory diagnosis. Understanding how CoVs interact with each other will provide insight into understanding pathogenesis and immunity of CoVs.

# BIBLIOGRAPHY


Department of Medical Biosciences, Faculty of Natural Sciences, University of the Western Cape, South Africa

Abdul-Rasool, S. and Fielding, B.C. (2010). "Understanding Human Coronavirus HCoV-NL63." Open Virol J **4**: 76-84.

Al Hajjar, S., Al Thawadi, S., et al. (2011). "Human metapneumovirus and human coronavirus infection and pathogenicity in Saudi children hospitalized with acute respiratory illness." Ann Saudi Med **31**(5): 523-527.

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

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

Altman, A. and Dixon, F.J. (1989). "Immunomodifiers in vaccines." Adv Vet Sci Comp Med **33**: 301-343.

Amersham (2001). GST Gene Fusion System Handbook, Amersham Biosciences**:** 116.

An, S., Chen, C.J., 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.

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

Arndt, A.L., Larson, B.J., et al. (2010). "A conserved domain in the coronavirus membrane protein tail is important for virus assembly." J Virol **84**(21): 11418-11428.

Atassi, M.Z. (1984). "Antigenic structures of proteins." Eur J Biochem **145**(1): 1-20.

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

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

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

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

Brockway, S.M., Clay, C.T., et al. (2003). "Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase." J Virol **77**(19): 10515-10527.

Burns, R. (2005). Immunochemical Protocols, Humana Press.

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

Canducci, F., Debiaggi, M., et al. (2008). "Two-year prospective study of single infections and co-infections by respiratory syncytial virus and viruses identified recently in infants with acute respiratory disease." J Med Virol **80**(4): 716-723.

Casais, R., Davies, M., et al. (2005). "Gene 5 of the avian coronavirus infectious bronchitis virus is not essential for replication." J Virol **79**(13): 8065-8078.

Chan, K.C., Tang, N.L., et al. (2005). "Absence of association between angiotensin converting enzyme polymorphism and development of adult respiratory distress syndrome in patients with severe acute respiratory syndrome: a case control study." BMC Infect Dis **5**(1): 26.

Chan, K.H., Cheng, V.C.C., et al. (2005). "Serological Responses in Patients with Severe Acute Respiratory Syndrome Coronavirus Infection and Cross-Reactivity with Human Coronaviruses 229E, OC43, and NL63." Clin Diagn Lab Immunol **12**(11): 1317-1321.

Chang, C.K., Hsu, Y.L., et al. (2009). "Multiple nucleic acid binding sites and intrinsic disorder of severe acute respiratory syndrome coronavirus nucleocapsid protein: implications for ribonucleocapsid protein packaging." J Virol **83**(5): 2255-2264.

Chang, L.Y., Chiang, B.L., et al. (2006). "Lack of association between infection with a novel human coronavirus (HCoV), HCoV-NH, and Kawasaki disease in Taiwan." J Infect Dis **193**(2): 283-286.

Che, X.Y., Qiu, L.W., et al. (2004). "Sensitive and specific monoclonal antibody-based capture enzyme immunoassay for detection of nucleocapsid antigen in sera from patients with severe acute respiratory syndrome." J Clin Microbiol **42**(6): 2629-2635.

Chen, C.Y., Chang, C.K., 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.

Chiu, S.S., Chan, K.H., 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.

Chou, P.Y. and Fasman, G.D. (1978). "Prediction of the secondary structure of proteins from their amino acid sequence." Adv Enzymol Relat Areas Mol Biol **47**: 45-148.

Cologna, R. and Hogue, B.G. (1998). "Coronavirus nucleocapsid protein. RNA interactions." Adv Exp Med Biol **440**: 355-359.

Corman, V.M., Eckerle, I., et al. (2012). "Detection of a novel human coronavirus by real-time reverse-transcription polymerase chain reaction." Euro Surveill **17**(39).

Corse, E. and Machamer, C.E. (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.

Curtis, K.M., Yount, B., et al. (2002). "Heterologous gene expression from transmissible gastroenteritis virus replicon particles." J Virol **76**(3): 1422-1434.

Dare, R.K., Fry, A.M., et al. (2007). "Human coronavirus infections in rural Thailand: a comprehensive study using real-time reverse-transcription polymerase chain reaction assays." J Infect Dis **196**(9): 1321-1328.

de Haan, C.A., Kuo, L., et al. (1998). "Coronavirus particle assembly: primary structure requirements of the membrane protein." J Virol **72**(8): 6838-6850.

de Haan, C.A., Smeets, M., et al. (1999). "Mapping of the coronavirus membrane protein domains involved in interaction with the spike protein." J Virol **73**(9): 7441-7452.

de Haan, C.A., Vennema, H., et al. (2000). "Assembly of the coronavirus envelope: homotypic interactions between the M proteins." J Virol **74**(11): 4967-4978.

