The use of recombinant nucleocapsid proteins as diagnostic antigens in human coronavirus cases

PhD

School of Biological Sciences

Rebecca Willmot

April 2020

Declaration of Authorship

Declaration:

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Statement of Ethics

All procedures performed in studies involving human participants were in accordance with the ethical standards of The University of Reading and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent was obtained from all individual participants involved in the study.

Abstract

Coronaviruses (CoV) primarily infect the upper respiratory and gastrointestinal tract of birds and mammals making them an important class of infections for agriculture, industry and human health. In 2003 an endemic of severe acute respiratory syndrome (SARS) resulted in approximately 8000 infections with a 10% mortality rate. This discovery added to four coronaviruses previously documented as being able to infect and cause disease in humans; OC43-CoV, KHU1-CoV, 229E-CoV and NL63-CoV. In 2012 Middle East respiratory syndrome coronavirus (MERS-CoV) emerged. Both SARS-CoV and MERS-CoV are thought to have originated from coronaviruses found in bats which were transmitted to man through different intermediate vectors. More recently, SARS-CoV-2 (commonly referred to as CoVid-19) emerged in December 2019 resulting in a world-wide pandemic. The coronavirus N protein is a ~45kDa protein found associated with the genome of the virus. Despite a common function in RNA binding CoV N proteins are antigenically distinct making them useful as diagnostic antigens for tests of seroconversion. To assess the MERS-CoV N protein as a diagnostic antigen a soluble full length His-tagged N protein was expressed using *E. coli*. Recombinant N protein was purified to homogeneity by IMAC chromatography and was observed as a single species of the predicted molecular weight with minimum breakdown. When used as a capture antigen in ELISA tests with a number of human CoV positive sera, recombinant MERS-CoV N protein was shown to react strongly with MERS-CoV positive sera but not with sera from other CoV infections. Similar data was obtained by western blot. These data suggest recombinant MERS-CoV N protein is a suitable antigen for serosurveillance. The expression of a number of other CoV N proteins to provide a mini array of N proteins for tests of a variety of human sera for their history of coronavirus infection will be described.

Acknowledgements

This thesis has been a long time in the making, and as such I have a number of people to thank for their help and support and mostly for their patience.

Reading University has been a huge part of my life science I first attended 12 years ago to start my undergraduate degree. During that time, I have met truly wonderful friends; Abi and Kate thank you for all the laughs and cake. I have looked up to and learnt from those with more experience; Silvia and Jay my laboratory mums. I have had the honour of being taught by inspirational scientist; Professor Ian Jones, Doctor Ben Neuman, Professor Mike Fry, Professor Kim Watson to name a few. And I have had the assistance of wonderful support staff; Kevan and the CSS ladies have proved invaluable. It is bittersweet to submit this thesis and have my affiliation with the university come to an end.

My passion for science and love for both learning and teaching will be taken with me wherever I go in life though. As a secondary school teacher; I look forward to inspiring generations to come to feel a similar love of all things microbial. I would like to say a huge thank you to all the wonderful students at The Misbourne who have given me the great honour of being their teacher the past two years. My form, in particular deserve a mention, a group of incredibly special teenagers who believe in me as much as I believe in them and have given me the confidence to submit. My students confirm that my transition from scientist to science teacher was absolutely the right one.

I'd like to thank all my friends who listen and support. Particularly to Ally, who has put up with more than most. To the friends who should still be here but aren't; Chantelle and Sarah, I hope you're looking down and are proud. Thank you to my dear friend Chris Davies, an exceptionally special person. Thank you to my brothers for letting me prove that not only are girls cleverer but also that Biology is the most interesting of the sciences! And the final, most important thank you, is to my parents. Without whom this thesis would not have been possible. Thank you mostly for getting me Monty the dog.

List of figures

List of tables

Company list

Company reference in text	Full company name	Location
Sigma	Sigma-Aldrich® Co LLC	Gillingham, Dorset, UK
Fisher	Fisher Chemicals	Loughborough, Leicester, UK
Abcam	Abcam plc	Cambridge, Cambridgeshire, UK
Dako	Dako, Agilent Technologies	Ely, Cambridgeshire, UK
Invitrogen	Invitrogen ThermoFisher LifeTechnologies Ltd	Paisley, Renfrewshire, UK
Oxoid	Oxoid Ltd	Basingstoke, Hampshire, UK
Geneflow	Geneflow Ltd	Lichfield, Staffordshire, UK
Lonza	Lonza group Ltd	Slough, UK
Millipore	EMD Millipore	Darmstadt, Germany
ThermoScientific	ThermoScientific	Loughborough, UK
Agilent Technologies	Agilent Technologies	California, USA
Kerafast	Kerafast Inc	Boston, USA
Melford	Melford Laboratories Ltd	Ipswich, Surrey, UK
Qiagen	Qiagen	Manchester, UK
Cambridge BioScience	Cambridge BioScience	Cambridge, UK
Source BioScience	Source BioScience, LifeSciences	Nottingham, UK
Marvel	Marvel Premier Food	UK
GE Healthcare	GE Healthcare	Little Chalfont, Buckinghamshire, UK
Boehringer Ingelheim GmbH	Boehringer Ingelheim GmbH	Ingelheim, Germany
DNA Technologies	DNA Technologies Inc	Leuven, Belgium
Satorius	Satorius Ltd	Epsom, Surrey, UK
Bio-Rad	Biorad	Hemel Hemstead, Hertfordshire, UK
Europa	Europa Bioproducts Ltd	Cambridge, UK
forMedium	forMedium	Norfolk, UK
Roche	Roche	Switzerland

Prologue

The original first draft of the abstract that I summited alongside my poster for the International Union of Microbiological Societies in 2014 included the following sentence "without the proof-reading ability of DNA polymerase, RNA viruses have an increased mutation rate. This, added to a large and unpredictable natural reservoir, means that it is likely that new, potentially epidemic or pandemic, coronaviruses may emerge in the future."

Indeed, post writing this thesis, a novel coronavirus outbreak (SARS-CoV-2, more commonly referred to as Covid-19) emerged in December 2019 (Velavan and Meyer 2020). The virus epi-centred in Hubei Province of the People's Republic of China and rapidly spread, resulting in a devastating worldwide pandemic.

The outbreak has meant a change to my current job role as Deputy Head of Science at a secondary school. As I am now adjusting to setting distance learning and am currently shielding, as per government advice, due to having underlying health conditions. It is for this reason that throughout the body of text only the 6 human coronaviruses known at the time of writing are largely mentioned.

The latest report from the world health organisation (WHO) published in December 2020 stated that there were 2,566 laboratory-confirmed cases of infection with MERS-CoV, including 882 deaths giving it a 34.4% fatality rate (WHO 2020). WHO reported SARS-CoV cases as totalling 8,422 including 916 deaths (fatality rate 10.9%) (WHO 2020). At the time of writing, August 2021, Covid-19's case numbers are estimated to be 207,784,507 resulting in 4,370,424 deaths (fatality rate 2.1 %) (WHO 2021). Due to increased understanding of the virus and the implementation of various disease transmission prevention strategies, alongside the introduction of vaccination programs incidences of new cases and deaths continue to decrease (WHO 2021). Clearly understanding coronaviruses in greater detail and working on vaccines, treatments and diagnostic kits has never been more topical.

Contents

Chapter 1: Introduction

1.1 Coronaviruses introduction

Coronaviruses (CoV) are Baltimore group four enveloped viruses with a positive-sense ribonucleic acid (RNA) genome. Coronaviruses are of the order Nidovirales, the family Coronaviridae and the subfamily Coronavirinae. Coronavirinae consists of four genera; Alphacoronavirus, Betacoronavirus, Gammacoronavirus and Deltacoronavirus. The genomic size of coronaviruses ranges from 26 kilobases to 32 kilobases, which is the largest of any RNA virus. The virions are spherical and approximately 70-120nm in diameter (Graham, Donaldson et al. 2013). The name coronavirus is derived from the Latin word "corona", meaning crown or halo, and refers to the characteristic appearance of virions under transmission electron microscopy (TEM), cryo-electron microscopy (cryo-EM), scanning electron microscopy (SEM) or atomic force microscopy (AFM). The projections seen are created by the viral spike (S) glycoproteins, as shown in Figure 1. S proteins are homotrimeric single-pass transmembrane (TM) proteins and it is these proteins that determine host tropism. Alongside the spike proteins; envelope (E), membrane (M) and nucleocapsid (N) proteins all contribute to the overall structure of coronaviruses (de Groot 2011) (King 2012).

1.1.1 Structure and proteins

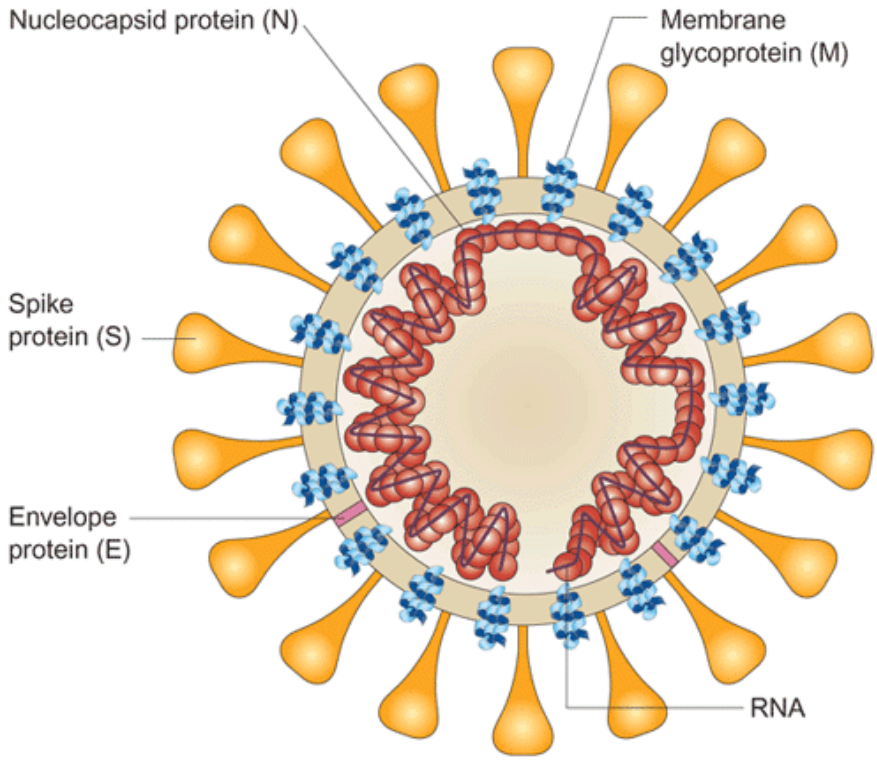


Figure 1: The basic structure of a coronavirus virion including the four main structural proteins; spike (S), envelope (E), membrane (M), nucleocapsid (N) (Peiris, Guan et al. 2004).

1.1.1.1 S (spike) protein (~180kD) (Graham, Donaldson et al. 2013)

The spike glycoprotein shown in Figure 1 contains a receptor-binding domain (RBD) used to facilitate viral entry into a cell. The RBD is the principle determinant of host range and yet it is poorly conserved among viruses; allowing host receptor usage to vary, as shown in Table 1 (Graham, Donaldson et al. 2013). The S glycoprotein has two subunit domains; the amino-terminal S1 domain, where the RBD is located and the carboxy-terminal S2 domain containing the putative fusion peptide, two heptad repeat domains and a TM domain (Graham and Baric 2010). S proteins can bind to sialic acid (9-O-acetyl neuraminic acid) on the host cell surface which gives the virus haemagglutinating ability (Huang, Dong et al. 2015) and antibodies against S protein are neutralizing (He, Lu et al. 2005).

Virus	Group	Receptor	May also bind
Human coronavirus 229E-CoV (229E-CoV)	Alpha	APN (Lachance, Arbour et al. 1998)	-
Human coronavirus NL63-CoV (NL63-CoV)	Alpha	ACE2 (Hofmann, Pyrc et al. 2005)	-
Severe acute respiratory syndrome coronavirus (SARS-CoV)	Beta	ACE2 (Hofmann, Pyrc et al. 2005) (Li, Moore et al. 2003)	DC-SIGN, DC-SIGNR, LSECtin (Gramberg, Hofmann et al. 2005)
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)	Beta	ACE2 (Harrison, Lin et al. 2020)	-
Middle East respiratory syndrome coronavirus (MERS-CoV)	Beta	DPP4 aka CD26 (Müller 2012) (Raj, Mou et al. 2013)	-
Human coronavirus OC43-CoV (OC43-CoV)	Beta	Unknown	Sialic acid (Kunkel and Herrler 2003)
Avian infectious bronchitis virus (IBV)	Gamma	Unknown	Sialic acid (Shahwan, Hesse et al. 2013)

Table 1: The coronaviruses used in this study along with their receptor. Table adapted from Journal of Virology (Graham and Baric 2010) APN= aminopeptidase N (aka CD13), ACE2= angiotensin converting enzyme 2, DC-SIGN= Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin, DC-SIGNR= dendritic cell-specific ICAM-grabbing non-integrin, ICAM= intercellular adhesion molecule, LSECtin= liver and lymph node sinusoidal endothelial cell C-type lectin, DPP4= dipeptidyl peptidase 4 and CD26= cluster of differentiation 26

1.1.1.2 HE (Haemagglutinin-Esterase) protein (65kD)

HE proteins are unique to the Betacoronaviruses subgroup (Huang, Dong et al. 2015). HE is a diamer and has been shown to be unessential for replication (Popova and Zhang 2002). The protein is used to create a reversible attachment to O-acetylated sialic acids on the host cell by acting initially as a lectin and later as a receptor-destroying enzyme (RDEs) in order to facilitate the release of viral progeny (Huang, Dong et al. 2015). Antibodies against HE protein can also neutralize the virus (Kasai, Morita et al. 1998).

1.1.1.3 M (membrane) protein (22-25kD) (Ujike and Taguchi 2015)

The M protein is the most abundant protein in the virus envelope (Ujike and Taguchi 2015). It consists of a short, glycosylated N-terminal ectodomain made up of three TM domains (referred to as tm1, tm2 and tm3 regions) and a long C-terminal cytoplasmic tail domain. The protein

provides the virion some rigidity and helps the nucleocapsid to attach to the membranes of internal structures such as the Golgi apparatus (Ujike and Taguchi 2015).

1.1.1.4 E (envelope) protein (9-12kD) (Wilson, McKinlay et al. 2004)

Found on the viral membrane the envelope protein is found around the nucleus and at the cell surface in infected cells. The E protein is important in virion assembly and morphogenesis. It has also been shown to have ion channel activity with immunopathological consequences (Verdiá-Báguenaa, Nieto-Torresb et al. 2012)

1.1.1.5 N (nucleocapsid) protein (42-50 kD) (Zuwała, Golda et al. 2015)

The main viral protein that this study is interest in is the N protein. The CoV N protein is usually phosphorylated. However, using the prokaryote *Escherichia coli* (*E. coli*) as an expression system results in limited eukaryotic post-translational modifications; meaning that phosphorylation will not occur (Khow and Suntrarachun 2012). The size of the N protein varies depending on virus and phosphorylation, see Table 2. The difference in expected molecular weight and observed molecular weight may be account for by the possible addition of a 6x polyhistidine tag. The 6x polyhistidine tag is often used to purify the recombinant proteins and has a molecular weight of 0.8kDa (Terpe 2003). The N protein is one of the most abundant viral proteins produced in an infected cell. It is a multifunctional protein with roles in replication, transcription and translation (Berry, Manasse et al. 2012). It has been shown to act as a viral suppressor of RNA silencing in mammalian cells (Cui, Wang et al. 2015), and can also induce apoptosis and reorganise actin in infected cells (Surjit, Liu et al. 2004).

As is typical of positive strand RNA viruses, coronaviruses do not incorporate the RNA polymerase into the virus particle; rather the polymerase is expressed after infection by using the positive sense genomic RNA as an mRNA (Hunt 2010).

Name	National Centre for Biotechnology Information (NCBI) GenBank reference	FASTA format sequence	Predicted molecular weight (KDa)	Recombinant protein molecular weight mentioned in literature (KDa)
Nucleocapsid protein [SARS coronavirus]	NP_828858.1	MSDNGPQSNQRSAPRITFGGPTDSTDNNQNGGR NGARPKQRRPQGLPNNTASWFTALTQHGKEELRFP RGQGVPINTNSGPDDQIGYYRRATRRVRGGDGKM KELSPRWYFYYLGTGPEASLPYGANKEGIVWVATEG ALNTPKDHIGTRNPNNNAATVLQLPQGTTLPKGFY AEGSRGGSQASSRSSSRSRGNSRNSTPGSSRGNSPA RMASGGGETALALLLLDRLNQLESKVSGKGQQQQG QTVTKKSAAEASKKPRQKRTATKQYNVTQAFGRRG PEQTQGNFGDQDLIRQGTDYKHWPQIAQFAPSAS AFFGMSRIGMEVTPSGTWLTYHGAIKLDDKDPQFK DNVILLNKHIDAYKTFPPTEPKKDKKKKTDEAQPLPQ RQKKQPTVTLLPAADMDDFSRQLQNSMSGASADS TQA	46	46 (Che, Qiu et al. 2005)
Nucleocapsid protein [SARS coronavirus 2]	YP_009724397	MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSG ARSKQRRPQGLPNNTASWFTALTQHGKEDLKFPRG QGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDL SPRWYFYYLGTGPEAGLPYGANKDGIIWVATEGAL NTPKDHIGTRNPANNAAIVLQLPQGTTLPKGFYAEG SRGGSQASSRSSSRSRNSSRNSTPGSSRGTSPARM AGNGGDAALALLLLDRLNQLESKMSGKGQQQQGQ TVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGPE QTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFF GMSRIGMEVTPSGTWLTYTGAIKLDDKDPNFKDQV ILLNKHIDAYKTFPPTEPKKDKKKADETQALPQRQK KQQTVTLLPAADLDDFSKQLQQSMSSADSTQA	46	47 (Rikhtegaran Tehrani, Saadat et al. 2020)
Nucleocapsid protein [Middle East respiratory syndrome coronavirus]	AHX71953.1	MASPAAPRAVSFADNNDITNTNLSRGRGRNPKPRA APNNTVSWYTGLTQHGKVPLTFPPGQGVPLNANST PAQNAGYWRRQDRKINTGNGIKQLAPRWYFYYTG TGPEAALPFRAVKDGIVWVHEHGATDAPSTFGTRN PNNDSAIVTQFAPGTKLPKNFHIEGTGGNSQSSSRA SSVSRNSSRSSSQGSRSGNSTRGTSPGPSGIGAVGG DLLYLDLLNRLQALESGKVKQSQPKVITKKDAAAAK NKMRHKRTSTKSFNMVQAFGLRGPGDLQGNFGDL QLNKLGTEDPRWPQIAELAPTASAFMGMSQFKLTH QNNDDHGNPVYFLRYSGAIKLDPKNPNYNKWLELL EQNIDAYKTFPKKEKKQKAPKEESTDQMSEPPKEQR VQGSITQRTRTRPSVQPGPMIDVNTD	45	44.5 (Song, Ha et al. 2015)
Nucleocapsid protein [Human coronavirus NL63-CoV]	AFV53152.1	MASVNWADDRAARKKFPPPSFYMPLLVSSDKAPYR VIPRNLVPIGKGNKDEQIGYWNVQERWRMRRGQR VDLPPKVHFYYLGTGPHKDLKFRQRSDGVVWVAKE GAKTVNTSLGNRKRNQKPLEPKFSIALPPELSVVEFE DRSNNSSRASSRSSTRNNSRDSSRSTSRQQSRTRSD SNQSSSDLVAAVTLALKNLGFDNQSKSPSSSGTSTPK KPNKPLSQPRADKPSQLKKPRWKRVPTREENVIQCF GPRDFNHNMGDSDLVQNGVDAKGFPQLAELIPNQ AALFFDSEVSTDEVGDNVQITYTYKMLVAKDNKNLP KFIEQISAFTKPSSIKEMQSQSSHVAQNTVLNASIPES KPLADDDSAIIEIVNEVLH	42	42.6 (Berry, Manasse et al. 2012)

Nucleocapsid protein [Human coronavirus OC43-CoV]	AAT84366.1	MSFTPGKQSSSRASSGNRSGNGILKWADQSDQFR NVQTRGRRAQPKQTATSQQPSGGNVVPYYSWFSG ITQFQKGKEFEFAEGQGVPIAPGVPATEAKGYWYR HNRRSFKTADGNQRQLLPRWYFYYLGTGPHAKDQ YGTDIDGVYWVASNQADVNTPADIVDRDPSSDEAI PTRFPPGTVLPQGYYIEGSGRSAPNSRSTSRTSSRAS SAGSRSRANSGNRTPTSGVTPDMADQIASLVLAKL GKDATKPQQVTKHTAKEVRQKILNKPRQKRSPNKQ CTVQQCFGKRGPNQNFGGGEMLKLGTSDPQFPILA ELAPTAGAFFFGSRLELAKVQNLSGNPDEPQKDVYE LRYNGAIRFDSTLSGFETIMKVLNENLNAYQQQDG MMNMSPKPQRQRGHKNGQGENDNISVAVPKSRV QQNKSRELTAEDISLLKKMDEPYTEDTSEI	49	50 (Che, Qiu et al. 2005) (Huang, Hsu et al. 2009)
Nucleocapsid protein [Human coronavirus 229E-CoV]	AGW80953.1	MATVKWADASEPQRGRQGRIPYSLYSPLLVDSEQP WKVIPRNLVPINKKDKNKLIGYWNVQKRFRTRKGK RVDLSPKLHFYYLGTGPHKDAKFRERVEGVVWVAV DGAKTEPTGYGVRRKNSEPEIPHFNQKLPNGVTVVE EPDSRAPSRSQSRSQSRGRGESKPQSRNPSSDRNH NSQDDIMKAVAAALKSLGFDKPQEKDKKSAKTGTP KPSRNQSPASSQTSAKSLARSQSSETKEQKHEMQKP RWKRQPNDDVTSNVTQCFGPRDLDHNFGSAGVV ANGVKAKGYPQFAELVPSTAAMLFDSHIVSKESGNT VVLTFTTRVTVSKDHPHLGKFLEELNAFTREMQQHP LLNPSALEFNPSQTSPATAEPVRDEVSIETDIIDEVN	43	44 (Che, Qiu et al. 2005)
Nucleocapsid protein [Infectious Bronchitis Virus]	AAB24054.1	MASGKATGKTDAPAPVIKLGGPRPPKVGSSGNAS WFQAIKAKKLNSPQPKFEGSGVPDNENFKTSQQHG YWRRQARFKPGKGRRKPVPDAWYFYYTGTGPAAD LNWGDSQDGIVWVAAKGADVKSRSNQGTRDPDK FDQYPLRFSDGGPDGNFRWDFIPLNRGRSGRSTAA SSAASSRPPSREGSRGRRSGSEDDLIARAAKIIQDQQ KKGSRITKAKADEMAHRRYCKRTIPPGYKVDQVFGP RTKGKEGNFGDDKMNEEGIKDGRVTAMLNLVPSS HACLFGSRVTPKLQPDGLHLKFEFTTVVPRDDPQFD NYVKICDQCVDGVGTRPKDDEPKPKSRSSRPATRTS SPAPRQQRLKKEKRPKKQDDEVDKALTSDEERNNA QLEFDDEPKVINWGDSALGENEL	45	45 (Fan, Ooi et al. 2005)

Table 2: The FASTA format sequence of the coronaviruses used in this study alongside their predicted and referenced molecular weights. Predicted molecular weights generated using "Protein molecular weight calculator" provided by Science Gateway (https://www.sciencegateway.org/tools/proteinmw.htm).

1.1.2 Genome

Coronavirus genes are arranged in the order 5'-rep1a-rep1b-S-E-M-N-3'. Some also contain the additional haemagglutinin esterase gene, as shown in KHU1-CoV and OC43-CoV in figure 2 (Kazi 2005). The replicase gene (rep) accounts for approximately two thirds of the genome and includes the two open reading frames (ORFs) ORF1a and ORF1b. Orf1a encodes a polyprotein of ~450-500kD and Orf1b a polyprotein of ~750-800kD. Translation of the entire rep region occurs using a -1 ribosomal frameshift signal located at the ORF1a/ORF1b junction. Orf1a encodes a 3C-like protease (3CIpro) similar to picornavirus 3C proteinases (Hegyi and Ziebuhr 2002). The CoV replicase also includes a series of domains with enzymatic activities including a large RNA-dependant RNA polymerase (RdRp) (Gorbalenya, Enjuanes et al. 2006).

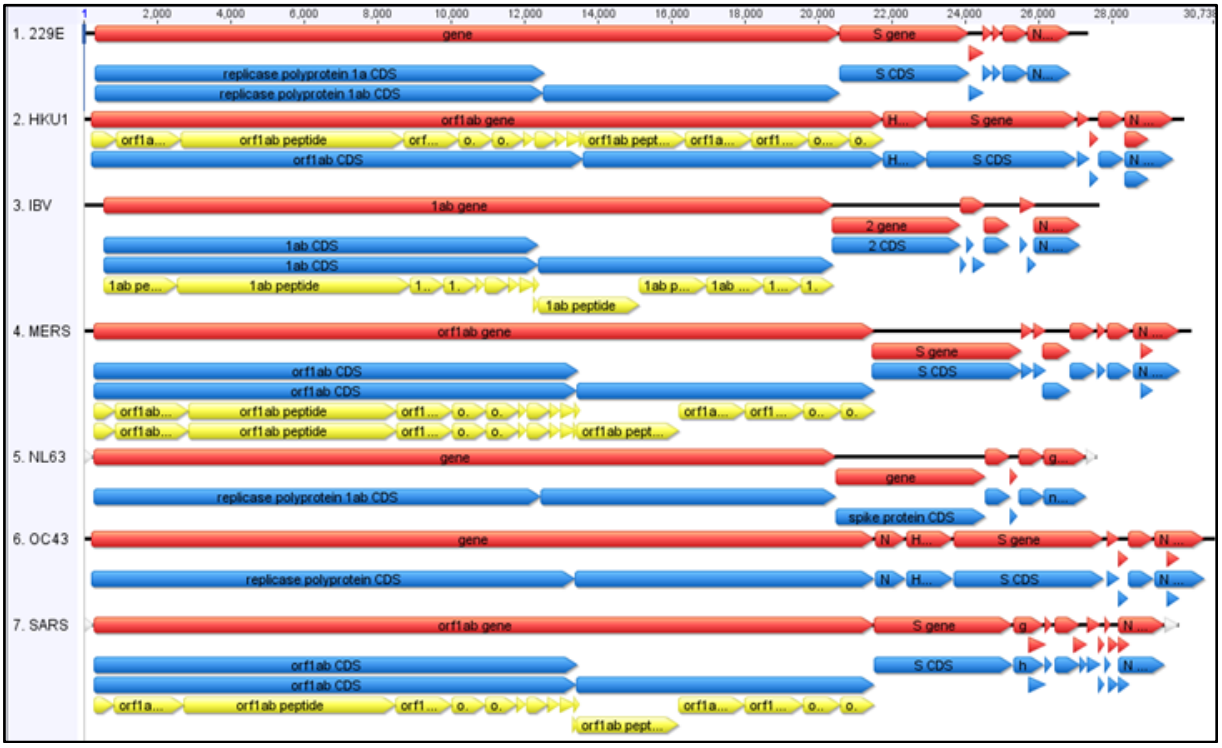


Figure 2: Schematic of the coding potential of six human coronavirus (HCoV) genomes and an avian coronavirus IBV (which will later be used as an experimental control). Colour coding indicates genes in red, coding sequences (CDS) in blue, mature peptides/open reading frames (ORFs) in yellow. Image created using Geneious version 8.0.5 (http://www.geneious.com (Kearse, Moir et al. 2012)). All sequences used were sourced from the National Centre for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov) (229E-CoV = NC_002645.1, KHU1-CoV = NC_006577.2, IBV= NC_001451.1, MERS-CoV= NC_019843.3, NL63-CoV = NC_005831.2, OC43-CoV = NC_005147.1 and SARS-CoV= NC_004718.3)

1.1.3 Replication

Once the viral particle enters the cell it is uncoated, and the genome is deposited in the cytoplasm. The Coronavirus genome has a 5' methylated cap and a 3' polyadenylated tail (Bouvet, Debarnot et al. 2010). This allows the RNA to attach to ribosomes for translation. As with all positive-stranded RNA viruses, replication of CoV genomes is mediated through the continuous synthesis of full-length negative stranded-RNA. This becomes the template for the synthesis of progeny virus genomes. The RNA-dependent RNA polymerase (RdRp) template switch allows for the production of 3′-coterminal nested set of subgenomic mRNAs (sgmRNAs) which are translated into virion structural proteins (SP) or to non-structural accessory proteins (NSP). There are up to 16 NPS which aid the replication process alongside the recruitment of host proteins (Wu and Brian 2010). Replication occurs within double-membrane vesicles (DMVs) (Angelini, Akhlaghpour et al. 2013). The viral N protein has also been shown to co-localize with the replication complex in DMVs and is thought to play a role in transcription. Studies have shown that only background levels of CoV RNA synthesis occurs in the absence of N protein (Almazan 2004). N proteins are important for viral replication, in the case of mouse hepatitis virus (MHV), a commonly used animal model for human coronaviruses, it has been shown the that N protein interacts with nsp3 a component of the replication/transcription complexes (RTCs) via two distinct regions, failure to do so results in impaired stimulation of genomic RNA and viral mRNA leading to decreased MHV replication and progeny production (Cong, Ulasli et al. 2020)

1.1.4 Mutation

RNA polymerase does not have the same proof-reading capability as DNA polymerase and, as such, it is prone to a high error rate during RNA replication. The large size of the coronavirus genome means that there can be several mutations in each progeny virus; these can be deletions, additions or substitutions. Recombination events also occur, which is unusual in non-segmented RNA viruses, but may be as a result of the discontinuous mode of RNA replication that occurs. Both these facts allow for the rapid evolution of the virus and can result in the introduction of new

strains. Although the mutation rates of coronaviruses are relatively high it is worth noting that CoVs express a 3′-to-5′ exoribonuclease in non-structural protein 14 (nsp14-ExoN) which has been shown to have an RNA proofreading function, one study showed that ExoN negative strains of SARS-CoV had a 16-fold increase in mutation rates, so some proofing is in place (Smith, Blanc et al. 2013).

1.1.5 Pathology

CoVs primarily cause upper respiratory tract infections in humans and fowls, and enteric infections in porcine and bovine species. HCoV infections have also been known to cause respiratory, enteric, hepatic, renal and even central nervous system problems (Rousset, Moscovici et al. 1984) (L Enjuanes 2008). Alongside Rhinoviruses, Coronaviruses are believed to cause a significant percentage of all common colds in human adults; with some estimates as high as 30% (Mesel-Lemoine 2012). The severity and location of a CoV infection varies depending on strain. 229E-CoV and NL63-CoV target epithelial cells and both can lead to respiratory infections and bronchiolitis. OC43-CoV is known to target the lower respiratory tract and can result in pneumonia alongside the common cold.

Animal Alphacoronaviruses include transmissible gastroenteritis coronavirus (TGEV), an enteropathogenic CoV that infects lung cells and the villus epithelial cells of the small intestine in pigs. Other Alphacoronaviruses include porcine respiratory CoV (PRCV) (Cox, Hooyberghs et al. 1990), porcine epidemic diarrhoea virus (PEDV) (Utiger 1995), canine coronavirus (CCoV) (Erles, Toomey et al. 2003) and feline coronavirus (FCoV) (Holzworth 1963).

Betacoronaviruses include murine CoVs. MHV has a high mortality rate and causes a progressive demyelinating encephalitis in mice. As such, it is often used as a murine model for multiple sclerosis. Commonly used laboratory strains infect the liver and brain and so are also used as animal models for hepatitis and encephalitis. KHU1-CoV is another Betacoronavirus which infects human alveolar cells. Severe acute respiratory syndrome corona virus (SARS-CoV) has been classified as a Betacoronavirus sub-group B and causes severe acute respiratory disease,

pneumonia, diarrhoea and can prove to be fatal. The Middle East respiratory syndrome coronavirus (MERS-CoV) is another Betacoronavirus (sub-group c) which leads to acute respiratory disease, kidney failure, multiple organ dysfunction and can cause fatalities.

Gammacoronaviruses include infectious bronchitis virus (IBV) which infects ciliated epithelia of the nose and trachea in chickens. It also targets their urogenital tract.

1.1.6 Immune response

The immune response is a complex, vast and multifarious system which is hard to view in its entirety. CoV infections have several known affects upon the host including up regulation of genes involved in inflammation, coagulation and the stress response. Inflammation response includes the up-regulation of interleukin 8 (IL-8). IL-8 induces chemotaxis in target cells, namely neutrophils, directing them towards the site of infection. IL-8 is also known to promote angiogenesis, the formation of new blood vessels from pre-existing vessels. Studies have shown a positive a correlation between the IL-8 levels and disease severity (Yoshikawa, Hill et al. 2009). Interferon gamma (IFN-γ) is also released by the host in response to CoV infections. IFN- γ promotes the differentiation of Th1 cells into cytotoxic CD8+ T-cells which can destroy infected cells. When a cell becomes infected several cellular transcription factors become activated, such as interferon regulatory transcription factor 3 (IRF-3) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB). These, in turn, activate the production of type I IFNs which bind to the IFN-α/β receptor (IFNAR) and initiate the Janus kinase and signal transducer and activator of transcription (JAK/STAT) pathway, causing activated transcription factors to translocate to the nucleus. The transcription factors activate genes containing interferon-stimulated response element (ISRE) in their promoters. These genes, once activated, enable the cell to enter a state that resists viral infection and can prevent virus replication (Garcia-Sastre and Biron 2006). Although the immune response is activated with the intention of being beneficial towards the host, many of these responses are general in their effects once initiated and collateral damage on non-infected cells can prove detrimental to the host.

More recently a number of studies have focused on antibody response to SARS-CoV-2 in order to try to ascertain whether the type and levels of patient antibodies can predict susceptibility to secondary infection and/or indicate immunity. Papers discussing the antibody response to SARS-CoV-2 published between 01/01/2020 and 26/06/2020 were comprehensively evaluated in a systematic review (Post, Eddy et al. 2020). The review included 150 papers. The studies were said to be of variable quality and had heterogeneity of methods, participants, outcome measured and assays used and lacked the characterisation of longer-termed patterns. As such, large, cross-national cohort studies using appropriate statistical analysis and standardised serological assays and clinical classifications remain needed. That being said, the review found that IgM was consistently detected before IgG in included studies, with peak titres detected at weeks two to five and declining over a further three to five weeks post-symptom onset, varying on the patient group. IgG titres peaked around weeks three to seven post-symptom onset then plateaued, generally continuing for at least eight weeks. Neutralising antibodies were detectable within seven to fifteen days following disease onset, increasing until days 14–22 before levelling and then decreasing. Titres of neutralising antibodies were lower in asymptomatic or clinically mild disease cases. Convalescent plasma has been used to source specific and potent neutralising antibodies; cross-reactivity but limited cross-neutralisation with other human coronaviridae was reported. The potential therapeutic use for such antibodies is further discussed in section 1.2.7.2.

1.2 Human coronaviruses

1.2.1 Brief overview

For many years, only two human coronaviruses were known; 229E-CoV and OC43-CoV. The discovery of SARS-CoV in 2003 added a third and by the end of 2004 three independent research laboratories reported the discovery of a fourth. It was named NL63-CoV (van der Hoek, Pyrc et al. 2004). In 2005, a fifth HCoV was discovered in Hong Kong, called KHU1-CoV (Vabret, Dina et al. 2006). In September 2012, a sixth human coronavirus was discovered and initially referred to as Novel Coronavirus 2012. It would later be called Middle East Respiratory Syndrome Coronavirus (MERS-CoV) (Zaki, van Boheemen et al. 2012). More recently, SARS-CoV-2 (commonly referred to as Covid-19) emerged in December 2019 resulting in a world-wide pandemic (Velavan and Meyer 2020).

1.2.2 Evolution

Bat coronaviruses are thought to be the gene source of Alphacoronavirus and Betacoronavirus, and avian coronaviruses as the gene source of Gammacoronavirus and Deltacoronavirus (Woo, Lau et al. 2012). The most recent common ancestor of Alphacoronavirus has been placed as ~2,400 BC, Betacoronavirus ~3,300 BC, Gammacoronavirus ~2,800 BC and Deltacoronavirus ~3,000 BC (Woo, Lau et al. 2012).

The virus continues to evolve. More recently identified strains of non-human coronavirus include a porcine enteric alphacoronavirus, swine acute diarrhoea syndrome coronavirus (SADS-CoV), similar to bat HKU2, responsible for the death of 25,000 piglets in China in 2016-17 (Gong, Li et al. 2017, Zhao, He et al. 2018). And following a coronavirus (FRCoV) being identified in ferrets (*Mustela putorius furo*) in 2006 (Wise, Kiupel et al. 2006), a 2017 study has identified ferret enteric coronavirus (FRECV) as a new species in the genus Alphacoronavirus (Li, Khalafalla et al. 2017).

1.2.3 Severe Acute Respiratory Syndrome (SARS)

In 2003, following the outbreak of SARS-CoV which had begun the previous year in Asia, the World Health Organization issued a press release stating that the causative agent had been identified as a novel coronavirus by a number of laboratories (WHO 2003). The epidemic resulted in over 8,000 infections, about 10% of which resulted in death (Li, Li et al. 2005). Angiotensin-converting enzyme 2 (ACE 2) was shown to be the cellular receptor for SARS-CoV (Li, Li et al. 2005). Once infected, a patient's first symptom was that of a fever; with temperatures being equal to, or exceeding, 38^{o}C (Xu 2003). Days 2-7 included flu-like symptoms such as chills, headache, muscle aches, sore throat, malaise, mucus etc. Various cohort studies have been conducted looking at the frequency of each symptom and the results of four of these are detailed in Appendix one. This developed into an upper respiratory tract infection resulting in a dry cough and shortness of breath. In some cases, as the infection progressed, pneumonia could be confirmed using a chest X-ray. The infection, if it develops further, results in respiratory failure and acute respiratory distress syndrome (ARDS) which can be fatal. The incubation of SARS-CoV is ~ 2–10 days. The body produces immunoglobulin M (IgM) as an antibody response during the acute phase of infection (~3 -12 weeks), IgG is produced later and peaks at week 12 (Xu 2003). SARS-CoV is thought to have originated in bats and spread through an intermediate animal known as a masked palm civet (*Paguma larvata)* (Guan, Zheng et al. 2003, Lau, Woo et al. 2005, Coleman and Frieman 2014).

1.2.4 Middle East Respiratory Syndrome (MERS)

The Middle East Respiratory Syndrome coronavirus (MERS-CoV) is a Betacoronavirus which was first reported on 24th September 2012 on Program for Monitoring Emerging Diseases (ProMED-mail) by Egyptian virologist Dr. Ali Mohamed Zaki in Jeddah (Zaki, van Boheemen et al. 2012). The largest outbreak of the disease occurring outside of the Middle East arose on the 20th May 2015 when a 68-year-old man was diagnosed with the disease in Korea (Lee 2015). The patient in question had recently returned from visiting Bahrain, Saudi Arabia and Qatar and it is thought that he picked up the virus whilst in the Middle East (Lee 2015). As shown in the epicurve in Figure 3, there was a second large outbreak of the virus in Korea. There have been 185 laboratory confirmed cases of MERS-CoV in Korea resulting in 36 deaths, there was also one case reported in China when a patient travelled there from Korea (Park, Perera et al. 2015)(WHO 2017). As of March 27th 2020 WHO has been notified of 2494 laboratory-confirmed cases of infection with MERS-CoV, including at least 858 related deaths, giving MERS a 34.4% mortality rate (WHO 2020). Fortunately sustained human-to-human transmission has not been well documented (Timothy M. Uyeki 2016). A total of 27 countries have reported cases of MERS-CoV. Bahrain, Iran, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, United Arab Emirates (UAE), and Yemen have all reported cases within the Arabian Peninsula. Travel-associated cases seen outside of this area have occurred in; Algeria, Austria, China, Egypt, France, Germany, Greece, Italy, Malaysia, Netherlands, Philippines, Republic of Korea, Thailand, Tunisia, Turkey, the United Kingdom (UK), and the United States of America (USA) as shown in a map produced in 2017 in Figure 3, although new cases have been reported since the image was produced these have all occurred in countries who have declared infections previously.

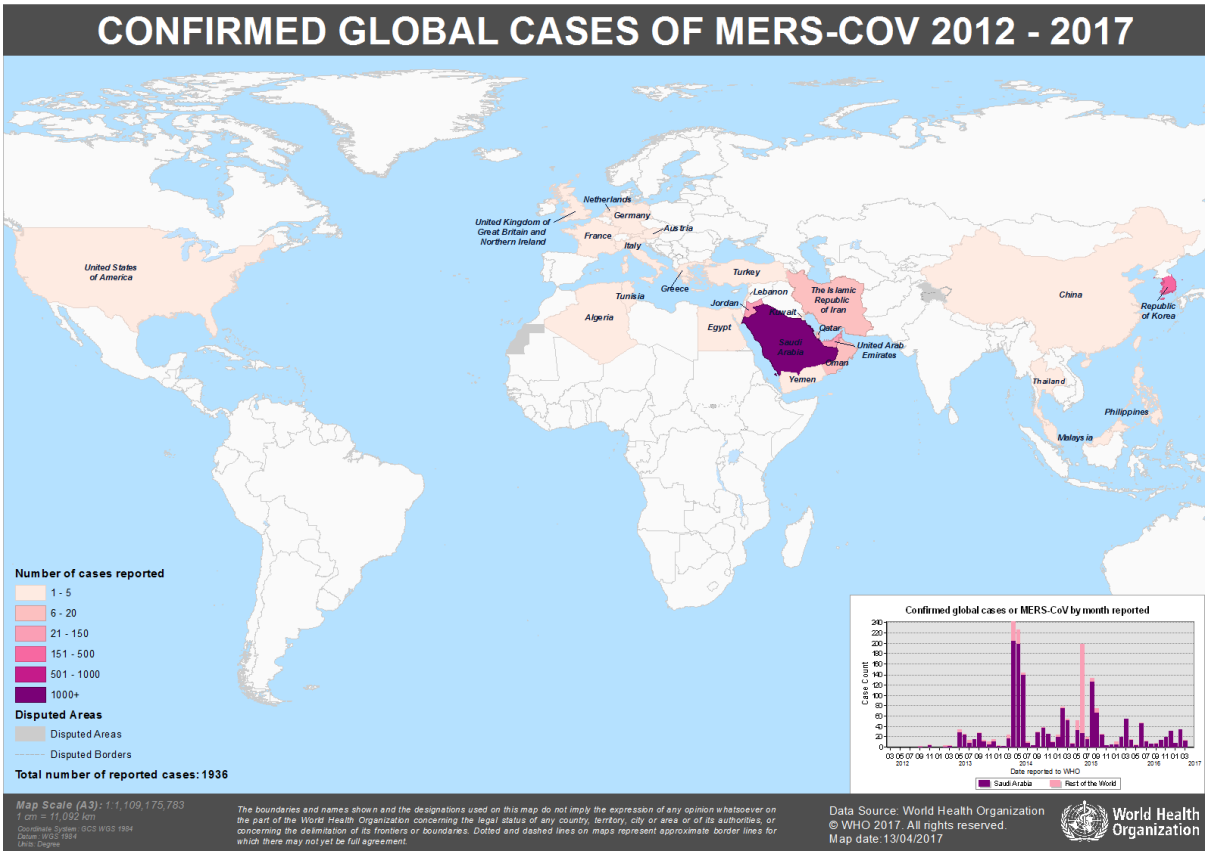


Figure 3: Middle East respiratory syndrome coronavirus map and epicurve: A global map showing the locations of the 1936 reported MERS-CoV cases up to 2017, colour coded by number of reported cases. Map produced 13/04/2017. (WHO 2017)

1.2.4.1 Signs and symptoms

In MERS-CoV infections, sufferers present with serious respiratory illness (fever, cough, shortness of breath and breathing difficulties). Pneumonia is common. Some patients also develop gastrointestinal symptoms, including diarrhoea and kidney failure. In people with immune deficiencies the disease may have an atypical presentation. Unlike most respiratory viruses MERS-CoV has a strong tropism for non-ciliated bronchial epithelial cells (Kindler, Jonsdottir et al. 2013). The virus has been shown to evade innate immune responses and to antagonise IFN production in infected cells (Raj, Mou et al. 2013). It was initially thought that MERS-CoV would use the same receptor as SARS-CoV (ACE2) (Li, Moore et al. 2003). However, studies showed that this was not the case and instead dipeptyl peptidase 4 (DPP4) acted as the cellular receptor (Müller 2012).

1.2.4.2 Origin

It is not certain whether the emergence of MERS-CoV occurred as a result of a single zoonotic event with subsequent human-to-human transmission, or if the fact that there are multiple geographic sites of infection is a result of multiple zoonotic events from a common unknown source. A study by led by Professor Ziad Memish estimated perhaps as many as 7 separate zoonotic transmissions may have occurred (Cotten, Watson et al. 2013).

1.2.4.3 Natural reservoir

1.2.4.3.1 Bats

The original source of the virus is thought to be bats. Bats have a widespread geographical distribution, are capable of flight and are the second largest group of mammals, behind rodents. These factors facilitate their role as natural viral reservoirs. Analysis has shown MERS-CoV to have sequence homology with both bat HKU4-CoV and HKU5-CoV, carried by the genera *Tylonycteris* and *Pipistrellus* respectively (Memish, Mishra et al. 2013). However, the study was not without problems as the collection of samples and cold chain transport was compromised. In the study 1003 bats were sampled from the Middle East, only one of which proved positive for MERS-CoV. This indicated that an intermediate vector was likely to be required to facilitate the

spread of the virus. This is similar to the progression of the SARS-CoV whereby *Rhinolophid* bats were identified at the viral reservoir and civets the intermediate vector (Lau, Woo et al. 2005). 229E-CoV is also believed to have a bat origin in *Hipposideros caffer* rubber bats (Pfefferle, Oppong et al. 2009, Drexler, Corman et al. 2014) as does NL63-CoV in *Perimyotis subflavus* bats (Huynh, Li et al. 2012).

1.2.4.3.2 Camels

Further investigative studies suggested that dromedary camels were likely to be the intermediate vector (Alagaili, Briese et al. 2014). In December 2013, a study conducted in the Middle East showed no evidence of previous MERS-CoV infection in sheep, goat, cattle or chickens; however, 90% of camels (280 out of 310 sampled) tested positive for previous exposure (Hemida, Perera et al. 2013). Between 2016 and 2018, one study collected a total of 4027 nasal swabs and 3267 serum samples from dromedary camels in Egypt, Senegal, Tunisia, Uganda, Jordan, Saudi Arabia, and Iraq (Kandeil, Gomaa et al. 2019). The study used RT-PCR and detected MERS-CoV RNA in nasal swab samples from Egypt, Senegal, Tunisia, and Saudi Arabia. Samples from all counties showed antibodies in microneutralization assays. Positive PCR samples were partially sequenced, and a phylogenetic tree was constructed. The study's results indicate the widespread distribution of the virus in camels, however the study highlighted the need for a systematic active surveillance and longitudinal studies for MERS-CoV to better understand the epidemiology of the disease and dynamics of viral infection (Kandeil, Gomaa et al. 2019). Ancestor analysis suggests that the jump of MERS-CoV from bats to camels may have occurred approximately 20 years ago in Africa, with the camels then being imported into to Arabian Peninsula (Corman, Ithete et al. 2014). Although camels do not often present with clinical signs of infection (Chu, Poon et al. 2014) more recent serological studies further backdated the presence of MERS-CoV antibodies in camels, suggesting that MERS-CoV has been present in dromedary camels for at least three decades in the Middle East and North and East Africa (de Wit, van Doremalen et al. 2016).

1.2.4.3.3 Other animals potentially susceptible

In vitro testing has been conducted in several studies in order to assess permissiveness of selected cell culture models, with the aim of identifying other potential other intermediate host species of MERS-CoV. One study looked at livestock, primate, rodent, insectivore and bat cell line susceptibility and saw efficient viral replication (>9.3 log10) virus RNA genome (equivalents/mL of cell culture supernatant) in goat kidney and lung cells, alpaca kidney cells and in dromedary umbilical cord cells (Eckerle, Corman et al. 2014). However, the same study found that the virus did not replicate in sheep, cow, rodent, or insect cells; even though all cell lines tested expressed the receptor DPP4 (Eckerle, Corman et al. 2014). Another animal study showed rhesus macaques (*Macaca mulatta)* were susceptible to infection but that hamsters, ferrets and mice were not (van Doremalen, Miazgowicz et al. 2014). The DPP4 recognizing RBD is found in the S1 N-terminal of the S protein and consists of ~240 residues spanning aa 367-606 (Hu, Ge et al. 2015). The study predicted that camels, goats, sheep, and cows could all potentially be used as intermediate host reservoir species for MERS-CoV (van Doremalen 2014). Horse DPP4 receptors have also shown to be permissive to infection (Barlan, Zhao et al. 2014). To summarise, published results have shown that MERS-CoV is able to infect and replicate in cell lines derived from humans, non-human primates (NHP), bat, swine, goat, horse, rabbit, civet, and camel cell lines, but not from mice, hamster, dog, ferret, and cat cell lines (Chan, Chan et al. 2013, Eckerle, Corman et al. 2014, van Doremalen, Miazgowicz et al. 2014, Haagmans, van den Brand et al. 2015). This count have potential agricultural implications and may affect disease outbreaks and spread in the future.

1.2.5 Diagnosis

Diagnostic testing of suspected cases during dangerous emerging coronavirus outbreaks can prove beneficial in containing the outbreak and can also provide information on the disease source and progression. Methods of detecting HCoV infections include isolating the virus by culture, visualization of the virions using TEM and detection of current or prior infection using serological techniques. Tissue culture and cell line selection is difficult when culturing the virus,

especially in relation to a previously unknown strain. TEM can be beneficial in that it does not require any reagents specific to the virus; however, it is not specific and may not be able to identify a virus past the family level (Goldsmith and Miller 2009). Presently the most used diagnostic techniques are based around the polymerase chain reaction (PCR), whereby specific or broadly covering primers can be used to detect viral presence. Real-time reverse transcriptase PCR (RT-PCR) can be used for rapid identification of MERS-CoV from patient samples (Corman, Eckerle et al. 2012). The primers used are called 'UpE' (targets elements upstream of the E gene) and '1A' (targets the ORF1a gene). However, there are some technical difficulties in using PCR based methods, mainly that the tests are conducted on patient samples with potentially live virus. This clearly has health and safety implications for any health care professional collecting samples and any laboratory worker conducting the screening. Detection rates by early quantitative RT-PCR (qRT-PCR) assays are often low as they rely on samples being taken during viral shedding and correct sample handling. Specificity can also be dubious as there is a chance for cross-contamination in laboratories processing multiple samples (Yam, Chan et al. 2003) (Patrick 2006). On 17th July 2015 emergency use authorization was given to a commercially developed RealStar MERS-CoV RT-PCR Kit U.S. (Altona Diagnostics GmbH, Hamburg, Germany) (Timothy M. Uyeki 2016) for the detection of MERS-CoV infection. Using immunofluorescence assays (IFA) has also been investigated, however antibodies against Betacoronaviruses have been found to cross-react, which limits the use of the procedure in diagnostics (Corman, Muller et al. 2012). There remains a lack of a readily available, rapid, simple and accurate test for the detection of MERS-CoV. This may be in part due to the limited availability of clinical specimens and MERS-CoV isolates from infected patients. Paired acute and convalescent serum samples would be needed in order to develop serological tests to confirm MERS-CoV infection in cases when viral shedding is not detectable and for population surveillance of both exposure and immunity to MERS-CoV infection.

1.2.5.1 Serological assays:

Diagnostic screening can also be conducted with the use of serological tests; using either virus-derived antigen or recombinant structural coronavirus proteins. These assays can be beneficial in that antibodies remain present and detectable over a long period; allowing transmission patterns to be assessed, the origin to be traced and allowing for the detection of asymptomatic infections. However serological testing can still be problematic. Validation of results relies on numerous previously highly characterised positive and negative sera samples (Meyer, Drosten et al. 2014). Gaining samples from patients can prove difficult for both logistical and ethical reasons. There are two main points to consider when serological testing takes place:

- Assay sensitivity: this is the number of positive samples that can be correctly identified
- Assay specificity: this is the number of negative samples that can be correctly identified. This is an important factor to consider in order to avoid false positive diagnoses.