DeDiego, M.L., Alvarez, E., 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.

Dijkman, R., Jebbink, M.F., et al. (2012). "Replication-dependent downregulation of cellular angiotensin-converting enzyme 2 protein expression by human coronavirus NL63." J Gen Virol **93**(Pt 9): 1924-1929.

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

Dominguez, S.R., Anderson, M.S., et al. (2006). "Blinded case-control study of the relationship between human coronavirus NL63 and Kawasaki syndrome." J Infect Dis **194**(12): 1697-1701.

Dominguez, S.R., Sims, G.E., et al. (2012). "Genomic analysis of 16 Colorado human NL63 coronaviruses identifies a new genotype, high sequence diversity in the N-terminal domain of the spike gene and evidence of recombination." J Gen Virol **93**(Pt 11): 2387-2398.

Donoghue, M., Hsieh, F., et al. (2000). "A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9." Circ Res **87**(5): E1-9.

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

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

Engvall, E. and Perlmann, P. (1971). "Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G." Immunochemistry **8**(9): 871-874.

Esper, F., Ou, Z., et al. (2010). "Human coronaviruses are uncommon in patients with gastrointestinal illness." J Clin Virol **48**(2): 131-133.

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

Esper, F., Weibel, C., et al. (2005). "Evidence of a novel human coronavirus that is associated with respiratory tract disease in infants and young children." J Infect Dis **191**(4): 492-498.

ExpTec. (2009). "Protein Yield or Protein Expression Level." Retrieved 19 October 2012, from http://www.exptec.com/Expression%20 Technologies/ Protein%20Yield.htm#Technologies_to_increase_yield_caused_by_post-induction_toxicity.

Fan, H., Ooi, A., 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.

Fang, X., Ye, L., et al. (2005). "Peptide domain involved in the interaction between membrane protein and nucleocapsid protein of SARS-associated coronavirus." J Biochem Mol Biol **38**(4): 381-385.

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

Fischer, F., Stegen, C.F., et al. (1998). "Analysis of constructed E gene mutants of mouse hepatitis virus confirms a pivotal role for E protein in coronavirus assembly." J Virol **72**(10): 7885-7894.

Forster, J., Ihorst, G., 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., Hartwig, N.G., et al. (2004). "A previously undescribed coronavirus associated with respiratory disease in humans." Proc Natl Acad Sci U S A **101**(16): 6212-6216.

Fouchier, R.A., Kuiken, T., et al. (2003). "Aetiology: Koch's postulates fulfilled for SARS virus." Nature **423**(6937): 240.

Freund, J., Stern, E.R., et al. (1947). "Isoallergic Encephalomyelitis and Radiculitis in Guinea Pigs After One Injection of Brain and Mycobacteria in Water-in-Oil Emulsion." The J Immunol **57**(2): 179-194.

Garnier, J., Osguthorpe, D.J., et al. (1978). "Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins." J Mol Biol **120**(1): 97-120.

Gaunt, E.R., Hardie, A., et al. (2010). "Epidemiology and clinical presentations of the four human coronaviruses 229E, HKU1, NL63, and OC43 detected over 3 years using a novel multiplex real-time PCR method." J Clin Microbiol **48**(8): 2940-2947.

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

Glowacka, I., Bertram, S., et al. (2010). "Differential downregulation of ACE2 by the spike proteins of severe acute respiratory syndrome coronavirus and human coronavirus NL63." J Virol **84**(2): 1198-1205.

Goding, J.W. (1996). Monoclonal Antibodies: Principles and Practice, Academic Press.

Graham, R.L. and Baric, R.S. (2010). "Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission." J Virol **84**(7): 3134-3146.

Haijema, B.J., Volders, H., 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.

Hajimorad, M.R. and Francki, R.I. (1991). "Some observations on the binding properties of alfalfa mosaic virus to polystyrene and its significance to indirect ELISA." Arch Virol **117**(3-4): 219-235.

Hancock, D.C. and O'Reilly, N.J. (2005). "Synthetic peptides as antigens for antibody production." Methods Mol Biol **295**: 13-26.

Harlow, E. and Lane, D. (1988). "A laboratory manual." New York: Cold Spring Harbor Laboratory **579**.

Hayes, A.J., Hughes, C.E., et al. (2008). "Antibodies and immunohistochemistry in extracellular matrix research." Methods **45**(1): 10-21.

He, R., Leeson, A., 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.

Hearn, M.T.H., Bishop, C.A., et al. (1979). "Application of Reversed Phase High Performance Liquid Chromatography in Solid Phase Peptide Synthesis." J Liq Chromatogr **2**(1): 1-21.

Hermanson, G.T. (2008). Chapter 20 - Antibody Modification and Conjugation. Bioconjugate Techniques (Second Edition). New York, Academic Press**:** 783-823.

Hiscox, J.A., Cavanagh, D., et al. (1995). "Quantification of individual subgenomic mRNA species during replication of the coronavirus transmissible gastroenteritis virus." Virus Res **36**(2-3): 119-130.

Hofmann, H., Pyrc, K., 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.

Holcroft, C.C. and Egan, S.M. (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.

Holmes, K.V. and Enjuanes, L. (2003). "Virology. The SARS coronavirus: a postgenomic era." Science **300**(5624): 1377-1378.

Hopp, T.P. and Woods, K.R. (1981). "Prediction of protein antigenic determinants from amino acid sequences." Proc Natl Acad Sci U S A **78**(6): 3824-3828.

Houen, G., Jakobsen, M.H., et al. (1997). "Conjugation to preadsorbed preactivated proteins and efficient generation of anti peptide antibodies." J Immunol Methods **206**(1-2): 125-134.

Houen, G., Olsen, D.T., et al. (2003). "Preparation of bioconjugates by solid-phase conjugation to ion exchange matrix-adsorbed carrier proteins." Bioconjug Chem **14**(1): 75-79.

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

Hurst, K.R., Kuo, L., 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.

ICTV. (2011). "ICTV Master Species List 2011 v2." Retrieved 29 October 2012, from http://talk.ictvonline.org/files/ictv_documents/m/msl/4090.aspx.

Jevsnik, M., Ursic, T., et al. (2012). "Coronavirus infections in hospitalized pediatric patients with acute respiratory tract disease." BMC Infect Dis **12**: 365.

Kahn, J.S. (2006). "The widening scope of coronaviruses." Curr Opin Pediatr **18**(1): 42-47.

Kaneshiro, N.K. and Zieve, D. (2012). "Croup." Retrieved 27 February 2013, from http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001955/.

Kleines, M., Scheithauer, S., 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." J Clin Microbiol **45**(3): 1032-1034.

Klumperman, J., Locker, J.K., et al. (1994). "Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding." J Virol **68**(10): 6523-6534.

Koetz, A., Nilsson, P., et al. (2006). "Detection of human coronavirus NL63, human metapneumovirus and respiratory syncytial virus in children with respiratory tract infections in south-west Sweden." Clin Microbiol Infect **12**(11): 1089-1096.

Kohler, G. and Milstein, C. (2005). "Continuous cultures of fused cells secreting antibody of predefined specificity. 1975." J Immunol **174**(5): 2453-2455.

Kolaskar, A.S. and Tongaonkar, P.C. (1990). "A semi-empirical method for prediction of antigenic determinants on protein antigens." FEBS Lett **276**(1-2): 172-174.

Kowalczyk, W., Monso, M., et al. (2011). "Synthesis of multiple antigenic peptides (MAPs)-strategies and limitations." J Pept Sci **17**(4): 247-251.

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

Kuo, L. and Masters, P.S. (2003). "The small envelope protein E is not essential for murine coronavirus replication." J Virol **77**(8): 4597-4608.

Lau, S.K., Woo, P.C., et al. (2004). "Detection of severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein in sars patients by enzyme-linked immunosorbent assay." J Clin Microbiol **42**(7): 2884-2889.

Lee, B.S., Huang, J.S., et al. (2010). "Production of antipeptide antibodies." Methods Mol Biol **657**: 93-108.

Lehmann, C., Klar, R., et al. (2009). "Kawasaki disease lacks association with human coronavirus NL63 and human bocavirus." Pediatr Infect Dis J **28**(6): 553-554.

Lehmann, C., Wolf, H., et al. (2008). "A line immunoassay utilizing recombinant nucleocapsid proteins for detection of antibodies to human coronaviruses." Diagn Microbiol Infect Dis **61**(1): 40-48.

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

Lindblad, E.B., Elhay, M.J., et al. (1997). "Adjuvant modulation of immune responses to tuberculosis subunit vaccines." Infect Immun **65**(2): 623-629.

Lu, R., Yu, X., et al. (2012). "Characterization of human coronavirus etiology in Chinese adults with acute upper respiratory tract infection by real-time RT-PCR assays." PLoS One **7**(6): e38638.

Luo, H., Chen, Q., 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.

Mackay, I.M., Arden, K.E., et al. (2012). "Co-circulation of four human coronaviruses (HCoVs) in Queensland children with acute respiratory tract illnesses in 2004." Viruses **4**(4): 637-653.

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

McBride, R., van Zyl, M., et al. (2014). "The coronavirus nucleocapsid is a multifunctional protein." Viruses **6**(8): 2991-3018.

McIntosh, K. (2005). "Coronaviruses in the Limelight." Journal of Infectious Diseases **191**(4): 489-491.

McIntosh, K., Dees, J.H., et al. (1967). "Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease." Proc Natl Acad Sci U S A **57**(4): 933-940.

Moes, E., Vijgen, L., 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 infectious diseases **5**(Journal Article): 6.