High seroprevalence of antibodies against coronaviruses responsible for the common cold, as well as cross-reactive against conserved regions of immunogenic coronavirus proteins, can result in false positive results. WHO advise the use of a viral neutralization test (VNT) to exclude any serological cross-reactions with other human or animal coronaviruses (WHO 2004).

1.2.5.2 Virus based serological assays:

These include immunofluorescence assays (IFAs), enzyme-linked immunosorbent assays (ELISAs) and Western-blot (WB) assays. The main disadvantage for virus-based serological assays is the necessity for the laboratories to be biosafety level three (BSL 3). IFAs rely on subjective interpretation of fluorescence staining patterns and are therefore hard to standardise. The process is also not appropriate for high-throughput screening. Also, ELISAs can only be effective if there is access to a well characterised serum collection to be able to determine the assay-specific cut-off value.

1.2.5.3 Recombinant protein-based assays:

The main advantage of using recombinant proteins as the base of an assay, as opposed to live virus, is the fact they negate the need for a BSL 3 laboratory. The general premise for these types of assays is that a single immunogenic viral gene is cloned into either a eukaryotic or prokaryotic plasmid to allow it to be expressed in either mammalian, insect, or bacterial cell culture. One downside of recombinant protein-based assays is the technical skill and time required for the various processes including; cloning, transfection, optimisation of expression and protein purification. There are two immunodominant proteins which are potential targets for screening in relation to coronavirus infection; the N and the S proteins.

- N protein: is the smaller of the two and its lack of glycosylation sites make it relatively easy to express and purify in large quantity from bacteria (Qiu, Shi et al. 2005). Although N protein has been noted as having high sensitivity it has also been reported as having cross-reactivity against other HCoVs when used in protein-based assays (Qiu, Shi et al. 2005) (Maache, Komurian-Pradel et al. 2006) (Che, Qiu et al. 2005).

- S protein: The majority of neutralizing antibodies are directed against the S protein (Buchholz, Bukreyev et al. 2004). The S protein can also be more useful when screening convalescent sera, as antibodies directed against it occur later in infection than those against the N protein (Tan, Goh et al. 2004, Woo 2005). However, there are technical difficulties in using bacteria to express the full-length protein and so often only fragments are used to conduct ELISAs and WB assays with, which may impair sensitivity and specificity (Meyer, Drosten et al. 2014).

1.2.5.4 Cross-reactivity and cross-neutralization amongst coronaviruses

When a coronavirus has been established in the population cross-reaction of antibodies can occur between immunogenic proteins of viruses that are closely related. Any common epitopes can go on to elicit, or re-activate, cross-neutralizing antibodies (Bradburne 1970). This function of the immune system is beneficial for the host; however, it can prove problematic in serological surveillance because it can lead to false positives. It is for this reason that a confirmation assay

such as a viral neutralization test (VNT) is sometimes recommended. However, this can prove labour-intensive and is not ideal in an outbreak situation when rapid detection plays a pivotal role in containing the disease.

1.2.6 Animal screening

If a viral outbreak has a zoonotic origin, the testing of livestock and wild animals can allow the animal reservoir and potential animal vector to be identified (Guan, Zheng et al. 2003, Hemida, Perera et al. 2013). This can allow human intervention to contain the virus by interrupting the chain of transmission. Studies can be impractical logistically and can also be complicated by the fact that animals may harbour unidentified coronaviruses that may cause a cross-reaction to any antigen being screened (Meyer, Drosten et al. 2014).

1.2.7 Therapeutic drugs

To date there are no commercially approved vaccines against, or therapeutic treatments for, MERS-CoV(Lu 2020). Although vaccines are undergoing clinical trials the drug development process generally takes over 10 years, making it impractical to develop novel anti-coronavirus drugs to treat an outbreak (Coleman, Sisk et al. 2016).

Alongside those mentioned in section 1.2.7.2, other therapeutics being research include substances derived from medicinal plants. One study investigated 15 such extracts for their anti-IBV properties prior and during infection and found the three most effective to be derived from *M. piperita*, *T. vulgaris* and *D. canadense* (Lelesius, Karpovaite et al. 2019)

1.2.7.1 Animal models

Before human clinical trials can take place for any potential countermeasure to MERS-CoV, proof of concept data must be obtained from *in vivo* studies of experimentally infected animals. However there is a lack of a small animal model that is naturally susceptible to MERS-CoV infection, as the virus does not infect mice, hamsters and ferrets (van Doremalen and Munster 2015). Transgenic mice expressing human DPP4 have been engineered to overcome this (Hao, Lv et al. 2019). Due

to the presence of receptors in their nasal epithelium, rabbits are also being investigated as potential hosts for MERS-CoV. A study showed high levels of viral RNA shedding from rabbits' noses following virus inoculation, however the rabbits did not develop clinical manifestations of the disease and produced only limited amounts of infectious virus from the nose. The was no transmission of the virus by contact or airborne routes observed in rabbits, limiting their use in MERS-CoV vaccine research (Widagdo, Okba et al. 2019). Due to the lack of availability of dromedaries, and the high cost of obtaining them in the United States, Alpacas are being investigated as a suitable proxy for camels (Crameri, Durr et al. 2016). The final problem with current animal models is the limited access to clinical samples and recent virus isolates to use upon them.

1.2.7.2 Potential therapeutic countermeasures

MERS-CoV S2 contains two heptad repeat regions needed for S protein-mediated membrane fusion, HR1 and HR2, it is therefore assumed that peptides that mimic these may prove beneficial in interfering with the viral-membrane fusion process (Wang, Hua et al. 2019). mABs targeting the RBD of the S protein have been shown to block viral attachment (Zhou, Yang et al. 2019). Heavy chain variable domain antibodies isolated from camelids, called nanobodies, have been used as therapeutics to block the RBD/receptor interaction (Zhao, He et al. 2018). A 2019 study showed that three peptides derived from the HR2 region in spike protein of BatCoV HKU4 exhibited potent inhibitory activity against MERS-CoV S-mediated cell-cell fusion and viral infection; making them another possible candidate for further development as antiviral agents against MERS-CoV infection (Xia, Lan et al. 2019).

Potential therapeutic drugs for MERS-CoV are shown in Table 3, these include repurposed food and drug agency (FDA) and European Medicines Agency (EMA) approved drugs that have nonspecific properties but that show efficacy against MERS-CoV *in vitro*. One study screened a library of over 1000 FDA-approved drugs in order to test for inhibition of coronavirus replication. Abelson (Abl) kinase inhibitors, including the anticancer drug imatinib, were shown to inhibit

both SARS-CoV and MERS-CoV *in vitro* by inhibiting Abelson tyrosine-protein kinase 2 (Abl2) and disrupting fusion of the virions at endosomal membranes (Coleman, Sisk et al. 2016). Another study screened FDA-approved drugs using a split-protein-based cell-cell fusion assay and found that Nafamostat, a serine protease inhibitor currently used to treat pancreatitis and disseminated intra-vascular coagulation (DIC), was found to be a potent inhibitor of S-mediated membrane fusion due to its ability to inhibit transmembrane protease serine 2 (TMPRSS2) (Yamamoto, Matsuyama et al. 2016). Investigating libraries of previously approved drugs could facilitate a rapid response to outbreaks and help curtail some of the delay in development and manufacturing time that a new product involves.

Source	Drug	Target	Anti–MERS-CoV activity	2016 status
NIAID Rocky Mountain Laboratories, Hamilton, MT, USA	Ribavirin + IFN	Polymerase + Immunomodulator	Active in cell culture and NHP	Approved for hepatitis C virus, compassionate use for MERS-CoV (Falzarano, de Wit et al. 2013) (Al-Tawfiq, Momattin et al. 2014, Omrani, Saad et al. 2014)
University of Hong Kong, Hong Kong	Interferon B1b	Immunomodulator	Active in cell culture	Preclinical development (Chan, Yao et al. 2015)
Hemispherix Biopharma, Philadelphia, PA, USA	Alferon N	Immunomodulator	Active in cell culture	Approved for human papillomavirus, orphan drug designation granted by the European Medicines Agency (Agency 2016)
Romark Laboratories, Tampa, FL, USA	Nitazoxanide	Host functions, glycosylation	Active in cell culture	Approved for cryptosporidia and giardia, in clinical trials for influenza virus (Stockman, Bellamy et al. 2006)
AbbVie, North Chicago, IL, USA	Lopinavir	Protease	Active in cell culture, NHP models	Approved for HIV (Chan, Yao et al. 2015)
BioCryst Pharmaceuticals, Durham, NC, USA	BCX4430	Polymerase	Active in cell culture and (Ad5)-DPP4 mouse	Clinical trial for Ebola virus (Warren, Wells et al. 2014)
Sarafianos Laboratory, Columbia, MO, USA	SSYA10–001	Helicase	Active in cell culture	Broadly active against coronaviruses (Adedeji, Singh et al. 2014, Wycoff, Maclean et al. 2015)
Planet Biotechnology, Hayward, CA, USA	Immunoadhesin (DPP4-Fc)	Spike/binding	Active in cell culture	Preclinical development (Wycoff, Maclean et al. 2015)
New York Blood Center, New York, NY, USA	HR2P-M2	Spike/fusion	Active in mouse models	Preclinical development (Channappanavar, Lu et al. 2015)
Loyola University, Chicago, Stritch School of Medicine, Maywood, IL, USA	Protease inhibitors	MERS-CoV PLpro, MERS-CoV 3CLpro5	Active in cell culture	Preclinical development (Kilianski, Mielech et al. 2013)
University of Maryland, College Park, MD. USA; Rega Institute, Katholieke Universiteit Leuven, Leiden, Belgium; NCATS; NIAID; University of Leiden, the Netherlands	FDA-approved drug screens	Multiple host targets	Active in cell culture; chloroquine and chlorpromazine are promising	Multiple screening efforts (de Wilde, Jochmans et al. 2014, Dyall, Coleman et al. 2014)

Table 3: MERS-CoV small molecule and biologics treatment candidates. NIAID=National Institute of Allergy and Infectious Diseases, IFN= interferon, NHP=nonhuman primate, DPP4= dipeptidyl peptidase-4, PLpro= papain-like protease, 3CLpro= 3C-like protease, NCATS= National Centre for Advancing Translational Sciences, NIAID= National Institute of Allergy and Infectious Disease and FDA= US Food and Drug Administration. Table sourced from CDC: Emerging Infectious Diseases (Timothy M. Uyeki 2016)

As well as using drugs as a treatment strategy, immunotherapeutics are also being studied. Immunotherapeutics work by altering the hosts' immune response either by stimulating, reducing or enhancing it. Immunotherapeutics often have less side effects than therapeutic drugs and have the advantage of being less likely to cause microbial resistance (Masihi 2001). Current immunotherapeutics under investigation for MERS-CoV consist of convalescent plasma as well as monoclonal or polyclonal antibody therapies, most of these have specific neutralizing activity against the spike protein, as seen in Table 4. Monoclonal antibodies can be isolated either from fully human convalescent blood or from transgenic animals which have the advantage of being able to be manufactured on a large scale with a high safety profile (Timothy M. Uyeki 2016). One study, shown in Figure 5, uses a transchromosomal (Tc) bovine production system to produce fully human polyclonal MERS-CoV immunoglobulin G (IgG) antibodies (Luke, Wu et al. 2016). In the study a γ-irradiated whole killed virus vaccine (Jordan strain) or a clade B spike protein nanoparticle vaccine (Al-Hasa strain) was used to infect two sets of Tc cows. The resulting Tc sera produced high ELISA and neutralising antibody titres *in vitro.* Two bovine human Tc IgG were purified; SAB-300 was produced after Jordan strain vaccination and SAB-301 from the Al-Hasa strain. SAB-301 was selected for *in vivo* and preclinical studies. SAB-301 was administered to transgenic Ad5-hDPP4 receptor-transduced mice, as a single dose either 12, 24 or 48 hours after MERS-CoV infection using the Erasmus Medical Centre 2012 strain. This rapidly resulted in viral lung titres near or below the limit of detection. The system could provide a means to rapidly produce a therapeutic to prevent and/or treat MERS-CoV infection and other emerging infectious diseases (Luke, Wu et al. 2016). Antibodies such as anti-C5aR1 antibody (Ab), can also be used to target the complement over-activation-induced immunopathogenesis resulting from a MERS-CoV infection (Jiang, Li et al. 2019). Fully human monoclonal antibodies usually have good safety profiles however problems may arise in the development of immunotherapeutics such as the risk of antibody-dependent enhancement of disease as well as the potential generation of escape mutant viruses that would be resistant to treatment.

Source	Drug	Target	Anti–MERS-CoV activity	Status
Multiple	IVIG	Spike, immune system	Unknown	Intravenous immunoglobulin is available and has been used for the treatment of >1 MERS-CoV patients with unknown clinical benefit (Stockman, Bellamy et al. 2006)
King Abdullah International Medical Research Centre, Riyadh, Saudi Arabia	Convalescent serum	Spike, immune system	Ad5-DPP4 mouse efficacy	A pilot clinical trial of convalescent plasma treatment of MERS-CoV patients is ongoing but not recruiting in Saudi Arabia (ClinicalTrials.gov. 2014)
Sanford Applied Biosciences, Sioux Falls, SD, USA	Transgenic bovine polyclonal	Spike	Ad5-DPP4 mouse and NHP studies	Preclinical development (Shultz 2014)
National Cancer Institute, NIH, Bethesda, MD, USA	M336, M337, M338	Spike	MERS-CoV neutralization	Preclinical development (Ying, Du et al. 2014)
Tsinghua University, Beijing, China	MERS-4, MERS-27	Spike	MERS-CoV neutralization	Preclinical development (Jiang, Wang et al. 2014)
Dana Farber Institute, Boston, MA, USA	3B11, 1F8, 3A1, 80R	Spike	MERS-CoV neutralization	Preclinical development (Tang, Agnihothram et al. 2014)
New York Blood Center, New York, NY, USA	Mersmab1	Spike	MERS-CoV neutralization	Preclinical development (Du, Zhao et al. 2014)
Regeneron Pharmaceuticals, Tarrytown, NY, USA	REGN3051, REGN3048	Spike	MERS-CoV neutralization and humanized DPP4 mouse studies	Preclinical development (Pascal, Coleman et al. 2015)
Juntendo University, Tokyo, Japan	2F9 and YS110	CD26	VLP neutralization	Preclinical development (Ohnuma, Haagmans et al. 2013)
Humabs Biomed SA, Bellinzona, Switzerland	LCA60	Spike	Ad5-DPP4 mouse	Preclinical development (Corti, Zhao et al. 2015)

Table 4: current MERS-CoV immunotherapeutic treatment candidates (MG= immunoglobulin, Ad5-DPP4= adenovirus 4 virus expressed dipeptidyl peptidase-4, NHP= nonhuman primate, DPP4= dipeptidyl peptidase-4 /CD26, dipeptidyl peptidase-4 and VLP= virus-like particle). Table sourced from CDC: Emerging Infectious Diseases (Timothy M. Uyeki 2016)

A 2019 study reported that using a combination of a MERS-CoV HR2 peptide mimic (HR2P-M2) alongside an RBD-targeting neutralising mAB (m336) worked more efficiently than either treatment alone (Wang, Hua et al. 2019).

1.2.8 Prevention

1.2.8.1 Physical prevention

Preventative strategies are being researched too. In the 2003 SARS-CoV outbreak, civet cats, the suspected reservoir of the virus, were culled to prevent further spread of the disease. Culling of camels in the Middle East would be culturally impractical as they are important for a variety of reasons including food, milk, transport and racing purposes (Timothy M. Uyeki 2016). Theoretically MERS-CoV infection could be prevented by vaccination, pre- or post-exposure antiviral chemoprophylaxis, or passive immunoprophylaxis of people deemed to be high risk.

1.2.8.2 Vaccines

1.2.8.2.1 Human vaccination

Animal data has shown that full protection against MERS-CoV requires both humoral and cellular immune responses, which are mainly induced by the S and N proteins. Although the N protein is more highly conserved, the crucial role of the S protein in viral entry makes it the main focus of vaccine research (Zhao, He et al. 2018). Candidate vaccines were reviewed during a symposium joint hosted by the International Vaccine Institute (ITI) and WHO in June 2018, the vaccines reviewed all targeted the full-length S protein but varied in vaccine approach including; live-attenuated, inactivated, VLP, and viral-vectored. At the time of writing, most of the potential MER-CoV vaccines shown in Table 5 are currently undergoing animal model analysis.

Vaccine development is not without complication, all vaccines must pay attention to ongoing assessment of antigenic evolution of circulating MERS-CoV strains in order to be effective (Agnihothram, Gopal et al. 2014). Another consideration is the potential for causing antibody-dependent enhancement of the disease upon virus challenge. A study using transgenic mice observed *"pulmonary Th2-immunopathology, associated with eosinophilic infiltration and increased pro-inflammatory cytokines in the lung"* after immunization with an inactivated MERS-CoV vaccine followed by a wild-type MERS-CoV challenge, similar observations have been seen in

a SARS murine model (Tseng, Sbrana et al. 2012); leading to an increase in focus on creating a vaccine which targets more discrete portions of the MERS-CoV S protein (Jiang, Li et al. 2019).

A team at the University of Madrid in Spain has successfully engineered a mutant virus lacking the structural E proteins. The mutant was proven to be replication competent and propagation defective therefore making it potentially a safe and promising vaccine candidate (Almazan, DeDiego et al. 2013). A human phase I clinical trial was started in January 2016 testing the safety of GLS-5300 (a DNA plasmid vaccine that expresses the MERS-CoV spike glycoprotein). The study is examining the results of GLS-5300 when administered at one of three dose levels following a three-injection vaccination regimen followed by electroporation. The study is also assessing immune responses over a 1 year period with respect to the generation of antibody and cellular responses (Trials.gov. 2016). Results are expected to be published in November 2020.

Source	Vaccine	2016 status
Novavax, Gaithersburg, MD, USA	Spike protein trimer in 40 nm particle; likely adjuvanted	Mouse immunogenicity shown (Arabi, Arifi et al. 2014)
NIAID/Vaccine Research Centre, Bethesda, MD, USA	Two candidate vaccine approaches: DNA spike prime-S1 protein boost and S1 prime-S1 boost	Mouse and NHP immunogenicity shown; NHP2 (macaque-radiological efficacy shown) (Wang, Shi et al. 2015)
GeneOne Life Science, Seoul, South Korea; Inovio Pharmaceuticals, Plymouth Meeting, PA, USA	DNA expressing spike; electroporation device	Mouse, NHP, and camel immunogenicity shown; NHP2 (viremia, lung pathology) (Muthumani, Falzarano et al. 2015) phase I study started (Trials.gov. 2016)
Greffex, Aurora, CO, USA	Fully deleted adenovirus packaging vector	Mouse immunogenicity (Greffex 2014)
Erasmus University Rotterdam, the Netherlands; University of Marburg, Marburg, Germany; Ludwig-Maximilians University, Munich, Germany	MVA vectored spike protein	Mouse immunogenicity and protection shown; clinical trials in planning stage (Song, Fux et al. 2013, Volz, Kupke et al. 2015)
New York Blood Centre, New York, NY, USA; Shanghai Medical College, Shanghai, China	Spike receptor-binding domain subunit vaccine	Recombinant protein containing the 377–588 aa fragment of the S1 subunit (Wang, Shi et al. 2015)

Table 5: Human vaccine candidates for MERS-CoV targeting spike protein (NHP= non -human primate, MVA= modified vaccinia Ankara and S1= portion of spike protein with the receptor binding domain). Table sourced from CDC: Emerging Infectious Diseases (Timothy M. Uyeki 2016)

1.2.8.2.2 Camel vaccination

Interrupting the transmission from dromedary camels to human is clearly an important countermeasure in preventing the spread of the disease (Adney, Wang et al. 2019). That being said, culling camels, similar to how badgers were culled to try and prevent the spread of bovine TB in the UK (Blowey, Gray et al. 2017) is impractical for the reasons mentioned in section 1.2.8.1.

However, prevention of camel-to-camel MERS-CoV transmission and reduction of spread from dromedaries to humans, remains worth investigating. An animal vaccination program has been successfully implemented in Australia whereby horse vaccination has interrupted the transmission chain between Hendra virus from fruit bats, to hoses, to humans (Middleton 2014).

Potential camel vaccines being researched are shown in Table 6, similar to candidate human MERS-CoV vaccines most of these focus on neutralization of the spike protein. Young camels have been reported as at a higher risk for MER-CoV infection and therefore may be a priority group when it comes to vaccination (Hemida, Chu et al. 2014, Khalafalla, Lu et al. 2015). Studies also show that camels can be re-infected meaning that a vaccination strategy may require multiple dosing and booster vaccinations to increase effectiveness overtime (Farag, Reusken et al. 2015, Timothy M. Uyeki 2016). A study was conducted using a modified vaccinia virus (Ankara vaccine) that expressed the MERS-CoV spike protein. The vaccine was administered intranasally and intramuscularly to dromedaries who subsequently had fewer signs of infection and lower MERS-CoV titres in the upper respiratory tract than their unvaccinated counterparts when challenged intranasally with MERS-CoV, (Haagmans, van den Brand et al. 2016).

Source	Vaccine	Status
USG/Academic Institution Consortium	Recombinant and inactivated whole virus	Camel vaccination
NIAID Rocky Mountain Laboratories, Hamilton, MT, USA/Colorado State University, Fort Collins, CO, USA	Spike protein subunit vaccine/Advax adjuvant (baculovirus expressed)	Camel and alpaca vaccination studies (Dodge 2014, Kupferschmidt 2015)
Erasmus University Rotterdam, Rotterdamn, The Netherlands; University of Marburg, Marburg, Germany; Ludwig-Maximilians University, Munich, Germany	MVA-vectored spike protein	Camel vaccination challenge studies (Kim, Okada et al. 2014, Kupferschmidt 2015)
Novavax AB, Uppsala, Sweden	Spike nanoparticles with adjuvant likely	In preclinical development
University of Pittsburgh, Pittsburgh, PA	Adenovirus vectored spike protein	In preclinical development (Kim, Okada et al. 2014)

Table 6: Camel vaccine candidates for MERS-CoV targeting spike protein (USG= US government, NIAID= National Institute of Allergy and Infectious Diseases, National Institutes of Health and MVA= modified vaccinia Ankara). Table sourced from CDC: Emerging Infectious Diseases (Timothy M. Uyeki 2016)

1.2.9 Summary

Although there are a number in preclinical and clinical trials, there are currently no approved vaccines against, or specific treatments for, any human coronaviruses, with the recent exception of SARS-CoV-2 (Prompetchara, Ketloy et al. 2020). The rate of reported MERS-CoV cases has decreased in time, and public interest has reduced, however the virus remains an important area of research due to its high mortality rate and the potential for a mutation to arise causing an increase in human to human transmission. Competition for funding, laboratory space and clinical expertise has arisen with other, more recent, emerging or re-emerging infectious diseases, such as the outbreaks of Ebola and Zika virus disease. The overall goal for clinical research in reference to MERS-CoV is the identification of effective therapies, the development of preventative vaccines, the optimization of diagnostics and the improvement of clinical management in order to improve survival rates and potentially eradicate the disease.

Research into potential medical countermeasures for MERS-CoV remains preliminary. The repurposing of FDA approved drugs remains a potential route to treat infection, however would be able to be implemented quicker if the mechanism of action for antiviral activity can be defined and understood, if there is no change to the route and drug form already approved, if the dosing does not exceed current approved dosing and suitable pharmacokinetic data supports the dose and finally if the risk-benefit profile is deemed acceptable.

To continue to make progress towards clinical developments animal models need to become standardised. Optimal challenge dose and route needs to be agreed on and greater access to NHP is required. Currently common marmosets are deemed the most suitable NHP as their onset and severity of disease requires a lower dose of therapeutic drug, however their small size makes repeat sampling difficult and models remain variable between laboratories (Timothy M. Uyeki 2016). Large animal models, including camels and camelids, such as alpacas, are being developed (Adney, Bielefeldt-Ohmann et al. 2016). These models could give greater insight as to the pathology and subsequent immunology caused by the disease in its natural host dromedaries.

There is currently a limited availability of current or recently circulating MERS-CoV strains. There is a need for more epidemiological, clinical, virologic and immunologic data to be shared to improve understanding of MERS-CoV pathogenesis. Only one study has investigated MERS-CoV infection in autopsy tissues of a patient who died from the disease (Ng, Al Hosani et al. 2016). Collaborations are needed to pool and systematically collect clinical specimens from MERS-CoV patients to allow for analysis on biomarkers, immunology and clinical illness, as well as long-term follow-up studies of survivors of severe disease. Studies into the differences between asymptomatic and fatal cases are needed. Effort should be made to partner with clinical trial networks in affected countries to evaluate safety and efficacy of investigational therapeutics,

WHO and the International Severe Acute Respiratory and Emerging Infection Consortium (ISAREIC) are collaborating in adapting standardized protocols for controlled clinical trials for MERS-CoV (Consortium 2017).

This study hopes to assist research developments by focusing on the need for improved diagnostic testing, both for current human coronaviruses and for the use during potential novel outbreaks in the future.

1.2.10 The SARS-CoV-2 pandemic

1.2.10.1 Introduction

As stated in the prologue, this thesis was initially written in 2017, the more recent global SARS-CoV-2 pandemic has required rapid development of vaccines and diagnostic testing kits to try and reduce the disease spread and burden. Whilst SARS-CoV-2 will appear omitted from the practical research element of this study, it is important to outline current diagnostic kits and vaccines for wider context. SARS-CoV-2 was first identified in Wuhan, China in December 2019 and declared a pandemic by the World Health Organization (WHO) in March 2020 (Triggle, Bansal et al. 2021). SARS-CoV-2 is a single-stranded, positive-sense, non-segmented enveloped RNA virus, approximately 29.9 kB in size with a diameter of 50–200 nm (Chen, Zhou et al. 2020). At the time

of writing, August 2021, there have been 207,784,507 cases reported to WHO resulting in 4,370,424 deaths (WHO 2021)

1.2.10.2 Diagnostic tests

The majority of diagnostic tests for SARS-CoV-2 are based on reverse transcription polymerase chain reaction (RT PCR) assays from nasopharyngeal samples which can be used to detect both symptomatic and asymptomatic cases and provide figures which can be used to calculate the R0 number (basic reproduction number). The R0 value is a number relating the average number of people an infected individual is likely to infect. Whilst RT PCR screening can be conducted on asymptomatic carriers of the virus, it is unlikely such individual would feel the need to have the test and therefore it is thought that such cases are not routinely captured in common testing strategies (Nikolai, Meyer et al. 2020). In addition to this, a number of studies now suggest that persistent positive RT-PCRs do not necessarily indicate the presence of replication-competent viruses(Alexandersen, Chamings et al. 2020, Rhee, Kanjilal et al. 2021).

Another commonly used option for SARS-CoV-2 detection are rapid antigen tests, commonly referred to as lateral flow tests. These show a lower sensitivity compared to the standard RT-PCR test however their specificity is generally reported to be high. Antigen tests offer a few advantages over RT-PCR in that they are quick, relatively low cost and can be done by the general public as point of care tests. Different organisations recommend different parameters of standards for the rapid antigen tests, WHO recommends a minimum performance requirement of ≥80% sensitivity and ≥97% specificity, while European Center for Disease Prevention (ECDC) advocates the use of tests with a performance closer to RT-PCR, ≥90% sensitivity and ≥97% specificity ((ECDC) 2020). The Department of Health and Social Care (DHSC) has agreed contracts with Omega Diagnostics and Global Access Diagnostics for the production of the LFT kits within the UK (Gov.UK 2021).

1.2.10.3 Track and trace and other control measures

As with any communicable disease understanding transmission route is vital so that measures can be put in place to minimise the spread of the disease. SARS-CoV-2 is pneumotropic and

spreads through respiratory secretions, droplets generated through coughing, sneezing and talking. Contaminated surfaces can also facilitate the spread of the virus, some studies estimate the SARS-CoV-2 viral particles can remain infectious for up to six days (van Doremalen, Bushmaker et al. 2020, Zhang, Zhang et al. 2020, Zhang, Li et al. 2020). Due to the nature of transmission, non-pharmaceutical interventions (NPIs) such as hand hygiene, face masks, sufficient indoor ventilation, social distancing, contact tracing, quarantining, community lock downs and social restrictions have played a vital role during the pandemic (Triggle, Bansal et al. 2021).

1.2.10.4 Therapeutics

In a similar manner to the two previous novel HCoV outbreaks of SARS-CoV and MERS-CoV, no established therapeutic measure was available to effectively treat SARS-CoV-2 and as such, attention was focused on repurposing currently available drugs including, but not limited to, Chloroquine and Hydroxychloroquine, remedesivir, dexamethasone, ivermectin, 3CLpro and PLpro inhibitors, famotidine, TMPRSS2 and furin inhibitors and convalescent plasma and targeted antibody therapy (Kaddoura, AlIbrahim et al. 2020). The mechanism of actions for some of these SARS-CoV-2 potential treatments can be seen in

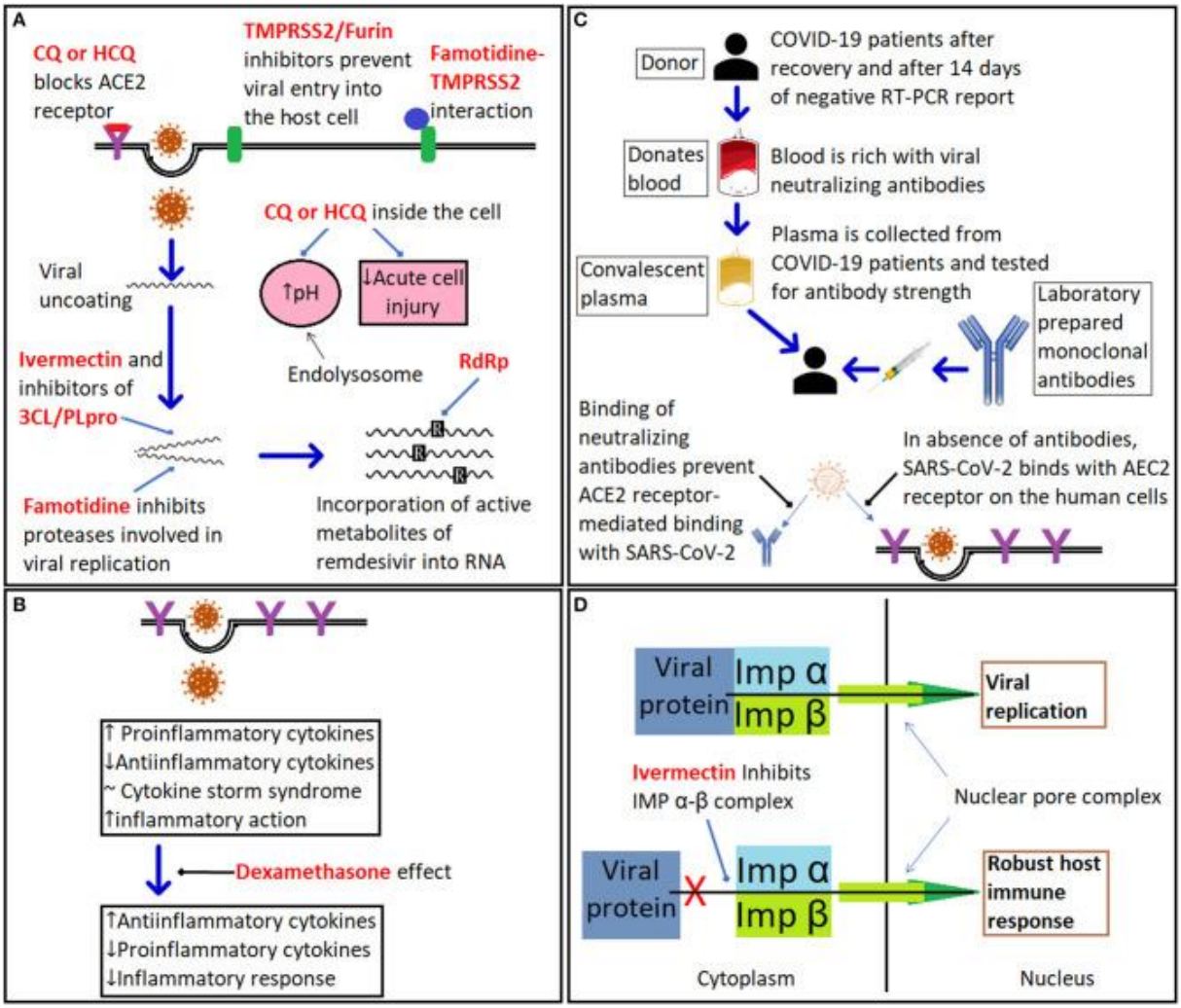


Figure 4, a more comprehensive list of potential and approved SARS-CoV-2 countermeasures can be seen in appendices three and four.

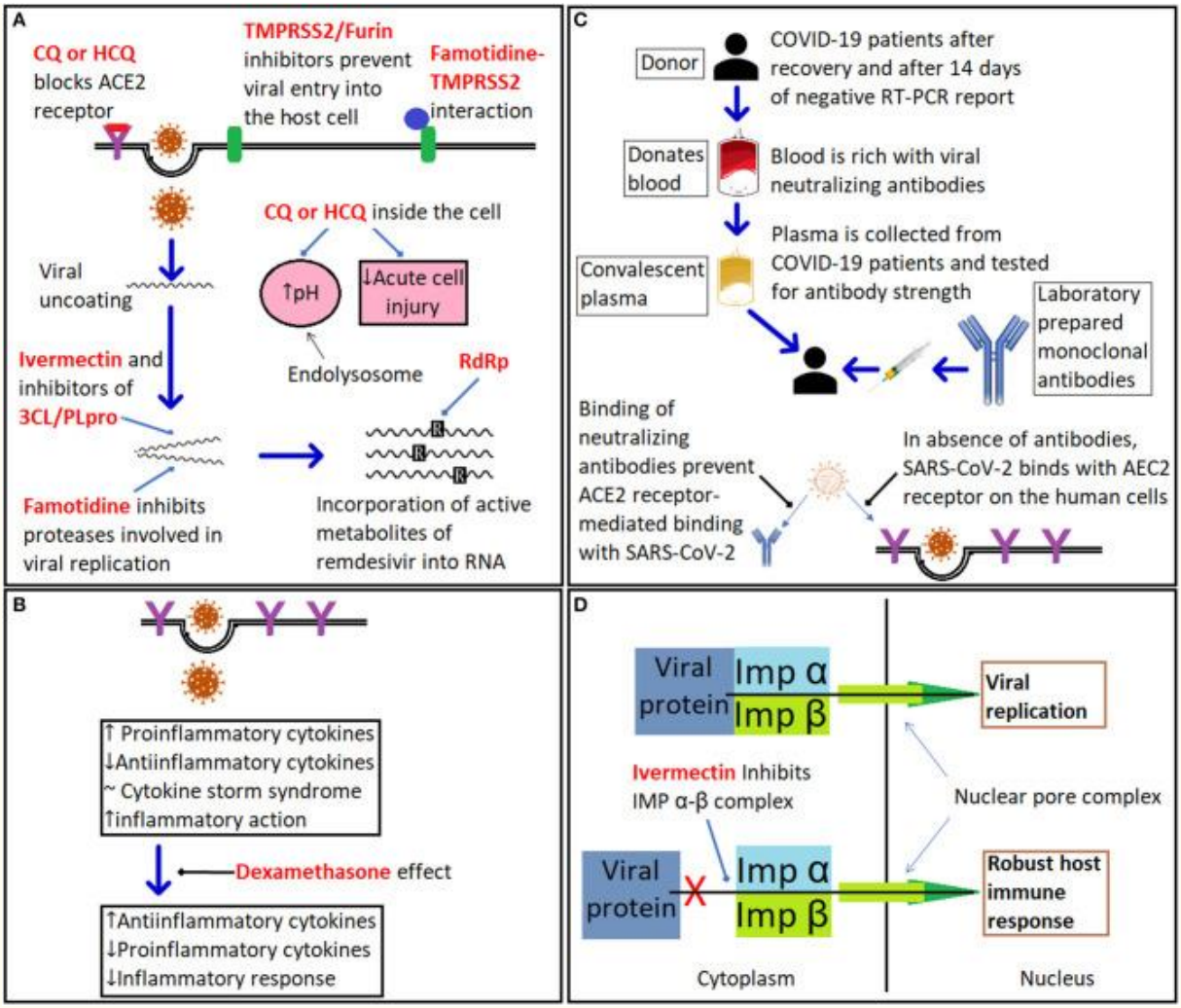


Figure 4: The mechanism of action of different therapeutics against COVID-19. (A) illustrates the mode of action of drugs targeting COVID-19 including chloroquine (CQ) and hydroxychloroquine (HCQ) and their multiple putative sites of action: (i). ACE2 receptor for SARS-CoV-2; (ii). increasing the pH of the endolysosome; and (iii). suppression of the immune response. Sites of action of TMPRSS2 inhibitors such as camostat, famotidine, and furin inhibitors are shown; famotidine is also a putative inhibitor of the 3CL/PLpro proteases; ivermectin is a putative TMPRSS2 inhibitor that also inhibits the importin (IMP) α-β complex and viral replication; while remdesivir inhibits viral RNA polymerase. (B) Dexamethasone suppresses expression of pro-inflammatory cytokines. (C) Summary of role of convalescent plasma and monoclonal antibody therapy. (D) Ivermectin inhibits the heterodimeric importin (IMP) α/β complex via binding directly to IMPα preventing nuclear import of key viral proteins. Figure taken from (Triggle, Bansal et al. 2021).

1.2.10.5 SARS-CoV-2 Vaccines

Since the start of the pandemic a high importance was placed globally upon discovering a vaccine against SARS-CoV-2 to prevent the spread of the virus and reduce the burden of disease. An up to date status can be found through accessing vaccine tracker sites (Corum, Grady et al. 2020). There are a number of reports comparing the leading potential and approved vaccine platforms (Krammer 2020, Sharma, Sultan et al. 2020), summarised in Table 7. The two vaccines most commonly used in the UK are the mRNA BioNTech/Fosun/Pfizer (BNT162) vaccine and the vector-based University of Oxford/AstraZeneca (ChAdOx1 nCoV-19) vaccine.

Vaccine platform	Candidate vaccine	Mechanism	Research institute	Planned route; doses	Published phase 1 & 2 results	Status of phase 3 clinical trials
Recombinant protein particles vaccines	NVX-CoV2373	Deployment of full length recombinant viral spike protein with Matrix M adjuvant	Novavax, USA	IM; Two: day 0 and day 21	(Keech, Albert et al. 2020)	Published 2021 (Heath, Galiza et al. 2021)
Inactivated	Three inactivated vaccines	Whole virus propagated into cell culture line then inactivated with organic compounds. Purified material adsorbed with aluminum adjuvant	Sinopharm, Wuhan/Beijing Institute of Biological Products, China	IM; Three: day 0, day 28, and day 56	(Xia, Duan et al. 2020)	Started Jul 2020, No published results Approved for emergency use
Vector based	ChAdOx1-S/AZD1222	Deployment of spike protein through replication deficient Chimpanzee adeno virus	University of Oxford/Astra Zeneca, UK	IM; Two day 0 and 28	(Folegatti, Bittaye et al. 2020, Ramasamy, Minassian et al. 2021, Voysey, Clemens et al. 2021)	Published Dec 2020 Approved for emergency use
	Ad5-vectored COVID-19	Deployment of spike protein through replication deficient type five	CanSino Biological Institute/Beijing Institute of Biotechnology,	IM; Single	(Zhu, Guan et al. 2020)	Started June 2020 No published results Approved

		adeno virus	China			for emergency use
	Gam-COVID-Vac (Sputnik V)	Deployment of spike protein through replication deficient type 5 and 26 adeno viruses	Gamaleya Research Institute, Russia	IM; Two: day 0 and day 21	(Logunov, Dolzhikova et al. 2020)	Started Sep 2020 No published results Approved for emergency use
mRNA	mRNA–127	Nanoparticle formulated mRNA encoding spike protein or receptors binging proteins	Moderna, USA	IM; Two: day 0 and 28	(Anderson, Rouphael et al. 2020) (Baden, El Sahly et al. 2021)	Published Dec 2020 Approved for emergency use
	BNT162b1		BioNTech/Fosun /Pfizer, Germany, USA	IM; Two: day 0 and day 21	(Mulligan, Lyke et al. 2020, Polack, Thomas et al. 2020)	Published Dec 2020 Approved for emergency use

Table 7: Details of prime potential or approved COVID-19 vaccines and their latest stages of development. IM= Intramuscular. Table taken from (Triggle, Bansal et al. 2021)

Ongoing genome sequencing of SARS-CoV-2 has shown multiple variants to be circulating globally (Koyama, Platt et al. 2020) some of which have proven to affect the infectivity of the virus and may have negative implications to vaccine effectiveness due to mutations in spike protein and subsequent neutralising antibodies. Alpha, beta, gamma and delta variants have been noted as variants of concern by WHO, with a further four (eta, iota, kappa and lambda) mentioned as variants of interest (WHO 2021).

1.3 Research objectives

1.3.1 Creating purified soluble N proteins for MERS-CoV, SARS-CoV, OC43-CoV, NL63-CoV, 229E-CoV and IBV

A previous study has investigated the use of N proteins of HCoV as a basis for serodiagnosis (McCrory 2009) however in the case of some of the viral strains, the proteins made remained insoluble and proved impossible to purify. Although they could be used in the form of crude lysates, the lack of equivalence among the antigens used meant that comparisons between sera banks had a degree of uncertainty. This research aims to overcome that issue by re-investigating the use of coronavirus N proteins as soluble antigens and increasing the range of proteins used to include MERS-CoV. This was achieved by investigating the effect of re-locating the His-tag used for purification from the C to the N terminus of the protein. One study found that on average, N-terminal fusion partners are preferable for optimal protein expression (Dyson, Shadbolt et al. 2004). Having all N proteins in equivalent soluble form would allow more accurate sero-prevalence tests to be conducted. The accession number and strains used in this study are as follows; IBV (FJ589731.1) isolate=IS/1045/03, NL63-CoV (NC_005831.2) strain= Amsterdam I, 229E-CoV (DQ243962.1) strain= HCoV-229E-20/1/04, OC43-CoV (KF530063.1) strain=OC43/human/USA/9612-48/1996, MERS-CoV (KF600623.1) isolate=Al-Hasa_7b_2013 and SARS-CoV (NC_004718.3) isolate=Tor2.

1.3.2 Using purified N protein in sera testing

Once generated, the N proteins will be used to conduct serosurveillance analysis. N proteins are used as they remain more conserved within CoV groups than S proteins (Meyer, Drosten et al. 2014). Enzyme linked immunosorbent assays (ELISAs) will be used to ascertain if there has been any prior infection with, or related to, the identity of N protein used in the test. Human sera can be obtained from a human serum bank provided by Public Health England (PHE) among others. This study aims to confirm the previously obtained data (McCrory 2009) and expand it to include a wider range of viruses.

Chapter 2: General materials and methods

This chapter covers the general materials and methods used in multiple experiments; more specific details are referred to in their relevant chapters.

2.1 Materials

2.1.1 Chemicals

All Chemicals used were of an analytical grade or higher and unless stated were purchased from Sigma or Fisher.

2.1.2 Antibodies:

2.1.2.1 Primary antibodies

- Anti-6X His tag antibody horseradish peroxidase (HRP) ab1187 (Abcam) developed in mouse, used at a 1:2,000 ratio, within manufacture's advisory range of 1:1,000 - 1:5,000.
- Anti-GFP antibody (HRP) ab184207 (Abcam) was used at a ratio of 1:5,000 which is within manufacturer's recommendation of 1:5,000 - 1:20,000

2.1.2.2 Secondary antibodies

- Anti-Chicken IgY (IgG) (HRP) (Sigma) was used 1:2,000 as advised by the manufacture.
- Anti-Mouse (HRP) (Dako) was used 1:1,000 as advised by the manufacture.
- Anti-Rabbit (HRP) (Dako) was used 1:3,000 as advised by the manufacture.
- Anti-Human polyvalent immunoglobulins (whole molecules) peroxidase was used 1:20,000 as advised by the manufacturer

2.1.2.3 Sera

Chicken, mouse and rabbit sera were supplied by the Pirbright institute.

Human sera was collected from volunteers within the laboratory.

2.1.3 Reagents used

The buffers and solutions used throughout this study are shown in Table 8.

Solution	Composition	Application
Phosphate buffered saline (PBS)	8% (w/v) NaCl 0.2% (w/v) KCl 1.15% (w/v) Na2HPO4 0.2% (w/v) KH2PO4 pH 7.4	Miscellaneous
Luria Agar (LA) with ampicillin	1% w/v) tryptone 0.5% (w/v) yeast extract 0.5% (w/v) NaCl 1.5% (w/v) agar 100µg/mL ampicillin	Bacterial cell culture and expression
Lysogeny Broth-Luria (LB) with ampicillin	1% (w/v) tryptone 0.5% (w/v) yeast extract 0.5% (w/v) NaCl 0.1% (w/v) glucose 100µg/mL ampicillin	Bacterial cell culture and expression
Super optimised broth with catabolite repression (S.O.C.) medium (Invitrogen)	2% (w/v) tryptone (Oxoid) 0.5% (w/v) yeast extract (Oxoid) 10mM NaCl 2.5mM KCl 10mM $MgCl_2$ 10mM $MgSO_4$ 20mM glucose	Bacterial cell culture and expression
Cell culture medium	Biowhittaker® Insect-XPRESS™ medium (Lonza) supplemented with 2% (v/v) foetal calf serum (Biosera)	Insect cell culture and infection
Tris-acetate, ethylenediaminetetraacetic acid (EDTA) buffer (TAE):	7mM Tris-HCl 2.5mM EDTA 0.025% (v/v) glacial acetic acid pH 7.7	Agarose gel electrophoresis
SDS lysis buffer	4% sodium dodecyl sulphate (SDS) 20% glycerol 10% 2-mercaptoethanol 0.004% bromphenol blue 125mM Tris-HCl	SDS PAGE
Running buffer	25mM tris base 192mM glycine 0.1% (v/v) sodium dodecyl sulfate (SDS)	SDS PAGE
Transfer buffer	25mM Tris base 19.2mM glycine 20% (v/v) methanol	Western blotting

Tris buffered saline (TBS)	50mM Tris 150mM NaCl pH 8.0	Western blotting
Tris buffered saline tween (TBST)	50mM Tris 150mM NaCl 0.2% (v/v) tween	Western blotting
Comassie staining solution	0.25% (w/v) coomassie brilliant blue R250 25% (v/v) methanol 8 % (v/v) glacial acetic acid	SDS-PAGE staining
Destaining solution	10 % (v/v) methanol 10% (v/v) glacial acetic acid	SDS-PAGE destaining
8 × sodium phosphate with NaCl buffer	160mM Na2¬HPO42H2O 160mM NaH2PO4H2O 4M NaCl	His-Trap column purification
Loading buffer:	20mM sodium phosphate 500mM NaCl 40mM imidazole	His-Trap column purification
Elution buffer:	20mM sodium phosphate 500mM NaCl 500mM imidazole	His-Trap column purification

Table 8: all of the solutions used in this study

2.1.4 DNA modification and restriction enzymes

The enzymes used were FastDigest enzymes (ThermoScientific). *NcoI* restriction enzyme recognizes C^CATGG sites, *XhoI* restriction enzyme recognizes C^TCGAG sites; both enzymes cut best at 37°C in 5-15 minutes using universal FastDigest Buffer.

2.1.5 DNA and protein markers

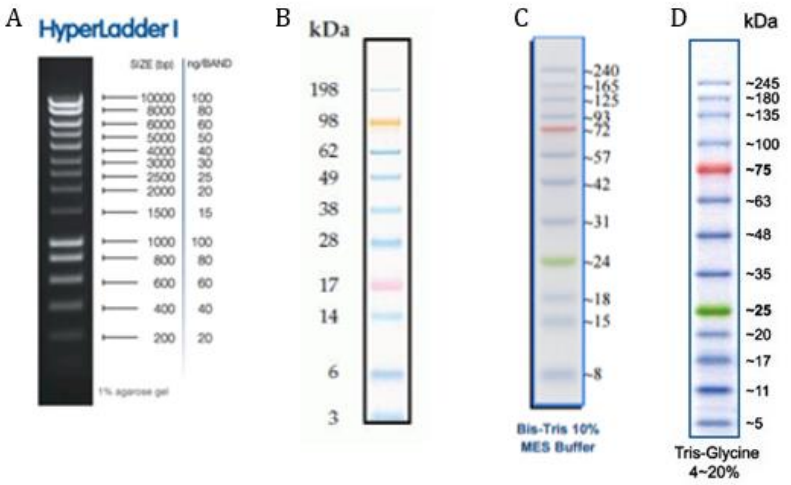


Figure 5: Four molecular gel ladders used. A) Hyperladder™I by Bioline used for DNA sizing in gel electrophoresis B) SeeBlue® Plus2 pre-stained protein ladder by ThermoFisher and C) BLUEeye protein ladder by Geneflow D) EXtended protein ladder by GeneOn Images taken from Bioline, Thermofisher, Geneflow and GeneOn's websites respectively.

Hyperladder™ I, purchased from Bioline Reagents Ltd (London, UK), is a ready-to-use molecular weight ladder consisting of 14 bands spanning from ~200 base pairs (bp) to 10,000bp used to establish size and concentration of DNA when run using gel electrophoresis, as shown in Figure 5.

Also shown in Figure 5 is SeeBlue® Plus2, purchased from Thermofisher (Hemel Hempstead, UK), a tricoloured protein ladder consisting of eight blue polypeptides, a purple myoglobin red band and an orange phosphorylase band. The markers span between ~ 5-200 kilodaltons (kDa) and are used to size proteins analysed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) during western blots and/or coomassie staining. Whilst using this marker in some blots the eighth band, the bovine serum albumin BSA marker, used to signal ~98kDa was found to cross react with the anti his antibody and therefore an alternative marker was looked into in order to produce clearer blots.

The final two ladders displayed shown in Figure 5 are BLUeye pre-stained protein ladder (Geneflow), a tricoloured protein standard displaying 12 proteins, these range from ~10 to 245kDa and Extended protein ladder (GeneOn)with 13 proteins ranging from 5 to 245kDa. These ladders produced no cross reactivity with the anti his antibody and were therefore used for the majority of the blots, depending on availability within the laboratory.

2.1.6 Antibiotics

Ampicillin sodium salt (Melford) was dissolved in nanopure water and underwent 0.22µm syringe filtration for sterilization purposes. 1,000 × stock solutions were made up to a concentration of 100mg/mL and kept as 1mL aliquots at -20^{o}C to prevent unnecessary thaw/freeze cycles.

Chloramphenicol (Boehringer Ingelheim) was used to make 1,000 × stocks (34mg/mL). This was done by dissolving 0.34g in 10mLs 100% ethanol before being 0.22µm filtered and kept at -20^{o}C in 1mL aliquots.

2.2 Methods

2.2.1 Prokaryotic generation methods

2.2.1.1 Baculovirus infection of Sf9 cells

DNA for cloning into prokaryotic expression vectors was generated by amplification of N genes already cloned into recombinant baculoviruses. Baculovirus stocks expressing OC43-CoV, NL63-CoV, 229E-CoV and IBV were used to infect a six well dish of Sf9 cells at 70% confluence and left for five days when cytopathic effects (CPE) of the cells were noticeable when viewed at ×40 magnification.

2.2.1.2 DNA extraction

The cells were carefully dislodged from the dish using a cell scrapper and centrifuged in 15mL falcons at 4,500 rpm at 4^{o}C for 10 minutes. The DNA was then eluted to a final volume of 200µL using a DNeasy kit (Qiagen), used as per the manufacturer's instructions.

2.2.1.3 PCR amplification of viral DNA plasmid

In order to amplify the N encoding DNA, as opposed to any cellular DNA, specific primers were designed and ordered from Integrated DNA Technologies, these are shown in Table 9

. The designed primers allowed for the insertion of a polyhistidine tag at the N terminal sequence as well as the addition of a stop codon at the C terminal sequence in order to prevent replication of a polyhistidine tag at the C terminal domain of the protein that would occur as a result of the sequence present in the pTriEx1.1 vector used.

Reference number	Product	Description	Section of sequence (5'-3')	Melting temperature (Tm) in °C
65978999	25nmole DNA oligo	229E-CoV_NR	CGCCTCGAGTTAGTTGACTTCATCAAT	63.3
65979000	25nmole DNA oligo	IBV_NR	GCGCTCGAGTTACAACTATTTTCACC	68.3
65979001	25nmole DNA oligo	OC43-CoV_NR	CGCCTCGAGTTATATTTCTGAGGTGTC	66.4
65979002	25nmole DNA oligo	NL63-CoV_NR	CGCCTCGAGTTAATGCAAAACCTCGTT	61.1
65979003	100nmole DNA oligo	229E-CoV_NF	CGCCCATGGGCAGCAGCCATCATCAT	72.9
65979004	100nmole DNA oligo	IBV_NF	CGCCCATGGGCAGCAGCCATCATCAT	73.9
65979005	100nmole DNA oligo	OC43-CoV_NF	CGCCCATGGGCAGCAGCCATCATCAT	70.7
65979006	100nmole DNA oligo	NL63-CoV_NF	CGCCCATGGGCAGCAGCCATCATCAT	73.3

Table 9: The 8 primers used alongside their corresponding melting temperature and sequences. Oligo= Oligonucleotides, F=forward primer and R=reverse primer.

Phusion® high-fidelity PCR master mix with high-fidelity (HF) buffer was purchased from New England Biolabs® Inc (Hitchin, Hertfordshire, UK). This was used in order to reduce error rates, as it has a *">50-fold lower error rate than that of Taq DNA Polymerase"* (McInerney, 2014). All PCR reactions were performed using a SensoQuest labcycler (Geneflow) or Bio-Rad T100™ Thermal cycler from Bio-Rad Laboratories Inc (Hemel Hempstead, Hertfordshire, UK).

Reactions were set up as follows:

10µL X2 Phusion® Master Mix

2µL forward primer (see table 3)

2µL reverse primer (see table 3)

1µL template

5µL autoclaved H_2O

Total= 20µL/reaction

Run conditions were programmed as below;

Number of cycles	Temperature (°C)	Time (seconds)	Step
1	98	30	Initial denature
30	98	10	Denaturing
	50	30	Annealing *
	72	45	Extension
1	72	300	Final extension
	4	∞	hold

Table 10: PCR run conditions used. *initially a higher annealing temperature of 58°C was used however this proved to be unsuccessful for all amplicons and so subsequently was lowered to 50°C.

2.2.1.4 Plasmid Miniprep

The pTriEx1.1 vector was gained using a pre-existing construct within the laboratory containing a MERS-CoV N protein insert. The glycerol stock was grown up overnight in LB containing ampicillin and the plasmid was extracted using a QIAprep spin miniprep kit (QIAGEN) used as per manufacturer's instructions; the bacterial cells were harvested by centrifugation at 10,000rpm for 15 minutes. The supernatant was discarded, and the cell pellet resuspended in 250µL of buffer P1 containing 0.1mg/mL RNase A. Lysis occurred with the addition of 250µL of Buffer P2, the mixture was left to incubate at room temperature (RT) for 5 minutes. 350µL of buffer N3 was added to neutralize the lysate and adjust to high-salt binding conditions required for binding of plasmid DNA to silica, all contents were mixed thoroughly by inverting the tube 4-6 times. Cell debris was pelleted by centrifugation at 13,000rpm for 10 minutes and the supernatant was applied to a QIAprep spin column containing a silica membrane. The QIAprep spin column was centrifuged at 10,000rpm for 1 minute. 500µL of buffer PB was added and centrifuged at 10,000rpm for 1 minute to remove any impurities within the column before 750µL of buffer PE was added and centrifuged at 10,000rpm for 1 minute. To remove all wash buffers an additional centrifugation at 10,000rpm for 1 minute was done. The plasmid DNA was eluted from the column

by the addition of 25µL of sterile water, incubation at RT for 2 minutes, and centrifugation at 10,000rpm for 2 minutes. The resulting plasmid DNA solution was then aliquot into 2 microcentrifuge tubes; one kept at -20^{0}C as a master stock while the other was kept at 4^{0}C as a working stock.

2.2.1.5 Double enzyme digest

Fast digest enzymes were used to remove the MERS-CoV N protein sequence and make the pTriEx1.1 plasmid available for the N protein sequences of OC43-CoV, NL63-CoV, 229E-CoV and IBV to be inserted. The fast digest enzymes were used both on the pTriEx1.1 MERS-CoV N protein vector as well as the OC43-CoV, NL63-CoV, 229E-CoV and IBV PCR fragments, to generate "vector" and "inserts" accordingly. The enzymes used were *NcoI* and *XhoI* as the sites flanked the N protein sequence. The reaction was set up as detailed in Table 11 and incubated for 30 minutes at 37^{o}C in order to separate the desired insert from the original vector. The double digest reaction was analysed by electrophoresis on a 1% agarose gel.

Reagent	Concentration/volume
NcoI	10 units
XhoI	10 units
10X fast enzyme digest buffer	2µL
DNA	10µL
Nuclease-free water	6µL
Total volume= 20µL	

Table 11: the reagents, volumes and concentrations used in double enzyme digest reactions

2.2.1.6 Gel electrophoresis

Gel electrophoresis was used for resolution of DNA fragments. Agarose powder (Bioline, UK) was weighed and mixed with TAE buffer to a concentration of 1.0% (w/v) and then completely melted in a microwave oven. After the solution had cooled to about 60^0C, 0.1µL/mL of gel red (Biotium Inc, USA) was added. The solution was poured into a casting tray containing a sample comb and left to solidify at RT. The sample comb was removed from the casting tray and the gel placed into the electrophoresis chamber with sufficient TAE buffer to cover. Samples were loaded with $1/6^{th}$ the volume of 6×loading buffer (Bioline). Gels were run at 100V until the tracking dye had migrated until the end of the gel and DNA was visualised and photographed using a G:BOX Chemi (Syngene) with GeneSys software.

2.2.1.7 Gel extraction

The pTriEx1.1 vector larger band (~5,300bp) produced by the double enzyme digest of the MERS-CoV plasmid was extracted, as were the ~1,500-1,600bp bands produced by the PCR reactions conducted on various constructs. Gel extraction was carried out using a Zymoclean gel DNA recovery kit (Cambridge BioScience); the target DNA band was identified using the blue light transilluminator and (as such no images were taken) excised from the gel. The gel slice was then placed in a 1.5mL microcentrifuge tube (Sarstedt, Germany) and weighed. 3 times the volumes of agarose dissolving buffer (ADB) was added to 1 volume of gel and the tube was incubated at 55^0C for 10 minutes in water bath to allow the gel to dissolve (CP instrument Co Ltd, UK). Next the solution was added to a Zymo-spin column and placed into a 2mL collection tube prior to centrifugation in a bench top microcentrifuge (MSE Micro Centaur, UK) at 10,000rpm for 1 minute. The flow-through was discarded before the column was washed twice by adding 200µl of wash buffer into the column, and centrifuging at 10,000rpm for 1 minute. 10µL of elution buffer was added into the column, incubated at RT for 2 minutes, before being centrifuged at 10,000rpm for 1 minute to elute the DNA from the column.

2.2.1.8 Ligation

T4 DNA ligase (Thermo Scientific) was used to ligate the insert into the newly extracted pTriEx1.1 vector. The manufacture's protocol for DNA insert ligation into vector DNA using sticky-end ligation was followed. In this case 1µL of vector was used with a 5:1 ratio insert to vector.

Reagent	Concentration/volume
Linear vector DNA	20-100ng
Insert DNA	1:1 to 5:1 molar ratio over vector
10X T4 DNA ligase buffer	2µL
T4 DNA ligase	1µL
Nuclease-free water	Up to 20µL
Total volume= 20µL	

Table 12: The reagents, volumes and concentrations used in the ligation reactions

A vector only sample was also made to check the occurrence of self-ligation. Once the reaction was prepared, it was incubated for 10 minutes at 22°C, before being ready for the initial transformation into Novablue singles (Millipore) *E. coli* cells.

2.2.1.9 Transformation

The transformation occurred using the heat shock method as laid out in Novagen competent cells protocol; cells were thawed on ice before being mixed gently with 2µL of the ligation reaction and incubated on ice for 30 minutes. Heat shock occurred at 42^{0}C for 30 seconds using a HAAKE B3 water bath (CP instrument, UK) before being cooled on ice for 2 minutes. 800µL of SOC medium was added and the mixture was incubated at 37^{0}C for 1 hour in a 250rpm-shaking incubator (Gallenkamp, UK) before being plated onto LB agar plates containing 100µg/mL of ampicillin. Plates were inverted and incubated at 37°C for transformed colonies to appear within 12-16 hours.

2.2.1.10 DNA sequencing

In order to ascertain that the ligation had been successful, and that the constructs' DNA were correct, the post-ligation transformed colonies were grown overnight and had their plasmids extracted using miniprep method.

Once the constructs had been confirmed correct subsequent transformation could take place in order to place the vectors into an inducible cell line to facilitate protein production. Reactions were conducted using purchased One shot® BL21(DE3) cells (Thermo Scientific) or 10µL of cells made competent in the laboratory (see section 2.2.1.13). Transformation for all cells, excluding the Novablue singles, followed the Invitrogen One shot BL21 manual. The Novablue singles transformation protocol followed that detailed in Novogen competent cells instead. The cells, once mixed SOC, were spread at two volumes of 50µL and 25µL onto plates containing chloramphenicol (34µg/mL) and ampicillin. Providing the 'vector only' control plate had few or no colonies present and the sample plates had grown, overnights were made as previously described.

2.2.1.11 Glycerol stocks

Glycerol stocks were made using fresh overnight cultures. This was done by adding 800µL of the overnight cultures to 200µL 50% (v/v) glycerol in a cryogenic tube, before being stored at -80°C. The rest of the overnight sample underwent a miniprep in order to isolate the constructs now containing a His-tag. The plasmid was transformed into BL21 cells for future protein production.

2.2.1.12 Isopropyl β-D-1-thiogalactopyranoside (IPTG) induction

IPTG triggers the transcription of the lac operon by binding to the lac repressor and can therefore be used to induce protein expression whereby the gene is under the control of the lac operator. IPTG (Melford) stocks were made up to 100mM by dissolving 0.23g in 10mLs of nanopure water and 1mL aliquots were then stored at -20°C until needed. Induction was carried out following the expression guidelines set out in Invitrogen One Shot BL21(DE3) user manual. The optical cell density (λ600nm = ~0.4) was reached after a 1:20 dilution of the overnight culture was grown for

approximately 2 hours at 37°C 225rpm and IPTG was added to a final concentration of 0.5mM and protein expression took place for 2-3 hours.

2.2.1.13 Competent cells

Expression of NL63-CoV, 229E-CoV and IBV were all low in Bl21 (see Figure 12) therefore other strains of *E. coli* were looked in to. 10mLs of LB was inoculated with 0.5mLs of an overnight culture. The four strains used were BL21-star(DE3)pLysS, C41(DE3), C43(DE3) (Thermo Scientific) and Rosetta™(DE3)pLysS. These were then grown to mid log phase (OD600 ~0.6) after being incubated at 37°C in a shaking incubator at 250rpm for approximately 90 minutes. Cells were then spun at 4,500 rpm for 10 minutes at 4°C and the pellets resuspended in 0.5mLs ice cold 100mM $CaCl_2$. Importantly cells were kept cold at all times. The cells were then incubated on ice for 30 minutes and pelleted in a microfuge run at 4°C, 13,000 rpm, for two minutes. The supernatant was discarded, and the cells resuspended in 0.1mLs ice-cold 100mM $CaCl_2$. Cells were kept on ice and used within 48 hours ready to be transformed. Plasmids were re-transformed into the four new cell types and the IPTG induction was conducted again.

2.2.2 Protein extraction, purification, concentration and storage

2.2.2.1 Cell lysis

Once the IPTG induced cells were pelleted the samples then needed to be lysed. Pellets were resuspended in 20-50 × initial volume in PBS. EDTA was also added in order to chelate ions. This was done because lysozyme is inhibited by any calcium ions that the cells may contain. However, the presence of EDTA in the sample proved to strip the column and so was not used thereafter. The lysis technique was done with the addition of Triton X100 to a final concentration of 0.1% (v/v) followed by using the cell disrupter at 50% amplification for 10 minutes; alternating 30 seconds on with 30 seconds of rest. The sample was always kept on ice to prevent heat denaturation of the proteins. Further to this, lysozyme (Sigma) was added to a final concentration of 1 mg/mL and the sample was left on ice for half an hour in order to hydrolyse the peptidoglycan

present in bacterial cell walls. Once the lysis had occurred a clarifying spin was done at 10,000rpm for twenty minutes at 4^{o}C in the centrifuge.

During the protein production process, other cell lysis techniques were tried and used including the use of the cell disruptor manufactured by Stansted Fluid Power Limited (Essex, England), cell lytic B and the sonicator was instead used for 1 minute at 60% amplification alternating 5 seconds on and 5 seconds off.

2.2.2.2 His-trap columns

The His-trap column works using immobilized metal affinity chromatography (IMAC). The column is prepacked with Ni Sepharose Fast Flow; a medium containing highly cross-linked agarose beads with an immobilizing chelating group, pre-charged with nickel ions (Ni^{2+}). The soluble histidine tags form complexes with the nickel ions, causing the tagged protein to become bound. This is a reversible process; the addition of imidazole out-competes the histidine for the nickel ions and the attached proteins become eluted.

2.2.2.2.1 Buffers

The 8 × phosphate buffer was initially made up to 90mL, then pH was adjusted to 7.4 using either 1 M NaOH or 5 M HCl. The buffer was then 0.45µm filtered. The imidazole (MW=68.08) was made up to a 1M solution by dissolving 13.6g in 200mL H_2O.

2.2.2.2.2 Run protocol:

The column was run in accordance with the manufacture's protocol. BioLogic LP system was used supplied by Bio-Rad. The program was set to run 0-100 % elution buffer over 60 minutes with a flow rate of 2.5mL/min. To account for tubing length a secondary step of 100% elution buffer was run at 5mL/min for 5 minutes at the end of the program. Samples were collected every 2 minutes.

2.2.2.2.3 Regeneration:

The column doesn't have to be stripped and recharged in between each individual use, however upon occasion the column may become hindered by contaminants and nickel ions can become depleted therefore it can become useful to regenerate the column using the following method.

5 – 10 × column volume of stripping buffer (binding buffer+ 0.05M EDTA)
5 – 10 × column volume H_2O
5 – 10 × column volume binding buffer
5 – 10 × column volume 1M NaOH leave at 4°C for 2 hours
5 – 10 × column volume H_2O
5 – 10 × column volume 1X binding buffer
Recharge with 5mLs $NiSO_4$
Wash column with H_2O

2.2.2.2.4 Storage:

For sterility, the column was stored in 20% ethanol at 4°C in between uses.

2.2.2.3 Protein concentration and storage:

All column fractions shown to be positive by western blot were pooled and concentrated using a vivaspin column (Satorius) with a molecular cut off weight (MCOW) of 30,000. The sample was spun until the concentration was ~1mg/mL (OD280=1.0) and then stored at -20°C.

2.2.3 Protein detection methods

2.2.3.1 SDS-PAGE

Proteins were resolved by SDS-PAGE to determine their size and purity. 10% precast polyacrylamide gels (Bio-Rad) were used. Samples were mixed with an equal volume of 2×SDS-loading buffer and heated to 100°C for 10 minutes. The denatured protein samples were vortexed and centrifuged (MS Microcentaur, Sanyo) for five minutes at 1,300rpm. Gels were run in 1×SDS running buffer at 200V until the bromophenol blue dye reached the bottom of the gel.

2.2.3.2 Western Blot Analysis

Proteins separated by SDS-PAGE were transferred to a polyvinyl difluoride membrane (PVDF) (Millipore). This was conducted using 6 sheets of filter paper (Whatman, USA) and 1 sheet of PVDF

membrane which was cut to fit the size of the gel. The PVDF membrane was initially submerged in methanol before being covered in the transfer buffer together with the filter paper. The apparatus was set up so with 3 sheets of filter paper on the anode plate of semi-dry blotting apparatus (ATTO cooperation, Japan) followed by the membrane, the gel, and the last 3 sheets of filter paper. The semi-dry blotting apparatus was then connected to the power station electrodes and electrophoresed for 1 hour and 20 minutes 35V/ 150mA to allow proteins from the gel to be transferred onto the membrane.

Membranes were then blocked in blocking solution overnight at 4^{o}C. The membrane was subsequently washed three times for 5 minutes with TBST before being probed with the primary antibody diluted in the blocking solution for one hour. Following another three 5-minute washes, membranes were incubated for one hour with the secondary antibody conjugated to horseradish peroxide (HRP) again diluted in the blocking buffer. Finally, membranes were washed for three times of 5 minutes and signal was detected using ChemiFast Chemiluminescence Substrate (Syngene) according to the manufacturer's instructions. Images were visualised and photographed using a G:BOX Chemi (Syngene) with GeneSys software.

2.2.3.3 Coomassie staining

Coomassie staining could be conducted alongside or instead of western blot analysis after SDS-PAGE has been conducted. The gel is extracted from the plastic container and left in the coomassie staining solution for >30 minutes at RT, the stain is then removed and replaced with the destaining solution which can be replaced as often as needs be for the bands to visualise and the remaining gel to destain. Images were visualised and photographed using a G:BOX Chemi (Syngene) with GeneSys software.

2.2.3.4 Mass spectrometry

To verify protein purification, matrix-assisted laser desorption-ionisation (MALDI) spectrometry was used. MALDI is an ionisation technique which uses a laser energy-absorbing matrix to create ions with little fragmentation and is commonly used in microbiology for the identification

of bacterial and fungal isolates (Israr, Bernieh et al. 2020). The process was outsourced to Protein and Nucleic Acid Chemistry (PNAC) Facility at the Biochemistry Department of the University of Cambridge and offers ease of use, high throughput, robustness, cost-effectiveness, rapid analysis and sensitivity (Israr, Bernieh et al. 2020).

2.2.4 Baculovirus expression materials and methods

2.2.4.1 Sf9 cell culture and cell counting

Sf9 cells (Invitrogen) were maintained in suspension at a concentration of between 3×10^5 to 2×10^6 cells/mL in Biowhittaker® Insect-XPRESS™ medium (Lonza). For small scale production, cells were cultured in 100mL spinner flasks with 25mL working volume and incubated at 28°C in a rotary shaker at 30rpm. Aseptic technique was used consistently when working with Sf9 cells and all culturing was carried out in a laminar flow (Peteric Ltd, UK). Sf9 cell confluence and viability was determined by using trypan blue stain followed by cell counting using a Countess automatic cell counter (Invitrogen, USA).

2.2.4.2 Transfection and production of recombinant baculovirus

Sf9 cells were transfected in order to generate new recombinant baculovirus. Cells at a 50% confluence level were used to produce a monolayer in a 6 well dish (~$1x10^6$ cells/well). Lipofectin reagent (Invitrogen) was used to transport the mixture of flashBACGOLD (Oxford Expression Technologies) and transfer vector into the cells. Reactions were made up to a total volume of 24µL using purified water. Each reaction used 1-2µL of flashBACGOLD, 8µL of Lipofectin and 100-500ng of the transfer vector. Positive control reactions were prepared in parallel to test transfections; these contained a Green Fluorescence Protein (GFP) plasmid (Laboratory strain; 1µL). The lipid-DNA complex was incubated at RT for 30 minutes. Cells were left to attach to the plastic during this time, before being washed with serum-free media prior to the addition of the 24µL reaction mix. The cells were incubated for ~3/5 days, or until visible signs of infection. After ~3/5 days, the supernatant of the each well was added to a fresh monolayer of Sf9 cells (~1×10^6 cells/ mL) to increase the viral titre. Passaged cells were incubated at 2°C for a further three to five days.

Passaged cells were harvested by centrifugation (Jouan GR422) at 1500rpm for 10 minutes at 4°C. The resultant supernatant containing the baculovirus was stored at 4°C.

2.2.5 Baculovirus amplification

Baculovirus was amplified in T175 stacker flasks. Flasks were seeded with a monolayer of 17.5×10^6 Sf9 cells and incubated with ~100-200µL of the relevant virus passage's supernatant (typically passage 2 or 3) and left at 28°C for three to five days. The cells were harvested by centrifugation at 4,500rpm for 30 minutes and the resultant supernatant was stored at 4°C.

2.2.6 Virus infection

Once a significant baculovirus titre was reached virus infections were carried out by infecting Sf9 cells (~1×10^6 cells/mL) cultured in media with 10% (v/v) recombinant baculovirus. Cells were cultured in T75 flasks with 10 mL working volume and incubated at 28°C for three to five days then harvested by low speed centrifugation. Pellets were stored at -20°C and supernatants were stored at 4°C following a clarifying spin.

2.2.7 ELISA

Protein samples were left to coat a flat 96 well plate (NUNC) overnight having been diluted to the required concentration in 0.1M $NaHCO_3$. Wells were washed three times for five minutes in TBST and nonspecific binding was reduced by incubation in the blocking buffer for one hour at RT. Wells were washed again for 3 times 5 minutes using TBST and incubated in 50µL of primary antibody or sera following the required dilution series, diluted in blocking buffer solution for one hour at RT. The wells were washed as before 3 times 5 minute and incubated with 50µL secondary peroxidase conjugate antibody diluted in blocking buffer solution for one hour at RT. The wells were washed a further three times five minutes in TBST. 50µL 3,3′,5,5′-tetramethylbenzidine (TMB) ELISA Stable Stop (Europa Bioproducts) was added to each well until a blue colour appeared. The reaction was stopped by addition of 50µL 0.5M H_2SO_4 to each well. The absorbance was read at 410nm using a GENios plate reader and Magellan software, and data was exported to excel.

Chapter 3: N protein expression and purification using a prokaryotic expression system

3.1 Prokaryotic expression system introduction

Prokaryotic expression systems are the most common expression systems used for protein production for a variety of reasons. Firstly, prokaryotic systems are simple and relatively inexpensive to conduct. They also allow large quantities of recombinant protein to be produced with a short generation time. The processes utilise well-known known mechanisms of transcription and translation. Finally, the simplicity of genetic modifications and the availability of many commercially adapted bacterial strains are additional advantages (Porowinska, Wujak et al. 2013).

The principle prokaryotic system is based upon the bacteria *E. coli.* In optimised media *E. coli* has a doubling time of approximately 20 minutes (Sezonov, Joseleau-Petit et al. 2007). The system is not without disadvantages. Proteins can be produced misfolded and remain biologically inactive in insoluble forms called inclusion bodies (Rosano and Ceccarelli 2014). Some synthesized proteins are also toxic to the bacterial host preventing cell cultures from reaching a high density (although this can be overcome by incorporating an inducible promotor which can be activated, only allowing the inserted gene to be transcribed once the culture has reached an optimum density). Prokaryotic cells lack the ability to carry out some eukaryotic post-translational modifications, such as glycosylation and phosphorylation, therefore if these are necessary for the target protein then a mammalian cell expression system may be preferred (Francis and Page 2010). Finally, the system requires aseptic techniques to be adhered to at all times to prevent cross-contamination of the culture.

3.2 T7 polymerase-based pET System

3.2.1 Overview

The pET system is the principal system developed for the cloning and expression of recombinant proteins in *E. coli* (Studier, Rosenberg et al. 1990). The gene of interest is cloned into a pET plasmid under the control of a bacteriophage T7 promoter. Gene expression is induced when a source of T7 RNA polymerase becomes available within the host cell. T7 RNA polymerase is highly selective and powerful; the host cell's resources are quickly switched to target gene expression. This system prevents gene expression until induction occurs. This is important if the desired protein were to have toxic effects towards the host cell. Initially target genes are cloned using hosts that do not contain the T7 RNA polymerase gene; in this study's case Novablue cells, a K-12 strain derivative produced by Novagen, were used. After which, plasmids are transferred into an expression *E. coli* strain containing a chromosomal copy of the T7 RNA polymerase gene under the control of the *lac* promoter. The addition of IPTG induces gene expression as detailed below.

3.2.2 Vectors

The vector is a circular plasmid that is taken up by *E. coli*. When an *E. coli* cell divides, it makes new copies of its own large host chromosome, encoding all the proteins required for cell function. During the bacterial binary fission, new copies of the smaller cloning vector are also replicated. Each progeny cell also contains one, or more often multiple, copies of the vector along with the host chromosome. The pET vectors were originally constructed by Studier and colleagues (Studier and Moffatt 1986, Studier, Rosenberg et al. 1990). Commercial vectors use a T7 promoter that is only transcribed by the RNA polymerase from T7 bacteriophage. The *E. coli* RNA polymerase cannot recognize the T7 promoter upstream of the target gene and so leaky gene expression is mostly avoided.

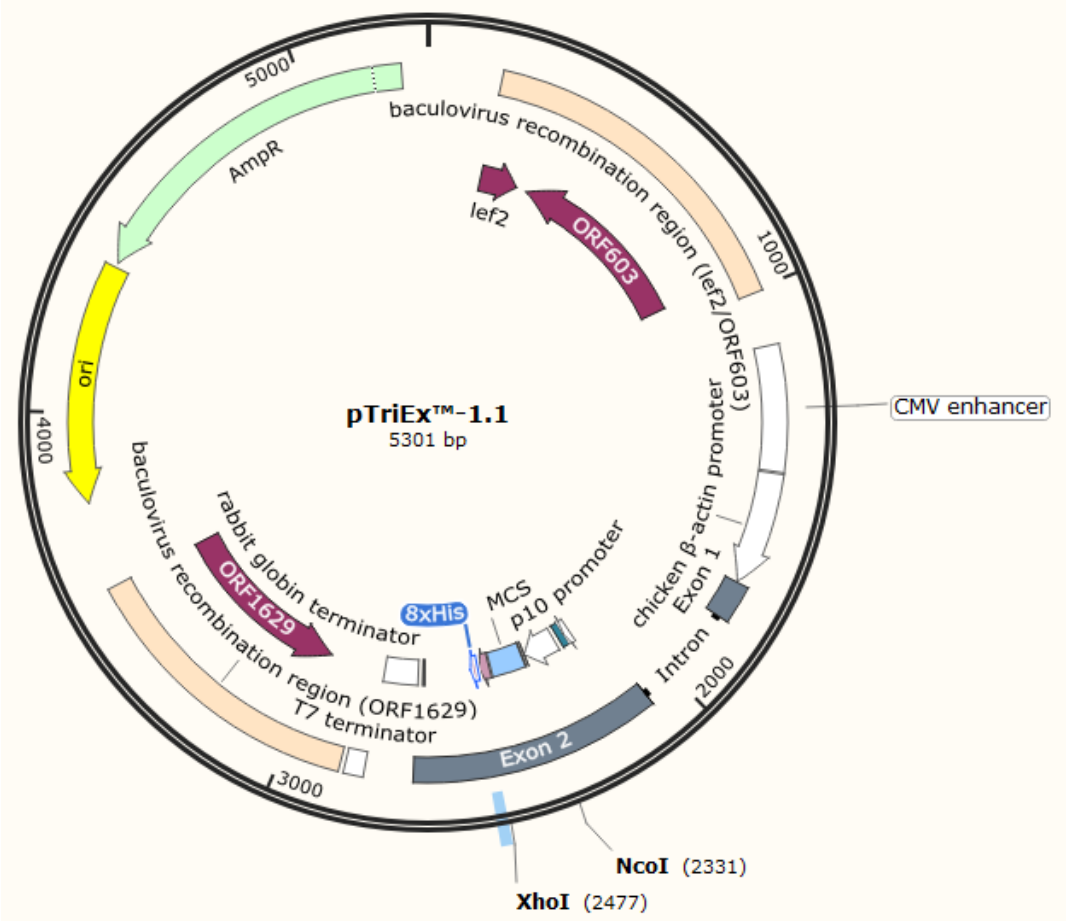


Figure 6: A plasmid map of the pTriEx1.1 vector used, showing the positioning of the *Xho*1 and *Nco1* sites as well as highlighting the plasmid's polyhistidine site in blue. Image generated using SnapGene Viewer (from GSL Biotech; available at snapgene.com)

pTriEx1.1 (Figure 1) was used as the T7 vector in this study. *Nco*I and *Xho*I were used for enzyme digestion within the multiple cloning site (MSC) for the insertion of target DNA. As Figure 6 shows, the *Xho*I site is situated just before the polyhistidine tag, highlighted in blue, and is generally used to fuse the His coding sequence to the incoming fragments to make a C-terminally His-tagged open reading frame. However, previous work in the host laboratory had suggested that an N-terminal His tag was beneficial for SARS-CoV N protein expression so in the cases described here, when the site was chosen the naturally occurring stop codon at the end of the N ORF was retained to prevent the addition of a polyhistidine tag to the C-terminal domain of the protein. Instead forward primers were used to insert the desired polyhistidine tag into the N-terminus of the N proteins to achieve an N terminal His tag.

3.2.3 *Lac* operon

As mentioned T7 DNA polymerase synthesis within the *E. Coli* is under the control of the *lac* promoter. Two enzymes are required for lactose metabolism to occur within *E. coli;* permease and β-galactosidase. Permease is the *lac*Y gene product which allows lactose to be transported into the cell. β-galactosidase is the lacZ gene product and cleaves the disaccharide molecule lactose into glucose and galactose, shown in Figure 7. The genes required for this metabolism process are negatively regulated via its repressor system, encoded by *lac* I (Jacob and Monod 1961). The *lac* operator site is the location of DNA where the *lac* repressor binds, normally located between the promoter and the *lac*Z gene. The *lac* promoter system can therefore be exploited as a molecular switch if a gene of interest is positioned after the promoter region, such that inactivating the repressor, would allow for its transcription.

3.2.4 The lac Repressor

In 1961 Jacob and Monod conducted research on gene regulation and looked into repressors acting as molecular switches (Jacob and Monod 1961). The lac repressor was first isolated by Gilbert and Müller-Hill five years later (Gilbert and Muller-Hill 1966). It is now known that residues 1-49 code for the N-terminal "headpiece" domain of the repressor which contains helix-turn-helix motifs essential for it to recognise the operator units (Friedman, Fischmann et al. 1995, Lewis, Chang et al. 1996). Residues 50-58 code a hinge-helix which connects the headpiece onto the main body of the repressor. If DNA is not present this region moves freely (Wade-Jardetzky, Bray et al. 1979) however when the operator is present it becomes ordered and binds to the central portion of the operator in the minor groove (Lewis, Chang et al. 1996). Residues 62-331 encode the core of the repressor which is divided into two subdomains. Crystal structures have shown that when lactose is absent the tetrameric LacI binds the operator sequence on DNA bending it 40 degrees, blocking the access of RNA polymerase to the promoter site and preventing transcription (Bell and Lewis 2000). When lactose binds to LacI it induces a conformational change in the protein structure leaving it unable to bind to the operator DNA sequence causing it to detach and make way for RNA polymerase and translation.

3.2.5 Effector molecules

Effector molecules alter the affinity of the repressor for the operator. Effector molecules are defined as inducers if their binding lowers the LacI's affinity for the operator and anti-inducers their binding stabilizes the repressor-operator complex. Allolactose is an analogue of lactose and is the natural inducer of the repressor molecule (Jobe and Bourgeois 1972). IPTG is known to be an effective inducer as it lowers the affinity that LacI binds to its operator 1000-fold (Barkley, Riggs et al. 1975).

3.2.6 IPTG

As shown in Figure 7 IPTG is structural mimic of lactose/allolactose that can also bind to the *lac* repressor and induce a similar conformational change that greatly reduces its affinity for DNA. Unlike lactose, IPTG is not part of any cellular metabolic pathway and the sulphur atom creates a chemical bond which is non-hydrolysable by the cell, as a result the cell neither metabolises nor degrades it, meaning the concentration of IPTG added remains constant, making it a more useful inducer of the *lac* operon than lactose itself (Daber, Stayrook et al. 2007).

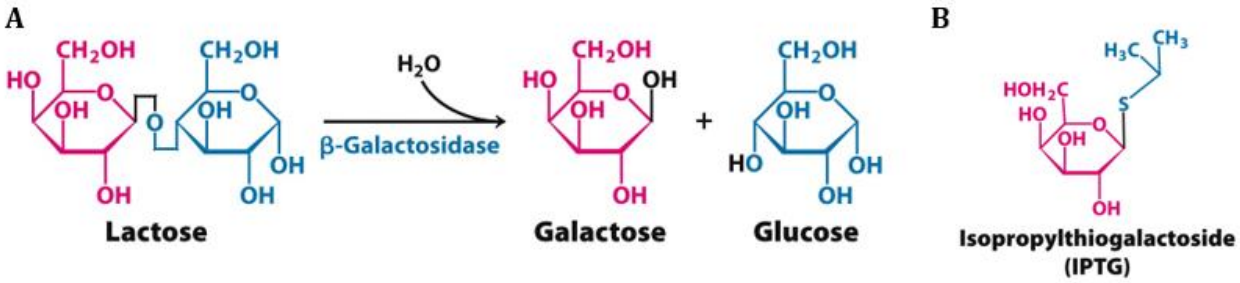


Figure 7: Structural similarities between IPTG and lactose. Figures modified from (Berg, Tymoczko et al. 2012)

3.2.7 Gene expression

The gene coding for T7 RNA polymerase has been engineered into many commercial *E. coli* strains under a modified *lac* promoter and repressor system. Cells denoted by "DE3" contain a λDE3 lysogen that carries the gene for T7 RNA polymerase under control of the *lac*UV5 promoter. With the addition of IPTG the *lac*I detaches from the operator DNA sequence, native *E. coli* RNA polymerase can attach and transcription and translation of T7 RNA polymerase can occur. The newly synthesised T7 RNA polymerase allows the gene of interest within the vector to be also transcribed. Some strains of *E. coli* add an additional gene coding for bacteriophage T7 lysozyme which inhibit T7 RNA polymerase and suppress basal expression of T7 RNA polymerase prior to induction (Stano and Patel 2004). All the steps are summarised in Figure 8.

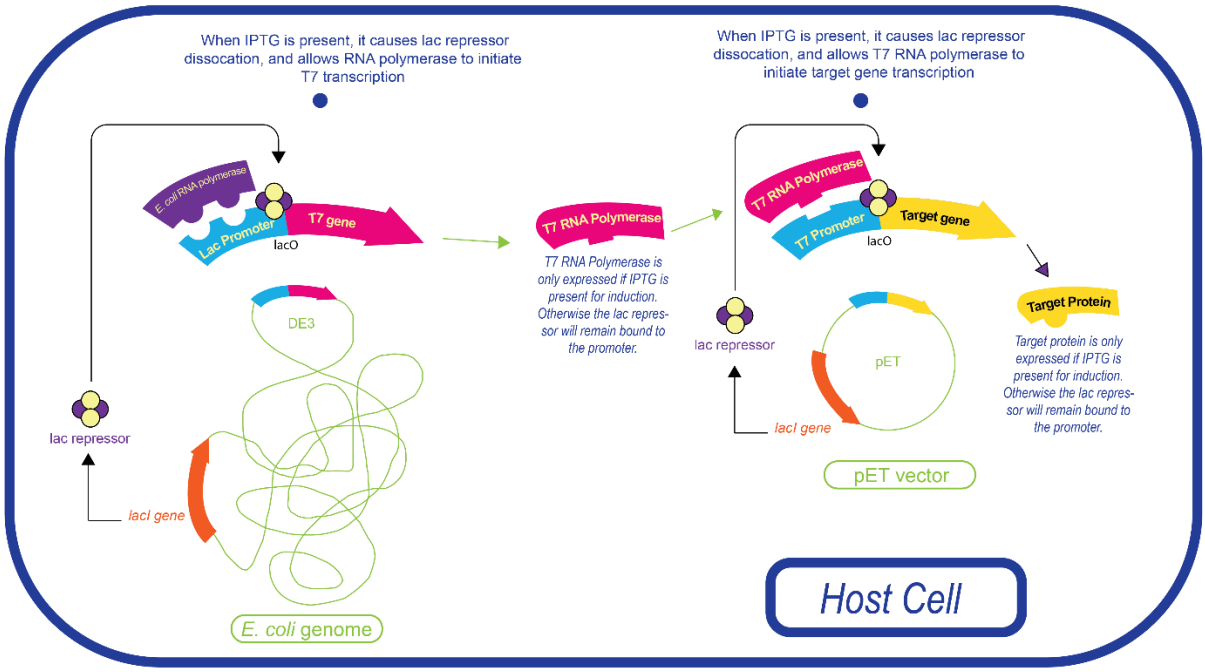


Figure 8: A schematic summarising the roles of IPTG, the pET expression vector, T7 RNA polymerase, the lac repressor and the lack operon functioning within the DE3 *E. coli* strain producing the protein from the gene of interest once induced (Martin 2020).

3.3 Prokaryotic expression materials

3.3.1 Materials/cell lines

The prokaryotic cells used were all variants of *E. coli.* They were sourced commercially, ready to be transformed as;

• Novablue singles™ competent cells (Millipore) were used in the initial cloning stage. One shot® BL21(DE3) cells from Thermo were then used for the initial screening of protein synthesis. The BL21 cells, and any strains containing the term "DE3", carry the lysogen λDE3, meaning that the strain carries a chromosomal copy of the T7 RNA polymerase gene under the control of the *lac*UV5 promoter, making them suitable for IPTG induction. Strains containing "pLysS" mean that the strain expresses T7 lysozyme, which is a T7 RNA polymerase inhibitor, used to suppress basal expression of T7 RNA polymerase prior to induction. Alternatively, the cells were made competent from strains available in the laboratory.

• One shot® BL21(DE3) Star pLys: The BL21-star cells are advantageous as they have been optimised to have high mRNA stability with the aim of increasing protein yield. They have a transformation efficiency of greater than $1x10^8$ cfu/µg plasmid DNA and also have low background expression in uninduced cells. The strain also reduces the degradation of heterologous proteins as it lacks both the Lon and OmpT proteases (Jiang, Oohira et al. 2002).

• OverExpress C41(DE3) and C43(DE3) (Thermo Scientific) C43 was derived from C41. Both strains are beneficial when expressing toxic and membrane proteins. They also carry a chloramphenicol-resistant plasmid that encodes T7 lysozyme and are also deficient in the Lon and OmpT proteases (Miroux and Walker 1996).

• Rosetta™(DE3)pLysS (Millipore): this strain contains a chloramphenicol resistance gene and is beneficial when trying to express proteins that contain codons rarely used in *E .coli* by supplying tRNAs for AGG, AGA, AUA, CUA, CCC, GGA (Doron 2015).

• LOBSTR (low background strain) (Kerafast) are a BL21(DE3) derivative designed specifically with His-tag affinity purification in mind. The strain eliminates the most abundant histidine-rich *E. coli* contaminant proteins *amM* and *slyD*, by containing modified copies of them. *SlyD* peptidyl-prolyl cis/trans isomerase has a 48 amino acid unstructured C-terminal tail containing 15 histidines (Andersen, Leksa et al. 2013).

• ArcticExpress(DE3) and ArcticExpress(DE3)-Ril (Agilent Technologies). Aggregates of insoluble, misfolded proteins are known as inclusion bodies and can be produced when *E. coli* undergoes a forced high-level expression of a heterologous protein. It is possible to purify aggregated protein using denaturation and re-folding steps however this takes additional time and can often lead to biologically inactive proteins being purified (Singh, Upadhyay et al. 2015). ArcticExpress cells are derived from Agilent's BL21-Gold competent cell line and aim to increase the yield of soluble proteins produced using low-temperature cultivation. The strain co-expresses the chaperonins Cpn10 and Cpn60 from the bacterium, *Oleispira Antarctica*, which have 74% and 54% amino acid identity to the *E. coli* GroEL and GroES chaperonins respectively. They are cold-adapted and used to aid protein refolding at temperatures of 4–12°C (Ferrer, Chernikova et al. 2003). ArticExpress-ril cells contain extra copies argU, ileY, and leuW used to encode tRNA that recognise specific codons AGA, AGG, AUA and CUA. This is done in order to prevent the limiting effect and stall to translation that depleted tRNAs can cause.

3.4 Prokaryotic Methods

3.4.1 Summary of prokaryotic system methods

Details of methods used during the expression of N proteins using the prokaryotic system are laid out in the methods section 2.2.1. To summarise; DNA for the OC43-CoV, NL63-CoV, 229E-CoV and IBV samples were obtained using previously constructed baculovirus stocks already available in the laboratory. PCR was conducted to amplify the N-protein sequences, adding a polyhistidine tag at the N terminal domain and a stop codon at the C terminal domain. Double digests using *Xho*1 and *Nco*1 were conducted to create an "insert" for each construct ready to be ligated into a similarly digested pTriEx1.1 vector. Agarose gels were run to verify molecular sizes and the appropriate bands extracted and ligated. The plasmids were transformed into Novablue singles, colonies were selected and grown overnight with ampicillin and the plasmids were obtained using Miniprep ready to be sequenced. Once sequencing was conducted and proved successful the plasmids were transformed into a variety of expression cell lines suitable for IPTG induction and protein expression. Once the proteins had been produced and the cells lysed, the protein required purification. This was done using a 5mL HisTrap FF Crude column (GE Healthcare). The presence of the N-proteins was confirmed using western blots and/or coomassie staining and the proteins were concentrated using Vivaspin columns and stored at -80^{o}C until needed.

3.5 Prokaryotic results

3.5.1 Initial cloning

The primers shown in Table 9 were used to conduct a PCR on extracted DNA templates obtained from previous baculovirus stocks already available in the laboratory encoding OC43-CoV, NL63-CoV, 229E-CoV and IBV N protein, under the conditions as described in Table 10. The PCR samples were then analysed by gel electrophoresis in order to visualise the expected band size of ~1,500 – 1,600kDa, construct dependant.

Previous work in the laboratory had meant that a MERS-CoV construct was already in the pTriEx1.1 vector with the polyhistidine tag correctly positioned at the N terminus and a stop codon preventing a polyhistidine tag addition at the C-terminal domain. The aim of the PCR was to amplify the other strains' DNA in order to be able to perform a double enzyme digest and subsequent ligation, so that all constructs could be positioned within the pTriEx1.1 vector with the polyhistidine tag located at the N terminal domain and not at the C terminal domain.

Bands denoted by the arrows in Figure 9 show clear success for the OC42, NL63-CoV and 229E-CoV samples. The slight smear at the bottom of the gel shown in Figure 9 is the result of a primer overload. The 1:10 dilution of the original primers referred to as "stock concentration" needed further diluting.

Although there was a slight band generated in the IBV sample when using $1:10^2$ DNA dilution, shown in lane 16, the concentration was low, and the preparation did not produce successful results in further use. Therefore, the PCR was repeated but with the stock concentration of primer undergoing a further 1:10 dilution first, with the hope to increase the yield of PCR product, results shown in Figure 10.

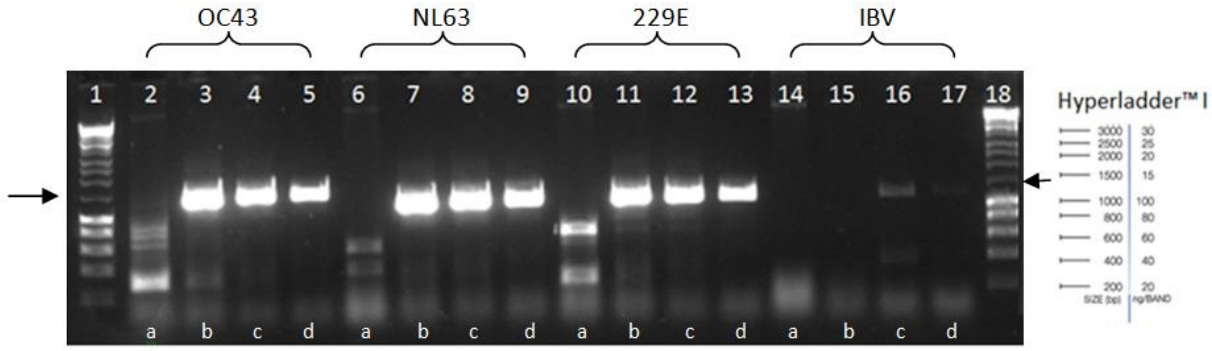


Figure 9: Gel electrophoresis showing PCR products using OC43-CoV, NL63-CoV, 229E-CoV and IBV templates, using a 1 in 100 dilution of stock primers. Lanes 1 and 18 contain Hyperladder I as the molecular weight marker, lane 2-5 OC43-CoV, 6-9 NL63-CoV, 10-13 229E-CoV and 14-17 IBV. Letters as marked on the bottom of the gel relate to the amount of template the PCR was conducted on, a= 1 µL of DNA template, b= 1:10 dilution, c= $1:10^2$ dilution, d= $1:10^3$ dilution. Arrows around the 1,500 bp mark indicate the size of predicted band.

The 1µL of crude DNA template proved to be too concentrated for all samples as denoted by lanes marked with the letter "a", however subsequent dilutions 1:10, $1:10^2$ and $1:10^3$ (marked "b" "c" and "d" respectively on the gel) produced a successful PCR product for all amplicons, except the IBV. The bands for OC43-CoV, NL63-CoV and 229E-CoV were successful and went on to be gel extracted for double enzyme digestion and plasmid ligation. Further work was conducted on the IBV template shown in Figure 10.

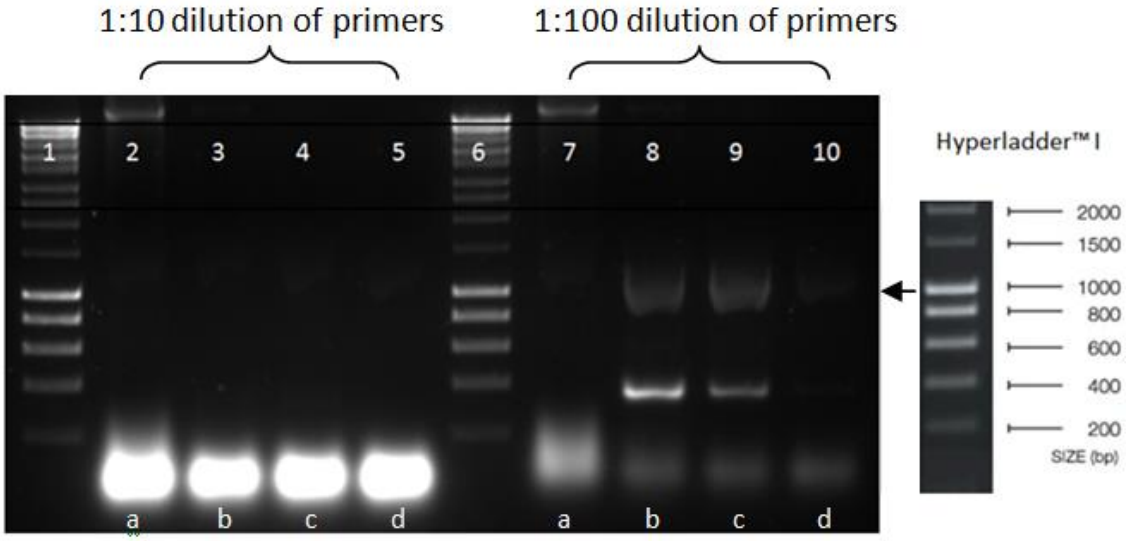


Figure 10: Gel electrophoresis showing results from second PCR conducted using IBV templates in order to try and produce a clear band. Hyperladder I was used in lane 1 and lane 6. 1 in 10 dilution of 1uM primers used in lanes 2-5 and $1:10^2$ dilution used lanes 7-10. Amount of DNA template used indicated by letters at the bottom of the gel, a= 1µL neat, b= 1:10 dilution, c=$1:10^2$ dilution and d=$1:10^3$ dilution.

The samples generated when conducting the PCR using a 1:10 dilution of primer were run in lanes 2-5 of Figure 10, however these proved unsuccessful as no bands were seen. The only clear two bands were seen in lane 8 and 9 produced by using a $1:10^2$ dilution of the primers with a 1:10 and $1:10^2$ dilution of the DNA template. Although the band is faint this was a successful amplification of a band approximating the target size expected and therefore the bands were gel extracted. Once gel extracted, all positive bands for IBV, OC43-CoV, NL63-CoV and 229E-CoV, were double digest with *XhoI* and *NcoI* (as described in section 2.2.1.5). The enzymes were heat inactivated by heating at 80°C for 10 minutes. The pTriEx1.1 vector double digested with the same enzymes, removing the resident MERS-CoV N gene insert and the larger band ~5.3Kbp representing linearized vector was gel extracted. The DNA concentration was measured by a nanodrop spectrophotometer.

The results of the nanodrop were as follows:
OC43-CoV N-gene insert 23.9 ng/µL
NL63-CoV N-gene insert 21.4 ng/µL
229E-CoV N-gene nsert 21.8 ng/µL
IBV N-gene insert 25.8 ng/µL

pTriEx1.1 vector 5.8 ng/µL

Subsequently each of the viruses N encoding DNA sequences were ligated into the pTriEx1.1 vector as described in section 2.2.1.8. Once ligated the samples were transformed into Novablue singles™ competent cells (EMD Millipore) and the plasmids obtained from the resulting colonies were extracted and sent for DNA sequencing. Once the sequence was deemed correct the plasmids were transformed into a variety of *E. coli* expression hosts to allow for protein production.

3.5.2 DNA sequences

See appendix two for sequencing data and basic local alignment search tool (BLAST) results. Two colonies/samples were sequenced per isolate. In all cases sequence A was selected for transformation of BL21 type strains, except in the case of IBV whereby B was used as the sequence obtained better results.

3.5.3 IPTG induction

As mentioned, the MERS-CoV construct already had the polyhistidine tag correctly situated at the N terminal domain and was situated within the pTriEx1.1 vector, therefore this was used as a control to test IPTG induction. This was done as described in section 2.2.1.12 in BL21(DE3) cells, the western blot results are shown in Figure 11.

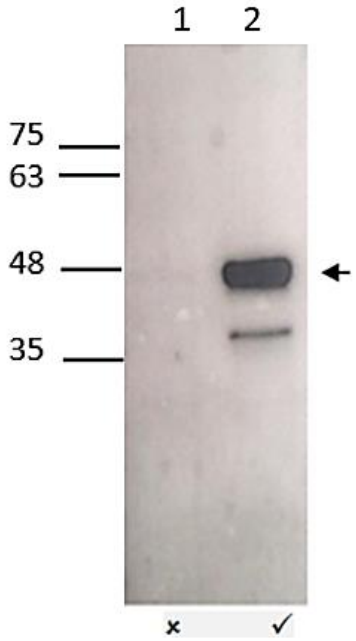


Figure 11: Western blot of a MERS-CoV N protein 3 hour 0.5mM IPTG induction using BL21 DE3 cells. The blot was conducted using an anti his antigen. BLUEeye pre-stained protein ladder was run and the corresponding molecular weights are marked alongside of the blot. Uninduced samples are marked with a cross under the blot in lane 1 and induced samples a tick in lane 2. The arrow indicates the band matching the expected molecular weight.

The band seen in lane 2 in the induced sample, which is not present in the uninduced sample in lane 1, shows the induction of a band of ~48kDa indicated by the arrow, which is near the predicted 45kDa size of the MERS-CoV N product. The smaller band also seen is likely to be minor breakdown of the protein which retains the His tag.

The other four constructed coronavirus N expressing strains underwent the same process of growth and IPTG indication using BL21 cells and the results were as shown in Figure 12.