Muller, A., Tillmann, R.L., et al. (2008). "Stability of human metapneumovirus and human coronavirus NL63 on medical instruments and in the patient environment." J Hosp Infect **69**(4): 406-408.

Muller, M.A., van der Hoek, L., et al. (2010). "Human coronavirus NL63 open reading frame 3 encodes a virion-incorporated N-glycosylated membrane protein." Virol J **7**: 6.

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

Nguyen, V.P. and Hogue, B.G. (1997). "Protein interactions during coronavirus assembly." J Virol **71**(12): 9278-9284.

Niman, H.L., Houghten, R.A., et al. (1983). "Generation of protein-reactive antibodies by short peptides is an event of high frequency: implications for the structural basis of immune recognition." Proc Natl Acad Sci U S A **80**(16): 4949-4953.

Ohara, K., Horibe, T., et al. (2011). "Characterization of antilytic peptide antibody: application for the detection of lytic-based hybrid peptide in serum samples." J Pept Sci **17**(7): 493-498.

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

Paddon, C.J., Vasantha, N., et al. (1989). "Translation and processing of Bacillus amyloliquefaciens extracellular RNase." J Bacteriol **171**(2): 1185-1187.

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

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

Posnett, D.N., McGrath, H., et al. (1988). "A novel method for producing anti-peptide antibodies. Production of site-specific antibodies to the T cell antigen receptor beta-chain." J Biol Chem **263**(4): 1719-1725.

Principi, N., Bosis, S., et al. (2010). "Effects of coronavirus infections in children." Emerg Infect Dis **16**(2): 183-188.

Promega (2009a). MagneGST Protein Prufication System USA, Promega Corporation

Promega (2009b). pFN2A (GST) Flexi® Vector. USA, Promega Corporation

Promega (2009c). Single Step (KRX) Competent Cells. USA, Promega Corporation

Pyrc, K., Berkhout, B., et al. (2005). "Molecular characterization of human coronavirus NL63." Recent Research in Development of Infection & Immunity **3**: 25-48.

Pyrc, K., Berkhout, B., et al. (2007a). "Antiviral strategies against human coronaviruses." Infect Disord Drug Targets. **7**(1): 59-66.

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

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

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

Qin, Q., Liu, Y.L., et al. (2005). "Construction of a transposon-mediated baculovirus vector Hanpvid and a new cell line for expressing barnase." J Biochem Mol Biol **38**(1): 41-48.

Risku, M., Lappalainen, S., et al. (2010). "Detection of human coronaviruses in children with acute gastroenteritis." J Clin Virol **48**(1): 27-30.

Rottier, P., Brandenburg, D., et al. (1984). "Assembly in vitro of a spanning membrane protein of the endoplasmic reticulum: the E1 glycoprotein of coronavirus mouse hepatitis virus A59." Proc Natl Acad Sci U S A **81**(5): 1421-1425.

Rubinstein, N.D., Mayrose, I., et al. (2008). "Computational characterization of B-cell epitopes." Mol Immunol **45**(12): 3477-3489.

Ryan, B.J. and Henehan, G.T. (2013). "Overview of approaches to preventing and avoiding proteolysis during expression and purification of proteins." Curr Protoc Protein Sci **Chapter 5**: Unit5 25.

Saiki, R.K., Scharf, S., et al. (1985). "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia." Science **230**(4732): 1350-1354.

Sambrook, J. and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Samuelson, J. (2011). Bacterial Systems. Production of Membrane Proteins, Wiley-VCH Verlag GmbH & Co. KGaA**:** 11-35.

Saravanan, P., Kataria, J., et al. (2003). "Detection of Infectious bursal disease virus by ELISA using an antipeptide antibody raised against VP3 region." Acta virologica **48**(1): 39-45.

Sastre, P., Dijkman, R., et al. (2011). "Differentiation between human coronaviruses NL63 and 229E using a novel double-antibody sandwich enzyme-linked immunosorbent assay based on specific monoclonal antibodies." Clin Vaccine Immunol **18**(1): 113-118.

Schaaper, W.M., Lankhof, H., et al. (1989). "Manipulation of antipeptide immune response by varying the coupling of the peptide with the carrier protein." Mol Immunol **26**(1): 81-85.

Schagat, T., Ohana, R., et al. (2008) "KRX Autoinduction protocol: a convenient method for protein expression." Promega Notes **98**, 16-18.

Schlom, J., Wunderlich, D., et al. (1980). "Generation of human monoclonal antibodies reactive with human mammary carcinoma cells." Proc Natl Acad Sci U S A **77**(11): 6841-6845.

Schulz, S., Rocken, C., et al. (2006). "Immunohistochemical detection of bombesin receptor subtypes GRP-R and BRS-3 in human tumors using novel antipeptide antibodies." Virchows Arch **449**(4): 421-427.