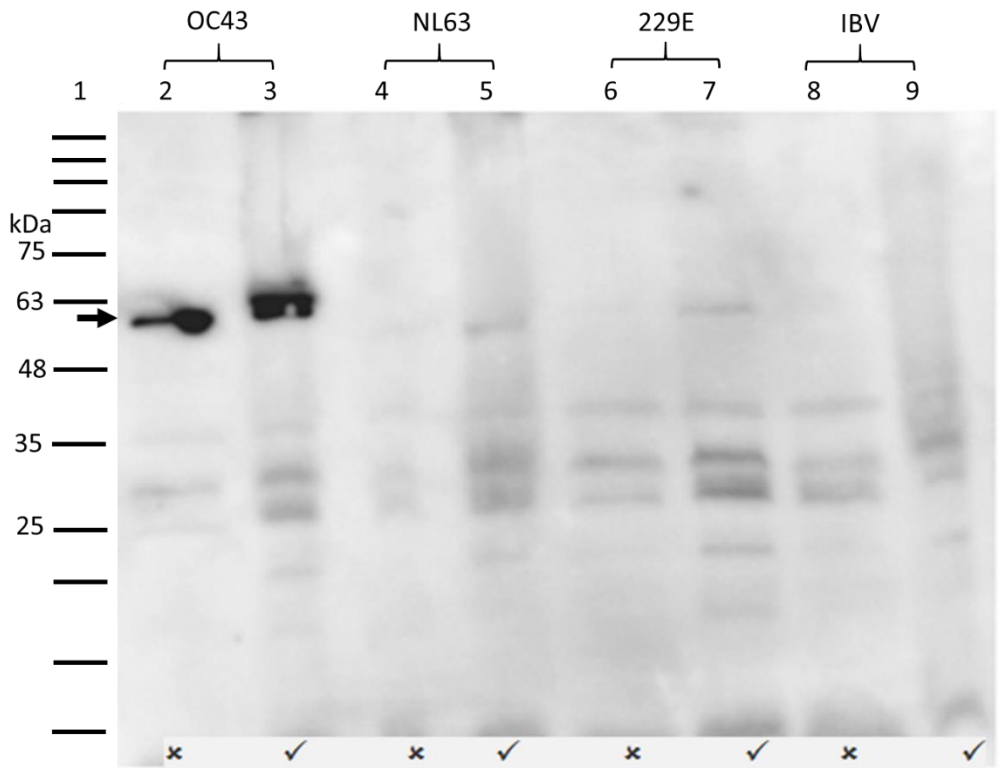


Figure 12: Western blot results using a his tag antibody on *E. coli* lysates expressing OC43-CoV, NL63-CoV, 229E-CoV and IBV N proteins. The samples were generated following 3 hour 0.5mM IPTG inductions conducted using Bl21 DE3 cells. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. Uninduced samples are marked with a cross under the blot and induced samples a tick. 0C43 samples were run in lanes 2 and 3, NL63-CoV in lanes 4 and 5, 229E-CoV in lanes 6 and 7 and IBV in lanes 8 and 9. The arrow indicates the band matching the expected molecular weight.

There is a polyhistidine signal even in the uninduced OC43-CoV sample shown in lane 2 but this is amplified post induction in lane 3 and is an indication of leaky expression. Faint bands can also be seen in lanes 5 and 7 for induced NL63-CoV and 229E-CoV respectively. These bands are around the expected weights of 42-50kDa. Although the OC43-CoV band appears higher as it approaches 63kDa rather than the 49kDa expected. However, this may be due to overloading of sample and/or poor gel running. IBV in lane 9 failed to produce a signal. As only the OC43-CoV sample was able to produce a strong polyhistidine signal when induced, other *E. coli* strains were tested to try and increase yields.

Firstly C41, C43, BL21-star and Rosetta strains were all screened with each construct as they were readily available within the laboratory.

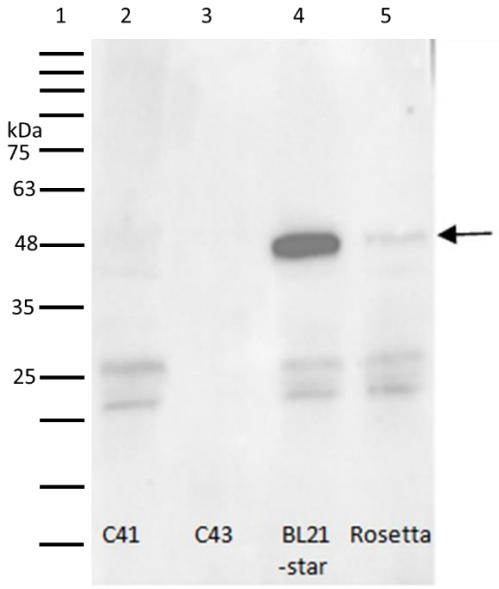


Figure 13: Western blot results using a his tag antibody on *E. coli* lysates expressing NL63-CoV N protein, using a variety of *E. coli* strains. The samples were generated following 3 hours 0.5mM IPTG inductions at 36°C. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. Inductions using C41 and C43 were run in lanes 2 and 3, BL21 cells were used for the induction shown in lane 4 and the Rosetta strain was displayed in lane 5. The arrow indicates the band matching the expected molecular weight.

Figure 13 shows the western blot produced when screening the NL63-CoV construct with C41, C43, BL21-star and Rosetta cell lines. The only cell line not to produce some level of his-expression in the case of NL63-CoV is C43; although C41 shown in lane 2 only appears to show breakdown product as opposed to the full-length protein. The most successful combination was with BL21-Star displayed in lane 4. Rosetta provided a slight signal seen in lane 5. The bands appear around the 48kDa mark which is slightly higher than the 43kDa expected but this may be due to the running of the gel.

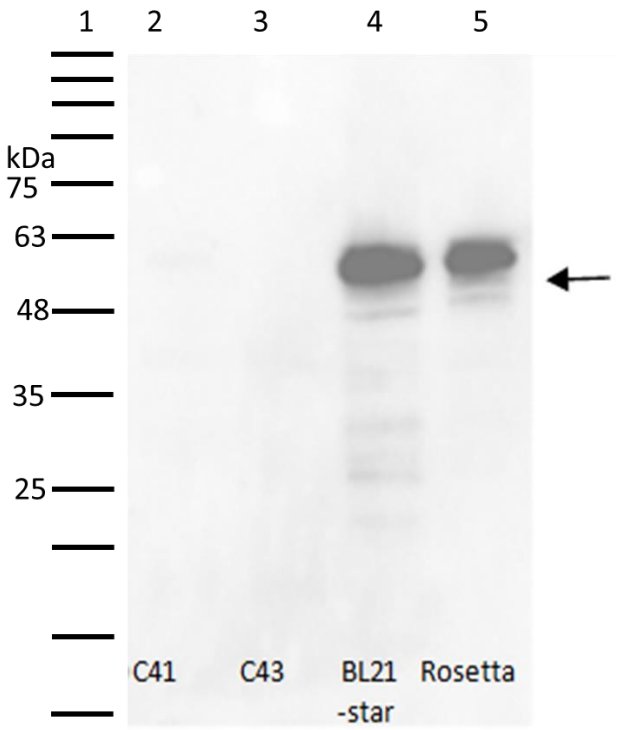


Figure 14: Western blot results using a his tag antibody on *E. coli* lysates expressing 229E-CoV N protein, using a variety of *E. coli* strains. The samples were generated following 3 hours 0.5mM IPTG inductions at 36°C. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. Inductions using C41 and C43 were run in lanes 2 and 3, BL21 cells were used for the induction shown in lane 4 and the Rosetta strain was displayed in lane 5. The arrow indicates the band matching the expected molecular weight.

Figure 14 shows the western blot produced when screening the 229E-CoV construct with C41, C43, BL21-star and Rosetta cell lines. As seen in lanes 2 and 3 of Figure 14, in the case of 229E-CoV neither C41 nor C43 produced a noticeable signal. . BL21-star produced the strongest signal in lane 4 and Rosetta had a good signal in lane 5. The bands appear around the 48kDa mark which is slightly higher than the 43kDa expected but this may be due to the running of the gel.

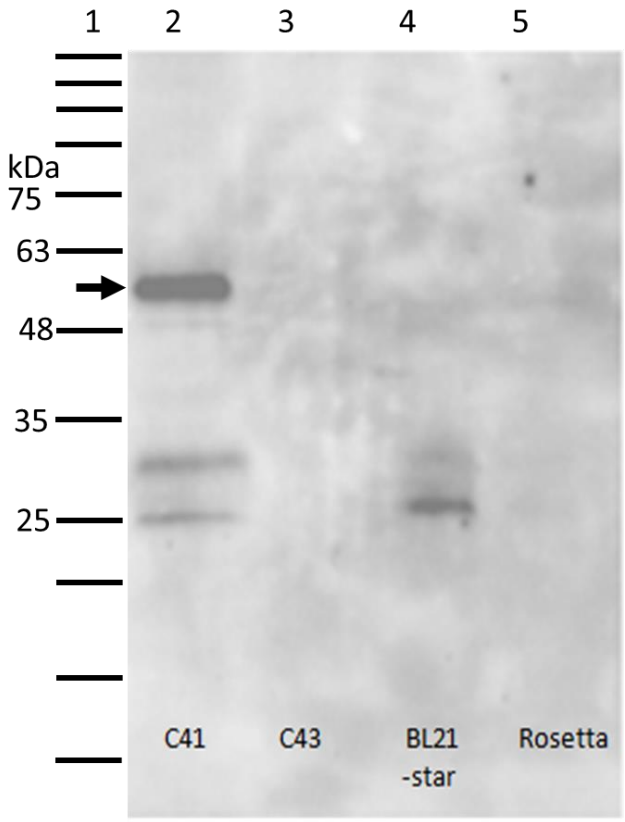


Figure 15: Western blot results using a his tag antibody on *E. coli* lysates expressing IBV N protein, using a variety of *E. coli* strains. The samples were generated following 3 hours 0.5mM IPTG inductions at 36°C. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. Inductions using C41 and C43 were run in lanes 2 and 3, BL21 cells were used for the induction shown in lane 4 and the Rosetta strain was displayed in lane 5. The arrow indicates the band matching the expected molecular weight.

Figure 15 shows the western blot produced when screening the IBV construct with C41, C43, BL21-star and Rosetta cell lines. The IBV construct showed a successful induction using C41 shown in lane two of Figure 15. The band appears around the 48kDa marker which is near the 45kDa expected. Lane 3 and 5, C43 and Rosetta respectively, show no signal. And only breakdown bands can be seen in lane 4 in the case of BL21-star.

It is worth noting that the sucessful production of IBV N protein using the C41 strain prove unreliable in subsequent inductions, as such chapter 4 examines IBV N protein expression in a eukaryotic system.

As such, the combination of strain and cell lines were as follows; OC43-CoV-BL21, NL63-CoV-BL21-Star, 229E-CoV-BL21-Star and IBV-C41.

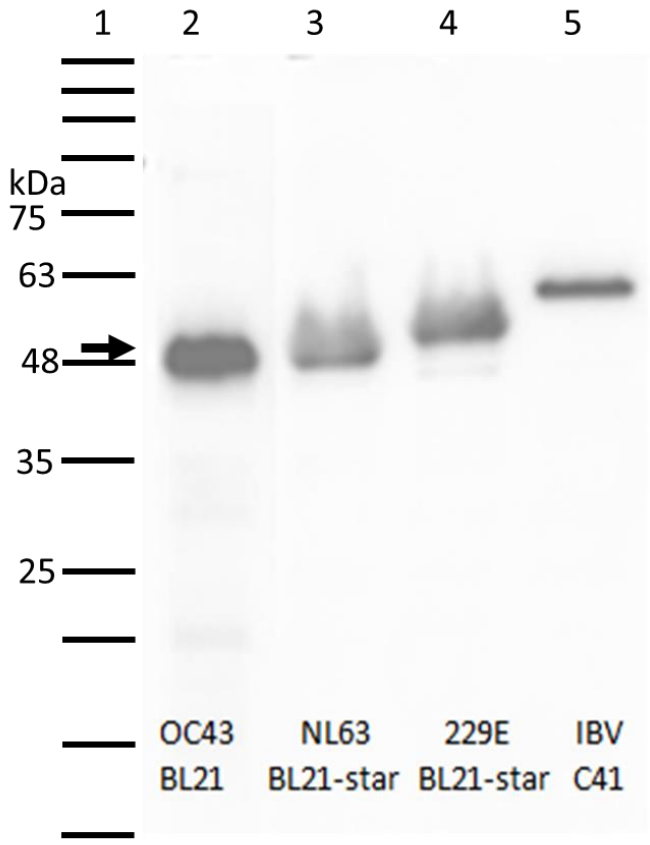


Figure 16: Western blot results using a his tag antibody on E. coli lysates expressing OC43-CoV, NL63-CoV, 229E-CoV and IBV N proteins, using a variety of E. coli strains, as written in the lanes at the bottom of the blot. The samples were generated following 3 hours 0.5mM IPTG inductions at 36°C. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. The arrow indicates the bands matching the expected molecular weight.

Figure 16 shows all combinations to be successful and capable of producing a strong band in the expected range 42-50kDa. OC43-CoV with BL21 cells is displayed in lane 2, BL21-star cells were used with NL63-CoV and 229E-CoV in lanes 3 and 4 respectively and IBV in C41 cells is shown in lane 5. The IBV signal in lane five appears above the expected 45kDa but this may be due to the running of the gel or ladder. The slight variation in size is minimal and is potentially because of the viral strains used or simply the running of the gel.

3.5.4 Column chromatography

After IPTG induction of the MERS-CoV-His tagged sample in BL21 DE3 cell line, lysis and IMAC was conducted as described in section 2.2.2.2, Figure 17 shows the western blot analyses of the results.

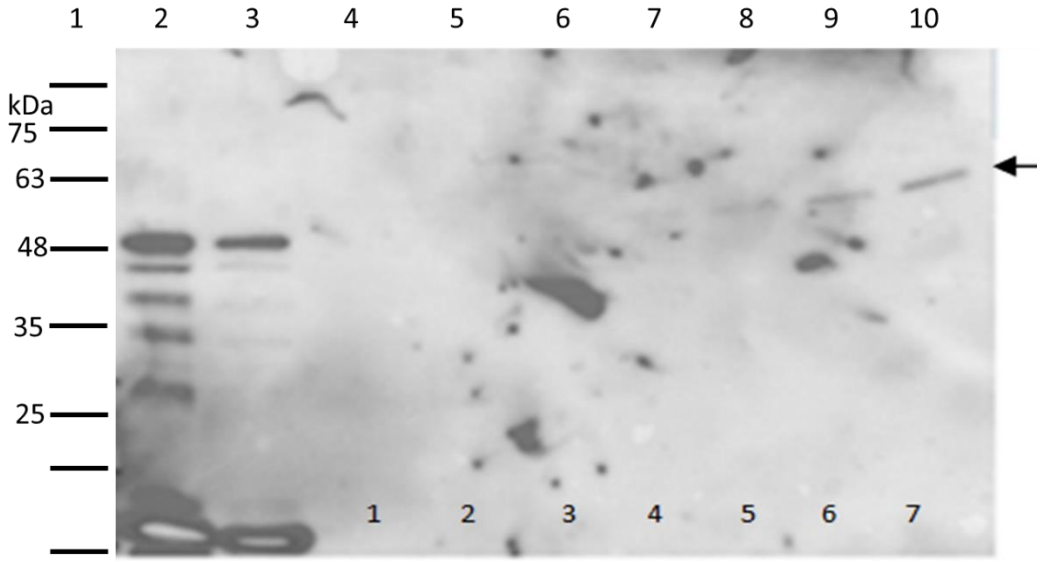


Figure 17: Western blot using his antibody, showing the results of IMAC on an IPTG induced MERS-CoV sample. Run conditions were 0-100 % elution buffer over 60 minutes with a flow rate of 2.5mL/min. Samples were collected every 2 minutes. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. The initial lysis load is in lane 2, the flow through is shown in lane 3. Tubes 1-7 are displayed in lanes 4-10 as marked at the bottom of the gel. The arrow indicates the bands matching the expected molecular weight.

A polyhistidine signal can be seen in lanes 8-10, indicating tubes 5, 6 and 7 contain the his-tagged protein. The three positive samples in tubes 5-7 (lanes 8, 9 and 10) were then pooled. The 15mLs was concentrated down to 500µL using a vivaspin column with a MCOW of 30kDa. The OD280 reading was 0.310 equating to approximately 1.55mg/mL. This was stored at -20^{o}C.

3.5.5 Large scale inductions

In order to produce a high yield of N-protein five large scale infections were set up (OC43-CoV in BL21 cells, NL63-CoV in BL21star cells and 229E-CoV also in BL21star cells) albeit with varying degrees of breakdown product, as shown in Figure 18. The signal from MERS-CoV is present but at a low level. IBV, however, has failed to produce a signal even in the C41 cell line it had previously been successful in (see Figure 15). Therefore, the pellets from OC43-CoV, NL63-CoV and 229E-CoV were lysed and purified and further investigation was conducted on IBV, see chapter 5.

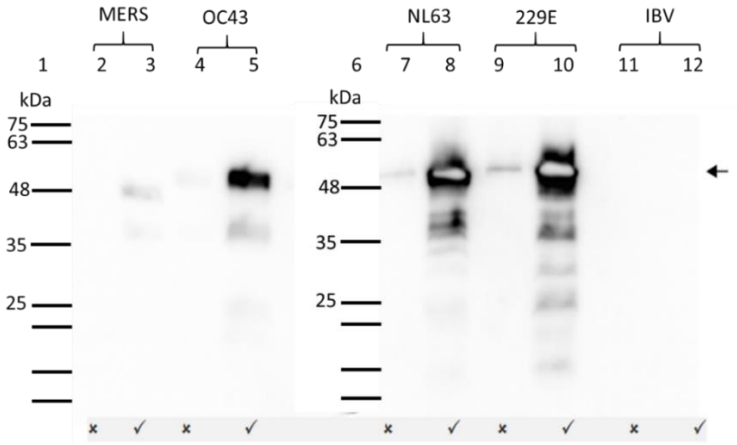


Figure 18: Western blot using his antibody on *E. coli* lysates of MERS-CoV, OC43-CoV, NL63-CoV, 229E-CoV and IBV N protein production. 1 litre IPTG induction were conducted on all five strains using 0.5mM IPTG inductions for 3 hours in 3L flasks at 36°C. BLUEeye pre-stained protein ladder was run in lane 1 and 6; the corresponding molecular weights are marked alongside of the blot. Uninduced samples are marked with a cross under the blot and induced samples a tick. MERS-CoV samples were run in lanes 2 and 3, OC43-CoV in lanes 4 and 5, NL63-CoV in lanes 7 and 8, 229E-CoV in lanes 9 and 10 and IBV in lanes 11 and 12. The arrow indicates the bands matching the expected molecular weight.

Figure 18 shows the western blot results from large scale inductions of all constructs and their corresponding expression strain. All constructs were able to produce signal around the expected molecular weight when induced (42- 50kDa) except for IBV as there is no band visible in lane 12. IBV induction was repeated on multiple occasions but the ability to produce signal at the expected molecular weight was unreliable and inconsistent, as a result further work was conducted in order to see if IBV yield could be increased and stability increased, this included the possibility of eukaryotic expression systems detailed in chapter 4 as well as structural manipulations described in chapter 5.

3.5.6 IMAC

After the majority of the proteins were successfully expressed the proteins went on to be purified and concentrated. Then pellets from the one litre preparations mentioned in section 3.5.5 were stored at -20°C until they were ready to be lysed and purified using the his trap column, protocol detailed in section 2.2.2.2. Once the sample was loaded onto the column a gradient of 0-100% elution buffer was run, samples were collected every 2minutes. Two peaks were seen on the graph. Western blot analysis was conducted on four samples from the first peak and two from the second in order to see which, if any, contained the polyhistidine tagged MERS-CoV N protein, the results of which are shown in Figure 19.

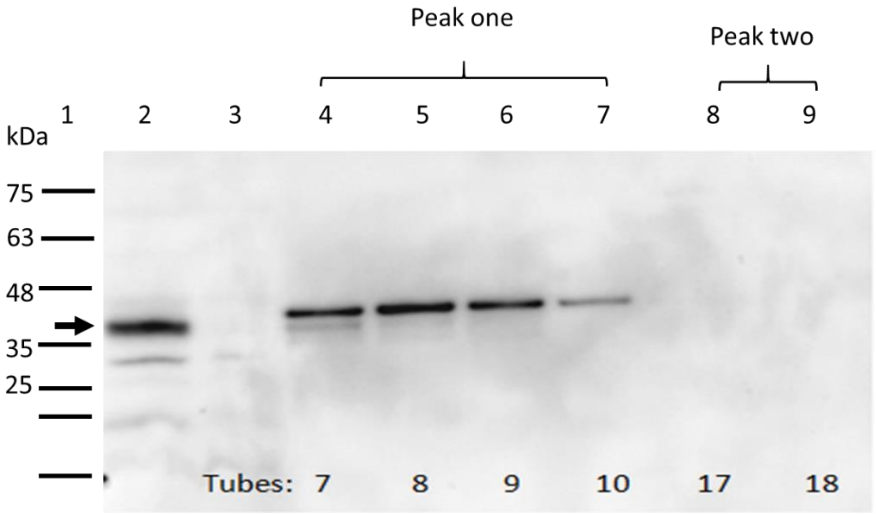


Figure 19: Western blot using his antibody, showing the results of IMAC on a 1L IPTG induced MERS-CoV sample. Run conditions were 0-100 % elution buffer over 60 minutes with a flow rate of 2.5 mL/min. Samples were collected every 2 minutes. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. The initial lysis load is in lane 2, the flow through is shown in lane 3. Tubes 7-10 are displayed in lanes 4-7 samples from the second peak, tube 17 and 18, are shown in lanes 8 and 9. The arrow indicates the bands matching the expected molecular weight.

Although the column run appears to be successful, in that the polyhistidine-tagged protein was able to successfully bind and be eluted in a defined peak, when comparing this signal's molecular weight on the blot (~30kDa) to the expected weight of ~46kDa it is clear that something has not gone as anticipated. The load sample shares this problem, but the original induction seen in Figure 18 did not. This implies that the cleavage occurred either during the freezing of the sample, or whilst it was being lysed.

3.5.7 Molecular weight

To see if the cleavage of the protein occurred before or after the pellet lysis, the remaining pellets from the 1litre infections were gently thawed and run on a gel to examine their molecular weight.

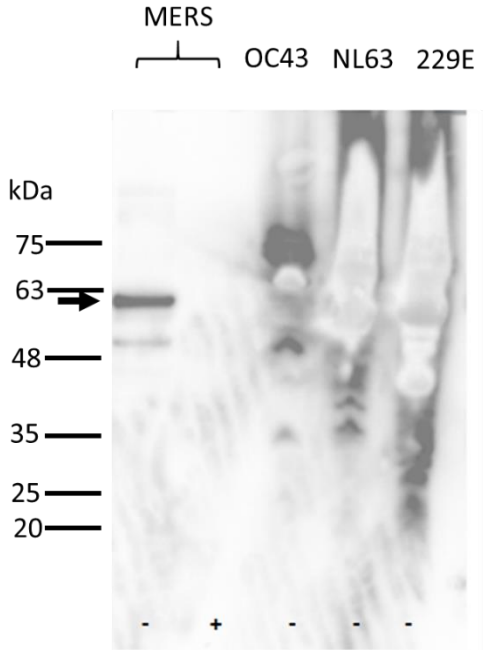


Figure 20: Western blot using his antibody, showing the molecular weight comparison pre and post -20°C storage of 4 types of N proteins. A "-"symbol at the bottom of the blot shows the sample is un-lysed whilst a "+" symbol indicates that lysis has occurred. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot.

Figure 20 is a western blot to show the molecular weights of the N proteins produced from the other 1L inductions that had not yet been lysed or purified. The signal has been all but lost in lane 3 probably due to the fact the lysed sample is far more dilute than the original frozen pellets and this was not adjusted for when the blot was run. More importantly it is clear to see all four samples maintained a higher molecular weight (~60kDa) even post -20°C storage. This would imply that the thawing process and/or the lysing process is either too harsh or allows for proteolytic cleaving of the protein. Subsequent lysis used protease inhibitor tablets (complete protease inhibitor cocktail tablets (Roche)).

3.5.8 All five proteins post IPTV induction, lysis, HisTrap™ column

With the new addition of protease inhibitors to the method to try and maintain the correct molecular weight of the proteins 200mLs of all samples in their optimized *E. Coli* strain were induced and lysed without being stored at -20°C. They were purified through the IMAC HisTrap column and concentrated using the Vivaspin columns (Epsom) before being run on a western blot as shown in Figure 21. Care was taken to keep all proteins chilled and the new methods and care proved beneficial in keeping the proteins intact.

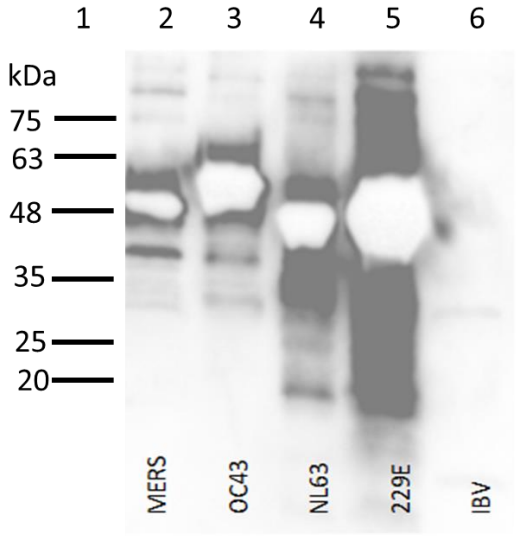


Figure 21: Western blot using his antibody conducted on all 5 N-proteins post purification and concentration. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot.

Although the IBV band is hardly visible in lane 6 it is still present. The concentrations of the proteins were all measured using the nanodrop in order to be standardised for future ELISA tests. MERS-CoV in lane 2 had a concentration of 1.55 mg/mL, OC43-CoV shown in lane 3 had a concentration of 1.19mg/mL, NL63-CoV shown in lane 4 had a concentration of 5.60mg/mL, 229E-CoV in lane 5 had a concentration of 4.49mg/mL and although a band remains faint and overshadowed by the 229E-CoV band next to it, IBV had a band in lane 6 with a concentration of 0.26mg/mL. Molecule weight was hard to assess due to the strength of signal however they all appear around the 48kDa marker.

3.5.9 Summary

This chapter shows the successful expression and purification of most constructs using the prokaryotic expression system. The exception being that of the IBV construct, which was hoped to be used as a negative control when it came to sera testing using ELISAs. Due to the fact that protein expression of the IBV construct at a consistent yield and the correct molecular sized proved inconsistent using prokaryotic expression, eukaryotic expression was also investigated shown in chapter 4. The prokaryotic system proved beneficial in terms of effectiveness for the remaining H-CoVs however.

Chapter 4: Eukaryotic N protein expression and purification

4.1 Baculovirus expression system introduction

4.1.1 Introduction to the baculovirus expression vector system (BEVS)

Having struggled to consistently produce a stable IBV N protein using the prokaryotic methods described in chapter 3 this chapter focuses instead on recombinant protein production in eukaryotic cells. Baculoviruses are viruses that infect insects and other athropods, particularly insects of the order Lepidoptera. Baculoviruses have double-stranded, circular, supercoiled DNA contained within a rod-shaped capsid. Two of the most common isolates used in gene expression are *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV) and Bombyx mori nuclear polyhedrosis virus (BmNPV) (LifeTechnologies 2011). AcMNPV was first used as an expression vector for human beta interferon in 1983 (Smith, Summers et al. 1983). Baculoviruses can be modified so that a redundant highly expressed, very late gene is replaced by a gene of interest to produce a recombinant virus that goes on to express the recombinant protein as part of its life cycle. The polyhedrin and p10 genes are both very late and have been successfully used for this purpose, shown in Figure 22. Because insect cells are eukaryotic they are able to carry out most of the post-translational modifications found in mammalian cell lines (Jones and Morikawa 1996) giving the system an advantage over prokaryotic cell lines in that respect. Linearised viral DNA is used in the system as it cannot initiate a viral infection unless rescued by the recombination event, as the virus DNA transfected into cells with a constructed transfer vector. This limits the production of background wild type virus alongside the recombinant and further enhances the efficiency of baculovirus expression (Kitts and Possee 1993, Zhao, Chapman et al. 2003). The DNA used today ensures essentially 100% recombination and a plaque assay to isolate recombinant viruses from the wild type background is no longer necessary.

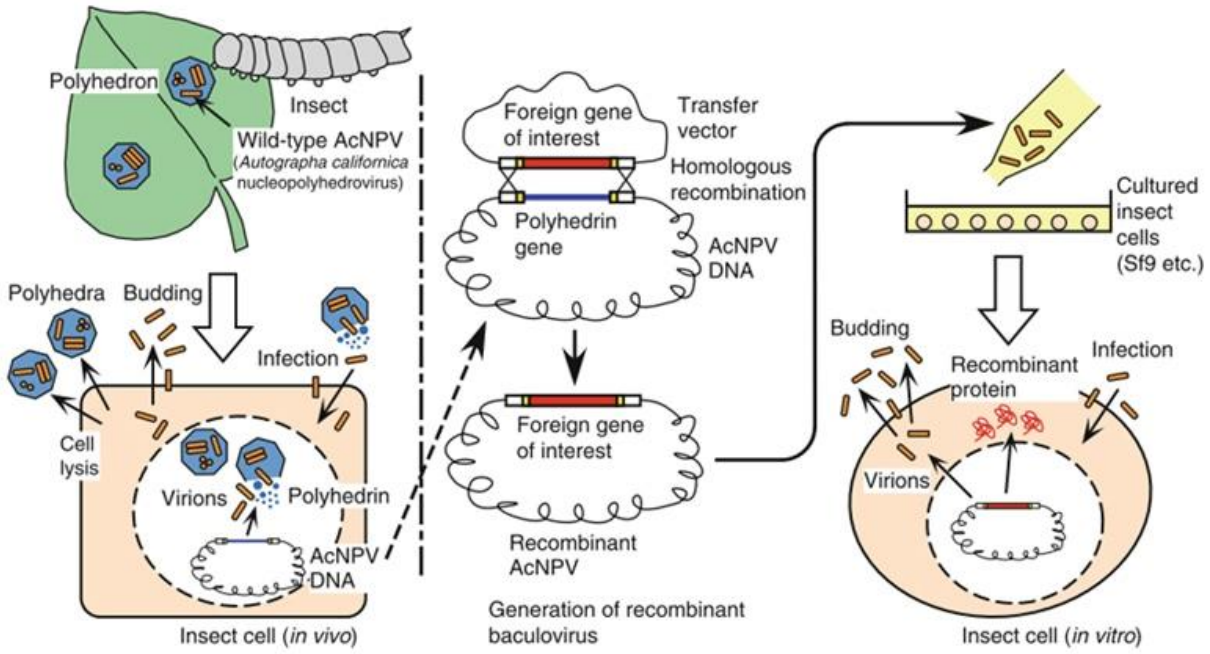


Figure 22: A pictorial overview of the BEV system (Yamaji 2011)

Advancements to the system came in 1993 whereby an alternative method for producing recombinant baculoviruses was developed using site-specific transposition in *E. coli* (Luckow, Lee et al. 1993). A baculovirus genome was constructed to contain a mini F replicon to allow replication in *E. coli* as a bacterial artificial chromosome (bacmid), an attTn7 target site for transposition and a kanamycin resistance gene. Transformation of these bacmid-containing cells with a transfer plasmid containing regions of the Tn7 transposon that flank a baculovirus promoter and gene of interest, in addition to a helper plasmid that provides the transposon functions, results in transposition of the gene and promoter into the attTn7 locus sited in the baculovirus genome. The addition of a kanamycin resistance gene meant this event can be isolated using antibiotic selection. Recombinant Bacmid DNA, with the gene of interest, can then prepared and transfected into insect cells where it initiates virus replication and growth. The technology has gained widespread usage through commercialisation as the Bac-to-Bac system (Invitrogen), however it is worth noting that instability has been noted in the final virus stocks with deletions mapped to the mini F replicon region still contained in the final recombinant baculovirus genome (Pijlman, van Schijndel et al. 2003).

4.1.2 Advantages of baculovirus expression

Using BEVS has many advantages as a eukaryotic expression system over prokaryotic systems or other eukaryotic methods such as mammalian cell lines. Safety is a major advantage as baculoviruses have a restrictive host range and are non-pathogenic to mammals and plants. As previously mentioned the fact that the baculovirus propagates in eukaryotic insect cells means that post-translational modifications can be made to the required proteins in a manner similar to that achieved using mammalian cells, with the exception of glycosylation which is restricted to high mannose types in insect cells (Shi and Jarvis 2007). AcNPV is used in cell lines derived from the fall armyworm *Spodoptera frugiperda* (Sf9) or from the cabbage looper *Trichoplusia ni* (TiNi). Both cell lines grow well in suspension cultures and so the system is easy to scale-up to produce high levels of recombinant gene expression. Sf9 cells were used for this study.

4.1.3 Disadvantages of baculovirus expression

Although eukaryotic cells are used, not all insect cell lines used in BEVS are able to produce glycoproteins with complex-type N-glycosylation patterns although Sf9 cells still can (Steele, Stone et al. 2017). Also, when infected with baculovirus cells lyse and die which limits protein production. To combat this expression plasmids have been transformed into cell lines coding for anti-apoptotic proteins to try and delay the cell death and lysis that baculovirus infection causes.

4.1.4 AcMNPV

AcMNPV is the most intensively studied and commonly used protein expression vector baculovirus (Chen, Zhong et al. 2013). AcMNPV has a double-stranded DNA genome that is 133.9kbp and contains at least 156 open reading frames (ORFs) (Chen, Zhong et al. 2013).

4.1.5 The flashBAC™ System

The flashBAC™ System (Oxford expression technologies) used within this study, uses the partial-deletion of an essential gene encoding ORF1629 from the AcMNPV genome in order to prevent non-recombinant virus from replicating within insect cells (Chen, Zhong et al. 2013). The complementing section of the gene is also present on the transfer vectors allowing reconstruction

of a viable copy by the recombination event. This eliminates the need to plaque-purify recombinant virus from parental virus. The chitinase gene (chiA) is also deleted which results in the production of a higher yield of secreted or membrane targeted recombinant proteins (Lu, Chapple et al. 2002). Additional non-essential genes are also deleted in flashBAC™ GOLD to enhance both the quality and the yield of proteins expressed. Insect cells are transfected with lipofectin, flashBAC™ DNA, and the transfer plasmid containing the gene of interest. Homologous recombination takes place within the insect cells and inserts the gene of interest under the control of the polyhedrin promoter. It also restores the function of ORF1629 allowing viral DNA to replicate and produce virus particles. The baculovirus produced by the replication of the recombinant virus genome can be harvested directly from the culture medium of transfected insect cells, forming a seed stock of recombinant virus. The recombinant virus is then passaged to a high enough titre whereby it can be used to produce a synchronous infection of a large culture to produce the protein of interest.

4.1.6 Similar studies

4.1.6.1 Baculovirus-expressed N protein in human sera studies

Similar to the proposed screening described in this thesis, a 2008 U.S. study has used ELISAs to screen human sera to ascertain a metropolitan population's exposure rates to several HCoVs (Severance, Bossis et al. 2008). The study used BEVS to produce the recombinant nucleocapsid proteins. The study looked at 229E-CoV, KHU1-CoV, NL63-CoV, and OC43-CoV and used feline coronavirus as a control. Instead of using a poly-histidine tag for purification, the study added a bovine polyomavirus (BVP) large T antigen peptide tag to the C terminus and fused the recombinant N proteins with glutathione S-transferase (GST) at the N terminus. This allowed verification that the full-length proteins, GST positive and BVP peptide positive by western blot, were being expressed. The recombinant CoV N proteins were expressed in *Trichoplusia ni* (High Five) insect cells and released by cell lysis. Interestingly the fusion proteins were not purified but used as crude lysates with a layer of casein-glutathione providing a capture layer for each antigen before being exposed to human sera (Sehr, Zumbach et al. 2001). Whether the capture of processed GST only or the fixed orientation of the fusion protein lessened sensitivity or sterically restricted antibody binding was not investigated. Using this format, sera were screened from 10 children aged between 2 and 4, and 196 adults between the ages of 18 and 65. The proportion of seropositive adults for each coronavirus were as follows: 229E-CoV, 91.3%; KHU1-CoV, 59.2%; NL63-CoV, 91.8%; and OC43-CoV, 90.8%. There was no evidence of a significant serological response to the feline coronavirus control. Significant associations of coronavirus seropositivity and antibody levels were tested with chi-square and regression analyses. The betacoronaviruses OC43-CoV and KHU1-CoV were significantly associated with race ($P \leq 0.009$ and $P \leq 0.03$, respectively). OC43-CoV levels were further elevated with an association to smoking status ($P \leq 0.03$). High NL63-CoV titres were seen to be associated with socioeconomic status ($P \leq 0.04$). In all four of the coronaviruses high-level immunoreactivity was significantly associated with the summer season ($P \leq 0.01$ to 0.0001). The study showed a generally high level of seroprevalence (90.8 to 91.8%) in 196 adults for three of the four human coronaviruses (229E-CoV, NL63-CoV,

and OC43-CoV) and a moderate exposure rate for the fourth (59.2% for KHU1-CoV). The results indicate that these viruses have an established presence within the population studied with race, socioeconomic and smoking status as risk factors for coronavirus exposure. Higher rates of seropositivity were observed in African Americans, smokers, and individuals of low socioeconomic status. All three of these risk factors have been previously shown to make individuals more prone to a variety of other respiratory illnesses (Nuorti, Butler et al. 2000, Chatila, Hoffman et al. 2006, Hegewald and Crapo 2007).

4.1.6.2 Baculovirus-expressed N protein in animal sera studies

BEVS expressed N proteins have also been used to screen animal sera. A 2001 study looked at turkey sera to check for prevalence of turkey coronavirus (TCV) using ELISAs (Breslin, Smith et al. 2001). The ELISAs detected antibodies specific for TCV and IBV, a closely related avian coronavirus, but did not detect antibodies specific for other avian viruses (such as avian influenza, avian reovirus, avian paramyxovirus 3, avian adenovirus 1, or Newcastle disease virus). The study deemed that baculovirus-expressed TCV N protein was a suitable source of antigen for ELISA-based detection of TCV-specific antibodies in turkeys. In a different study, the sensitivity and specificity of a competitive ELISA for the detection of TCV antibodies were determined by comparison with the indirect fluorescent antibody test (IFAT) (Guy, Smith et al. 2002). Again, the ELISA detected antibodies specific for TCV and the closely related IBV. Sensitivity and specificity of the ELISA relative to IFAT were 92.9% and 96.2%, respectively, which was deemed high. The study also concluded that the ELISA was a rapid, sensitive, and specific serologic test for detection of TCV antibodies in turkeys.

4.1.6.3 Other uses of baculovirus-expressed CoV proteins

4.1.6.3.1 SARS-CoV S protein vaccine

Other CoV proteins have also been expressed using BEVS, namely the spike protein. Recombinant proteins are useful not only for diagnostics in terms of sera screening and ELISAs, but also for

potential vaccines. One such BEVS S protein has been engineered as a potential vaccine against SARS-CoV when administered with an adjuvant (McPherson, Chubet et al. 2016).

4.1.6.3.2 MERS-CoV virus-like particles (VLP)

MERS-CoV VPLs made following co-expression of MERS-CoV M, E and S proteins have been expressed in insect cells and shown to induce specific humoral and cellular immunity in rhesus macaques (Wang, Zheng et al. 2017). Electron microscopy and immunoelectron microscopy showed that the MERS-CoV VLPs were structurally similar to the native virus. When Rhesus macaques were inoculated with MERS-CoV VLP and alum adjuvant, they produced virus-neutralizing antibodies titres, specific IgG antibodies against the receptor binding domain (RBD) and a T-helper 1 cell (Th1)-mediated immunity. The data showed that MERS-CoV VLP have excellent immunogenicity in rhesus macaques and could be a promising vaccine candidate. The structural equivalence of the recombinant VLPs and virus particles as well as their performance as candidate vaccines suggested the insect cell background allowed authentic protein folding of all components and that any individual protein expressed similarly should be an excellent diagnostics antigen.

4.1.6.3.3 IBV VLP

VLP have also been studied in regard to their use with animal CoVs. Baculovirus-derived IBV VLPs have been produced with a membrane, envelope and the recombinant spike proteins and had their humoral immune responses measured in chickens (Xu, Wu et al. 2016). The IBV VLP were generated through the co-infection with three recombinant baculoviruses separately encoding the recombinant M, E or S genes. After subcutaneous injections into chickens the results of immunogenicity tests demonstrated that the efficiency of the VLP was comparable to that of the inactivated M41 virus vaccine in eliciting IBV-specific antibodies and neutralizing antibodies.

4.2 Baculovirus expression materials and methods

4.2.1 Materials and growth conditions

There are a variety of growth conditions to consider when using the BEVS:

4.2.1.1 Foetal bovine serum (FBS)

FBS is a growth supplement used in insect cell culture medium, even in cases where serum free media is used. FBS provides cells with growth-promoting factors such as amino acids, peptides, and vitamins. It is used to promote growth, provides shear force protection, protect against proteolytic degradation and environmental toxicities, and it can contribute to cellular attachment. There are negatives to its use; it may cause excessive foaming in bioreactors, its addition it may introduce contamination, its use increases cost and complexity of downstream processing, there may be fluctuations in price, quality, and availability, and there is a chance that FBS may cause suboptimal cell growth, toxicity and result in a decreased product yield. In the case of baculovirus growth it has been found that the presence of serum increases the viability of virus on long term storage (Maranga, Coroadinha et al. 2002).

4.2.1.2 Serum free media (SFM)

SFM may be favoured as opposed to FBS for the negative reasons listed previously. SFM contain optimized concentrations of amino acids, carbohydrates, vitamins, and lipids. The media can reduce or eliminate the effect of rate-limiting nutritional restrictions or deficiencies. Many commercially available medias have been optimized to support faster insect cell population doubling times and high saturation cell densities. Optimized SFM formulations offer numerous advantages over using FBS; they eliminate the cost of FBS, they can increase cell and product yields, using them can improve sterility by reducing the risk of contamination and they provide lot-to-lot consistency. For these reasons and the negative reason listed for the use of FBS, SFM Insect-XPRESS™ (Lonza) was used throughout the studies described here. However, to maintain long-term virus stocks FBS was added to 2% final volume. With this addition and the insect-

XPRESS™ used Sf9 cell densities in excess of 4×10^6 cells/mL were achieved. It also contains L-glutamine and supports superior production of recombinant proteins when using the BEVS.

4.2.2 Environmental Factors

As well as medium and serums, there are other environmental factors to be considered when cultivating insect cells including, temperature, pH, osmolarity, aeration and shear forces.

4.2.2.1 Temperature

25°C to 30°C is the preferred range for growth and infection of most cultured insect cells, with the optimal range 27°C to 28°C. (Shao-Hua, Hong-Liang et al. 1998).

4.2.2.2 pH

pH during insect cell cultivation affects both cellular proliferation and production of both viral and recombinant proteins. A slightly acidic pH range of 6.2 to 6.9 is advised for most lepidopteran cell lines and is controlled by the use of phosphate buffers in commercial media (Zitzmann, Sprick et al. 2017).

4.2.2.3 Osmolality

360 to 375 milliosmoles per kilogram (mOsm/kg) is the optimal osmolality of medium for use with most insect cell lines (Manual 2011) and is provided by the use of commercial media.

4.2.2.4 Aeration

When growing is suspension insect cells need adequate transfer of dissolved oxygen by either passive or active methods for optimal cell proliferation and recombinant protein expression. On a larger scale, not used herein, bioreactor systems that use active or controlled oxygenation systems require dissolved oxygen at 10% to 50% of air saturation (Manual 2011).

4.2.2.5 Shear Forces

Suspension culture techniques generate mechanical shear forces. The total shear stress is contributed to by a variety of factors including the mechanism used to create motion, the size and

velocity of any bubbles and the turbulent action at the culture surface. For large suspension batch cultures of insect cells, shear force protection is advised. Serum concentrations between 5% and 20% are recommended to be used as specialised shear force protectant products (Lynn 2007).

4.2.2.6 Summary of growth conditions

For this study the insect cells were from the caterpillar *Spodoptera frugiperda* (SF9 cells) (Invitrogen). Sf9 cells are a clonal isolate of the *Spodoptera frugiperda* cell line IPLB-Sf21-AE, they were originally established from ovarian tissue of the fall armyworm (Vaughn, Goodwin et al. 1977). They were grown in Insect-XPRESS™ media (Lonza) and incubated at 28^{o}C without the addition of antibiotics. Cells were maintained in suspension until needed, at a density of 3×10^{5} to 3×10^{6} cells/mL and incubated in small batch cultures in a rotary shaker at 99rpm. These were used to allow infection with, and amplification of, recombinant baculovirus stocks and were also used for protein expression.

4.2.3 Transfection using BEVS

Sf9 cells were transfected in order to generate new recombinant baculovirus. Cells at a 50% confluence level were used to produce a monolayer in a 6-well dish (~$1x10^{6}$ cells/well). Lipofectin reagent (Invitrogen) was used to transport the mixture of flashBACGOLD (Oxford Expression Technologies) and transfer vector into the cells. Reactions were made up to a total volume of 24 µL using purified water. Each reaction used 1-2µL of flashBACGOLD, 8µL of Lipofectin and 100-500ng of the transfer vector. The lipid-DNA complex was incubated at RT for 30 minutes. Cells were left to attach to the plastic during this time, before being washed with serum-free media prior to the addition of the 24µL reaction mix. The cells were incubated for ~5 days, or until visible signs of infection. The supernatant recovered from centrifuging harvested cells (13,000rpm for 15 minutes at 4^{o}C) was used to infect fresh cells for subsequent passages to increase the viral titre.

4.2.4 Protease inhibitors

One problem with the BEVS when expressing recombinant protein is sometimes the level of proteolytic degradation. Although a proteasome has not been noted within insect cells,

ubiquitination of proteins produced within them has been seen (Low, Doherty et al. 1995). The baculovirus AcMNPV encodes a protein that has 76% identity with the eukaryotic protein ubiquitin called v-ubi (Guarino 1990). A frame-shift mutational study showed that the viral ubiquitin was a nonessential protein but reported however that it may confer a slight growth advantage (Reilly and Guarino 1996). Protease inhibitors are often used within protein production to prevent degradation. The use of the proteases inhibitors was studied on the insect cell line BTI-TN-5B1-4 (High Five™) and Sf9 cells (Martensen and Justesen 2001). The study observed that recombinant proteins migrated in SDS-PAGE in agreement with poly-ubiquitinated forms of the protein, suggesting a ubiquitin/proteasome degradation pathway. The study looked at the effect of adding proteasome inhibitors to both the growth medium of recombinant baculovirus-infected High Five insect cells and to the lysis buffer, to establish the most efficient way to inhibit proteolytic activity. It was shown that the use of cocktail protease inhibitors such as Complete™ (Roche Molecular Biochemicals, Hvidovre, Denmark) was insufficient, as the yield of protein produced in High Five cells was still impaired with 20%–50% of the protein being degraded. The study screened multiple protease inhibitors: LLL (Z-Leu-Leu-Leu-CHO) also known as MG132; ALL (N-Acetyl-Leu-Leu-Met-CHO); Lactocystin (clasto-Lactocystin β-lactone); Leupeptin; E-64 (N-[N-(L-3-trans-carboxyoxirane-2-carbonyl)-L-leucyl]-agmatin); Chloroquine; TPCK (tosyl-phenylalanin chloromethyl keton); Pepstatin; Pefabloc SC (4-(2-Aminoethyl)-benzenesulfonyl-fluoride, HCl (AEBSF)); Complete (cocktail of inhibitors); P8340 (inhibitors for mammalian cell extracts) and P8849 (inhibitors for poly-histidine tagged proteins). As well as testing both High Five and Sf9 cells, two different recombinant baculoviruses were used each producing a different non-secreted protein (p46 and human tryptophanyl-tRNA synthetase (hWRS)).

The study concluded by recommending the addition of 20μg/mL E-64 to the growth medium 24 hours after infection, together with 100μM Leupeptin in the lysis buffer when using High Five cells in order to avoid extensive degradation of non-secreted recombinant proteins. This differed to Sf9 cells, whereby treatment with the proteasome inhibitor MG132 was recommended. It is a specific

inhibitor of proteasome activity in eukaryotic cells (Lee and Goldberg 1998). It was shown that although the addition of MG132 did lead to a reduction to protein degradation, higher concentrations lead to a considerable reduction to recombinant protein yield. Without the addition of MG132 52% of the total p46 expressed was degraded, with 25μM this marginally reduced to 42%, however expression of the full-length protein reduced to 93%. At 50μM there was a marked improvement to degradation as only 3% of the protein was degraded, however overall expression was limited to 60%. The final concentration looked at was 125μM in which no protein was expressed. For this reason, when screening with MG132 during this study the concentration used will be 50μM in order to have the lowest amount of degradation without significantly jeopardising yield.

4.2.5 Protein production

A six well dish infection was conducted over five days to ascertain whether it was possible to produce an IBV-his signal using the BEVS in Sf9 cells. A recombinant baculovirus expressing GFP was used for a control infection. A "Cells only" sample was used as a negative control. A purified his-tagged human scavenger receptor class B protein (SCARB2) was used as a positive control for the western blot. All secondary antibodies were used 1:2,000. The WB was done in two stages, the first to ensure the infections were efficient and the second to probe for IBV N expression.

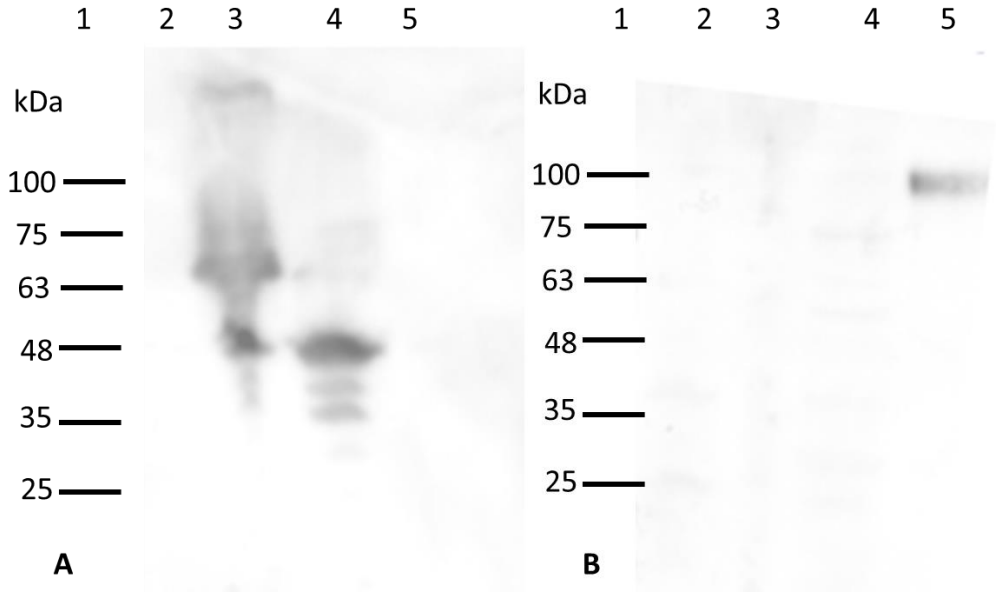


Figure 23: Western blot of baculovirus expression trials. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. Lane 2 - cells only sample, lane 3- IBV N sample, lane 4 - GFP control, lane 5 – SCARB2 control. Blot A) used primary antibody against P39 1:250 with secondary anti-mouse, blot B) used human sera as a primary antibody source and a human antibody as secondary.

Figure 23 blot A was probed with an antibody to the baculovirus major capsid protein P39 and shows that both the IBV N protein and GFP infected tracks have strong baculovirus expression in lanes 3 and 4 respectively. Blot B was probed with an anti-Human SCARB2 antibody identifying purified SCARB protein in tack 5. As infection was clearly achieved a second blot was probed for the presence of the his-tag using an anti-his antibody, as shown in Figure 24.

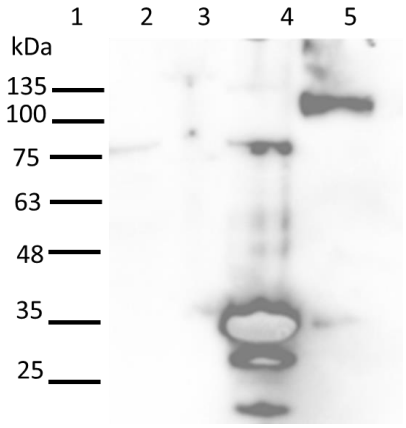


Figure 24: Western blot showing the polyhistidine signal produced by each sample generated from BEV expression using an anti-his antibody at 1:2000. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. Lane 2 contained a "cells only" sample, lane 3 the IBV samples, lane 4 a GFP control and lane 5 a SCARB control.

The western blot, Figure 24, shows a poly-histidine signal produced by the SCARB2 protein at ~135kDa and a strong signal from the GFP protein and a possible faint signal for the adjacent IBV sample in lane 3. A repeat of this WB with reduced levels of GFP control was done to ensure a better balance of signal by the chemiluminescence imager, lane order was changed to try to prevent an overriding signal from the GFP sample (Figure 25) to allow the his-tag signal from the IBV sample to be visualised in lane 2.

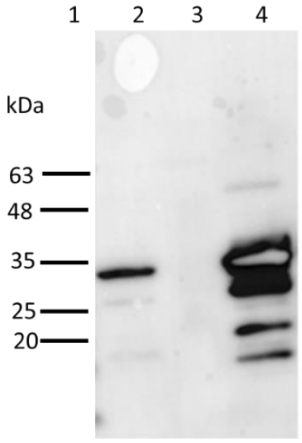


Figure 25: Western blot showing the polyhistidine signal produced by each sample using an anti-his antibody 1:2000 and secondary mouse antibody. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. Lane 2 contained the IBV sample, lane 3 a cells only negative control and finally lane 4 a GFP control following an additional 1:100 dilution from the sample used in Figure 24. Lanes 2 and 4 both show the successful production of a polyhistidine signal in both the IBV and GFP samples, the control lane of cells only in lane 3 remains blank.

4.2.6 Protease inhibitor effects

As discussed in section 4.2.4 protease inhibitors can be beneficial whilst using the BEVS. 50μM of MG132 was added to Sf9 cells one-hour post infection and the results were as shown in Figure 26.

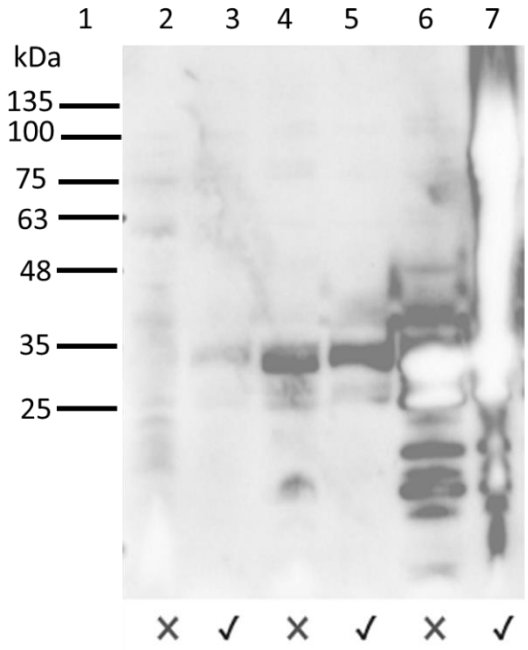


Figure 26: Western blot showing the polyhistidine signal produced by each sample using an anti-his antibody 1:2000. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside of the blot. Lanes 2 and 3 contained a cell only sample, lanes 4 and 5 the IBV expression and lanes 6 and 7 a GFP control. The ticks and crosses under each lane denote the addition of, or lack of, 50 μM of the protease inhibitor MG132 respectively. The MG132 was added one-hour post infection. Lane 5 shows a cleaner banding pattern to that of lane 4 showing that the protease inhibitor worked for the IBV example. Lane 7 was overloaded so no comparison to lane 6 with the GFP sample could be made.

Figure 26 shows that MG132 works well as a protease inhibitor in terms of reducing the amount of degradation shown with the IBV sample in lanes 4 and 5; however, the signal is still not at the expected size to be an intact IBV N protein signal (45kDa). If the BEVS was to be used to produce the remainder of the CoV N proteins the addition of MG132 may prove beneficial in reducing degradation, however effects on yield would need to be observed and the concentration optimised accordingly.

4.3 Baculoexpression summary

The results in this chapter show that the BEVS is suitable for the producing of the CoV N proteins required for this study however the IBV N protein would have to be extracted and purified via IMAC and tested on ELISAs. Transfections would need to be carried out on the other viral strains too. The process takes a longer amount of time than that of the prokaryotic techniques described in chapter 3 and would also potentially incur additional costs as protease inhibitors may need to be used. As such work was conducted to see if there was a better way of utilising the prokaryotic system to consistently produce IBV N protein without high levels of break down, detailed in chapter 5.

Chapter 5: IBV and the generation of a truncated IBV construct

5.1 IBV introduction

First described in the 1930's in the USA, and seen in 1948 in the UK, avian infectious bronchitis virus (IBV) is an important pathogen within the global commercial poultry industry (Jones 2010). IBV can be a cause of economic loss due to reduced egg and meat productivity, the need to slaughter sick birds and a high mortality rate (Najafi, Ghalyanchi Langeroudi et al. 2017) which can reach 70% when the virus infects a farm for the first time (Cavanagh 2007). The total adverse effects make IBV the biggest single cause of infectious disease-related economic loss in the UK (Jones 2010). IBV is a Gammacoronavirus and infects avian hosts, principally chickens (*Gallus gallus*), however other bird species are thought to be infected too (Cavanagh 2007). The virus primarily affects the respiratory tract, including the trachea and lungs, however strains differ in both virulence and tropism and other organs such as the alimentary tract, spleen, ovaries and kidneys can also be targeted (Najafi, Ghalyanchi Langeroudi et al. 2017). Chicks of all ages are susceptible to infection and symptoms include coughing, sneezing, gasping and nasal discharge (Liu, Zhang et al. 2009). In younger chicks, death may arise due to the primary infection or as a result of a secondary bacterial infection. Many vaccines are available however disease control remains problematic for a variety of reasons. The main issue is the virus' ability to generate antigenic diversity. Among the large number of IBV serotypes the S1 amino acid sequence, the receptor binding domain of the spike protein, may differ by 20 to 25% although sometimes this figure can be as high as 50% (Jones 2010) (Lai and Cavanagh 1997). Studies have shown that as little as 2-3% difference (10 to 15 residues) within the S1 sequence can result in a change of serotype producing a lack of cross-neutralization using convalescent sera (Jones 2010). This diversity results in the continuous emergence of new serotypes or variants and complicates the design of appropriate control programs due to the low degree of cross-protection among IBV serotypes.

Whilst this study is not focusing on the use of recombinant IBV nucleocapsid as a diagnostic antigen as avian sera are not being screened. The IBV N protein will be used as a negative control when looking for sero-reaction to other coronaviruses in human sera. Any sero-reactivity to the IBV N protein will be used as a cut-off, and anything at this level and below will be deemed a negative response. Dedicated work on the expression of IBV N for the purpose was necessary as at the time it was conducted both the prokaryotic methods and eukaryotic methods described in chapters 3 and 4 respectively, were unable to provide a consistent source of stable recombinant IBV N protein. Of all the coronavirus N proteins expressed, IBV N proved to be the most problematic with either poor yields or evident breakdown of the full-length protein to smaller fragments. The problem was eventually overcome by use of expression in commercially bought BL21-star *E. coli* strain. This chapter details the expression work conducted in the interim on IBV N protein and details the construction of a novel truncated construct, designed to remove a potential protease sensitive region in the middle of the protein was constructed and tested.

5.2 IBV results

5.2.1 Additional bacterial strains tested

Initially the *E. coli* strains screened in chapter three for robust expression of soluble coronavirus N protein included BL21, BL21-Star, C41, C42 and Rosetta. The strongest IBV signal was seen post induction using C41 cells, however the results remained inconsistent in terms of both yield and stability, and as such two additional *E. coli* strains compatible with the T7 based expression system were also tested.

5.2.1.1 ArcticExpress cells

ArticExpress cells (Agilent Technologies) are designed to increase the yield of soluble protein produced in *E. coli* by using low temperature cultivation to reduce the problem caused by inclusion bodies. However as shown in Figure 27, although the strain was able to produce a signal with the control MERS-CoV N construct, no signal was seen for IBV N.

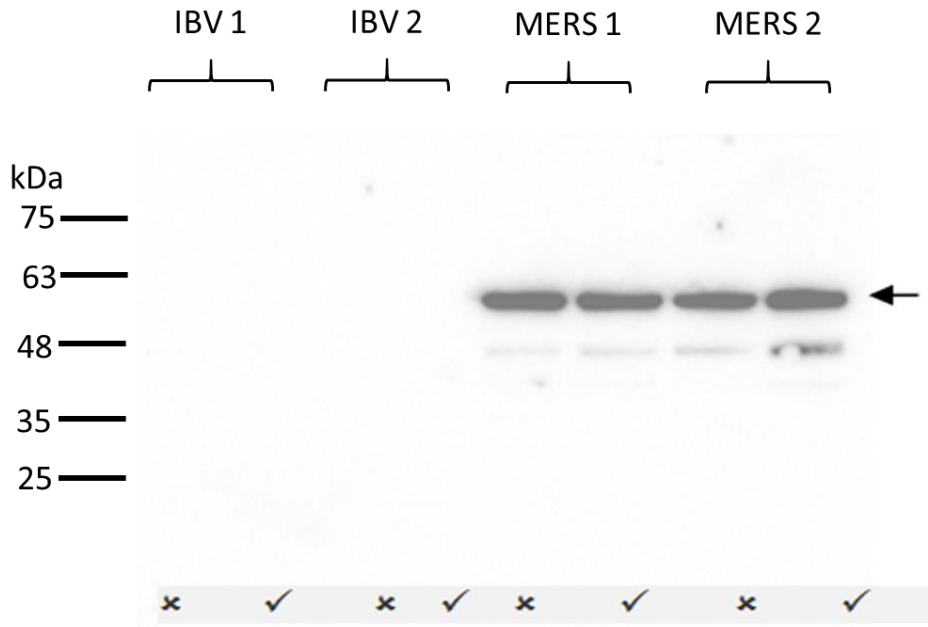


Figure 27: Western blot conducted on *E. coli* lysates from an 0.5 mM IPTG inductions of ArticExpress cells for 24 hours at 10oC for both MERS-CoV N protein and IBV N protein expression. Primary antibody was anti-his and the secondary anti-mouse-HRP conjugate. BLUEeye pre-stained protein ladder was run in lane 1; the corresponding molecular weights are marked alongside of the blot. Uninduced samples are marked with a cross under the blot and induced samples a tick. IBV samples are in lanes 2 and 3 and repeated with a different initial overnight sample in 4 and 5; neither provided a positive result. MERS-CoV inductions were run in lanes 6 and 7 and repeated in 8 and 9. MERS-CoV showed strong signal even in the un-induced in 6 and 8.

5.2.1.2 Mass spectrometry

Further analysis of the lower molecular weight band seen in Figure 15 was done to ascertain if the protein was undergoing a specific cleavage event and, if so, at which sites the breakdown was occurring. Following SDS-PAGE the IBV samples were stained with coomassie blue and the band was cut out and set to the Protein and Nucleic Acid Chemistry (PNAC) Facility at the Biochemistry Department of the University of Cambridge for mass spectrometry. All buffers were filtered, and the gel was only handled with gloves to avoid risk of contamination. The results, detailed in appendix five, showed the bands in question not to be related to IBV nucleocapsid protein but instead to be three *E. coli* contaminant proteins; peptidylprolyl cis-trans isomerase A (PPIA), 30S ribosomal protein S4 and GTP cyclohydrolyase 1 (GTPCH). Thus, low level IBV N expression may be masked by high background *E. coli* protein binding to the IMAC column.

5.2.1.3 LOBSTR and LOBSTR-ril

As a result of the contamination seen by the mass spectrometry, LOBSTR (low background strain) cells were tested as the induction hosts for the expression of IBV N. LOBSTR cells (Kerafast) are designed specifically for polyhistidine-tag affinity purification of proteins expressed in *E. coli,* as they reduce the presence of naturally histidine-rich contaminant proteins (Andersen, Leksa et al. 2013). However, Figure 28 shows that protein induction at the molecular weight indicative of N protein expression was again only successful with the control MERS-CoV N-protein transformation and not the IBV construct.

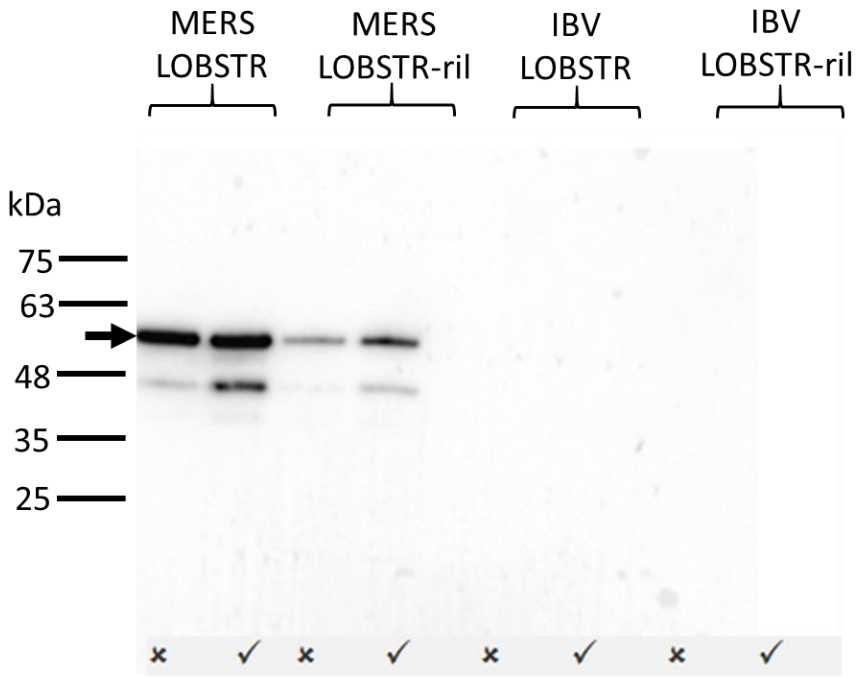


Figure 28: Western blot of *E. coli* lysates using primary anti-his antibody and anti-mouse HRP conjugate as a secondary antibody. 0.5 mM IPTG inductions were conducted using LOBSTR and LOBSTR-ril cells for 3 hours. BLUEye pre-stained protein ladder was run in lane 1; the corresponding molecular weights are marked alongside of the blot. Uninduced samples are marked with a cross under the blot and induced samples a tick. Neither LOBSTR nor LOBSTR-ril provided a successful IBV induction as shown in lanes 6-9. MERS-CoV was able to be successfully induced as shown by bands in lanes 2-5.

5.2.2 IBV temperature

Similar to the use of ArcticExpress cells, the optimized combination of IBV N expression in C41 was assessed at two different temperatures, 37°C and 28°C. However as shown in Figure 29, the reduced temperature showed no real benefit. Even when induction was conducted at 37°C the yield was very low and a band was only visible on a high contrast negative of the image, again showing the inconsistency of results seen when using C41 to express IBV N protein.

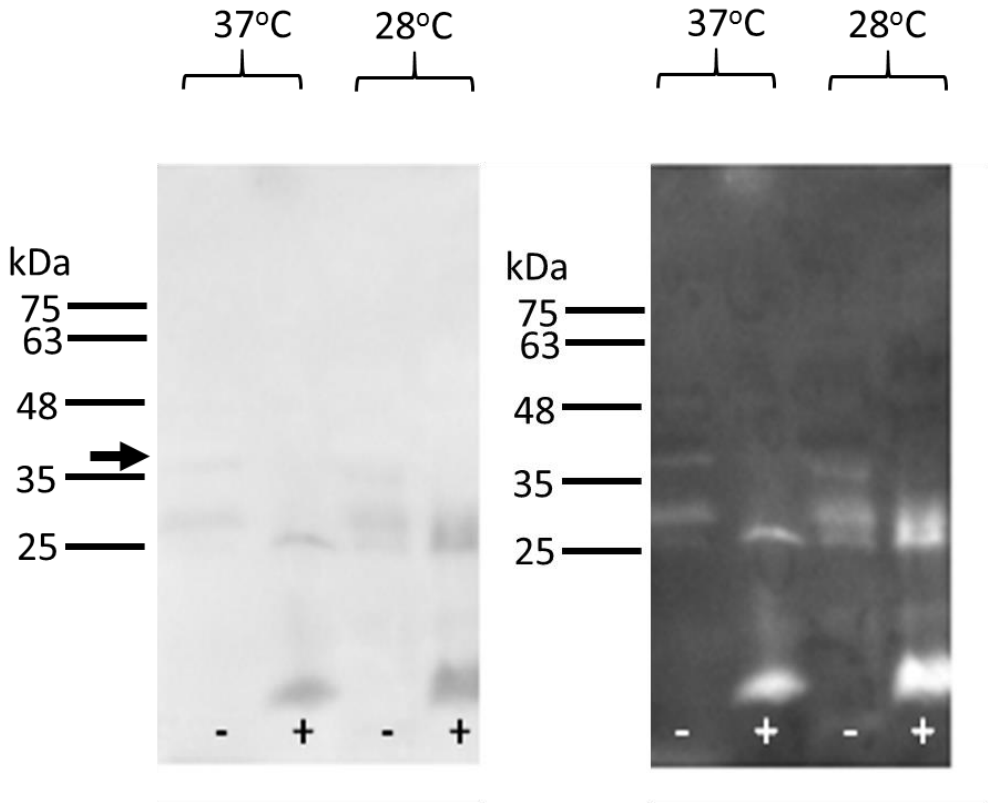


Figure 29: Western blot of *E. coli* lysate using primary anti-his antibody and anti-mouse HRP as a secondary antibody. 0.5mM IPTG inductions were conducted on IBV N protein C41 cells for 3 hours. Induction at 37°C shown in lanes 2 and 3 and a lower temperature of 28°C in lanes 4 and 5. A negative repeat of the image is shown to the right in order to help visualize bands. BLUEeye pre-stained protein ladder was run in lane 1; the corresponding molecular weights are marked alongside of the blot. All samples are post induction. The – at the bottom of the blot indicates that the sample is yet to be lysed whereas the + symbol indicates the lysis has taken place. The arrow marks where bands would be expected, however neither temperature reveals a positive result.

5.2.3 Autoinduction media

As a way of making the process more efficient an autoinduction medium (Formedium) was trialled on all the N expressing strains constructed to date. This media has been formulated to grow all strains until a density at which natural induction of the *lac* promoter takes place. This is achieved by having a limited glucose concentration which is preferentially metabolised by the *E. coli* during growth. Once this is depleted in mid- to late- log phase, the bacteria switches to metabolising lactose and the *lac* promoter is activated, which induces expression of the chromosomally encoded T7 RNA polymerase and any genes under T7 promoter control.

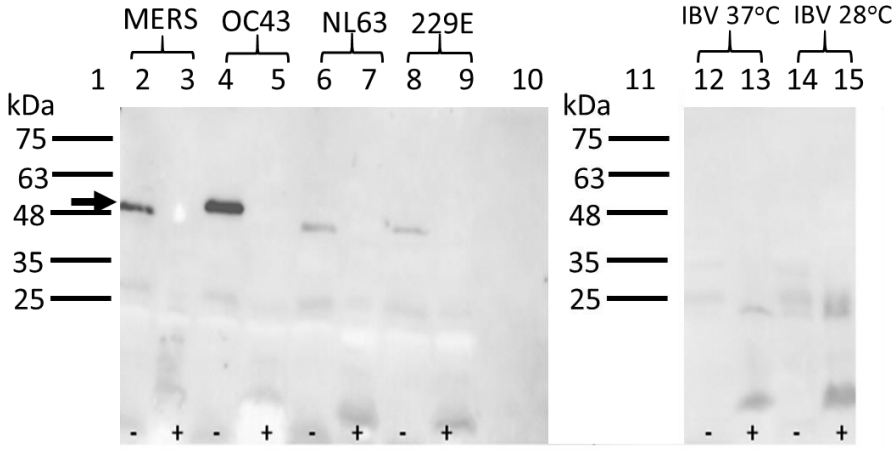


Figure 30: Western blot on *E.coli* lysate using primary his antibody and anti-mouse as a secondary antibody, conducted using lysate from all 5 HCoV N constructs when grown using autoinduction media. Cultures were incubated for 3 hours. BLUEeye pre-stained protein ladder was run in lane 1; the corresponding molecular weights are marked alongside of the blot. All samples are post induction. The –ve at the bottom of the blot indicates that the sample is yet to be lysed whereas the +ve symbol indicates the lysis has taken place. MERS-CoV shown in lanes 2 and 3 used BL21 cells, as did OC43-CoV seen in lanes 4 and 5. NL63-CoV seen in lanes 6 and seven used BL21-Star, as did 229E-CoV shown in lanes 8 and 9. Lane 10 was left blank. IBV at two different temperatures was done using C41. No signal remained for any samples post lysis.

Figure 30 shows that although four of the proteins were expressed after autoinduction had occurred (MERS-CoV, OC43-CoV, NL63-CoV and 229E-CoV) IBV was not one of them. In all cases signal appears lost post-lysis either as a result of volumes not being accounted for in preparation for SDS-PAGE or possibly as a result of inclusion bodies. These results show that neither varying the temperature of the induction nor the use of autoinduction media helped to make IBV N protein expression more efficient. As a result of these findings a dedicated new IBV construct was designed following an analysis of the N protein secondary structure.

5.3 IBV expression and purification

5.3.1 Comparative IBV expression

To clarify which cell line provided optimal expression of N protein all five T7 compatible *E. coli* strains used to date were screened alongside each other. The ArcticExpress was induced as per manufacturer's recommendation of 10°C for 24 hours, instead of the 37°C for 3 hours protocol used for all other strains. Figure 31 again shows C41 cells to produce the strongest band by western blot with an anti His antibody. However as mentioned in section 3.5.5 Large scale inductions this was unable to be replicated at a larger scale and the second band seen between the 20-25kDa marker in the majority of successful constructs shows breakdown of the protein is occurring even with the use of protease inhibitors.

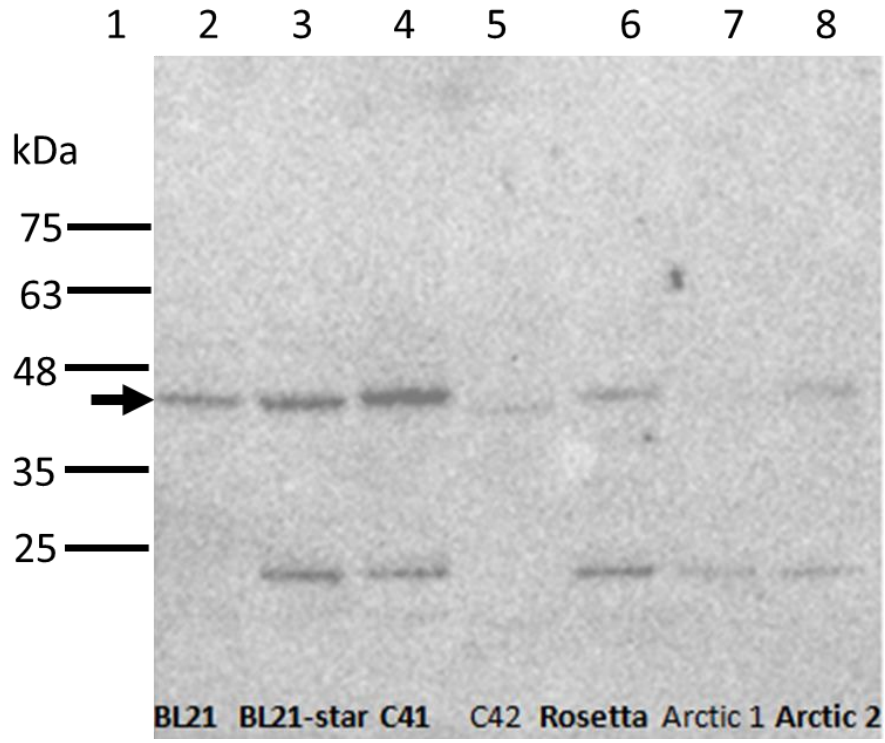


Figure 31: Western blot on *E.coli* lysate using primary his antibody and anti-mouse as a secondary antibody. Lysate gained from all strains of *E.coli* used to express IBV N protein. BLUEeye pre-stained protein ladder was run in lane 1; the corresponding molecular weights are marked alongside of the blot. All inductions were done using 0.5 mM IPTG and took place for 3 hours at 37°C except the ArticExpress ones shown in lane 7 and 8 which were conducted at 10°C for 24 hours. The band shown appears around the expected 45 kDa mark. Bold inductions went on to be pooled, lysed and manually put through the his trap column (BL21, BL21-star, c41, Rosetta and ArcticExpress 2 shown in lanes 2,3,4,6 and 8 respectively. C42 and ArticExpress 1 in lanes 5 and 7 were discarded.

5.3.2 Immobilized Metal Affinity Chromatography (IMAC)

Any single N proteins expressed post IPTG induction should be identical regardless of which *E. coli* strain was used to express it, therefore pooled IPTG induced IBV samples were lysed as described in the methods (2.2.2.1 Cell lysis), before being applied, washed and eluted by syringe through the IMAC column, each eluted fraction being 5mLs in volume.

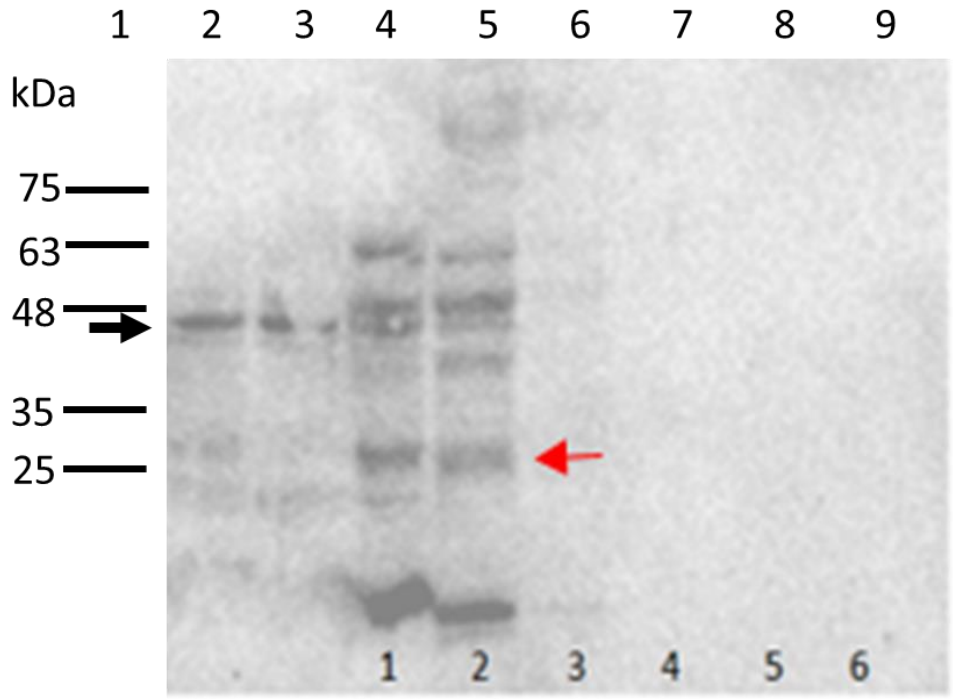


Figure 32: Western blot on *E.coli* lysate using primary his antibody and anti-mouse as a secondary antibody, showing the results of IMAC on an IPTG induced pooled IBV samples. BLUEeye pre-stained protein ladder was run in lane 1 and the corresponding molecular weights are marked alongside. The initial lysis load is in lane 2, the flow through is shown in lane 3. Fractions 1-6 are displayed in lanes 4-9 as marked at the bottom of the gel. Signal can be seen in lanes 4 and 5, indicating that tubes 1 and 2 contain the protein. Bands appear just under 48kDa, showed by a black arrow. Possible breakdown bands are indicated with a red arrow around the 25kDa mark.

Although there is signal present in the flow through (lane 3), fractions 1 and 2 contain signal of the correct MW too, albeit with other proteins. A main signal appears just under the 48kDa molecular weight marker, which is close to expected 45kDa, in the 10mL combined fractions of tubes 1 and 2.

5.3.3 Testing part purified IBV N-protein with positive sera

Positive animal sera was used to test bind to the IBV N-protein to show its utility as an ELISA antigen, the results are shown in Figure 33. The ELISA was conducted in duplicates and the averages were plotted. The mouse monoclonal antibody gave the strongest signal, as expected of a purified MAb. Weaker although still positive binding was seen in the chicken and rabbit sera. Unfortunately, no data on titre or strain was given with the gifted sera and the animals themselves may not have produced high titre antibodies to the virus used as immunogen.

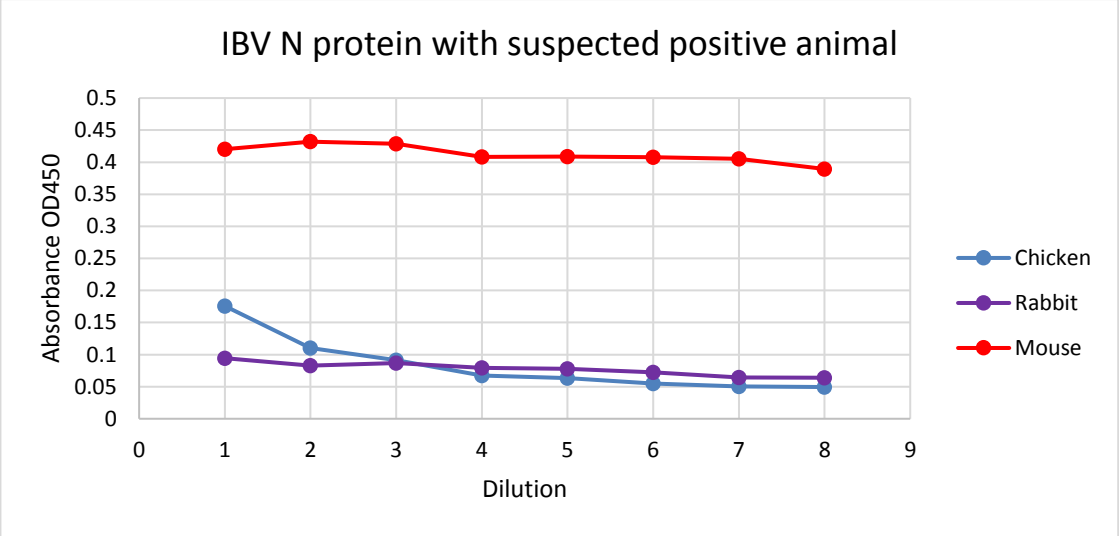


Figure 33: ELISA results of suspected positive animal sera against recombinant IBV protein. Sera used is colour coded; chicken in blue, mouse in red and rabbit in purple. The dilution series started with the stock sera being diluted 1:500 and followed a two-fold dilution series. As such 1= 1:500, 2=1:1,000, 3=1:2,000, 4=1:4,000, 5=1:8,000, 6=1:16,000, 7=1:32,000 and 8=1:64,000. The results show that only the mouse sera produced a reaction. Error bars not shown due to lack of repeats.

As such, the ELISA was repeated with the mouse sera starting with an initial dilution of 1: 10,000. This was done in duplicate as indicated by test one and test two shown in Figure 34. The drop-off in absorbance as dilution of sera increases indicates a true reaction. The concentration of mouse sera may have been much higher than that of the chicken and rabbit or it may be that the mouse sera is monoclonal whereas the chicken and rabbit sera are polyvalent.

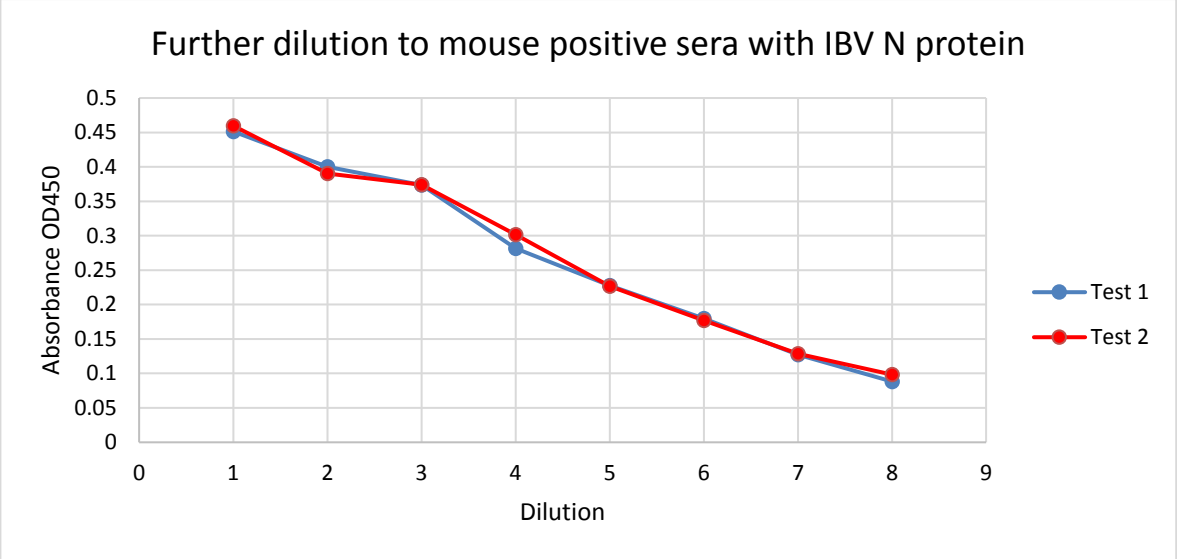


Figure 34: ELISA result of further diluting the mouse sera against recombinant IBV protein. The test was done in duplicate, test one results are in blue and test two results are in red. The dilution series started with the stock sera being diluted 1:10,000 and followed a two-fold dilution series. As such dilution 1=1:10,000, 2=1:20,000, 3=1:40,000, 4=1:80,000, 5=1:160,000, 6=1:320,000 and 7=1:640,000 and 8=1:1,280,000. The results shows the expected drop-off of absorbance with increased dilution indicating that the mouse sera has a true reaction to the recombinant IBV N protein. Error bars not shown due to lack of repeats.

As the purpose of the recombinant IBV N protein is to be a negative control when screening human sera, an ELISA was also conducted using known positive human coronavirus sera with the recombinant IBV N protein, the results are shown in Figure 35.

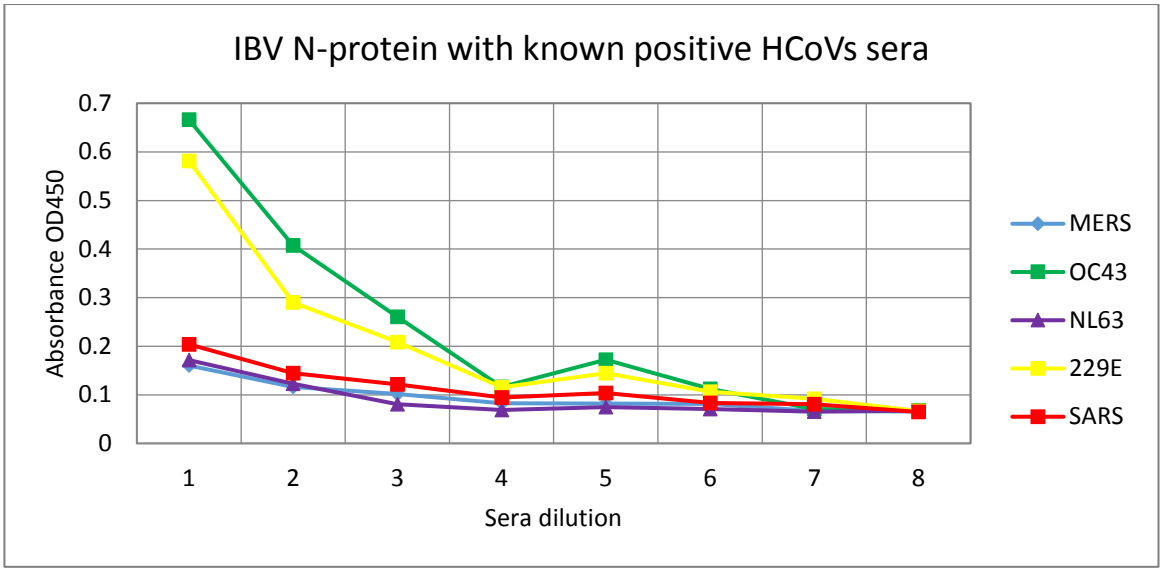


Figure 35: ELISA result of human sera known to be positive for HCoVs against recombinant IBV N protein. Sera used is colour coded; MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow, SARS-CoV in red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. As such 1= 1:50, 2=1:100, 3=1:200, 4=1:400, 5=1:800, 6=1:1,600, 7=1:3,200 and 8=1:6,400. Although no reaction should be seen OC43-CoV and 229E-CoV produce strong signals, it is worth noting that due to the volume of sera available 10 times less was used during the animal sera screening. Error bars not shown due to lack of repeats.

Ideally Figure 35 would show that the recombinant IBV N protein produced no reaction when screened with the known positive human sera. However, OC43-CoV and 229E-CoV sera appear to react with IBV N protein suggesting that the use of IBV N as a method to recognise nonspecific binding may not be valid. There is only one positive sera for each HCoV to test on and it may be that the concentration or stickiness of the sera is high enough to produce false positives. The sera would need to be screened with both recombinant IBV N protein and its corresponding HCoV protein to see if the IBV result would interfere with the overall interpretation of results. In any event, screening on IBV N was only envisaged in cases where there was apparent reaction with multiple sera. In order for IBV to play the role of a negative control first a more stable recombinant protein would need be constructed, as a result protein structure was further examined.

5.4 Protein structure

5.4.1 N protein structure introduction

The N proteins of different coronaviruses share regions of homology and can be divided into five parts; the N terminal flexible arm, the N terminal domain (NTD), the middle-disordered region referred to as the linker region (LKR), the C terminal domain (CTD) and the C terminal flexible tail (Gui, Liu et al. 2017). Three domains (NTD, LKR and CTD) have been shown in different CoVs to bind with viral RNA (McBride, van Zyl et al. 2014). Most of the literature available focuses on the SARS-CoV N protein structure however the homology of the proteins allows for structural similarities to be inferred for IBV.

5.4.1.1 N Terminal Domain

The NTD, also referred to as the RBD (RNA-binding domain), varies in sequence and length among CoVs. The NTD of IBV-CoV has been mapped to aa 19-162 as seen in part a of Figure 36 (Jayaram, Fan et al. 2006). Common characteristics of CoV N protein NTDs include predicted secondary structures including a central β-sheet platform bordered by α-helices (Jayaram, Fan et al. 2006), a basic RNA binding groove along the β-sheet platform and an extended β-hairpin (McBride, van Zyl et al. 2014). The NTD is rich in aromatic and basic residues and the 3-dimensional (3D) structure is said to resemble a right hand with a hydrophobic palm, basic figures extending beyond this and an acidic wrist (Fan, Ooi et al. 2005, Tan, Fang et al. 2006). The β-hairpin finger-like projections are flexible and positively charged and are thought to bind RNA by neutralizing the DNA phosphate groups, whilst at the same time the RNA base moieties can come into contact with exposed aromatic residues located in the hydrophobic palm (Fan, Ooi et al. 2005). There are two prominent regions within loops corresponding to residues 22 to 23 and 74 to 86; they protrude from the globular core resulting in a "U" shaped monomer (Jayaram, Fan et al. 2006). Site-directed mutagenesis studies have identified Tyr-94 and Arg-76 as critical residues for RNA binding (Tan, Fang et al. 2006). Tyr-94 is located in strand β3 of the four-stranded anti-parallel β sheet and Arg-76 is located close by at the base of the extended flexible hairpin loop (Tan, Fang et

al. 2006). However as no single mutation is able to completely disrupt RNA binding it is thought that other aromatic/basic residues at the surface of the NTD also contribute to nucleic acid binding (Tan, Fang et al. 2006). A 2006 study by Jayaram, Fang et al was able to use limiting amounts of trypsin to identify two stable independent domains of IBV N protein (Jayaram, Fan et al. 2006). The treatment showed two major cleavage sites at residues 19 and 219 and two secondary sites at 27 and 226 as seen in Figure 36 (Jayaram, Fan et al. 2006). The study referred to the constructs produced from residues 19 to 162 aa as the NTD and from 219 to 349 aa the CTD; they found that the NTD was monomeric at moderate protein concentrations but even at low concentrations the CTD formed a dimer (Jayaram, Fan et al. 2006). The study went on to conduct x-ray crystallography of the Gray strain of IBV which diffracted to 1.3-Å resolution. The structure of the Gray strain of IBV has been found to be quite similar to the Beaudette strain previously reported, as used in this study, with the exception of five additional residues in the NTD (Jayaram, Fan et al. 2006). Biochemical studies have located the RNA binding site in the N-terminal domain with the minimal region being mapped to residues 177 to 231 in MHV (corresponding to residues 136 to 190 in IBV) (Nelson, Stohlman et al. 2000, Fan, Ooi et al. 2005).

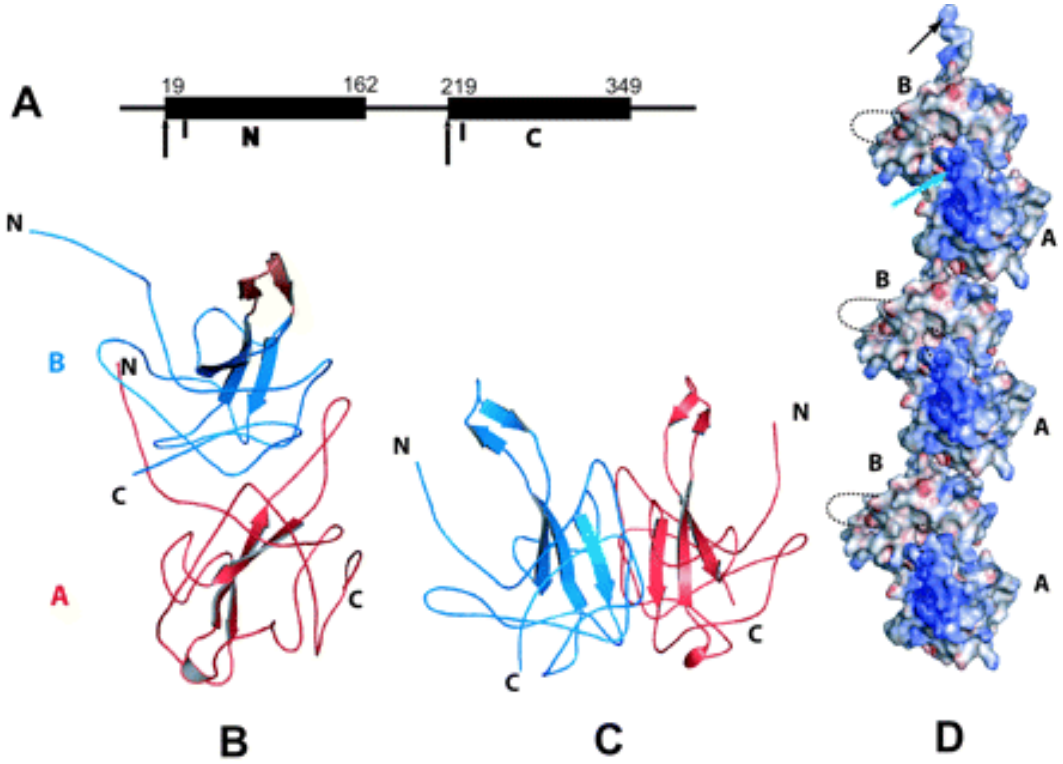


Figure 36: The structural domains of the IBV N protein and the NTD RNA binding domain (A) Schematic diagram showing the major (arrow) and minor trypsinization sites (short vertical line) seen in the Jayaram, Fan et al study. The locations of the N- (residues 19 to 162) and C-terminal domains (residues 219 to 349) are depicted as black rectangles. (B) Ribbon representation of the 1.3-Å structure of the NTD Gray strain asymmetric homodimer each monomer labelled A and B. The LKR is coloured orange. (C) The NTD of the Beaudette strain determined by Fan et al. (Jayaram, Fan et al. 2006). (D) Electrostatic potential surface of the linear array of NTD dimers molecules A and B that form the dimers are indicated. The N-terminal arm is indicated by a black arrow and the region corresponding to the internal arm, rich in basic residues, a cyan arrow. The LKR in the B molecule is indicated by a dotted line. Image from (Jayaram, Fan et al. 2006)

5.4.1.2 Intrinsically disordered regions (IDRs)

The NTD and the CTD are interspersed with areas termed intrinsically disordered regions (IDR), meaning they lack a tertiary structure and have no fixed 3D shape in the native form. The IDRs account for almost half of the molecule (Chang, Hsu et al. 2009). In SARS-CoV N protein there are three main IDRs, one before the NTD (aa 1-44), one after the CTD (aa 366-422) and a central one referred to as the LKR (aa 182-247). One study used electrophoretic mobility shift assays (EMSA), small-angle X-ray scattering (SAXS) and nuclear magnetic resonance (NMR) to examine the RNA binding behaviour of all three disordered regions of the SARS-CoV N protein. It was shown that all three regions are involved in RNA binding (Chang, Hsu et al. 2009). The study showed that the

presence of either the first SARS-CoV IDR (aa 1-44) or the LKR (aa 182-247) increased the apparent binding affinity to ssRNA three to fourfold over that of the NTD alone (aa 45-181) (Chang, Hsu et al. 2009). There was a six to eightfold increase in binding affinity seen with the inclusion of either the LKR or the C-terminal IDR (aa 366-422) in the construct of the CTD (aa 248-365). This is likely to be due to the dimeric state of the CTD which has two disordered regions attached whereas the NTD only has one. The presence of IDR allows for extended conformation of the N protein which can increase the collision radius with RNA. In SARS-CoV, both the middle and C-terminal IDRs have been implicated in the oligomerization of the N protein (Chang, Sue et al. 2006, Luo, Chen et al. 2006) and the middle IDR (the LKR) is also associated with N protein functionality and N-M interaction (McBride, van Zyl et al. 2014).

5.4.1.3 Linker region (LKR)

The LKR, also referred to as the SR-domain due to the high frequency of serine and arginine residues, is intrinsically disordered (Chang, Sue et al. 2006, Hurst, Koetzner et al. 2009). The region is flexible, capable of direct RNA interaction *in vitro* and is known to be involved in cell signalling (You, Dove et al. 2005, Chang, Hsu et al. 2009, Chang, Chen et al. 2013). Unlike the structural CTD and the NTD, the disordered regions of the different coronaviruses share little sequence homology. However, they do have similar physiochemical properties; all Coronavirus N protein LKRs start with the SR-rich region, followed by a helix, and end with a region rich in basic residues (Jayaram, Fan et al. 2006, Chang, Hsu et al. 2009). One study showed the RNA binding affinity of the LKR was comparable with that of the CTD and NTD (Chang, Hsu et al. 2009). Studies have shown the LKR to be directly involved in N protein oligomerization; mutational deletions to the area has been shown to abolish N protein self-multimerization (He, Dobie et al. 2004). The LKR contains phosphorylation sites and it is proposed that hyperphosphorylation of the area reduces the total positive charge of the protein and leads to enhanced oligomerization of di-domain constructs (Chang, Chen et al. 2013). However, some studies oppose this idea and report that the LKR interferes with oligomerization if it is phosphorylated (Peng, Lee et al. 2008) or when the CTD is also present (Luo, Ye et al. 2005). The LKR is prone to degradation and limited

structural information is available due to its flexible nature and high positive charge (Chang, Hsu et al. 2009). Inclusion of the LKR has been shown to increase ssRNA binding affinity and also has a noticeable effect on the apparent Hill coefficient, which is a measure used to quantify the degree of interaction between ligand binding sites (Weiss 1997, Chang, Hsu et al. 2009).

5.4.1.4 SR-rich region

The beginning of the LKR contains a Serine and Arginine rich area (Chang, Chen et al. 2013). The SR-rich region has been shown to carry out a number of protein-protein interactions and play a part in self-association (He, Dobie et al. 2004, Chang, Hsu et al. 2009). The region has multiple phosphorylation sites and contains the highest density of positive charges within the LKR and making it important for RNA binding (Surjit, Kumar et al. 2005).

5.4.1.5 CTD disordered region

The IDR located at the CTD has also been shown to participate in the oligomerization of the N protein (Luo, Chen et al. 2006) and to bind to nucleic acid (Chang, Hsu et al. 2009). The highly charged nature of both the LKR and the CTD IDR hinders N protein to N protein interaction due to charge repulsion between the domains. However this may be neutralized by the binding of nucleic acids, allowing two protein molecules to come into contact and oligomerise (Chang, Hsu et al. 2009).

5.4.1.6 Summary of IDR

The presence of the IDRs is advantageous when it comes to the formation of the RNP. The LKR's flexibility allows more freedom for different parts of the N protein to interact with each other, resulting in specific packaging of the helical RNP molecule. The ID regions could also play a role in optimizing the interaction of the RNA molecule with all the other segments of the N protein (Chang, Hsu et al. 2009).

5.4.1.7 C-terminal domain (CTD)

The CTD, sometimes referred to as the dimerization domain (DD), is hydrophobic and helix rich. It has been mapped to aa 219-340 for IBV (Fan, Ooi et al. 2005). The CTD is a tightly intertwined dimer with twofold symmetry (Jayaram, Fan et al. 2006). It has two β strands and one α helix from one monomer making extensive contacts with the other monomer. The CTD dimer has a rectangular shape. There is a concave floor consisting of an antiparallel βsheet (β1A-β2A-β2B-β1B) contributed by monomers A and B, surrounded by several α helices and one short $_3$10 helix. Several biochemical studies map the dimerization domain of the full-length protein to its C-terminal domain (Surjit, Liu et al. 2004, Yu, Gustafson et al. 2005). One study conducted *in vitro* oligomerization using cross-linking techniques showed that residues 218-329 within the CTD of IBV play a major role in the proteins multimerization (Fan, Ooi et al. 2005).

In its monomeric form the CTD folds into a structure with a large cavity in the centre, rendering it unstable and making oligomerization necessary for a stable conformation (Yu, Oldham et al. 2006).

Reverse genetic-complementation assays have shown the CTD to be responsible for N protein interaction with M proteins within the viral membrane (Kuo and Masters 2002). It is thought that this interaction is principally electrostatic in nature; involving the basic C terminus of the M protein and an acidic β-sheet floor in the CTD dimer (Luo, Wu et al. 2006).

A 2005 study by Tang, Wu et al., used monoclonal antibodies (mAb) directed to specific epitopes along the N protein structure of SARS-CoV and concluded that the C-terminal fragment (amino acids 214 – 422) is buried within the entire N protein (Tang, Wu et al. 2005) this could explain why locating the polyhistidine tag at the C terminus in previous work undertaken in the laboratory was less successful in producing detectable soluble protein, whereas relocating the polyhistidine tag to the N terminal domain proved successful.