Severance, E.G., Bossis, I., et al. (2008). "Development of a nucleocapsid-based human coronavirus immunoassay and estimates of individuals exposed to

coronavirus in a U.S. metropolitan population." Clin Vaccine Immunol **15**(12): 1805-1810.

Shao, X., Guo, X., et al. (2007). "Seroepidemiology of group I human coronaviruses in children." J Clin Virol **40**(3): 207-213.

Shimizu, C., Shike, H., 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.

Shinnick, T.M., Sutcliffe, J.G., et al. (1983). "Synthetic peptide immunogens as vaccines." Annu Rev Microbiol **37**: 425-446.

Siu, Y.L., Teoh, K.T., et al. (2008). "The M, E, and N structural proteins of the severe acute respiratory syndrome coronavirus are required for efficient assembly, trafficking, and release of virus-like particles." J Virol **82**(22): 11318-11330.

Smuts, H. and Hardie, D. (2006). "Human bocavirus in hospitalized children, South Africa." Emerg Infect Dis **12**(9): 1457-1458.

Smuts, H., Workman, L., et al. (2008). "Role of human metapneumovirus, human coronavirus NL63 and human bocavirus in infants and young children with acute wheezing." J Med Virol **80**(5): 906-912.

Stewart-Tull, D.E.S. (1995). The Theory and Practical Application of Adjuvants, Chichester, UK: John Wiley & Sons.

Stills, H.F., Jr. (2005). "Adjuvants and antibody production: dispelling the myths associated with Freund's complete and other adjuvants." ILAR J **46**(3): 280-293.

Strynadka, N.C., Redmond, M.J., et al. (1988). "Use of synthetic peptides to map the antigenic determinants of glycoprotein D of herpes simplex virus." J Virol **62**(9): 3474-3483.

Sturman, L.S., Holmes, K.V., et al. (1980). "Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid." J Virol **33**(1): 449-462.

Sun, Z.F. and Meng, X.J. (2004). "Antigenic Cross-Reactivity between the Nucleocapsid Protein of Severe Acute Respiratory Syndrome (SARS) Coronavirus and Polyclonal Antisera of Antigenic Group I Animal Coronaviruses: Implication for SARS Diagnosis." J Clin Microbiol **42**(5): 2351-2352.

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

Surjit, M., Liu, B., et al. (2004). "The SARS coronavirus nucleocapsid protein induces actin reorganization and apoptosis in COS-1 cells in the absence of growth factors." Biochem J **383**(Pt 1): 13-18.

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

Tam, J.P. (1988). "Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system." Proc Natl Acad Sci U S A **85**(15): 5409-5413.

Tang, T.K., Wu, M.P., et al. (2005). "Biochemical and immunological studies of nucleocapsid proteins of severe acute respiratory syndrome and 229E human coronaviruses." Proteomics **5**(4): 925-937.

Thornton, J.M., Edwards, M.S., et al. (1986). "Location of 'continuous' antigenic determinants in the protruding regions of proteins." EMBO J **5**(2): 409-413.

Timani, K.A., Ye, L., et al. (2004). "Cloning, sequencing, expression, and purification of SARS-associated coronavirus nucleocapsid protein for serodiagnosis of SARS." J Clin Virol **30**(4): 309-312.

Towbin, H., Staehelin, T., et al. (1979). "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications." Proc Natl Acad Sci U S A **76**(9): 4350-4354.

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

Vabret, A., Dina, J., et al. (2009). "[Human coronaviruses]." Pathol Biol (Paris) **57**(2): 149-160.

Vabret, A., Dina, J., et al. (2006). "Detection of the new human coronavirus HKU1: a report of 6 cases." Clin Infect Dis **42**(5): 634-639.

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

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

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

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

van Elden, L.J., van Loon, A.M., et al. (2004). "Frequent detection of human coronaviruses in clinical specimens from patients with respiratory tract infection by use of a novel real-time reverse-transcriptase polymerase chain reaction." J Infect Dis **189**(4): 652-657.

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

Vlasova, A.N., Zhang, X., et al. (2007). "Two-Way Antigenic Cross-Reactivity between Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Group 1 Animal CoVs Is Mediated through an Antigenic Site in the N-Terminal Region of the SARS-CoV Nucleoprotein." J Virol **81**(24): 13365-13377.

Voller, A., Bertlett, A., et al. (1978). "Enzyme immunoassays with special reference to ELISA techniques." J Clin Pathol **31**: 507-520.

Voss, D., Pfefferle, S., et al. (2009). "Studies on membrane topology, N-glycosylation and functionality of SARS-CoV membrane protein." Virol J **6**: 79.

Wang, G.Z., Tang, X.D., et al. (2011). "Multiple antigenic peptides of human heparanase elicit a much more potent immune response against tumors." Cancer Prev Res (Phila) **4**(8): 1285-1295.

Wang, Y., Wu, X., et al. (2004). "Low stability of nucleocapsid protein in SARS virus." Biochemistry **43**(34): 11103-11108.