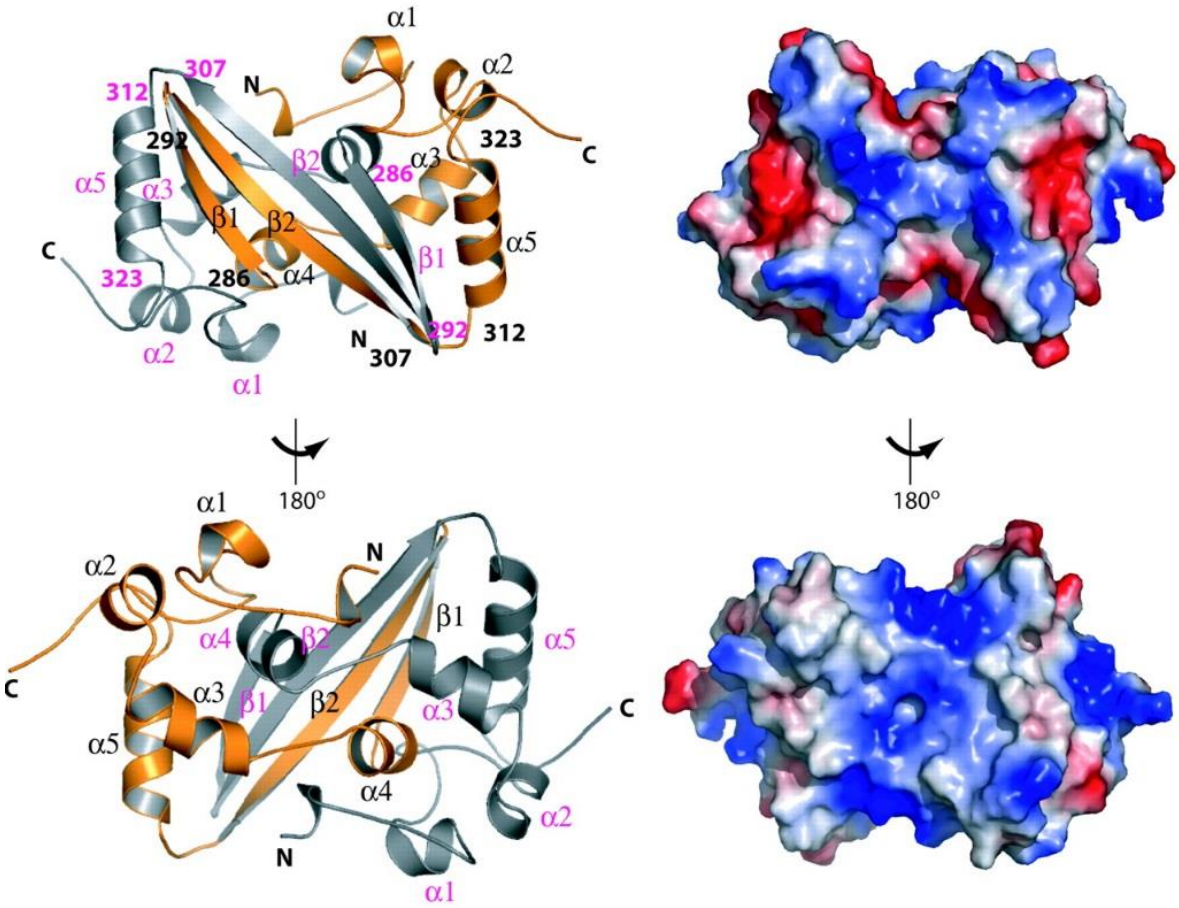


Figure 37: Structure of the CTD dimerization domain. The left panel shows a ribbon representation of the "front" and "back" of the CTD dimer related by a rotation of 180° about the vertical axis, and the right panel shows the electrostatic potential surface of the dimer in the same orientations. Positively charged surfaces are represented in shades of blue and the negatively charged surfaces in shades of red. Left: The intertwined CTD dimer is formed by exchanging two β strands and one α helix between the two monomers. The two monomers, shown in yellow and grey, are related by a noncrystallographic twofold axis of symmetry. The β strands from both monomers form an extended antiparallel β-sheet floor flanked by several α helices. Secondary structural elements are labelled. Right: a large patch of positively charged residues (blue) that could be involved in RNA binding is visible on one of the faces of the CTD protein (bottom). (Jayaram, Fan et al. 2006)

5.4.1.8 Tertiary structure

The N protein packages the viral genome into a helical ribonucleocapsid and is vital for viral self-assembly (Chang, Chen et al. 2013). There are two main activities involved in packaging the viral genome with structural proteins to produce the RNase resistant ribonucleoprotein (RNP) complexes (Jayaram, Fan et al. 2006). One is the interaction between protein and nucleic acid, the other is the ability of the complex to oligomerise (Zlotnick 2005). Coronavirus N proteins preform both of these functions (McBride, van Zyl et al. 2014).

As mentioned N proteins are dimeric, *in vitro* these dimers have shown a tendency to form tetramers and oligomers with higher molecular weights (Yu, Gustafson et al. 2005). X ray crystallography of the CTD has shown that two dimers create a butterfly-shaped tetramer which then come together to form an octamer (Chen, Chang et al. 2007). The SARS-CoV N protein dimer is formed principally by insertion of the β-hairpin of one subunit into the cavity of the opposite subunit, this causes four β-strands of two subunits to form an anti-parallel β-sheet that is covered by two long alpha helices (Yu, Oldham et al. 2006). The interaction is highly stable as a result of the extensive hydrogen bonds between the two hairpins alongside the hydrophobic interactions between the beta-sheet and the alpha helices (Chang, Sue et al. 2005). The octamer's structure is maintained through hydrophobic interactions and hydrophilic contacts among the four dimers; inter-dimer hydrogen bonds help stabilize the octamer further (Chen, Chang et al. 2007). The CTD has also been shown to cause oligomerization in 229E-CoV via its C terminal IDR (Lo, Lin et al. 2013). Disruption of N protein self-association via oligomerization could be a possible target area for drugs to combat CoV infections (Lo, Lin et al. 2013).

5.4.2 Structure of other positive-stranded ssRNA viral N proteins

Other positive-stranded ssRNA viruses have had their nucleocapsid proteins studied. Crystal structures are available for porcine reproductive and respiratory syndrome virus (PRRSV) in the Arteriviridae family (Doan and Dokland 2003), West Nile virus in the Flaviviridae family (Dokland, Walsh et al. 2004), and Sindbis virus and SemLiki Forest virus in the Togaviridae family (Choi, Lu et al. 1997) (Choi, Tong et al. 1991). A systematic structural homology search (Holm and Sander 1998) showed the coronavirus N protein CTD closely resembles the N protein of PRRSV. The PRRSV CTD (aa 73-123) has a similar dimeric structure and displays self-association as seen in the IBV CTD (Doan and Dokland 2003). Although the CTD fold is shared with the arterivirus PRRSV, the NTD fold is seen only in the coronavirus N proteins (Fan, Ooi et al. 2005). The N protein of PRRSV is shorter and the NTD appears to be largely disordered (Doan and Dokland 2003). It is thought that members of the Coronaviridae and Arteriviridae families (order Nidovirales) share a common mechanism of filamentous nucleocapsid formation, as their structures appear similar. However, it is assumed that there are adaptations in order to interact specifically with their respective genomes. Structural differences occur with flaviviruses and togaviruses, which display icosahedrally symmetric exteriors and are not pleomorphic like coronaviruses (Zhang, Corver et al. 2003, Mukhopadhyay, Kuhn et al. 2005), probably reflect variations in their replication strategies and assembly pathways (Jayaram, Fan et al. 2006). The two distinct domains whereby one dimerises and the other interacts with the viral gene, remains a common N protein feature across all positive-strand ssRNA viruses (Jayaram, Fan et al. 2006).

5.4.3 Protein stability

5.4.4 Predicted protein disorder plots

Other studies have noted that recombinant coronavirus N protein expressed in *E. coli* can be highly susceptible to proteolysis (Fan, Ooi et al. 2005, Jayaram, Fan et al. 2006, Zuwała, Golda et al. 2015). As discussed in section 5.4.1.3 one of the main areas of disorder is the flexible LKR. This can be further seen when using protein predicted disorder software PrDOS (Protein disorder prediction system) (Ishida and Kinoshita 2007) to plot the disordered areas of the amino acid sequence of IBV N, shown in Figure 38. Similar patterns showing a disordered central region are seen for all of the CoV N protein sequences, results not shown.

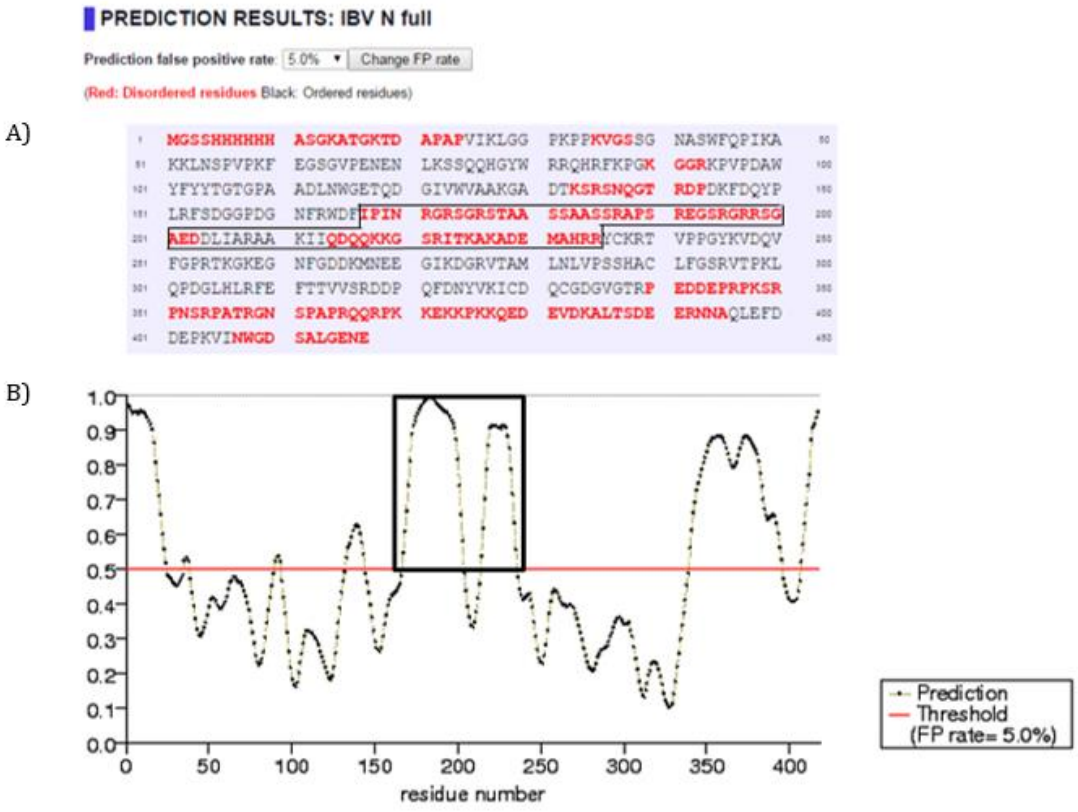


Figure 38: Predicted protein disorder plots for the full-length IBV construct, created using PrDOS (Ishida and Kinoshita 2007). Section A shows the amino acid sequence, disordered residues are shown in red, the block box indicates the main area of disorder being focused on from Isoleucine 167 to Arginine 235. Section B shows the disorder plot with the false positive rate threshold set to 5.0%, again the area of disorder being focused on corresponding to the section marked out in section A is boxed.

5.5 Delta IBV

5.5.1 Introduction to delta IBV

To improve IBV N expression levels and minimise degradation, it was decided to delete the disordered region to see if it would produce a more stable truncated protein construct, which would remain antigenic. Amino acids 167 to 235 were chosen to be removed creating a construct IBVΔ167-235 which will be simply referred to as ΔIBV. The original full length IBV sequence was 418 amino acids and had an expected molecular weight of 46kDa, whereas ΔIBV is 349 amino acids long with an expected molecular weight of 39kDa. A truncated recombinant nucleocapsid protein has been successfully used for diagnostic purposes already in the case of OC43-CoV (Blanchard, Miao et al. 2011). The study created three sections of OC43-CoV N protein, (N1= aa1-119, N2=120-332 and N3=333-448 of MWs 17kDa, 25kDa and 17.5kDa respectively) to be used as antigens in immunoassays (Blanchard, Miao et al. 2011). A total of 15 acute and 11 convalescent sera from OC43-CoV patients were screened against all three constructs. All three constructs showed reactivity against antibodies in the convalescent sera, however detection by acute-phase sera was limited. The predominant response (11 out of 11) was seen against N3, followed by a lesser response to N2 and N1 (4 and 3 out of 11 respectively). The study saw cross-reactivity of full-length recombinant OC43-CoV N protein with convalescent 229E-CoV and SARS-CoV sera. However, the absorbance seen was reduced by 88-90% when the sera was screened against the N3 construct as opposed to the full-length N protein (Blanchard, Miao et al. 2011). Although this study is not planning on separating N proteins into three fragments it does imply that the antigenic reactivity is not limited to the central domain of the protein but that is lies rather in the CTD and therefore ELISA results should not be substantially affected by the removal of the central region.

After varying PCR conditions to try and produce the truncated IBV construct using the QuikChange II site-mutagenesis proved unsuccessful, the sequence was instead designed and ordered in from Life Technologies Limited. The ΔIBV sequence is 1047bp long, 207bp smaller than the full-length sequence of 1254bp, shown by PCR results in Figure 39.

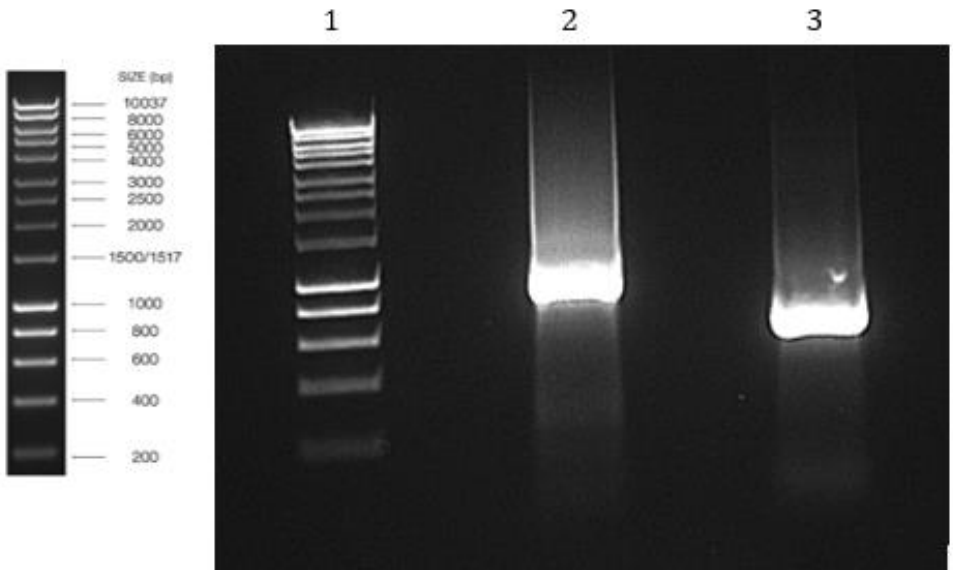


Figure 39: Gel electrophorisis showing the size difference between the full length IBV insert and the ΔIBV. Hyperlader I was used in lane 1, the lane labelled 2 contains the full length IBV and lane 3 ΔIBV. Although the lanes were overloaded and did not run at the ~1200bp and ~1000bp expected the difference seen between the two bands indicated the 200bp deletion was successful (~1000bp and ~800bp) the gel molecular sizes may be better assessed with a more dilute loading and/or longer running of the gel for a greater resolution.

The ΔIBV synthetic DNA was digested by restriction enzyme *NcoI* and *XhoI* and ligated into the pTriEx1 vector, similar digested. Following transformation and isolation, the correct construct was characterised and digested by both enzymes next to all previously isolated N expression constructs digested similarly. Results for the double digest to show vector and inserts for NL63-CoV, OC43-CoV, 229E-CoV, IBV and ΔIBV are displayed in Figure 40.

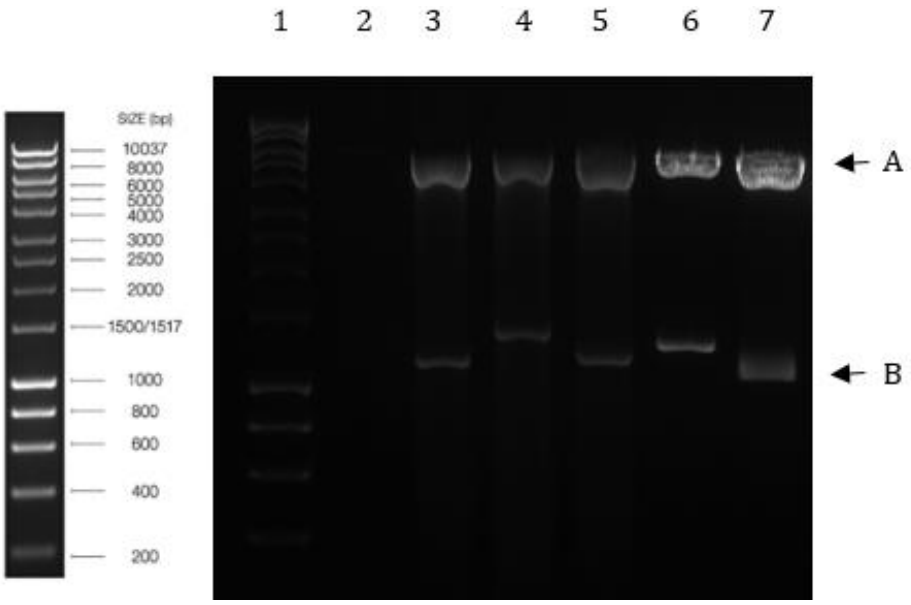


Figure 40: Gel electrophoresis showing five constructs after a double enzyme digest using Nco1 and Xho1. HyperLadder I was loaded in lane 1, lane 2=blank, lane 3= NL63-CoV, lane 4=OC43-CoV lane 5=229E-CoV lane 6= IBV and lane 7=ΔIBV. Bands denoted by the letter A show the larger his-tagged vector pTriEx1 (5301bp) and bands shown by the letter B are the viral-specific inserts, lane 6 and 7 show the ~200bp deletion lost by the mutation to IBV.

Based on published sequences found on NCBI the size of each insert band, should be 1,131bp for NL63-CoV, 1,344bp for OC43-CoV, 1,344 for 229E-CoV and 1,167 for IBV. As shown in Figure 40 the bands lie in between the 1,000bp and 1,500bp size markers with OC43-CoV being greater than the rest. Similar correct band sizes were observed for MERS-CoV and SARS-CoV (data not shown). Figure 40 shows the ΔIBV construct has been successfully inserted into the pTriEx1 plasmid and is therefore ready to be transformed into a suitable *E. coli* strain for IPTG induction.

Whilst work was being undertaking to construct the ΔIBV construct, DNA encoding for SARS-CoV N protein was sourced and underwent the same processes detailed in section 3.4, results not shown, such that purified SARS-CoV N protein could be screened alongside the previous 5 H-CoV N proteins already generated.

5.5.2 Delta IBV expression and purification

Also whilst work was done on creating the truncated IBV protein, the laboratory purchased some SoluBL21 cells and found that to be the preferred T7 compatible expression host at consistently expressing all coronavirus N proteins, as shown Figure 42. Therefore, SoluBL21s were also used to express ΔIBV as shown in Figure 41. The corresponding band appears around the expected 39kDa mark, smaller than the full length expected 46kDa, although appears slightly overloaded. The reduced protein size of the ΔIBV construct seen in lane 16 was accompanied by apparently less break protein down as fewer lower MW bands are visible.

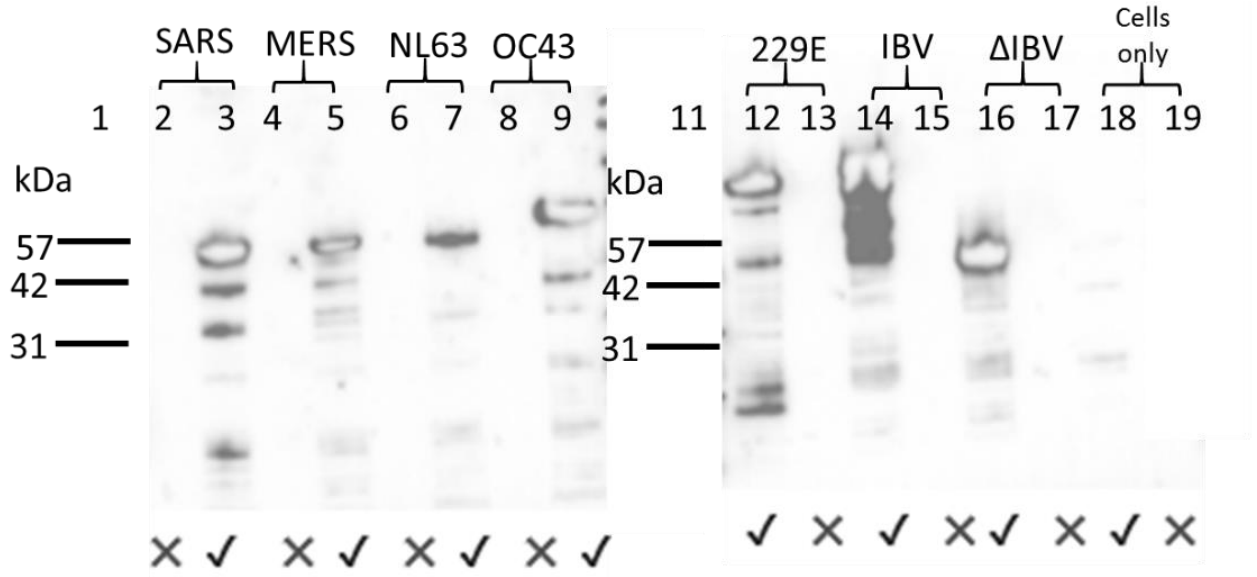


Figure 41: Western blot on *E.coli* lysate using primary his antibody and anti-mouse as a secondary antibody. Lysates generated from eight 0.5mM IPTG 50mL inductions using SoluBL21 cells. Lanes 1 and 11 contain BLUEeye pre-stained protein ladder and the corresponding molecular weights are marked alongside of the blot. Crosses and ticks underneath the blot indicate whether the sample is uninduced or induced respectively. Lanes 2 and 3 contain the SARS-CoV recombinant N protein induction, lanes 4 and 5 MERS-CoV, lanes 6 and 7 NL63-CoV, lanes 12 and 13 229E-CoV, lanes 14 and 15 IBV, lanes 16 and 17 ΔIBV and lanes 18 and 19 a cell only control. All bands produced were at the expected molecular weight and inductions were deemed successful.

5.5.3 Delta IBV comparison to full length IBV ELISA

ELISAs were conducted in order to see if the ΔIBV N protein retained the antigenicity of the full-length protein. Methods for purifying IBV and ΔIBV N protein via IMAC were the same as detailed in section 2.2.2Protein extraction, purification, concentration and storage, results not shown. The same mouse, chicken and rabbit sera used for the original IBV N protein ELISAs, section 5.3.3, were used. The ELISA was conducted in duplicates and the average results plotted in Figure 42.

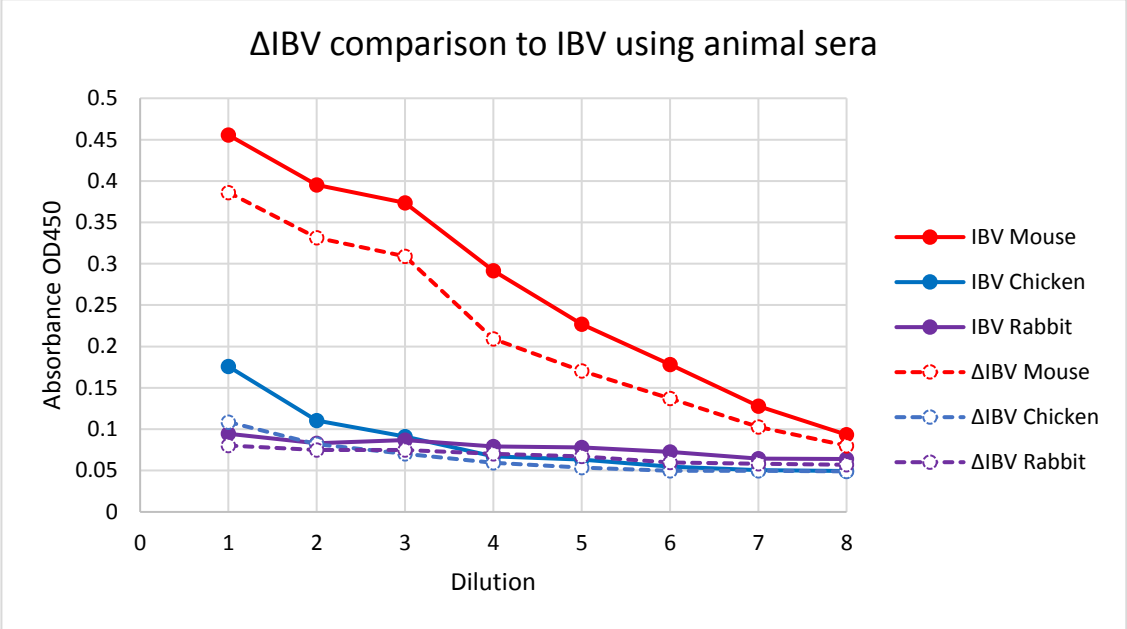


Figure 42: ELISA results showing the average results of suspected positive animal sera against recombinant full-length IBV N protein and the mutant ΔIBV N protein. Sera used is colour coded; chicken in blue, mouse in red and rabbit in purple. The results for the full-length construct are marked with solid lines and markers and the ΔIBV with hollow markers and dotted lines. The dilution series started with the stock sera being diluted 1:500 and followed a two-fold dilution series. As such 1= 1:500, 2=1:1,000, 3=1:2,000, 4=1:4,000, 5=1:8,000, 6=1:16,000, 7=1:32,000 and 8=1:64,000. The results show that only the mouse sera produced a reaction and also that the ΔIBV N protein is able to produce a similar pattern to the full-length construct with a slightly lower absorbance. Error bars not shown due to lack of repeats.

The results in Figure 42 show that the ΔIBV is able to bind test sera to produce similar absorbance readings and titration patterns as the full length IBV N construct. Similar mutations could be undertaken in the other 5 viral strains to improve stability however due to time limitations this was not carried out.

5.6 ΔIBV discussion

This chapter looked to overcome the problem of protein break down and inconsistent expression of the IBV N protein that was seen prior to the purchase and use of SoluBL21 cells. Two additional prokaryotic cell lines (Artic Express and LOBSTR) were investigated but were unable to produce a band on westerns blot, as seen in sections 5.2.1.1and 5.2.1.3 respectively. Autoinduction media was also not deemed beneficial.

Removing the disordered flexible LKR region, amino acids 167 to 235, produced an IBV N protein construct that was 349 amino acids long with an expected molecular weight of 39 kDa. Post protein expression and purification the protein was shown to be able to elicit comparable ELISA results to the full-length protein, see figure Figure 42, and deemed a success.

The concept of using truncated recombination nucleocapsid proteins for diagnostic purposes is not new, as already mentioned Blanchard *et al* in 2011 looked at doing this for OC43-CoV N protein (Blanchard, Miao et al. 2011).

Yu, Lee *et al* used truncated SARS-CoV N protein to conduct ELISAs; the team found the (N)Delta(121) truncated protein was able to provide them with better levels of sensitivity and specificity when screening health care workers in Vietnam compared with the full length N protein based ELISA; making it useful for large scale epidemiological studies (Yu, Le et al. 2005)

In 2008 Mu, Niu *et al* worked not on a truncated N protein but a fusion of a truncated N and truncated S protein for SARS-CoV (Mu, Niu et al. 2008). The study showed that the truncated S-N SARS-CoV fusion protein was a suitable immunodiagnostic antigen as well as a potential vaccine candidate when animal trials were conducted.

Similar studied using truncated proteins have been done looking at animal CoVs also, such as one carried out by Chang, Peng *et al* who used a truncated S protein when conducting ELISAS detecting antibodies against porcine epidemic diarrhoea virus (Chang, Peng et al. 2019). The full-length ELISA had a sensitivity of 97.8% and a specificity of 94% however this was improved on when the

truncated version (S^{1-501}) of the protein was used gaining a sensitivity of 98.9% and a specificity of 99.1% (Chang, Peng et al. 2019).

Further work could be done to see if specificity and sensitivity values could be generated for the IBV and ΔIBV used in this study. It would be interesting also to produce truncated versions of the five HCoVs used in this study too to see if they produced the benefits aforementioned in previous studies. However due to time constraints and the successful production of IBV N using SoluBL21 cells, sera screening was prioritised.

Chapter 6: ELISAs results

6.1 ELISA introduction

6.1.1 Previous ELISA studies on HCoV

Recombinant N-proteins expressed in *E. coli* have previously been used as antigens for ELISA screening in respect to both SARS-CoV (Shi, Yi et al. 2003, Lau, Woo et al. 2004) and MERS-CoV (Chen, Chan et al. 2015). The studies commend ELISA as a technique, noting that it is cost-effective, sensitive and user-friendly as well as highlighting the fact that the use of recombinant N protein as opposed to virus lysates makes the screening technique safer. One study even found that the ELISA they conducted on positive nasopharyngeal aspirate and faecal specimens picked up positives that the RT-PCR had missed (Lau, Woo et al. 2004).

6.1.2 Previous ELISA studies on animal CoV

Similar studies have used recombinant CoV N protein in animal sera studies looking at IBV in turkeys (Gomaa, Yoo et al. 2008), Bovine Coronavirus (BCV) in cows (Cho, Hoet et al. 2001), Porcine epidemic diarrhoea virus (PEDV) in swine (Wang, Jiyuan et al. 2015), porcine deltacoronavirus (PDCoV) in swine (Su, Li et al. 2016) and BatCoV HKU9 in bats (Lau, Poon et al. 2010). Screening of animal sera with the recombinant HCoV N proteins generated within this study could be conducted to see if there are any circulating coronavirus strains that show cross-reactivity to the HCoV N proteins produced.

6.2 ELISA results on known positive sera

6.2.1 Screening of recombinant HCoV N protein with positive sera

Known positive sera was gifted by Public Health England and used as positive controls to test that the recombinant N proteins created were structurally able to bind to their respective antibodies. The purified proteins were coated onto ELISA plates at 5µgs/mL and the ELISAs were conducted as described in section 2.2.7. The data obtained shows clear sero-conversion to produce antibodies to the cognate viral antigen, suggesting that purified N protein can be used for serosurveillance. It should be noted that most people have been serially infected with the common coronaviruses so reactivity may be more complex when general population sera are tested, as indicated by the minor reaction to NL63-CoV N protein by more than one sera shown in Figure 43.

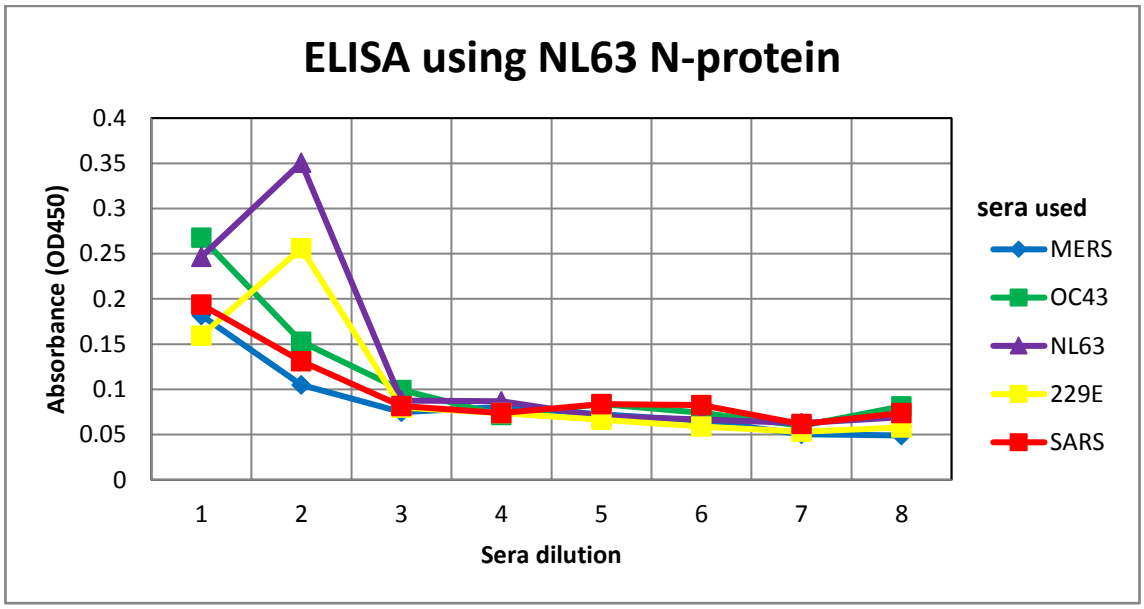


Figure 43: ELISA results showing known positive human sera against recombinant NL63-CoV protein. Sera used is colour coded; MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and SARS-CoV in red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. As such 1= 1:50, 2=1:100, 3=1:200, 4=1:400, 5=1:800, 6=1:1,600, 7=1:3,200 and 8=1:6,400. Error bars not shown due to lack of repeats. Reactivity can be seen with 229E-CoV which is a HCoV known to cause the common cold. It is possible that the person from whom the 229E-CoV positive sera was collected had also had a prior infection or exposure to NL63-CoV. The highest absorbance was seen with the NL63-CoV sera, which was as expected.

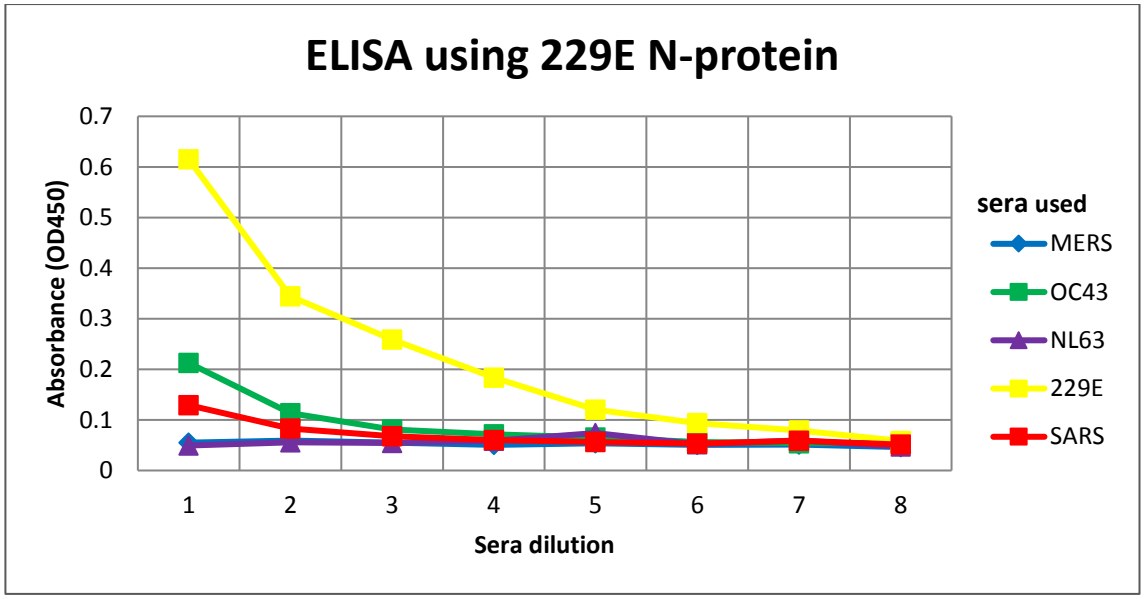


Figure 44: ELISA results showing known positive human sera against recombinant 229E-CoV protein. Sera used is colour coded; MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and SARS-CoV in red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. As such 1= 1:50, 2=1:100, 3=1:200, 4=1:400, 5=1:800, 6=1:1,600, 7=1:3,200 and 8=1:6,400. Error bars not shown due to lack of repeats. The ELISA shows a strong positive reaction to the 229E-CoV sera with minimal cross-reaction.

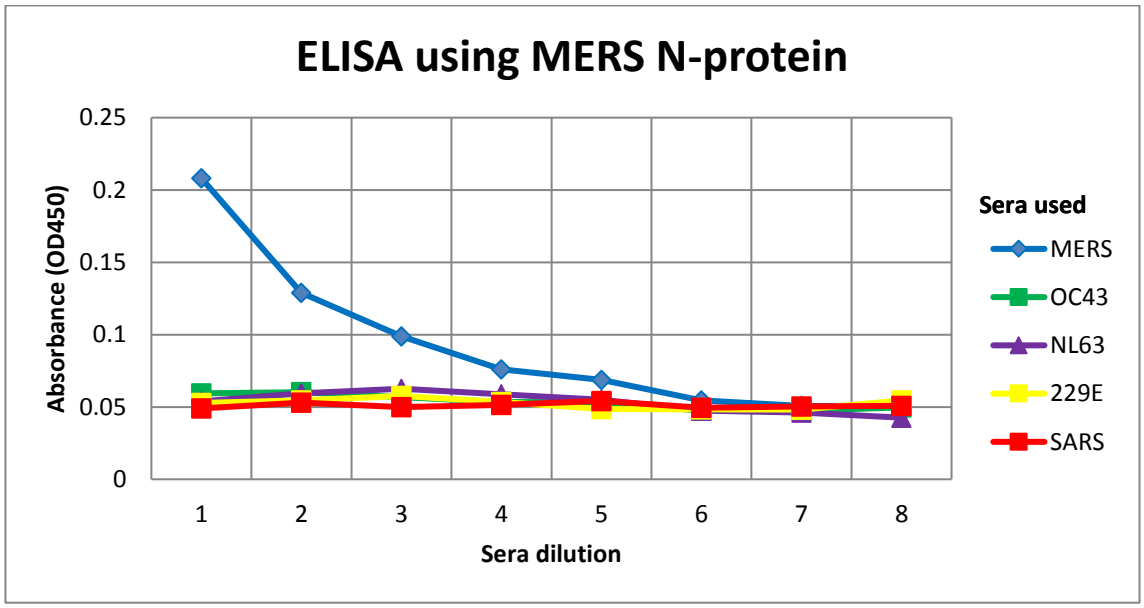


Figure 45: ELISA results showing known positive human sera against recombinant MERS-CoV N protein. Sera used is colour coded; MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and SARS-CoV in red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. As such 1= 1:50, 2=1:100, 3=1:200, 4=1:400, 5=1:800, 6=1:1,600, 7=1:3,200 and 8=1:6,400. Error bars not shown due to lack of repeats. This shows there to be strong positive reaction to the MERS-CoV positive sera with no cross reactivity.

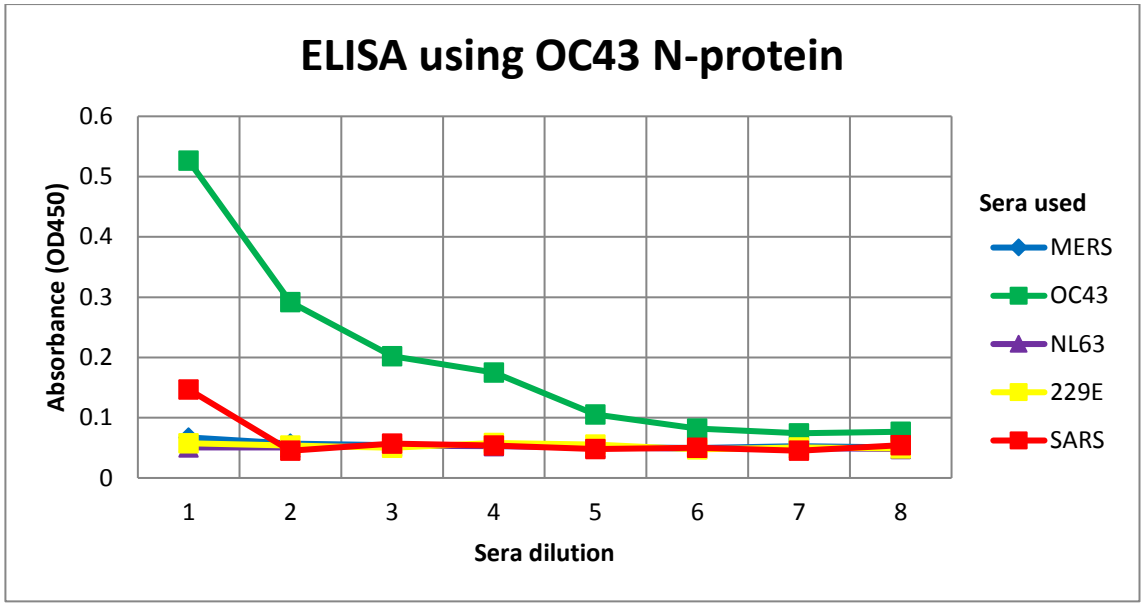


Figure 46: ELISA results showing known positive human sera against recombinant OC43-CoV N protein. Sera used is colour coded; MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and SARS-CoV in red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. As such 1= 1:50, 2=1:100, 3=1:200, 4=1:400, 5=1:800, 6=1:1,600, 7=1:3,200 and 8=1:6,400. Error bars not shown due to lack of repeats. The ELISA shows strong positivity to the OC43-CoV sera and minimal reactivity to the other sera.

The MERS-CoV and SARS-CoV sera used were both UK index cases and the 229E-CoV, OC43-CoV and NL63-CoV were convalescent serum. The data obtained show clear sero-conversion in the known status sera to the cognate viral antigen. Figure 43 shows that the recombinant NL63-CoV N protein produces the highest OD450 readings when screened with the NL63-CoV-positive sera. Similarly Figure 44 shows a clear high reading between the 229E-CoV-positive sera and the 229E-CoV recombinant N protein, Figure 45 and Figure 46 show similar positive results to the corresponding N proteins with both the MERS-CoV and OC43-CoV sera respectively. These results suggest that purified N protein can be used for serosurveillance.

In the case of the NL63-CoV ELISA in Figure 43 some minor reactivity with a 229E-CoV positive serum was also apparent. Both NL63-CoV and 229E-CoV are Alphacoronavirus so there is a possibility that this is a cross-reaction, although the reaction seen is more likely due to the fact that members of the general population will have been serially infected with the common coronaviruses. This means that reactivity may be more complex when general population sera are tested (Gorse, Patel et al. 2010).

In order to generate a cut-off value, more known-positive sera could generate a mean $OD_{450,}$ - whereby standard deviation values could be obtained. Often the cut-off OD_{450} readings of the ELISAs are defined as the mean and two standard deviations; these values may vary between Coronavirus strains (Lau, Woo et al. 2004). Any values that were known to be negative but shown to give positive results above said cut-off value, could be used to ascertain the specificity of the test. Similarly, any sera known to be positive that were unable cross the cut-off thresh-hold would help to calculate the sensitivity of the ELISA. Ideally there would be more than one serum example for positive cases to screen with, in order that the ELISAs cut-off values could be calculated and sensitivity and specificity calculated. However unfortunately in this study none are available.

6.2.2 Nickle coated plates

One negative element in the current ELISA process is the need for the recombinant N proteins to be purified using IMAC using a 5mL HisTrap FF Crude column and subsequently concentrated using vivaspin column with an appropriate MCOW. The use of nickel-coated 96-well plates could eliminant the need to purify and concentrate the protein before using it to coat the plate; making the overall ELISA procedure more time-effective and potentially reducing any loss to the yield that the various processes each risk. Pierce™ Nickel Coated Plates (Fisher Scientific) were purchased. The nickel surface of the plates enables metal-chelate binding of polyhistidine-tagged proteins. Detergents used in the lysis of the cells do not inhibit binding as they do with plain polystyrene. The crude bacterial lysate containing the polyhistidine-tagged fusion proteins can be added directly to the plates with no need for prior purification or concentration. OC43-CoV was selected at random for the nickel coated plate to be used upon to see if it proved beneficial, results shown in Figure 47.

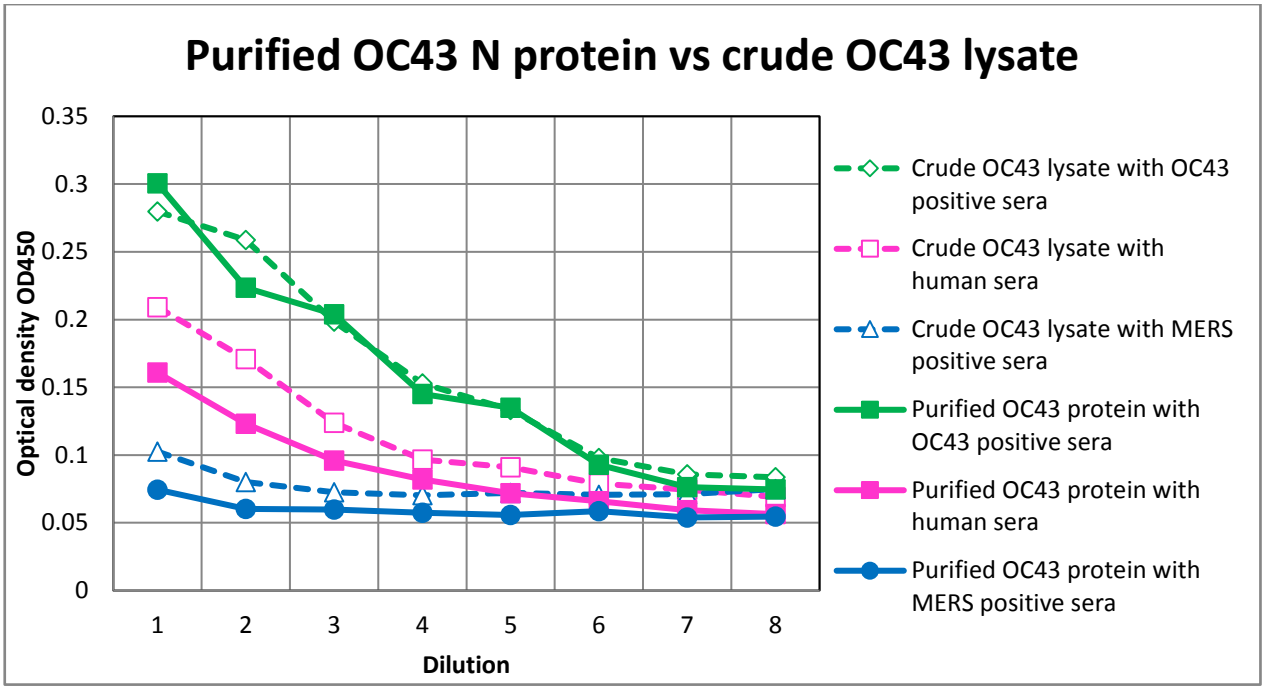


Figure 47: ELISA results showing known positive OC43-CoV, know positive MERS-CoV and an unknown human sera being screened on both purified OC43-CoV recombinant N proteins and crude induced cell lysate coated on the 96-well plate. Sera used is colour coded; OC43-CoV in green, MERS-CoV in blue and the human sera in pink. The results seen when the sera were screened with the crude lysate coating the plate are shown by hollow markers and dashed lines. The results seen when the sera were screened with purified N protein are shown with full markers and straight lines. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. As such 1= 1:50, 2=1:100, 3=1:200, 4=1:400, 5=1:800, 6=1:1,600, 7=1:3,200 and 8=1:6,400. Error bars not shown due to lack of repeats. The results show reactivity to the OC43-CoV sera as expected and the human sera shows some reactivity. The crude lysate was shown to produce a similar pattern to that of the purified lysate.

The results shown in Figure 47 show that the crude lysate is able to produce similar patterns to the purified N proteins. The true positive readings shown in green by the OC43-CoV remain the closest in values, and in the case of the human and MERS-CoV-positive sera the crude lysate produces slightly higher OD_{450} readings. This may be due to higher concentration of any potential contaminants that would cause a false positive result. However, the point remains that the crude lysate produces similar readings to that of the purified sample with a much quicker and easier production process. The cost of the nickel-coated plates would need to be off-set against the cost of the his-tag columns, buffers, viva spin columns and process time taken and as such would most likely be more cost-effective.

6.2.3 Pre-exclusion incubation

Another potential change to the pre-existing ELISA method used in this report would be pre-exclusion of the sera being screened against all antigens/ recombinant N proteins aside from the one being investigated to minimise any non-specific results. Any antibodies that carry out non-specific binding in the sera would bind to the recombinant N protein mixture added to the sera, preventing them from binding to the specific N protein coated on to the plate being screened. This should reduce non-specific background levels. However, this was unable to be successfully conducted in the laboratory, possibly due to the age of the sera tested as a positive result was unable to be achieved with or without the pre-exclusion step, data not shown. Future work could look into this as a principle with a fresh positive set of sera to screen with.

6.2.4 Summary

The results shown in this chapter demonstrate that the recombinant N proteins generated are sufficiently antigenic to be able to proceed with the screening unknown sera samples. Although the nickel-coated plate has been shown to be effective both in terms of cost and time saving, the previously generated and purified N protein are of sufficient quantity to use and therefore there is no need to purchase additional nickel-coated plates.

6.3 ELISAs conducted on unknown sera: diabetic and non-diabetic sera

6.3.1 Introduction

Having generated soluble versions of all six CoV recombinant N-proteins to a reasonable level of purification, the next step essential to their sue as a diagnostic reagent was to test their effectiveness as coating antigens by conducting a small-scale serum screening by ELISA. Sera samples were taken from eight colleagues working in the laboratory. As this was a limited experiment with few test subjects, no records were made of the age or gender of each person. However, it is worth noting that subject 1 is an immunocompromised individual and that subject 5 is unusual in that the individual has had a multi-national upbringing. The simple hypothesis under test here was that the recombinant N proteins would be a suitable protein set to distinguish recent CoV infection. That is, that different reactivity would be found per individual, consistent with random exposure to a number of viruses in the general population.

6.3.2 Materials and methods

6.3.2.1 Serum preparation

Whole blood samples were collected using single-use finger lancing devices on washed hands. The blood was left at RT for thirty minutes to allow it to clot. The clot was removed by centrifugation at 13,000 rpm for 10 minutes at 4^{o}C. The resulting supernatant is designated the serum and was immediately pipetted into a clean polypropylene tube and kept at 4^{o}C until needed.

6.3.2.2 Plate set up

The ELISAs were run as detailed in section 2.2.7. The sera underwent a 2-fold dilution series, recorded on the x axis of the data graph as 1/[log2] and was started at a dilution of 1:50 with eight subsequent dilutions - a range of 1:50 to 1: 6,400.

6.3.3 Results

The ELISAs were run in duplicate and the graphs plotted use the mean of the two, to reduce the impact of any anomalous results. Discriminating between non-infected and infected individuals

(D=0 and D=1 respectively) requires a discrimination or cut-off value (c) to be ascertained to define positive and negative test results. Assuming that a higher marker value would be associated with infection, individuals with a diagnostic test value (T) equal to, or higher than, c would be classified as infected (positive test T+) whereas individuals with a test value lower than c would be classified as non-infected (negative test T-). Ideally the test would correctly identify an infection individual (true positive, TP) and a person disease-free (true negative, TN), however classifying results in this manner is not free from error as the test may classify a healthy person incorrectly (false positive, FP) or it may fail to detect a diseased individual (false negative, FN). Errors of classification would need to be quantified before a diagnostic test could be routinely applied (López-Ratón, Rodríguez-Álvarez et al. 2014). This quantification of accuracy is usually expressed by the terms sensitivity (Se) and specificity (Sp) previously discussed in section 1.2.5.1.

There are several methods that can be used to help determine optimal cut-off values however most require known positive and negative individuals in order to ascertain the cut-off value that will best discriminate between the two (López-Ratón, Rodríguez-Álvarez et al. 2014). This is particularly challenging where the infection is common as finding a true negative is difficult. Moreover, it cannot be ruled out in closely related infections, such as is the case with the human coronaviruses, that a level of cross reactivity would occur, further complicating the choice of a cut-off value.

A general formula for a cut off value is:

$$Cut\ off = a.\bar{X} + f.SD$$

Whereby $\bar{X}$ is the mean and SD is the standard deviation of the independent negative control readings and a and f are two multipliers which can be set arbitrarily depending on the test design (Lardeux, Torrico et al. 2016). If a=2 and f=0 then the role of the standard deviation is not included in the cut-off value, as such it would be calculated as twice the mean absorbance obtained from the negative control/s.

Other methods for determining cut-off values are more complex with some using the (*1-α*)th percentile of the one-tailed Student *t*-distribution with (*j-1*) degrees of freedom (Frey, Di Canzio et al. 1998) or the additional use of positive controls alongside negative ones (Pan, Rosenberg et al. 1992). Alternatively, there are a variety of software packages that can be used; DiagnosisMed (Luo and Xiong 2012), pROC (Robin, Turck et al. 2011) and Epi (Carstensen and Hills 2013) and OptimalCutpoints (López-Ratón, Rodríguez-Álvarez et al. 2014).

In the case of human coronavirus infections, it is reasonable to suppose that no reactivity would be expected for IBV as this virus does not circulate in humans. Hence any reactivity to IBV N could be considered a bottom threshold for human serum reactivity with proteins of this type, that is, RNA binding proteins with shared properties of charge and molecular mass. Accordingly, the cut-off values were worked out to 3 significant figures for each serum using IBV as a negative value at 1:50 dilution and setting *a*=2 and *f*=0. Any CoV with OD450 readings equal to, or greater than, *c* is displayed in Table 13.

Sera number	Cut-off value (*c*)	Positive test (T+)	*C* calculated at 2nd dilution	T+ *C* calculated at 2nd dilution
1	0.30	229E-CoV and NL63-CoV	N/A	N/A
2	0.40	None	N/A	N/A
3	0.18	OC43-CoV and NL63-CoV	0.13	None
4	0.35	229E-CoV and NL63-CoV	N/A	N/A
5	0.32	None	0.25	SARS-CoV
6	0.40	NL63-CoV	N/A	N/A
7	0.27	None	N/A	N/A
8	0.20	None	N/A	N/A

Table 13: Cut-off values based on the 8 ELISAs conducted and subsequent positive results for 8 sera screened

In two examples (serum 3 and serum 5) human judgement deemed it necessary to observe results seen at subsequent dilutions and so the cut-off values were calculated using the second dilution 1: 10^2. Using twice the IBV negative control values seen at this dilution, showed serum 5 to in fact

have a positive result with SARS-CoV and discounted the OC43-CoV and NL63-CoV positive results that serum 3 first appeared to have.

However, although SARS-CoV circulated in the population following its emergence on 2003, that circulation was restricted, and the virus has since become extinct, so the apparent positivity was observed in one individual (serum 5) is suspect. MERS-CoV positivity was not seen in any of the sera, but NL63-CoV and 229E-CoV positivity were both observed in two individuals (sera 1 and 4, Figure 48 and Figure 51 respectively). OC43-CoV positivity was not observed in any individuals.

The graphs shown from Figure 48 through to Figure 55 show the ELISA results per each individual serum along with the cut-off value.

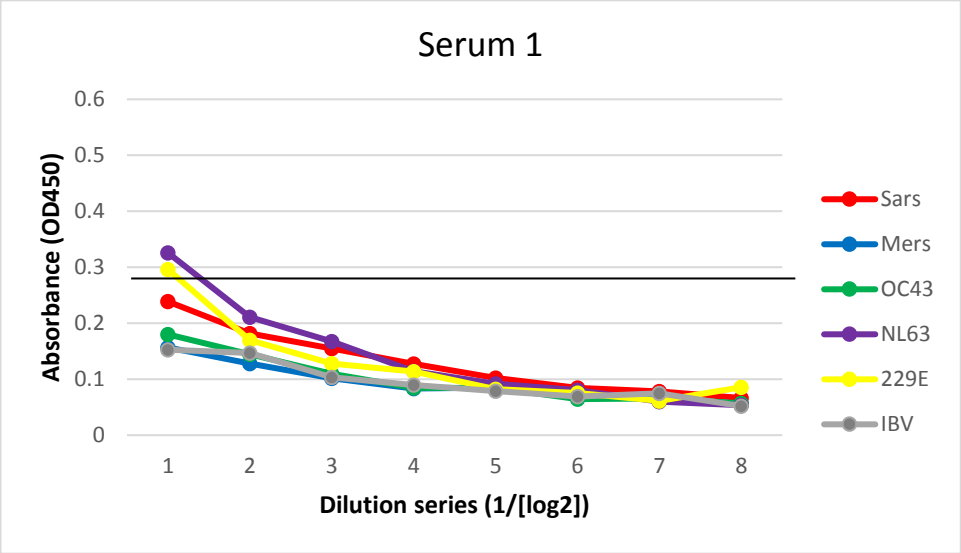


Figure 48: ELISA results of sample 1, human serum reactivity on recombinant CoV proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started at 1:50 and followed a two-fold dilution series. The line y=0.30 indicates the suggested cut-off value. Error bars not shown due to lack of repeats.

The ELISA results shown in Figure 48 suggest low level of reactivity on all proteins with possible prior infection with 229E-CoV and NL63-CoV consent with the highest reactivity to these proteins. However maximum OD value are low.

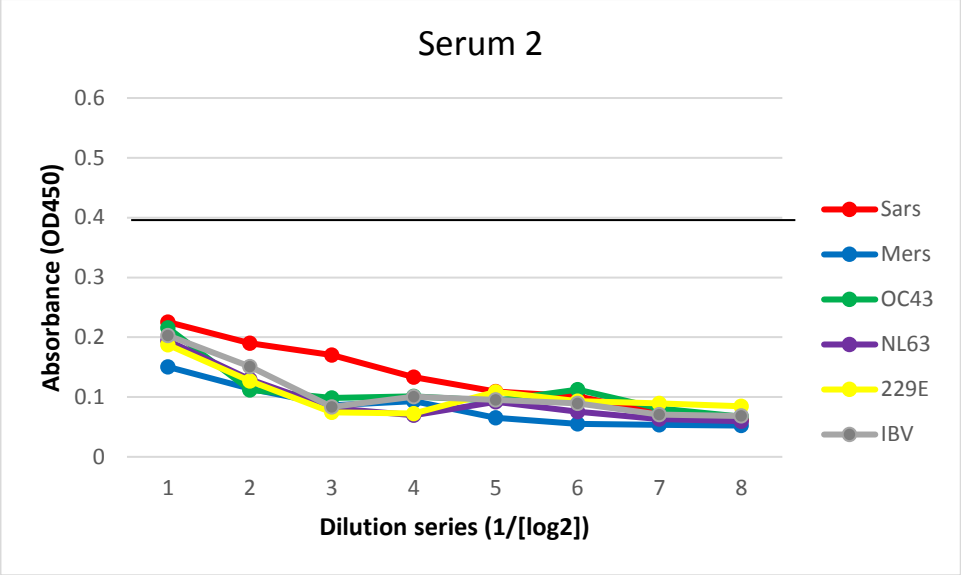


Figure 49: ELISA results of sample 2, human serum reactivity on recombinant CoV proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The line y=0.40 indicates the suggested cut-off value outside of the graph's plot area. Error bars not shown due to lack of repeats.

A low level of absorbance (OD450 <0.25) was also shown by the ELISA results for serum two, Figure 49. Reactivity on SARS-CoV N protein was the highest, but the bunching of the data does give confidence that this is anything more than a high background reaction. The pattern of antigen recognition overall suggests true HCoV infection to be unlikely.

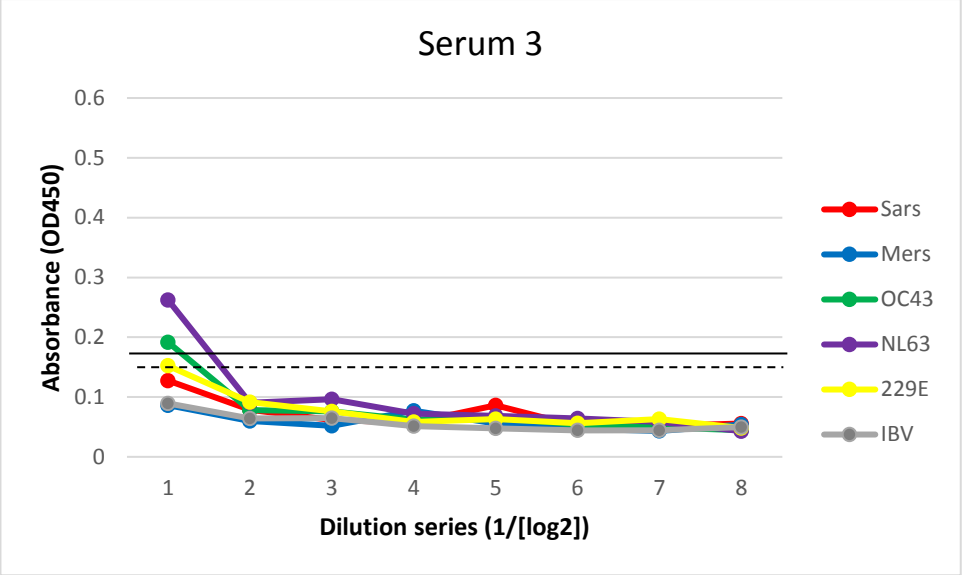


Figure 50: ELISA results of sample 3, human sera against recombinant CoV proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The solid line y=0.18 indicates the suggested cut-off value based on the first dilution series' values, the dashed line y=0.13 shows the cut-off value based on the values in the second dilution. Error bars not shown due to lack of repeats.

Serum 3 also showed poor overall reactivity with the set of N proteins, Figure 50, with the highest reactivity with the N protein of NL63-CoV. However, the titration fails to follow the dilution curve again suggesting probably background.

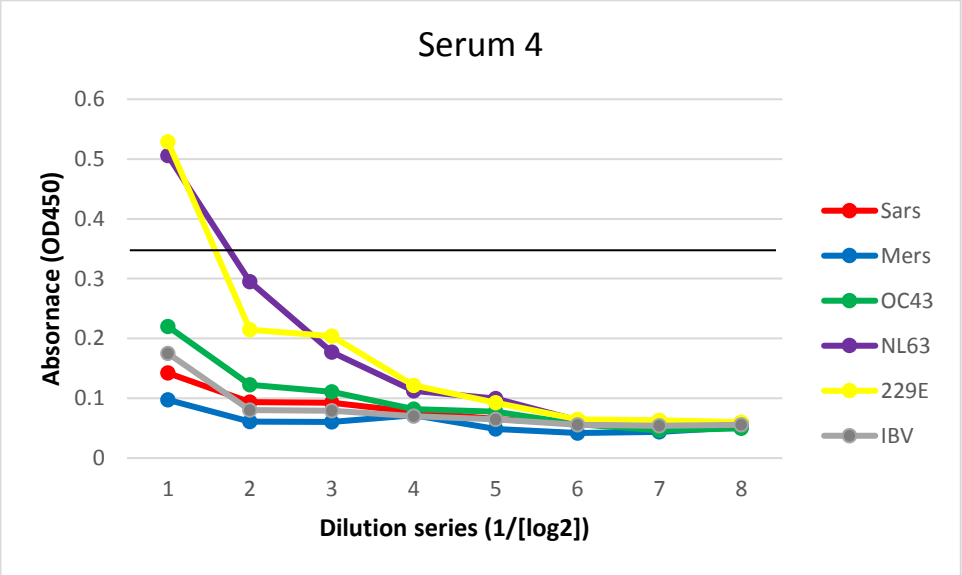


Figure 51: ELISA results of sample 4, human sera against recombinant CoV proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The solid line y=0.35 indicates the suggested cut-off value. Error bars not shown due to lack of repeats.

In contrast to serum samples 1-3, the ELISA result for serum sample 4, Figure 51, appear more convincing. The maximum OD is significantly higher for reactivity with the N proteins of both 4 229E-CoV and NL63-CoV and the sera titres follow the dilution series, clearly remaining above reactivity on the other N proteins to a significant dilution (1: 400).

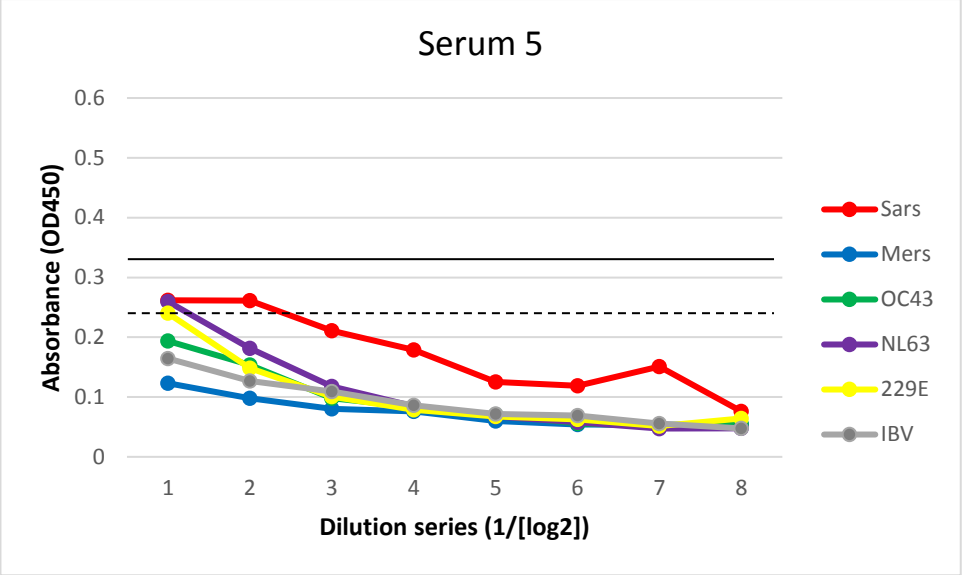


Figure 52: ELISA results of sample 5, human sera against recombinant CoV proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The solid line y=0.32 indicates the suggested cut-off value based on the first dilution series' values, the dashed line y=0.25 shows the cut-off value based on the values in the second dilution. Error bars not shown due to lack of repeats.

Serum 5 was unusual in that the predominant reactivity was to SARS-CoV N protein, Figure 52. However, the titration does not follow the dilution series and overall reactivity was low. Individual 5 is a foreign national who had visited a town which had known SARS cases. However, the individual is also asthmatic and prone to respiratory infections. Overall, the reactivity to SARS-CoV N appeared anomalous and unlikely to be valid.

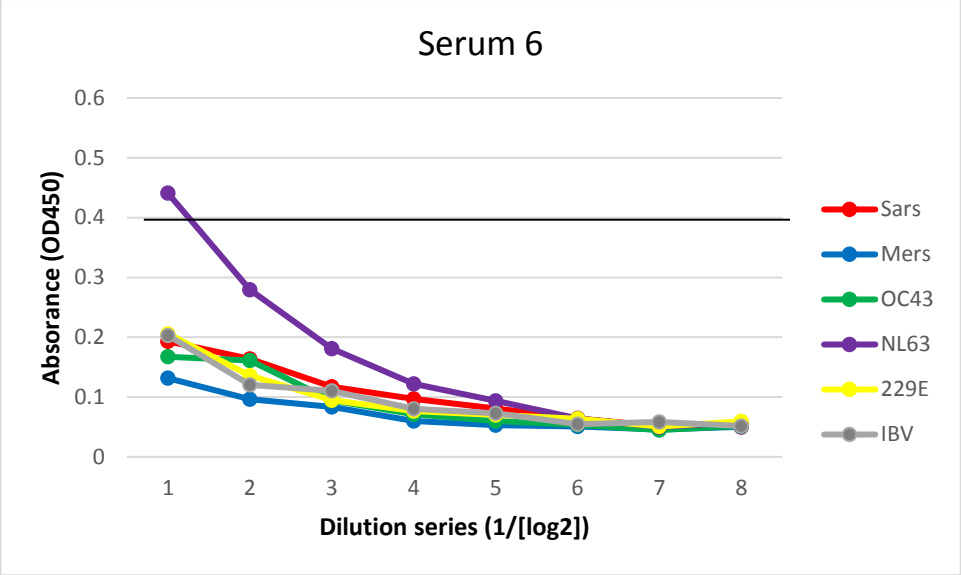


Figure 53: ELISA results showing sample 6 human sera against recombinant CoV proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The solid line y=0.40 indicates the suggested cut-off value. Error bars not shown due to lack of repeats.

Serum sample 6, the ELISA data presented n Figure 53 show convincing evidence for recent prior infection with NL63-CoV. The maximum OD values are reassuringly high and follow the titration curve. In addition, general reactivity on remaining CoVs show only low absorbance levels.

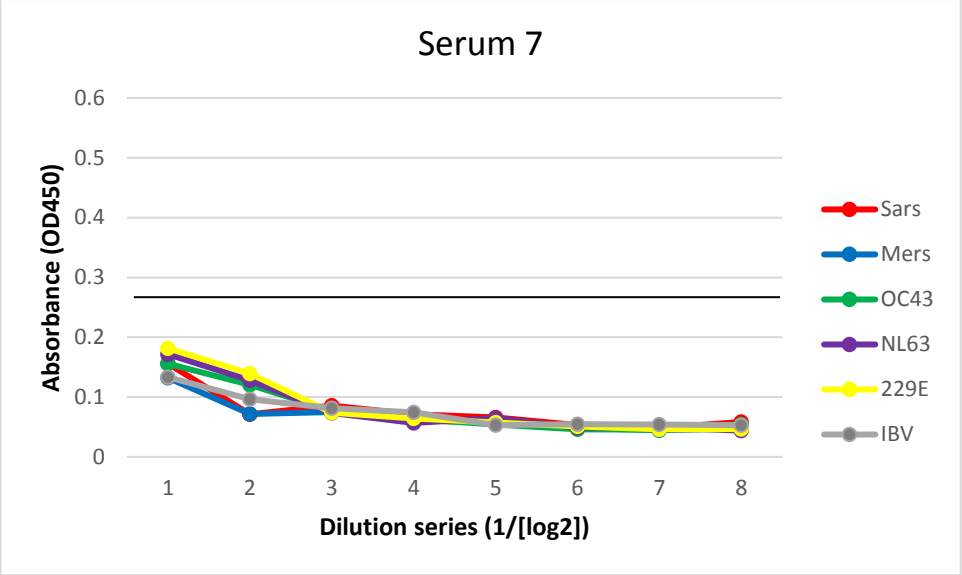


Figure 54: ELISA results showing sample 4 human sera against recombinant CoV proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The solid line y=0.27 indicates the suggested cut-off value outside of the graph's plot area. Error bars not shown due to lack of repeats.

Serum 7 had very low OD readings throughout (OD450 <0.2) the ELISA, Figure 54. There was no convincing reactivity to suggest recent prior HCoV infection.

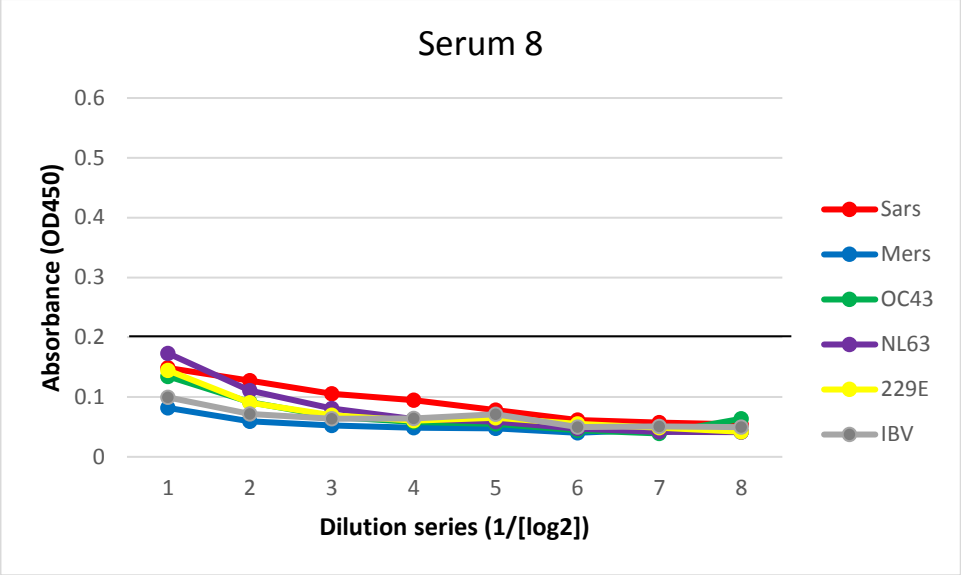


Figure 55: ELISA results showing sample 8 human sera against recombinant CoV proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The solid line y=0.20 indicates the suggested cut-off value. Error bars not shown due to lack of repeats.

Serum 8 was also generally poorly reactive with any N protein by ELISA, Figure 55. There was no indication of recent HCoV infection in this individual.

6.3.4 Discussion

The preliminary ELISAs results, using 8 serum samples of unknown status with regard to previous CoV infection indicate that the use of recombinant CoV N proteins as diagnostic antigens for serum survey is justified. The general level of background binding was low (OD=~0.2) and the shape of the binding curves by such sera generally did not follow the dilution series.

A true more statistical analysis of the data was not justified as the sample size tested would need to be far greater in order to improve significance.

Other comparisons between the data sets would also need to be made e.g. if incidence rates in males compared to females, or above the age of "x" verses under. Of the 8 samples, serum 4 is suggested to have had prior infection with 229E-CoV and NL63-CoV. Interestingly these viruses are more closely related to each other than to any other human coronavirus suggesting serum cross-reactivity could be possible. However, serum 6 showed clear reactivity to only NL63-CoV indicating that seroconversion to only this virus is possible. It has since been shown that more complex "double sandwich" ELISA assays may be required to distinguish between these activities (Sastre, Dijkman et al. 2011).

Nevertheless, recombinant CoV N proteins were able to differentiate between prior CoVs infection or not and so a slightly larger sera set and variables were examined as shown in Section 6.3.5.9.1, which details the differences in N protein coronavirus reactivity observed between sera of type one diabetics compared to non-type one diabetics and chapter eight details the differences in N protein coronavirus reactivity observed among sera previously screened for influenza infection.

6.3.5 Comparing type one diabetic sera to non- type one diabetic sera

6.3.5.1 Introduction

Type 1 diabetes mellitus (DM) (T1DM) is a lifelong autoimmune condition that arises from a lack of insulin production due to autoimmunity damage and the loss of pancreatic islet β- cells; resulting in increased blood glucose levels (hyperglycaemia) (Katsarou, Gudbjornsdottir et al. 2017). It is estimated that 5-10% of all people living with diabetes are classified as having T1DM (Maahs, West et al. 2010) whereas the more common type of diabetes, due to a combination of resistance to insulin action and inadequate compensatory insulin secretory response, is type 2 diabetes mellitus (T2DM). It is estimated that T1DM is diagnosed in almost 90,000 children a year world-wide; the incidence rate is increasing globally although it varies greatly between countries (Diaz-Valencia, Bougneres et al. 2015). The highest incidence rates are reported in Scandinavian countries, followed by other European countries (such as the UK), then North America and Australia. T1DM diagnosis is rare in some Asian countries such as China, Korea and Japan (Diaz-Valencia, Bougneres et al. 2015). Differences in incidence rates may be linked to genetic susceptibility and environmental and lifestyle factors, including hygiene and childhood infections. The high incidence recorded in developed countries has led to a "hygiene hypothesis" that suggests lack of exposure to infectious agents increases risk and prevalence of autoimmune conditions (Borchers, Uibo et al. 2010). However, incidence rates may appear low in less economically developed counties in part due to a lack of diagnostic resources as opposed to a true reduced incidence rate.

6.3.5.2 Diabetes and virology

Onset of T1DM within populations often follows a seasonal pattern which has suggested a viral aetiology (Adams 1926). A variety of viruses have been examined as potential inducers of T1DM; cytomegalovirus (CMV) (Pak, Eun et al. 1988), parvovirus (Guberski, Thomas et al. 1991, Kasuga, Harada et al. 1996), encephalomyocarditis virus (Craighead and McLane 1968), retroviruses (Conrad, Weissmahr et al. 1997), rotaviruses (Honeyman, Coulson et al. 2000), rubella (Gale

2008) and mumps (Goto, Takahashi et al. 2008). However, none of the associations have been proven but the differing ages of diagnosis in monozygotic twins do suggest that nongenetic variables must also play a role in T1DM pathogenesis (Redondo, Jeffrey et al. 2008).

6.3.5.3 Enteroviruses

The most robust association with viruses and T1DM involves enterovirus species; single-stranded positive sense RNA (ssRNA) viruses belonging to the picornavirus family (Katsarou, Gudbjornsdottir et al. 2017) (Yeung, Rawlinson et al. 2011) (Oikarinen, Martiskainen et al. 2011). Some strains of enterovirus have been shown to induce or accelerate T1DM in animal models (Coppieters, Boettler et al. 2012). Epidemiological and clinical studies have supported the role of enterovirus, especially Coxsackie B4 and B3 virus, in the development of T1DM in genetically predisposed individuals (Flyvbjerg 2010) (Clements, Galbraith et al. 1995) (Banatvala, Bryant et al. 1985) (Gamble, Kinsley et al. 1969). The rate of progression from islet autoimmunity to T1DM has been shown to be significantly increased following detection of enteroviral RNA in serum (Stene, Oikarinen et al. 2010). There is thought to be an association between the occurrence of T1DM and peaks of enterovirus infection and other evidence linking the two include; the detection of anti-enterovirus antibodies, enterovirus RNA, and the capsid protein VP1 in blood, small intestine biopsies, and autopsy pancreas specimens of individuals with T1DM (Hober and Sane 2010). Such samples are difficult to obtain however as the target organ in T1DM, the pancreas, has a fairly inaccessible anatomical location (Coppieters and von Herrath 2009) and pancreas samples from recently diagnosed T1DM individuals are limited due to improved clinical management of the condition.

Various mechanisms can explain the role of enterovirus in the pathogenesis of T1DM as summerised by (Stene, Oikarinen et al. 2010): *"(1) persistent infection of pancreatic beta cells provoking cell damage and release of sequestered antigens inducing an autoimmune response; (2) molecular mimicry (partial sequence homology) between the 2C viral protease and the epitope GAD65 (Glutamic Acid Decarboxylase) which is a major target antigen in T1DM (Kaufman, Erlander*

et al. 1992) and between the VP1 viral capsid protein and the IA2 protein; (3) bystander activation of autoreactive T cells; (4) thymus infection; and (5) loss of regulatory T cells." However, a direct causative correlation has not been shown for any of these possibilities and an active mechanism still needs to be established. In contrast, some studies report no such correlations suggesting that the notion that any potential association with virus infection is not absolute but depends on genetic susceptibility or other environmental factors (Fuchtenbusch, Irnstetter et al. 2001, Graves, Rotbart et al. 2003).

6.3.5.4 Immunity in T1DM

Most researchers conclude that there is clinical evidence pointing to the higher prevalence of infectious diseases among individuals with DM (Geerlings and Hoepelman 1999, Muller, Gorter et al. 2005, Vardakas, Siempos et al. 2007). The main pathogenic mechanisms involved in DM infections are: the hyperglycaemic environment which may increase the virulence of some pathogens; lower production of interleukins in response to infection; reduced chemotaxis and phagocytic activity, immobilization of polymorphonuclear leukocytes; glycosuria, gastrointestinal and urinary dysmotility (Joshi, Caputo et al. 1999).

6.3.5.5 Respiratory infections and T1DM

Respiratory tract infections are responsible for a significant number of medical appointments by persons with DM compared to those without DM (Barros, Cartagena et al. 2005, Peleg, Weerarathna et al. 2007, Cano, Iglesias et al. 2010). The most frequent respiratory infections associated with DM are caused by *Streptococcus pneumoniae* and influenza virus (Joshi, Caputo et al. 1999, Barros, Cartagena et al. 2005). It is estimated that individuals with DM are six times more likely to need hospitalization during influenza epidemics than non-diabetic patients (Peleg, Weerarathna et al. 2007). Individuals with diabetes are also at a higher risk of contracting tuberculosis than individuals without DM (Harries, Lin et al. 2011, Restrepo, Camerlin et al. 2011).

6.3.5.6 Coronavirus and T1DM

Studies examining risk factors associated with MERS-CoV infection often include DM, although no reference is made to distinguish between type1 and type2 (Alraddadi, Watson et al. 2016). Similarly DM has been noted as being a risk factor during SARS-CoV infection (Booth, Matukas et al. 2003) (Chan, Ng et al. 2003). Plasma glucose levels and diabetes have been reported as independent predictors for mortality and morbidity in SARS patients (Yang, Feng et al. 2006).

6.3.5.7 Experimental aim and hypothesis

Although the causative virological agent for the initiation of T1DM is more likely to be one of those described, for example an enterovirus, as opposed to a coronavirus, the immunocompromising effects that hyperglycaemia and T1DM brings could mean a correlation between T1DM and CoV infection still exists. As a small serum set from six confirmed T1DM was available, it was screened for CoV N reactivity as before and the results compared with those of six non-T1DM individuals working in the laboratory. The hypotheses under test were:

H0: There is no association between diabetic and non-diabetic status and coronavirus recombinant N-protein reactivity (the proportion is the same for T1DM individuals vs non-diabetic individuals)

H1: There is an association between diabetic/ non-diabetic status and coronavirus recombinant N-protein reactivity (the proportion is different for T1DM individuals vs non-diabetics individuals)

6.3.5.8 Materials and methods

ELISAs were conducted as described in 2.2.7.

6.3.5.9 Results

Cut-off values (c) were determined per sera using the formula:

$$Cut\ off = a.\bar{X} + f.SD$$

Whereby $\bar{X}$ is the mean and SD is the standard deviation of the independent negative control readings and a and f are two multipliers which can be set arbitrarily. For this study, the cut-off values were set to 3 significant figures for each serum using IBV as a negative value at 1:50 dilution and setting a=2 and f=0. A line has been drawn on each graph showing Y=c.

6.3.5.9.1 Diabetic vs non-diabetic sera screened on all six purified N proteins displayed per individual

6.3.5.9.2 Non-diabetic sera screening displayed per individual

Figure 56 shows the serum from non-diabetic individual 1 to show sign of infection against SARS-CoV. This is the same individual whose results were displayed in Figure 52 previously. The results from non-diabetic individual number 2 show no sign of prior CoV infection as all values remain below the cut-off value shown. Prior infection with OC43-CoV may be likely in the non-diabetic individual 3 sample. Non-diabetic individual number 4 shows no sign of prior CoV infection as all values remain below the cut-off value. The results for non-diabetic individual number 5 also all remain below the cut-off level, suggesting that prior CoV infection is unlikely. Finally, the results of non-diabetic individual number 6 suggest prior infection with both 229E-CoV and NL63-CoV at the first dilution series, using the data from the second dilution ($1:10^2$) only shows 229E-CoV reactivity.

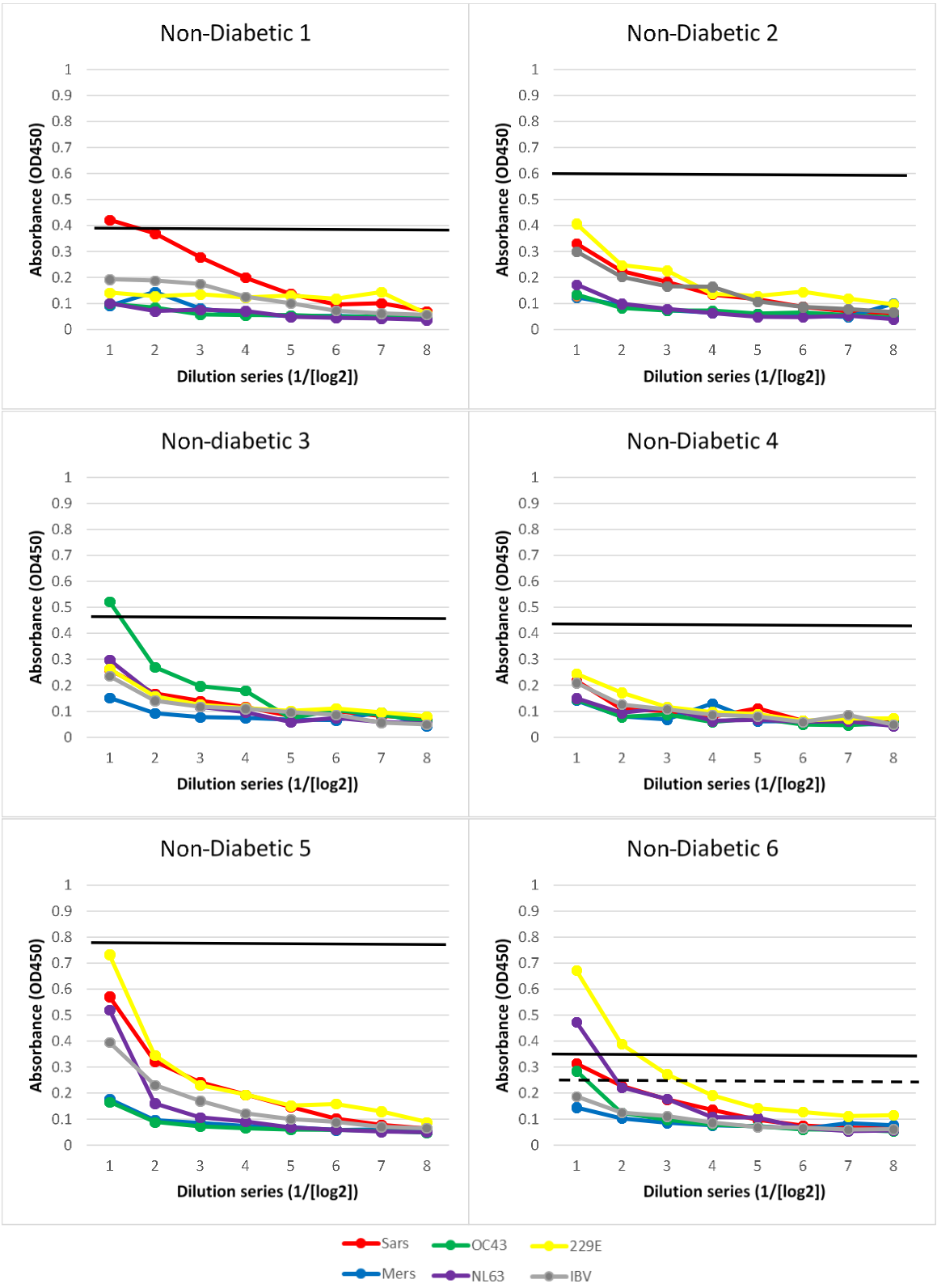


Figure 56: ELISA results of all non-diabetic sera against recombinant N proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The line seen on each graph denotes the cut-off value; for serum 1 y=0.38, for serum 2 y=0.60, serum 3 y=0.47, serum 4 y=0.42, serum 5 y=0.79 and for serum 6 the solid line seen on the graph denotes the cut-off value using data from the first dilution y=0.37, the dashed line y=0.25 is the cut-off value gained using data from the second dilution. Error bars not shown due to lack of repeats.

6.3.5.9.3 Diabetic sera screening displayed per individual

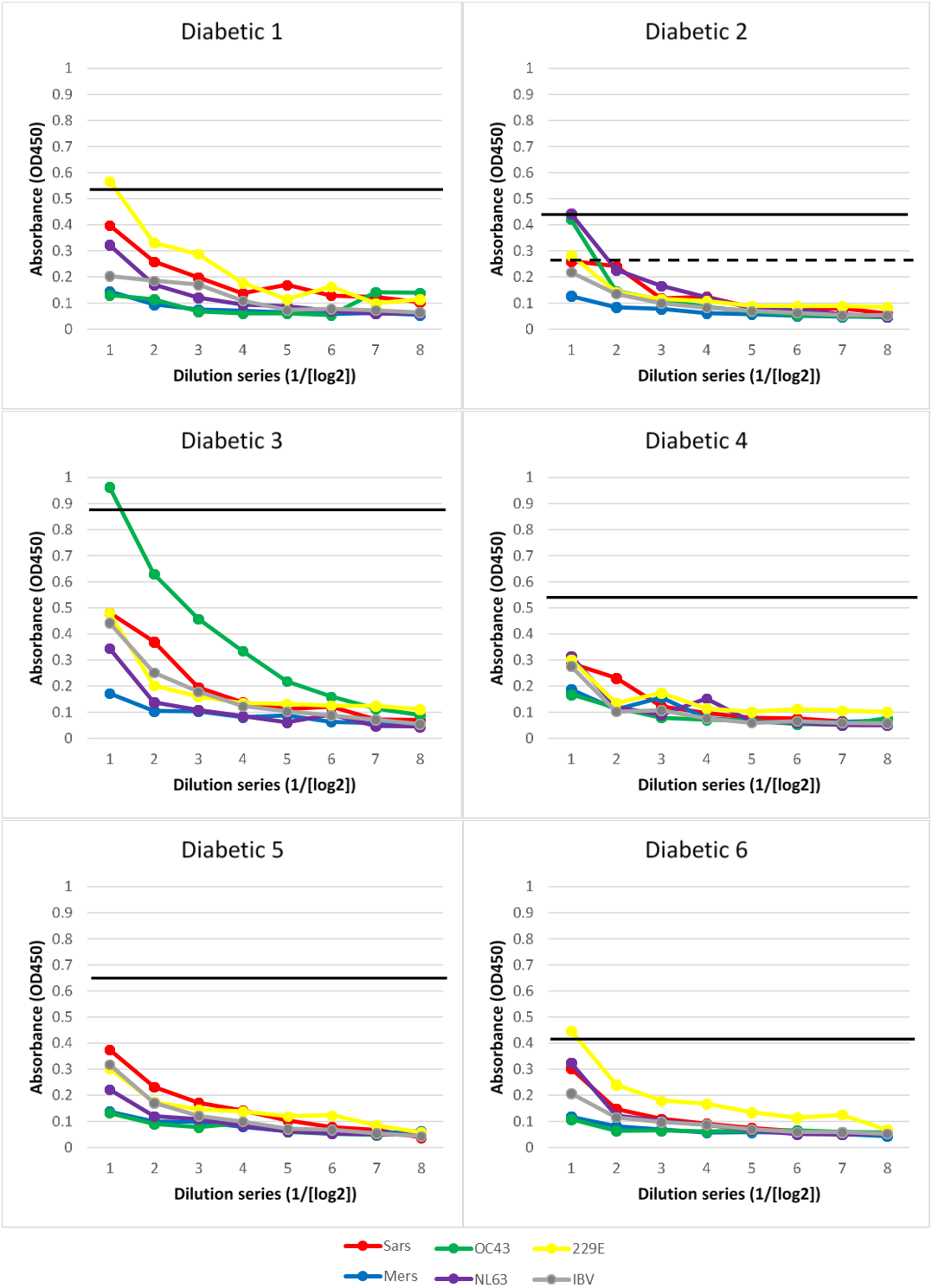


Figure 57: ELISA results of all diabetic sera against recombinant N proteins. The recombinant N protein used is colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. The line seen on each graph denotes the cut-off value; for serum 1 y=0.55, for serum 2 the solid line seen on the graph denotes the cut-off value using data from the first dilution y=0.44, the dashed line y=0.27 is the cut-off value gained using data from the second dilution, serum 3 y=0.89, serum 4 y=0.55, serum 5 y=0.64 and for serum 6 y=0.41. Error bars not shown due to lack of repeats.

The results from diabetic individual number one shown in Figure 57 indicate a prior 229E-CoV infection. The results of diabetic individual number 2 suggest prior infection with both NL63-CoV at the first dilution however the gradient steeply drops and using the data from the second dilution ($1:10^2$) shown on the graph with a dashed line, shows no CoV reactivity. The results from diabetic individual number three indicate a prior OC43-CoV infection. The results for diabetic individual number 4 all remain below the cut-off level, suggesting that prior CoV infection is unlikely. The results for diabetic individual number 5 also all remain below the cut-off level, suggesting that prior CoV infection is unlikely. Finally, the results from diabetic individual number six indicate a prior 229E-CoV infection.

Next the results are displayed based on recombinant-N protein with and without individual sera cut-off levels imposed and with diabetic status colour coded.

The only individual showing SARS-CoV reactivity shown in Figure 57 is a non-diabetic individual, this is an unusual case already mentioned in section 6.3.3.

6.3.5.9.4 Diabetic vs non-diabetic sera screened on all six purified N proteins displayed with colour coding for diabetic status

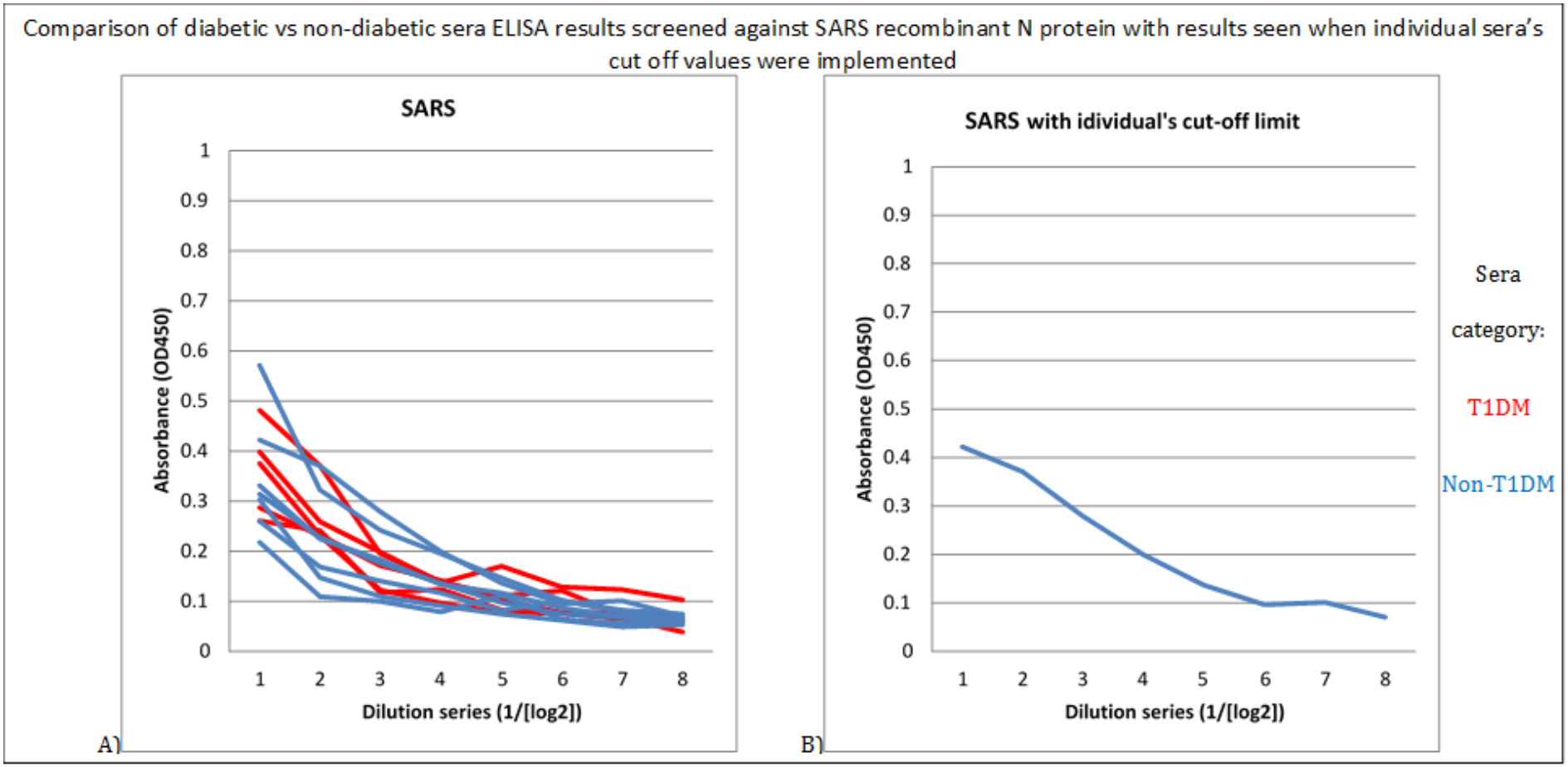


Figure 58: ELISA results comparing diabetic vs non-diabetic sera screened against recombinant SARS-CoV N protein with results seen when individual sera's cut off values were implemented. Sera obtained from non-diabetics are coloured blue, sera obtained from a diabetic are coloured red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 58 shows that once cut-off limits were imposed only one nondiabetic individual showed signs of prior infection to SARS-CoV.

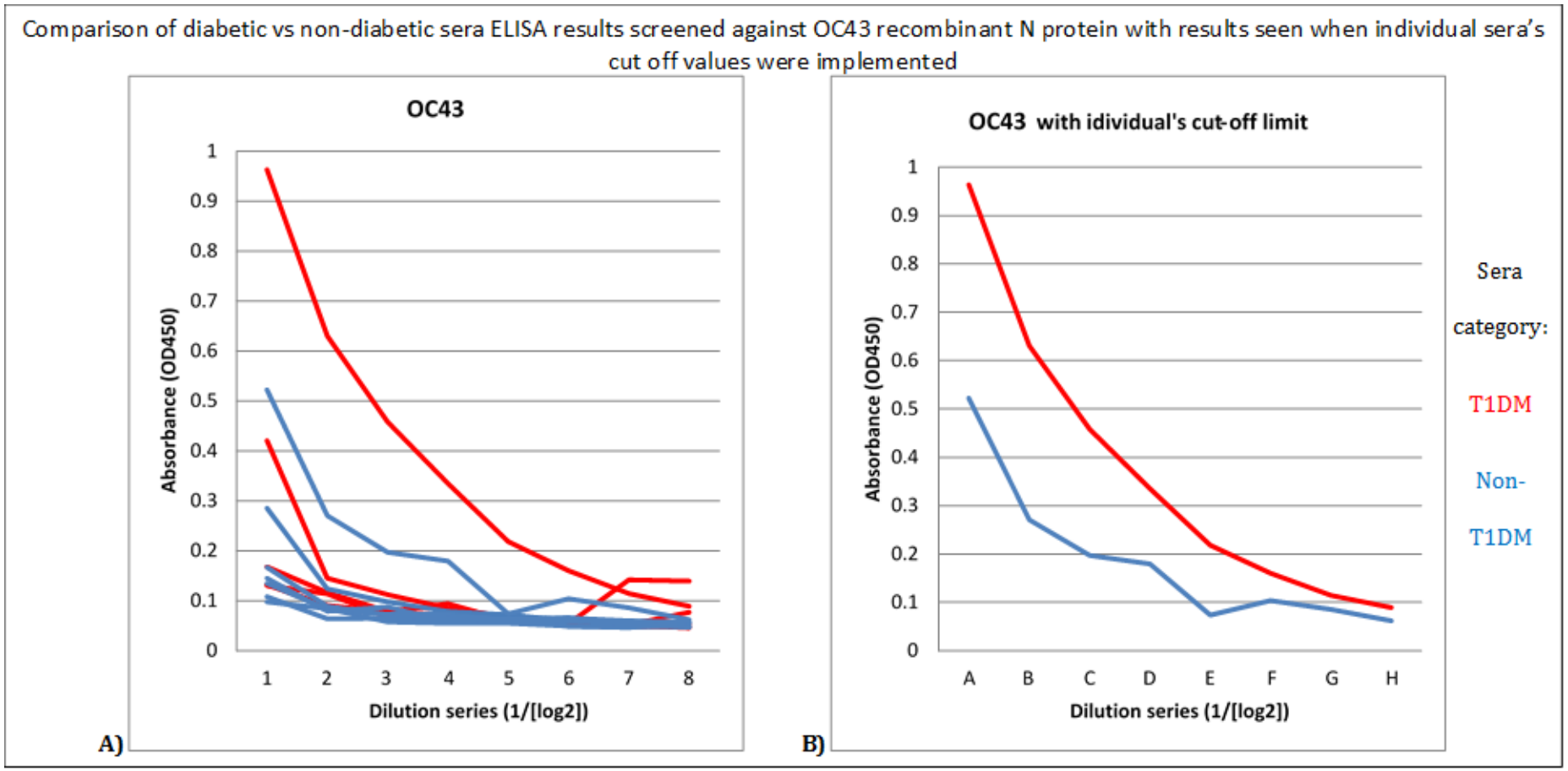


Figure 59: ELISA results comparing diabetic vs non-diabetic sera screened against recombinant OC43-CoV N protein with results seen when individual sera's cut off values were implemented. Sera obtained from non-diabetics are coloured blue, sera obtained from a diabetic are coloured red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 59 shows that once cut-off limits were imposed only one diabetic and one nondiabetic individual showed signs of prior infection to OC43-CoV.

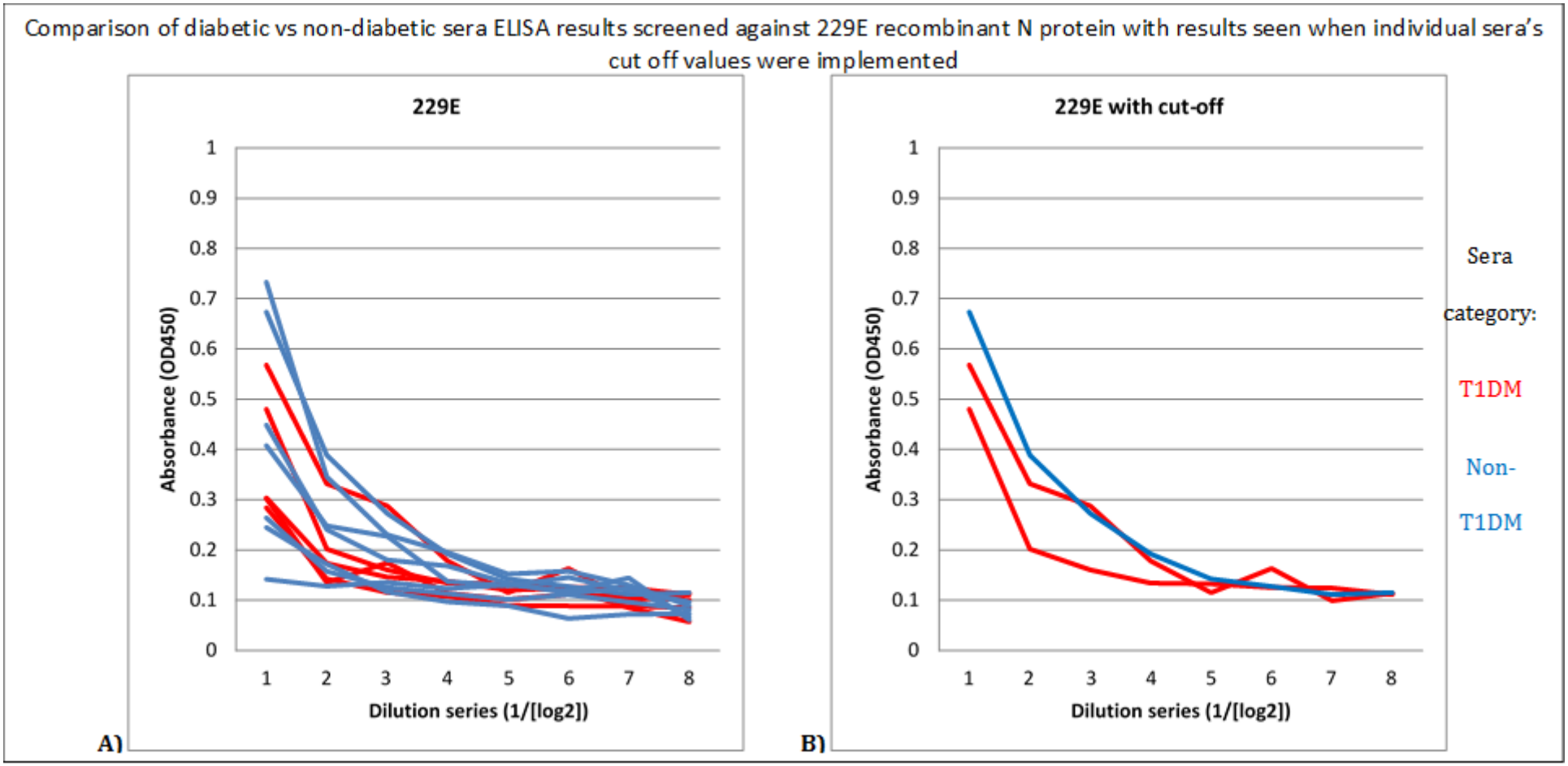


Figure 60: ELISA results comparing diabetic vs non-diabetic sera screened against recombinant 229E-CoV N protein with results seen when individual sera's cut off values were implemented. Sera obtained from non-diabetics are coloured blue, sera obtained from a diabetic are coloured red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 60 shows that once cut-off limits were imposed two diabetic and one nondiabetic individuals showed signs of prior infection to 229E-CoV.