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

Woo, P.C., Lau, S.K., et al. (2004a). "False-positive results in a recombinant severe acute respiratory syndrome-associated coronavirus (SARS-CoV) nucleocapsid enzyme-linked immunosorbent assay due to HCoV-OC43 and HCoV-229E rectified by Western blotting with recombinant SARS-CoV spike polypeptide." J Clin Microbiol **42**(12): 5885-5888.

Woo, P.C., Lau, S.K., et al. (2004b). "Detection of specific antibodies to severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein for serodiagnosis of SARS coronavirus pneumonia." J Clin Microbiol **42**(5): 2306-2309.

Woo, P.C., Lau, S.K., et al. (2005b). "Differential sensitivities of severe acute respiratory syndrome (SARS) coronavirus spike polypeptide enzyme-linked immunosorbent assay (ELISA) and SARS coronavirus nucleocapsid protein ELISA for serodiagnosis of SARS coronavirus pneumonia." J Clin Microbiol **43**(7): 3054-3058.

Wu, K., Chen, L., et al. (2011). "A virus-binding hot spot on human angiotensin-converting enzyme 2 is critical for binding of two different coronaviruses." J Virol **85**(11): 5331-5337.

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

Xin, C., Yong, Z.Z., et al. (2012). "Human coronavirus NL63 in hospitalized children with respiratory infection: a 2-year study from Chongqing, China." Indian Pediatr **49**(10): 825-828.

Yang, Y., Xiong, Z., 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.

Yazynin, S., Lange, H., et al. (1999). "A new phagemid vector for positive selection of recombinants based on a conditionally lethal barnase gene." FEBS Lett **452**(3): 351-354.

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

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

Yu, X., Lu, R., et al. (2012). "Etiology and clinical characterization of respiratory virus infections in adult patients attending an emergency department in Beijing." PLoS One **7**(2): e32174.

Zaki, A.M., van Boheemen, S., et al. (2012). "Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia." N Engl J Med **367**(19): 1814-1820.

Zarei, O., Irajian, G.R., et al. (2011). "Peptide-based Polyclonal Antibody Production against P110 Protein of Mycoplasma genitalium." Avicenna J Med Biotechnol **3**(2): 79-85.

Zhang, G., Hu, Y., et al. (2012). "High incidence of multiple viral infections identified in upper respiratory tract infected children under three years of age in Shanghai, China." PLoS One **7**(9): e44568.

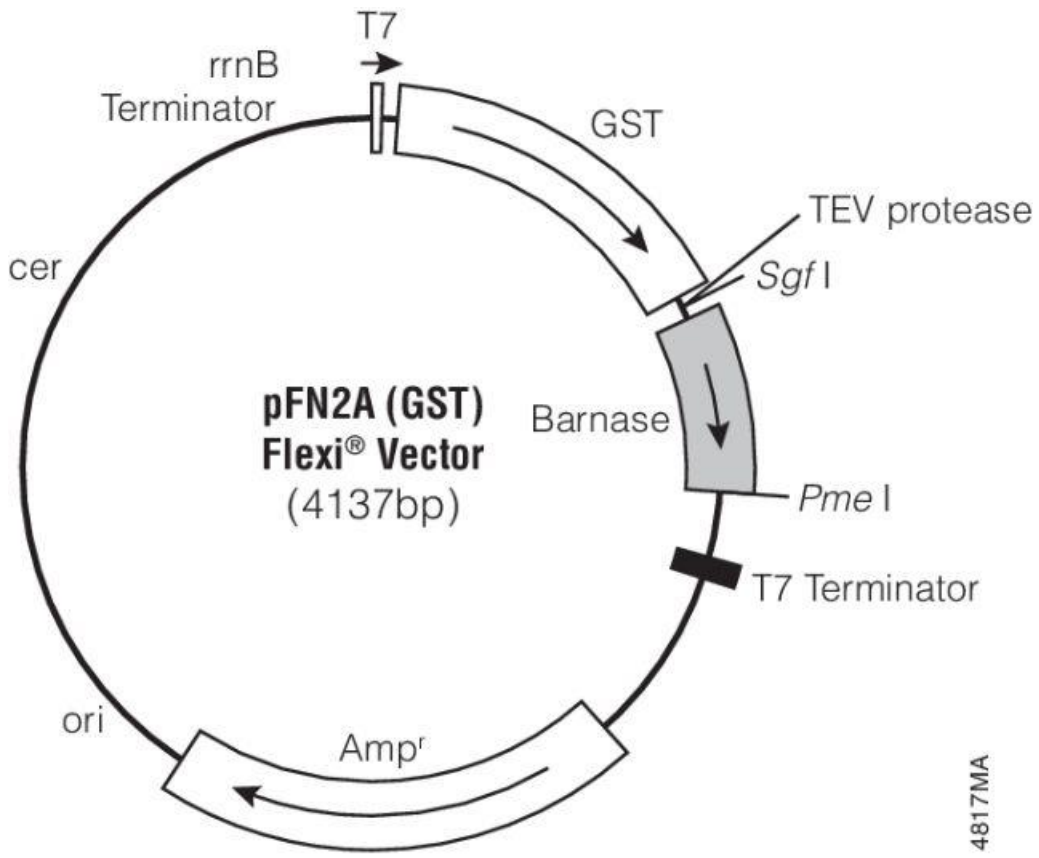
Zhao, J., Wang, W., et al. (2007). "Comparison of immunoglobulin G responses to the spike and nucleocapsid proteins of severe acute respiratory syndrome (SARS) coronavirus in patients with SARS." Clin Vaccine Immunol **14**(7): 839-846.

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

# APPENDIX


Department of Medical Biosciences, Faculty of Natural Sciences, University of the Western Cape, South Africa

# Appendix 1





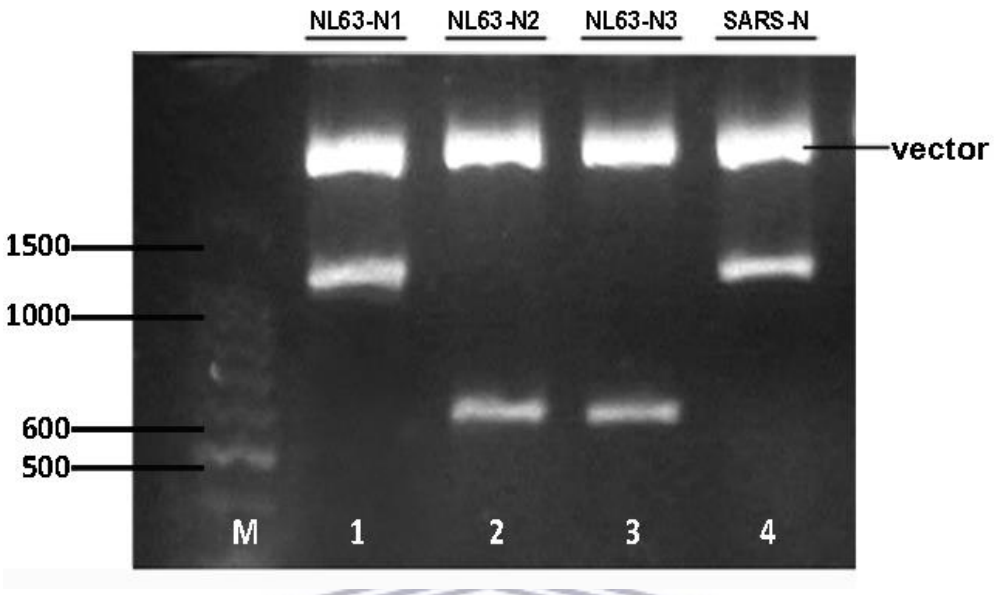

**pFN2A (GST) Flexi vector circle map and sequence reference points.** The vector contains a T7 promoter for bacterial or *in vitro* protein expression of a protein-coding region. The vector appends an N-terminal GST coding region that can be used to purify the expressed protein. The GST tag contains a TEV protease site for removal of the tag after purification. The vector also contains the lethal barnase gene for positive selection of the insert, an ampicillin-resistance gene for selection of the plasmid and unique endonucleases restriction sites *SgfI* and *PmeI* that allows easy insertion or transfer of the sequence of interest. [Technical Manual pFN2A (GST) Flexi® Vector Systems. 2009. Promega Corporation].

# Appendix 2

MASVNWADDR AARKKFPPPS FYMPLLVSSD KAPYRVIPRN LVPIGKGNKD
EQIGYWNVQE RWRMRRGQRV DLPPKVHFYY LGTGPHKDLK FRQRLDGVVW
VAKEGAKTVN TSLGNRKRNQ KPLEPKFSIA LPPELSVVEF EDRSNNSSRA
SSRSSTRNNS RDSSRSTSRQ QSRTRSDSNQ SSSDLVAAVT LALKNLGFDN
QSKSPSSSGT STPKKPNKPL SQPRADKPSQ LKKPRWKRVP TREENVIQCF
GPRDFNHNMG DSDLVQNGVD AKGFPQLAEL IPNQAALFFD SEVSTDEVGD
NVQITYTYKM LVAKDNKNLP KFIEQISAFT KPSSVKEMQS QSSHVAQNTV
LNASIPESKP LADDDSAIIE IVNEVLH

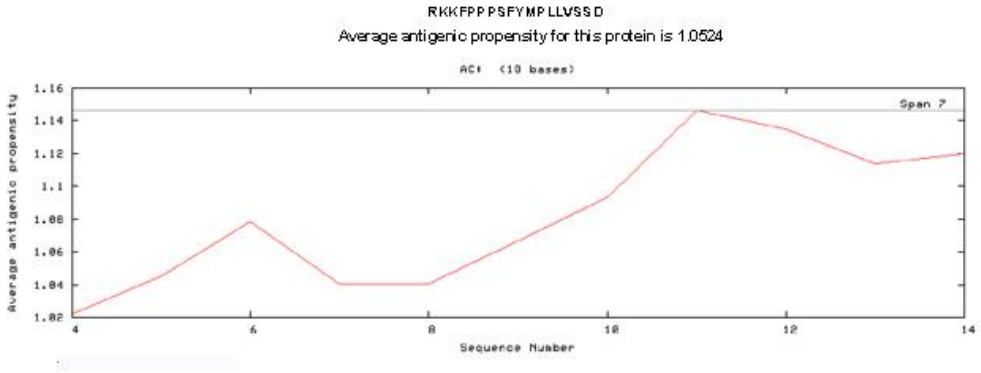
**HCoV-NL63 N protein sequence as per NCBI Genbank.** Single-letter amino acid code for full sequence of HCoV-NL63 N protein.

# Appendix 3



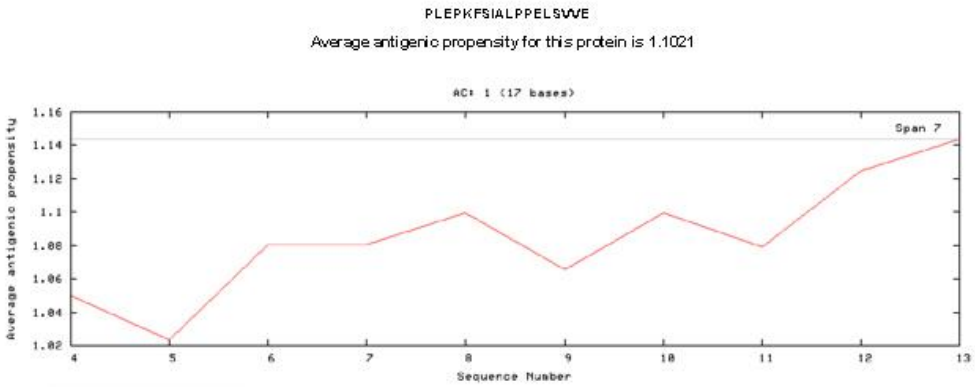

**Restriction enzyme digest of isolated SARS-CoV and HCoV-NL63 N plasmid DNA from miniprep lysate.** Lane M: 100bp DNA ladder. Lane 1: HCoV-NL63 full-length N gene (N1) (~1100bp); Lane 2: N-terminal fragment of HCoV-NL63 N gene (N2) (~550bp); Lane 3: C-terminal fragment of HCoV-NL63 N gene (N3) (~550bp); Lane 4: SARS-CoV full-length N gene (SARS-N) (~1200bp).

# Appendix 4



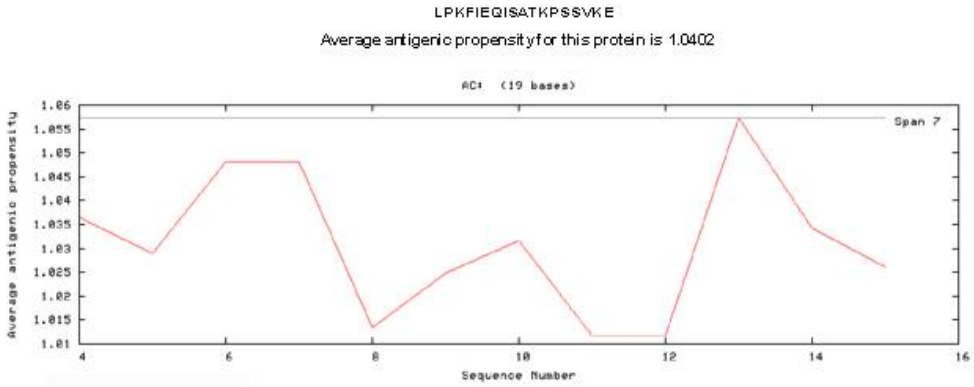

**Average antigenic propensity plot for sequence n=1.** Antigenic propensity plot of sequence n=1 (RKKFPPPSFYMPLLSSD; aa 13-30). The average antigenic propensity for this sequence is 1.5024.

# Appendix 5




**Antigenic Propensity plot of Sequence *n*=5.** Antigenic propensity plot of sequence n=5 (PLEPKFSIALPPELSVVE; aa 122-139). Average antigenic propensity of 1.1021.

# Appendix 6




**Average antigenic propensity plot of sequence n=11.** Antigenic propensity plot of sequence n=11 (LPKFIEQISAFTKPSSVKE; aa 319-337). Average antigenic propensity for this sequence is 1.0402.