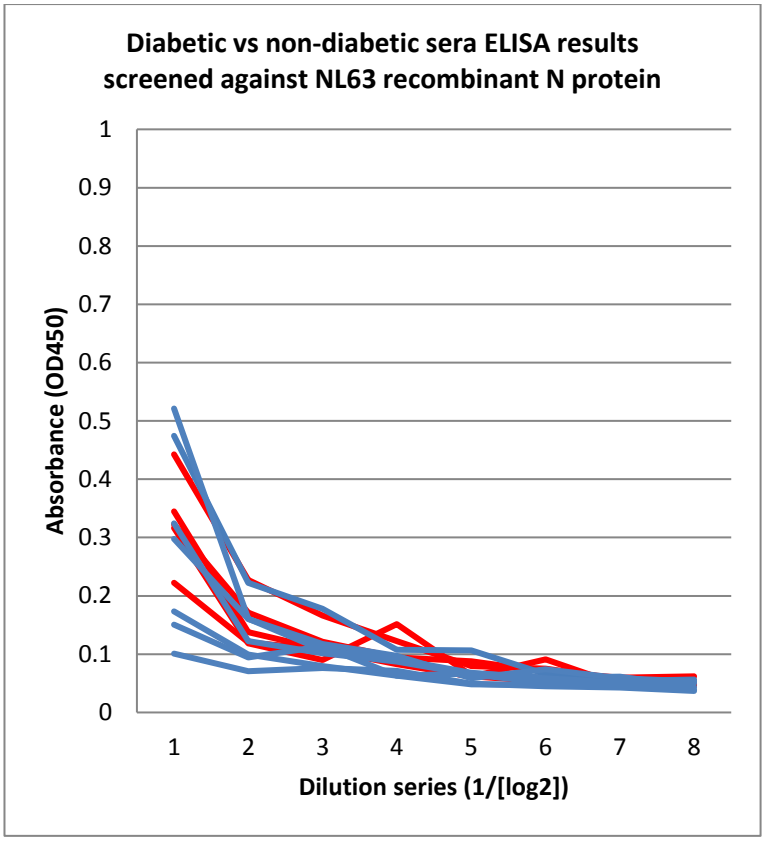


Figure 61: ELISA results comparing diabetic vs non-diabetic sera screened against recombinant NL63-CoV N protein. Sera obtained from non-diabetics are coloured blue, sera obtained from a diabetic are coloured red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

In the case of NL63-CoV shown in Figure 61, once cut-off limits were imposed no sera showed indication of prior infection.

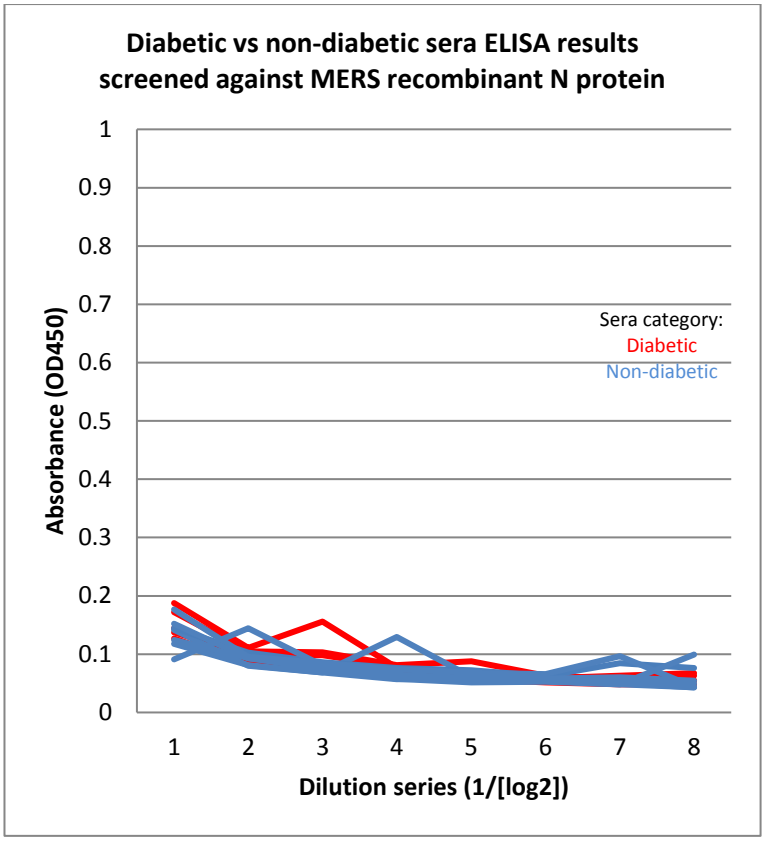


Figure 62: ELISA results comparing diabetic vs non-diabetic sera screened against recombinant MERS-CoV N protein. Sera obtained from non-diabetics are coloured blue, sera obtained from a diabetic are coloured red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

As expected in the case of MERS-CoV shown in Figure 62, once cut-off limits were imposed no sera showed indication of prior infection. Potentially the low OD readings seen in Figure 62 imply that using recombinant MERS-CoV N protein may have proven useful in determining cut-off limits, instead of, or as well as, recombinant IBV N protein, shown in Figure 63.

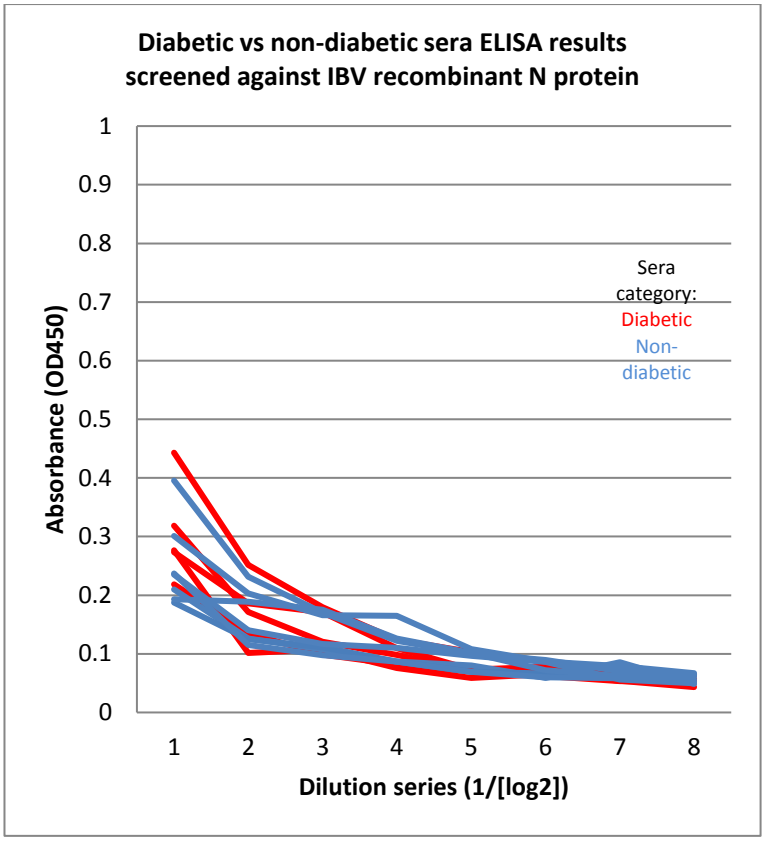


Figure 63: ELISA results comparing diabetic vs non-diabetic sera screened against recombinant IBV N protein. Sera obtained from non-diabetics are coloured blue, sera obtained from a diabetic are coloured red. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 63 shows the data obtained when screening sera from T1DM positive and negative individuals against recombinant IBV N protein, used as a negative control in order to ascertain cut-off values for the HCoV recombinant N proteins being screened. Interestingly values seen in dilutions 1 and 2 of IBV appear much higher than those seen in MERS-CoV, Figure 62, suggesting that a different cut-off value would have been calculated using data from the MERS-CoV recombinant N proteins as a negative control either instead of, or averaged with, IBV. Without being able to test known positive and negative sera though it is unable to see if this would be beneficial or not.

There were no technical problems encountered whilst conducting the trial. It does not appear that being a type one diabetic has any effect on likelihood to have a coronavirus infection as overall there were three positive results for each group; Non-diabetic individuals 1, 3 and 6 showed

reactivity to SARS-CoV, OC43-CoV and 229E-CoV respectively, and diabetic individuals number 1 and 6 showed 229E-CoV reactivity whilst diabetic number 3 showed OC43-CoV reactivity. Although there were twice as many diabetics as non-diabetics showing 229E-CoV reactivity in such a small data set it is unlike to prove statistically significant. In order to truly accept the null hypothesis statistical analysis needs to take place.

6.3.5.10 Statistical analysis of data

In order to know if the null hypothesis can be accepted or rejected the level of statistical confidence needs to be measured.

6.3.5.10.1 Selecting which statistical test to use

Binomial testing can be conducted on the data obtained as all results are either diabetic positive or negative and recombinant Coronavirus N protein reactivity positive or negative. Chi-square or G-tests could be used to test for statistical significance however these work best on a larger sample size (Kim 2017). Fisher's exact test of independence will be used instead.

6.3.5.10.2 Fisher's exact test

Fisher's exact test is used when there are two nominal variables and is easily conducted on 2x2 contingency tables whereby rows (R) contain data sets and columns (C) contain outcomes (McDonald 2014). The null hypothesis states that the relative proportions seen in one variable are independent of the second variable, so that the probability of getting the observed data are the same for both groups. In the two screenings done in this report the null hypotheses would be that being T1DM has no correlation to the likelihood of a coronavirus infection, and in the second case it would be that prior influenza infection has no effect on the likelihood of a coronavirus infection. In contrast, the hypothesis would be that there is a correlation and T1DM status does impact the coronavirus reactivity results seen, either in a causative suggestion or a preventative suggestion, although the test simply states if such correlation exists.

The test calculates the probabilities of all possible combinations of numbers in an R×C contingency table, then adds the probabilities of those combinations that are as extreme or more extreme than the observed data. The test better suits smaller data sets as when R and C get larger, and as the total sample size increases, the number of possible combinations dramatically increases, and the time taken for a computer to calculate the results may be unreasonable.

	Data set 1	Data set 2	Row total
Outcome 1	a	b	a + b
Outcome 2	c	d	c + d
Column total	a + c	b + d	a + b + c + d (=n)

Table 14: Example of a 2x2 contingency table

Fisher showed that the probability of obtaining any such set of values was given by the hypergeometric distribution, this is a discrete probability distribution used to describe the probability of successes in a fixed number of draws, without replacement, from a finite population of size, whereby each draw is either a success or a failure (Fisher 1922, Rivals, Personnaz et al. 2007).

$$p = \frac{\binom{a+b}{a}\binom{c+d}{c}}{\binom{n}{a+c}} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{a!b!c!d!n!}$$

The above equation is used to calculate the p value, $\binom{n}{k}$ is the binomial coefficient and the symbol "!" indicates the factorial operator, the letters correspond to the numbers that would be observed in Table 14. Often the result is calculated using a spreadsheet or a website, in this study GraphPad Software was used (GraphPad 2017). Two-tailed testing was carried out, which is where the probability of getting deviations as extreme as the observed, but in the opposite direction, is also calculated.

6.3.5.10.3 2x2 contingency tables for each coronavirus recombinant N protein

SARS-CoV			
	Reactivity	No Reactivity	Total
Diabetic	0	6	6
Non-Diabetic	1	5	6
Total	1	11	12
MERS-CoV			
	Reactivity	No Reactivity	Total
Diabetic	0	6	6
Non-Diabetic	0	6	6
Total	0	12	12
NL63-CoV			
	Reactivity	No Reactivity	Total
Diabetic	0	6	6
Non-Diabetic	0	6	6
Total	0	12	12
OC43-CoV			
	Reactivity	No Reactivity	Total
Diabetic	1	5	6
Non-Diabetic	1	5	6
Total	2	10	12
229E-CoV			
	Reactivity	No Reactivity	Total
Diabetic	2	4	6
Non-Diabetic	1	5	6
Total	3	9	12

Table 15: The six 2x2 contingency tables for each recombinant coronavirus N protein using T1DM positive and negative sera

6.3.5.11 Discussion

For all of the tables listed in Table 15, the following outcome was observed "The two-tailed P value equals 1.0000, so the association between rows (groups) and columns (outcomes) is considered to be not statistically significant". This allows the null hypothesis to be accepted and states that there is no association between diabetic and non-diabetic status and coronavirus recombinant N protein reactivity seen in this study of a limited number of individual sera. Further work will be needed on a larger data set to test the null hypothesis and see if diabetics are at a higher risk of coronavirus infections.

6.4 ELISAs conducted on unknown sera: influenza sera set gifted from Germany

6.4.1 Influenza sera introduction

The use of a limited set off sera (12 in all) used in chapter 7 had shown that the use of purified N proteins was generally acceptable as a coating antigen for diagnosis of recent coronavirus infection with readings well about background for cases of 229E-CoV and OC43-CoV. No cases of NL63-CoV were found in this limited analysis. A single individual had an apparently strong reaction to the SARS-CoV N protein which, while not impossible to discount based on the history of the SARS epidemic, is unlikely to be real. Accordingly, the same format was used to interrogate a larger sera set (n=24) supplied by a collaborator group in Hannover, Germany. While the previous study sought a relationship between T1DM and coronavirus infection the novel aspect of this set of sera was that they had been screened for influenza infection, albeit only by the use of bedside diagnostics. Sera 1-11 were deemed influenza virus infection negative and sera 12-24 influenza virus infection positive. However it is worth noting that a study conducted a meta-analysis and found the pooled sensitivity and specificity of such tests were 62.3% and 98.2% respectively (Chartrand, Leeflang et al. 2012). As such the reliability of said bedside diagnostic is questionable.

It was of interest to know if influenza infection was more or less likely in those recently infected with coronaviruses. A coronavirus infection would stimulate a number of antiviral responses which might provide partial protection form influenza infection.

On the other hand, individuals who are prone to reparatory infections, such as asthmatics, might be more susceptible to both viruses. In addition, the laboratory was able to provide reagents to confirm influenza serum reactivity and to discriminate between H1 and H3, the circulating strains at the time. Thus, confirmation of recent influenza status to accurately differentiate between

positive and negative influenza infection was done, followed by a screen on recombinant coronavirus N proteins as before.

6.4.2 Influenza virus introduction

Influenza viruses are segmented negative-sense single stranded RNA enveloped viruses of the family *Orthomyxoviridae* (Arbeitskreis Blut 2009). There are three genera of influenza, A and B consisting of eight genome segments and C with seven. Table 16 shows a list of the proteins corresponding to the RNA segments along with their functions. Seroconversion in this study focused on the use of the virus surface haemagglutinin glycoprotein (HA), used to anchor the virus to the cellular surface. The second glycoprotein, the neuraminidase (NA), used to digest host sialic acid for the release of viral particles from the host cell (Fields, Knipe et al. 2007, Labella and Merel 2013) was not used. Influenza B is only known to infect humans and seals (harbour seals (*Phoca vitulina*) and grey seals (*Halichoerus grypus*) (Osterhaus, Rimmelzwaan et al. 2000)). Influenza A has been shown to infect a variety of warm-blooded animals, including birds, swine, horses and humans (Webster 2002). Whereas coronaviruses are thought to have their natural reservoir in bats, aquatic birds serve as the natural reservoir for all known subtypes of influenza A virus (Webster, Bean et al. 1992). Influenza A viruses are subdivided by antigenic characterization of the haemagglutinin (HA) and neuraminidase (NA). Eighteen HA and eleven NA subtypes are known (Tong, Zhu et al. 2013). The nomenclature system follows the pattern H(x)N(y) including the host of origin, geographical location, strain number, and year of isolation (Fouchier, Munster et al. 2005). Influenza B viruses are not divided into subtypes.

RNA segment	Protein	Protein function
PB1	Transcriptase	Cap elongation
PB2		Cap binding
PA		Protease activity
HA	Haemagglutinin	Anchoring the cell
NP	Nuclear protein	RNA binding and transport
NA	Neuraminidase	Virus release
M1/M2	Matrix proteins	M1= major component of virion M2= ion channel
NS1/NS2	Non-structural proteins	NS1=RNA transport and translation NS2=nuclear export of virion RNAs

Table 16: The RNA segments of influenza along with the protein they encode and the protein function.

6.4.2.1 Mutation

Influenza viruses can undergo two types of genetic changes allowing the HA and NA to evade previously acquired immunity; antigen drift and antigenic shift (also known as re-assortment). The accumulation of point mutations during replication occurs in influenza viruses because, like coronaviruses, their RNA polymerase complex has limited proofreading activity (Taubenberger and Morens 2008). Antigenic drift occurs when the accumulation of point mutations in the antigenic portions of surface glycoproteins limit or prevent antibody binding, allowing the virus to evade any pre-existing immunity the host may have. Influenza virus' genes have high mutation rates ($\sim 1\times10^{-3}$ to 8×10^{-3} substitutions per site per year) (Chen and Holmes 2006). Antigen shift is a term used to define re-assortment that occurs whereby there is an exchange of whole genome segments, which might result in influenza viruses which have a selective advantage compared with their parent viruses (Fields, Knipe et al. 2007). Re-assortment can only occur when a cell is simultaneous infected with different influenza A viruses resulting in hybrid viruses.

6.4.2.2 Influenza infection

Influenza virus infection is associated with high morbidity and mortality (Taubenberger and Morens 2008), especially in the elderly, infants and people with certain chronic diseases. Influenza virus is amongst the most common cause of human respiratory infection (Fields, Knipe et al. 2007). Annual influenza epidemics are estimated to result in about 3 to 5 million cases of severe illness, and about 250,000 to 500,000 deaths worldwide (Clem and Galwankar 2009). Influenza infection symptoms include; fever, coughs, headache, inflammation of the upper respiratory tract, inflamed glands, malaise, weakness and catarrh production.

6.4.2.3 Pandemics

Alongside seasonal influenza mortalities numerous pandemics have been reported within the past century with varying degrees of severity; all as a result of an influenza A virus. Spanish influenza occurred in 1918 due to H1N1 resulting in an estimated 40 million deaths worldwide (Patterson

and Pyle 1991, Mills, Robins et al. 2004). Asian flu occurred in 1957-1958 caused by H2N2 and was estimated at to have killed 1-2 million (Neumann and Kawaoka 2015). Hong Kong influenza arose in 1968-1969 as a result of H3N2 and was estimated to have been responsible of 0.75-1 million deaths (Fauci 2006). Avian influenza H5N1 caused a minor human outbreak in 1996, often referred to as bird flu (Wan 2012). H1N1 re-emerged in 1977-1978 and was referred to as Russian flu (Donaldson, Rutter et al. 2009). Finally, in 2009-2010 there was a worldwide flu pandemic often referred to as swine flu caused by H1N1/09 (Cheng, To et al. 2012), a virus that still circulates today.

6.4.2.4 Vaccination

Viral antigenic drift in influenza A means that the vaccine's effectiveness can often become compromised. Accordingly, the vaccine requires updating, alongside revaccination of at-risk individuals (Gensheimer, Meltzer et al. 2003). WHO set up an influenza surveillance program in 1947 to monitor and advise on relevant strains (Heymann and Rodier 2004). Vaccination against circulating strains of influenza A and B with inactivate or live attenuated vaccines are often used as a countermeasure against the disease; prophylactic or therapeutic drugs are also available (Couch 2000). There are a variety of influenza vaccine types shown in Table 17. Trivalent vaccines contain two subtypes of influenza A and one type B virus. Trivalent vaccines for use in 2019-2020 have been recommended to contain B/Colorado/06/2017-like virus of the B/Victoria/2/87-lineage as their influenza B component (WHO 2019). The 2019-2020 quadrivalent vaccine consists of an A/Brisbane/02/2018 (H1N1)pdm09-like virus; an A/Kansas/14/2017 (H3N2)-like virus; a B/Colorado/06/2017-like virus (B/Victoria/2/87 lineage); and a B/Phuket/3073/2013-like virus (B/Yamagata/16/88 lineage) (WHO 2019).

Supplier	Name of product	Vaccine type	Age indications
AstraZeneca UK Ltd	Fluenz Tetra Quadrivalent	LAIV (live attenuated influenza vaccine) supplied as nasal spray suspension	From 24 months to <18 years of age
GSK	FluarixTM Tetra QIVe	(standard egg grown quadrivalent influenza vaccine), split virion, inactivated	From 6 months
MASTA	Quadrivalent Influenza vaccine QIVe	(standard egg grown quadrivalent influenza vaccine), split virion, inactivated	From 6 months
Mylan	Quadrivalent Influenza vaccine Tetra MYL QIVe	(standard egg grown quadrivalent influenza vaccine), supplied as surface antigen, inactivated	From 3 years
	Quadrivalent Influvac sub-unit Tetra	QIVe (standard egg grown quadrivalent influenza vaccine), supplied as surface antigen, inactivated	From 3 years
Sanofi Pasteur Vaccines	Quadrivalent Influenza vaccine QIVe	(standard egg grown quadrivalent influenza vaccine), split virion, inactivated	From 6 months
	Trivalent Influenza Vaccine, High-Dose	TIV-HD (standard egg-grown trivalent influenza vaccine), split virion, inactivated	65 years of age and over
Seqirus UK Ltd	Flucelvax® Tetra QIVc	(cell-grown quadrivalent influenza vaccine) supplied as surface antigen, inactivated, prepared in cell cultures	From 9 years
	Fluad® aTIV	(adjuvanted trivalent influenza vaccine) supplied as surface antigen, inactivated, adjuvanted with MF59C.1	65 years of age and over

Table 17: Influenza vaccines for the 2019/20 influenza season. Table sourced from (PublicHealthEngland 2019)

Table 17 summarises the 2019-2020 influenza vaccines available in the UK. Different age demographics are recommended different vaccine types, the micrograms per dose of ovalbumin content varies between vaccine type, some are therefore more suited to severe egg allergy suffers than others.

6.4.2.5 Antiviral drugs

Antiviral drugs should ideally be administered within 48 hours of symptom onset and fall into three categories. Firstly, matrix 2 ion channel blockers (amantadine and rimantadine) which are effective against influenza A viruses, but resistant viral strains develop rapidly. Secondly,

neuraminidase (NA) inhibitors, zanamivir and oseltamivir, which are effective against both influenza A and B viruses although resistance has again been noted (Monto, McKimm-Breschkin et al. 2006). Both classes of drugs are effective in preventing influenza when administered prophylactically (Monto, Robinson et al. 1999, Hayden 2006, Hayden and Pavia 2006). Finally the use of polymerase inhibitors, such as baloxavir marboxil, pimodivir and favipiravir are also being investigated (Naesens, Stevaert et al. 2016, Stevaert and Naesens 2016, Uyeki, Bernstein et al. 2019). Baloxavir has already been approved in the US and Japan (Hayden and Shindo 2019).

6.4.2.6 Influenza and serology diagnosis

Serological tests differ to those of genome detection, such as PCR, as they eliminate the need for live virus samples to be obtained, whilst also being able to detect prior infections where the antibody titre has remained high. Virus-specific antibodies against influenza can be identified in patient's sera either as total antibody levels or by antibody sub-classes (IgA, IgM, IgG). An anti-IgM response indicates a current or very recent infection whereas the IgG response includes immunological memory of previous infections. A variety of serological methods can be used in confirm current or prior infection including; complement-binding reaction (CBR), immunofluorescence test (IFT), haemagglutination inhibition test (HIT), the neutralisation test and, as used in this study, ELISAs (Arbeitskreis Blut 2009). Initially the use of ELISAs to detect influenza positive sera was thought of as having good specificity however a lower sensitivity than other methods which detect live virus or viral genome; this notion is being challenged more recently by a variety of studies all examining detection of avian influenza using ELISAs (Jensen, Andersen et al. 2017, Wibowo, Tarigan et al. 2017, Zhang, Hou et al. 2017).

6.4.3 Influenza and Coronavirus experimental aim

When gifted to the laboratory sera 1-11 were noted as being influenza virus negative and sera 12-24 influenza sera positive. Influenza positive serum is assumed to be positive for either H1 or H3, to coincide with the circulating strains at the time. The sera was firstly used to ascertain an optimal sera dilution for distinguishing between the two subgroups, before it was used for

screening with the recombinant coronavirus N proteins. Primarily the set of sera is beneficial for its larger sample size and its wholly different geographical origin. Also although influenza viruses and CoVs differ in terms of replication, immune stimulation, and overall lethality (Ng, To et al. 2006) in order to establish a successful respiratory infection, both viruses have been shown to overcome the effectors produced by interferon stimulated genes (ISGs) who's functions are to impede viral pathogenesis (Menachery, Eisfeld et al. 2014). The signalling cascade produced by type I IFN provides a first line of defence against viral pathogens and initiates transcription of ISGs with antiviral, immune modulatory, and cell regulatory functions (Katze, He et al. 2002). A 2002 study conducted by Katze *et al* used models of human airways, transcriptomics and proteomics datasets to compare ISG response after highly pathogen H5N1 avian influenza virus, 2009 H1N1 pandemic influenza virus, SARS-CoV and MERS-CoV infections. It was noted that each virus antagonised the ISG response with distinct approaches, sharing some similarities and some differences. H5N1 actively manipulated the ISG response with both up and downregulation of ISG subsets and H1N1 produced a strong, uniform induction, whereas SARS-CoV and MERS-CoV both successfully delayed ISG expression until after viral titres had peaked. Similarities were seen between the highly virulent H5N1 and MERS-CoV, as both were able to downregulate ISG subset expression using altered histone modification. As a result, it may be hypothesised that having infection with influenza virus may increase likelihood of a coronavirus infection, or vice versa. The null hypothesis is that having prior infection with either has no result on the likelihood of infection with the other.

6.4.4 Influenza materials and methods

Sf9 cells were used to produce the recombinant H1 and H3 influenza proteins using the same method of BEV used in section 4.1.

ELISAs were conducted in a similar manner to that described in chapters 6 and 7 however the wells were pre-coated with 50µL of snowdrop (*Galanthus Nivalis*) lectin from at a final concentration of 10µg/mL using sodium bicarbonate to dilute and left overnight at 4^{o}C. Snowdrop

lectin is a mannose-specific lectin, mannose is the most common carbohydrate in insect cell expressed glycoproteins, including the HA (Dormitzer, Andrews et al. 2013). This eliminates the need for purification steps as once the recombinant HA proteins are expressed, the Sf9 cells simply need to be lysed and clarified to produce lysate suitable for use on the lectin coated 96-well plates.

Cell pellets were resuspended in 4mLs PBS and lysed using sonication for 3 minutes on ice (pulse 2 seconds on 2 seconds off, amplification 60%). Triton X100 was added to the final concentration of 0.1% and the lysate subjected to a clarifying spin at 13,000rpm for 15 minutes to remove the nuclei. The clear cell lysate was either stored in 1mL aliquots at -80^{o}C for future use or diluted 1:10 with PBS and used immediately for ELISA. After the snowdrop lectin had coated the plates overnight, they were washed three times with 1XTBST for 5 minutes each and then 200µL of lysate was added per well and left at RT for 1 hour before removal and washing. Thereafter the ELISA was conducted as described in section 2.2.7. ELISAs were conducted using varying dilutions of the sera in order to see which dilution would best differentiate between those deemed influenza virus negative (serum 1-11) and influenza positive (serum 12-24). The initial dilution tested was 1:10 with a threefold dilution series.

The cut-off value this time were drawn simply using the highest negative sera result (sera 1-11) although clearly if this were to be conducted for diagnostic purpose a more statistical method to determine cut-off value would need to be used.

6.4.5 H1 and H3 protein ELISAs

6.4.5.1 Overview

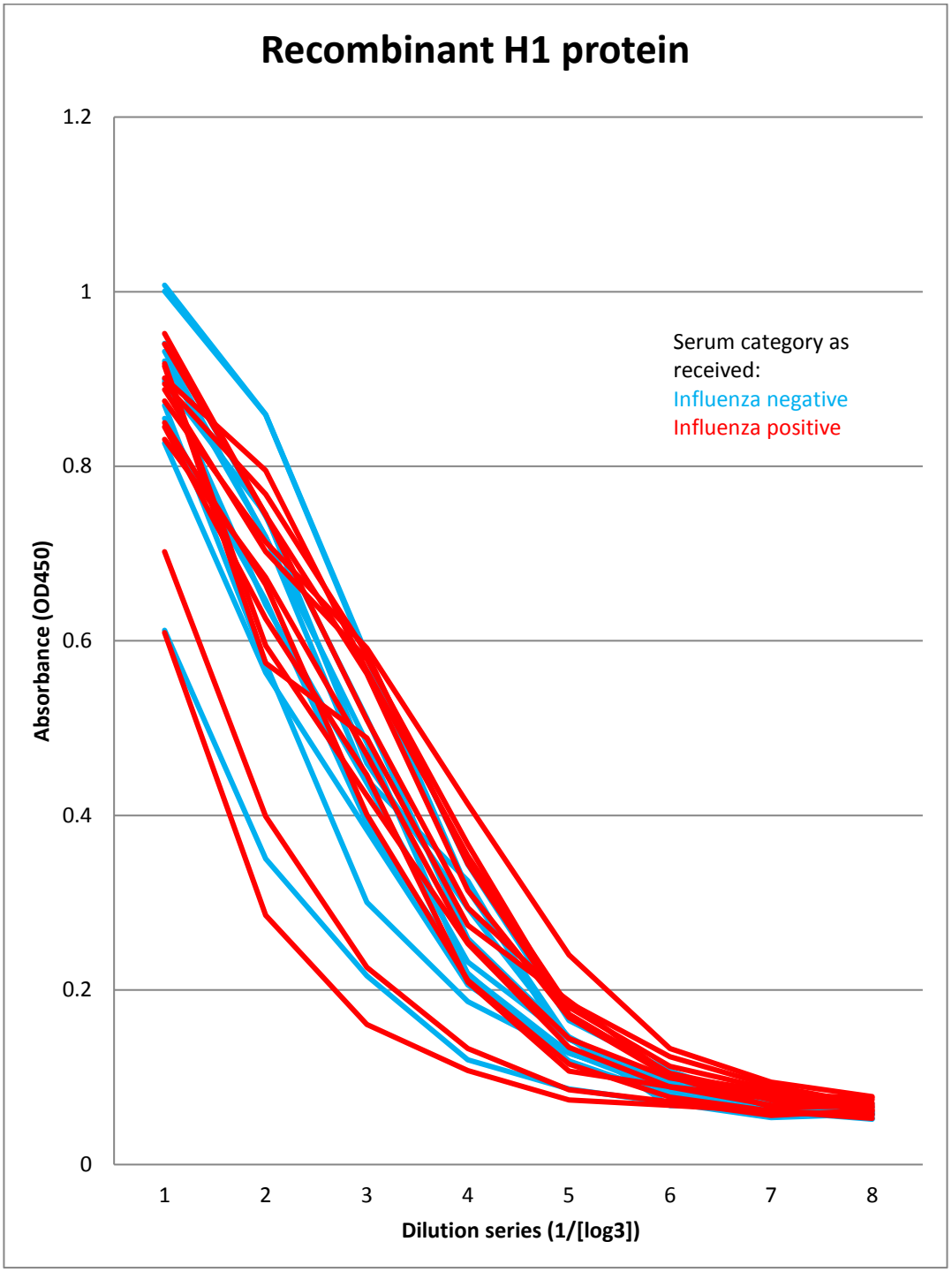


Figure 64: ELISA results showing the comparison of influenza positive vs influenza negative sera screened against recombinant influenza H1 protein. Sera deemed influenza negative are coloured blue, sera deemed influenza positive are coloured red. The dilution series started with the stock sera being diluted 1:10 and followed a three-fold dilution series. Error bars not shown due to lack of repeats.

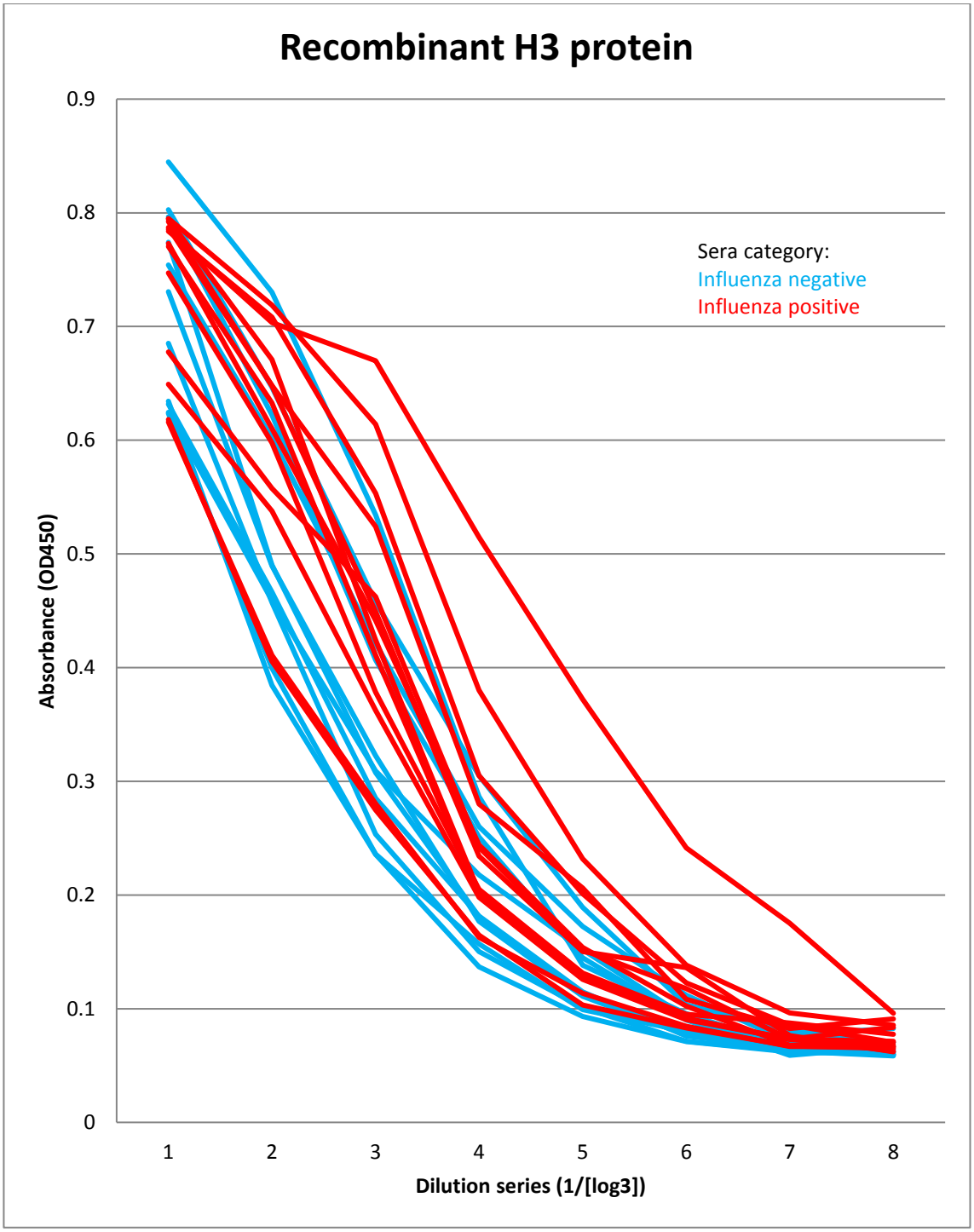


Figure 65: ELISA results showing the comparison of influenza positive vs influenza negative sera screened against recombinant influenza H3 protein. Sera deemed influenza negative are coloured blue, sera deemed influenza positive are coloured red. The dilution series started with the stock sera being diluted 1:10 and followed a three-fold dilution series. Error bars not shown due to lack of repeats.

Both the graphs seen showing the ELISA results using recombinant H1 as the antigen, Figure 64, and using recombinant H3 as the diagnostic antigen, shown in Figure 65, show that the initial dilution of $1:10x3^0$ provides a wide range of OD450 readings unlike to discriminate between influenza positive and negative sera and that readings from $1:10x3^5$ onwards appear too dilute. For a more detailed look each concentration was examined individually.

6.4.5.2 Determining optimal dilution factor for differentiating

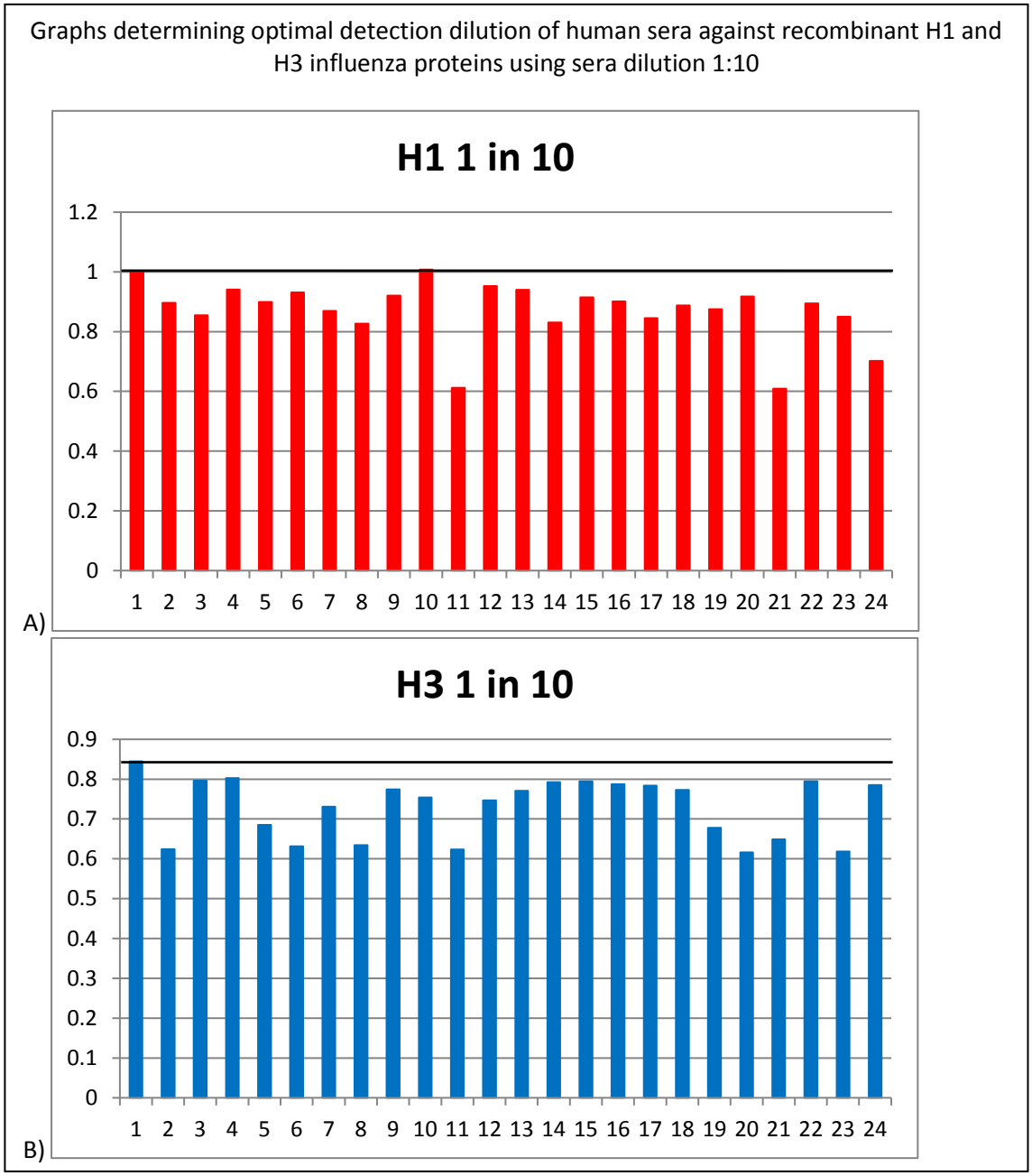


Figure 66: ELISA results obtained using a 1:10 sera dilution, conducted using negative and positive influenza sera and recombinant H1 and H3 proteins. Sera 1-11 deemed influenza negative and sera 12-24 deemed influenza positive. Figure a) results using recombinant H1 protein in red b) results using recombinant H3 protein in blue. Error bars not shown due to lack of repeats.

When the sera is used to conduct the ELISA at a 1:10 dilution, shown in Figure 66, none of the positive influenza sera are seen as being above the cut-off value for either H1 or H3. The 1:10 dilution is clearly not effective at determining between influenza positive and negative sera.

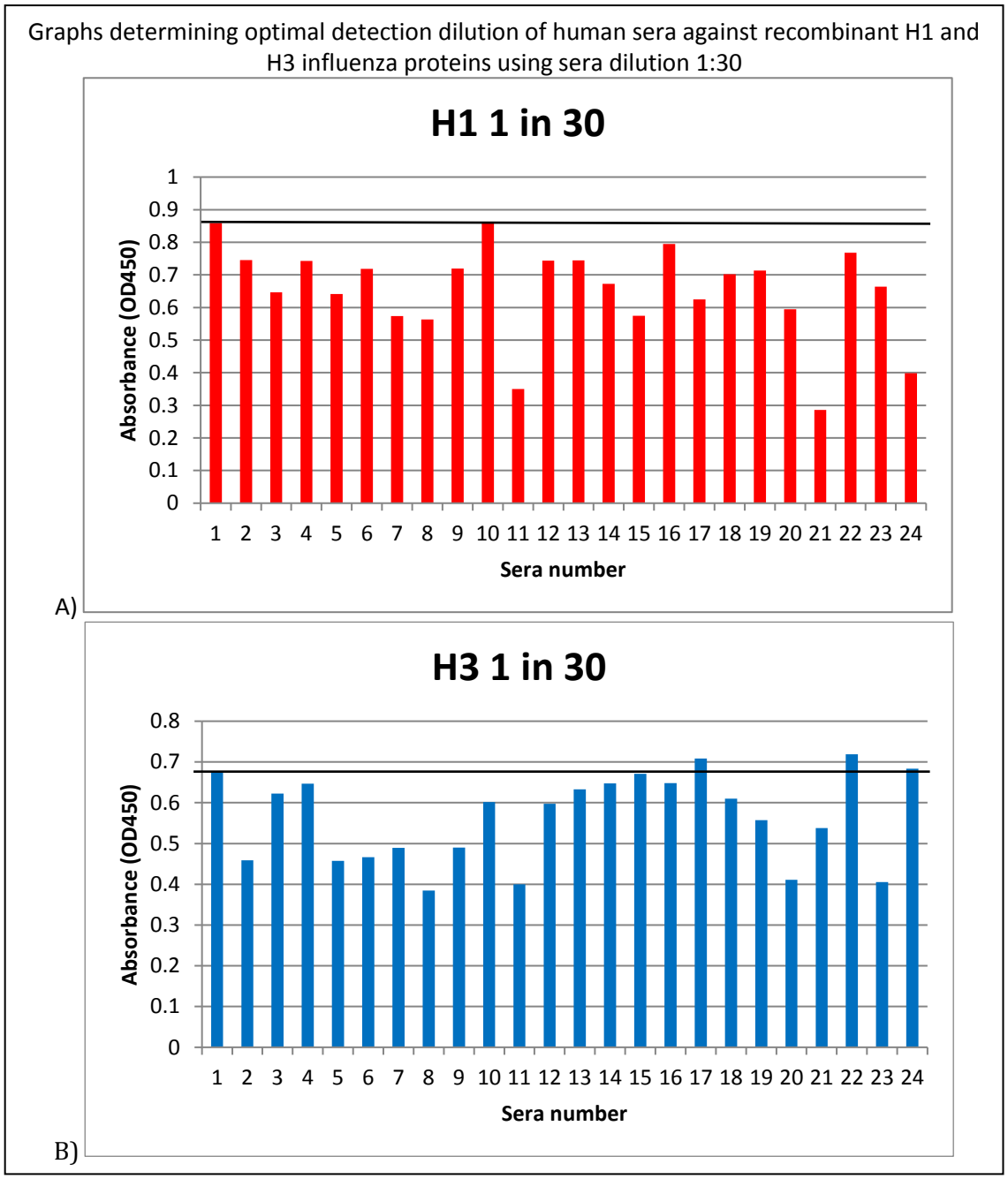


Figure 67: ELISA results obtained using a 1:30 sera dilution, conducted using negative and positive influenza sera and recombinant H1 and H3 proteins. Sera 1-11 deemed influenza negative and sera 12-24 deemed influenza positive. Figure a) results using recombinant H1 protein in red b) results using recombinant H3 protein in blue. Error bars not shown due to lack of repeats.

When the sera is used to conduct the ELISA at a 1:30 dilution, shown in Figure 67, none of the positive influenza sera are seen as being above the cut-off value for H1, however serum 17,22 and 24 exceed the cut off line for H3. The 1:30 dilution is also clearly not effective at determining between influenza positive and negative sera as only 3 out of the 13 positive sera were detected.

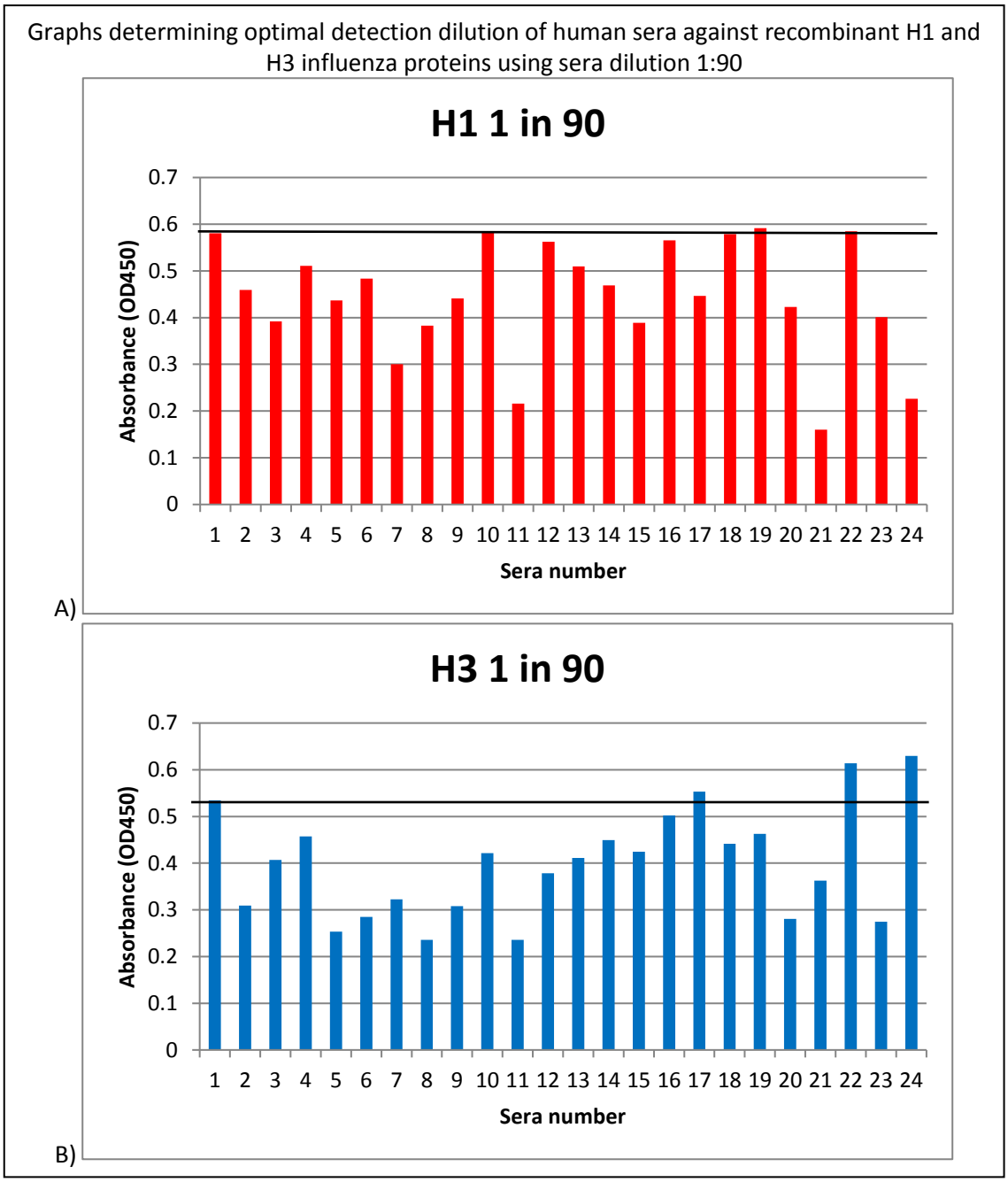


Figure 68: ELISA results obtained using a 1:90 sera dilution, conducted using negative and positive influenza sera and recombinant H1 and H3 proteins. Sera 1-11 deemed influenza negative and sera 12-24 deemed influenza positive. Figure a) results using recombinant H1 protein in red b) results using recombinant H3 protein in blue. Error bars not shown due to lack of repeats.

When the sera is used to conduct the ELISA at a 1:90 dilution, shown in Figure 68, only serum 19 is above the cut-off value for H1. Sera 17, 22 and 24 again exceed the cut-off line for H3 this time in a more defined manner than that seen in the 1:30 dilution shown in Figure 67. The 1:90 dilution is not effective at determining between influenza positive and negative sera as again only 3 out of the 13 positive sera were detected.

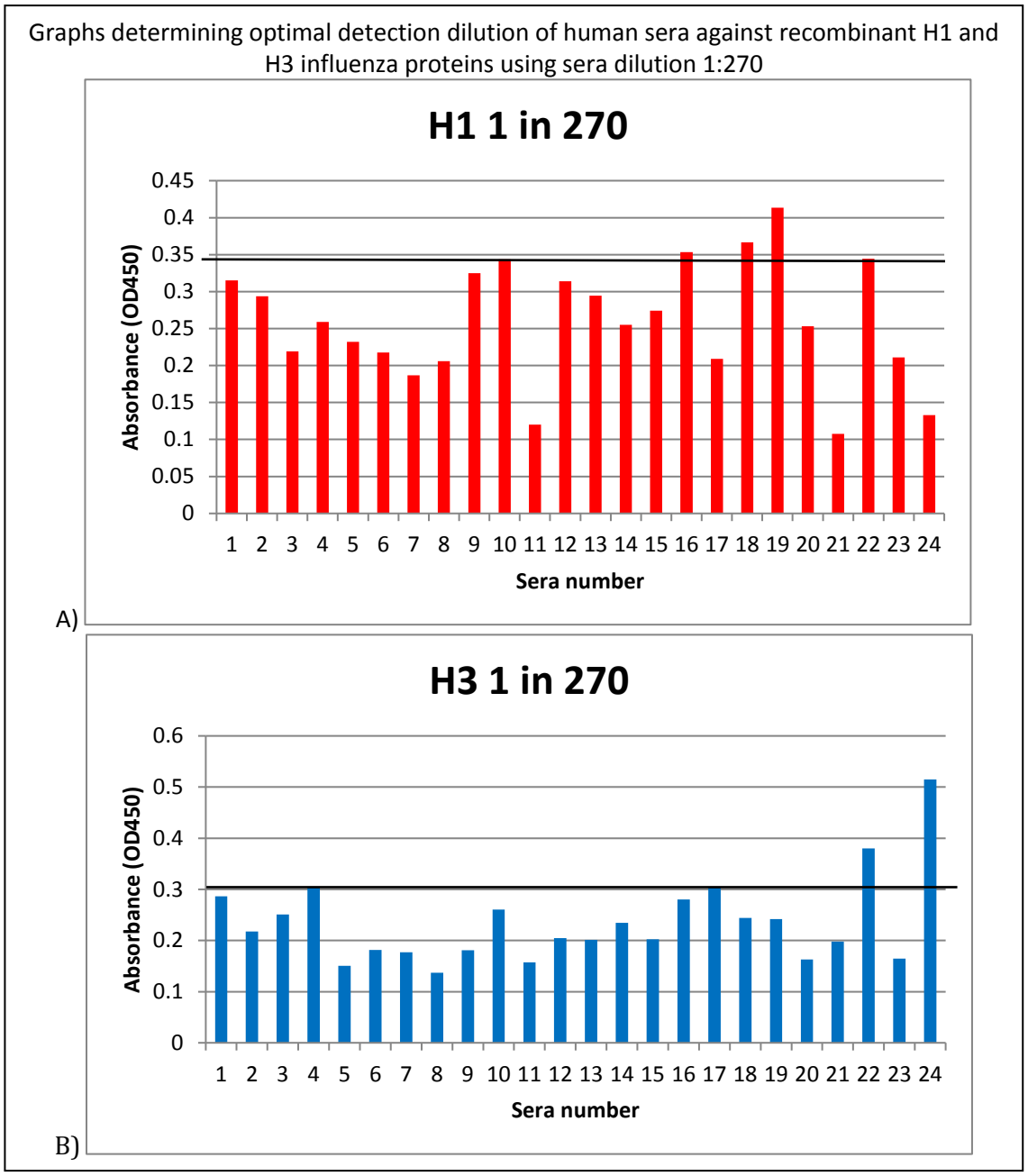


Figure 69: ELISA results obtained using a 1:270 sera dilution, conducted using negative and positive influenza sera and recombinant H1 and H3 proteins. Sera 1-11 deemed influenza negative and sera 12-24 deemed influenza positive. Figure a) results using recombinant H1 protein in red b) results using recombinant H3 protein in blue. Error bars not shown due to lack of repeats.

When the sera is used to conduct the ELISA at a 1:270 dilution, shown in Figure 69, for the first time positive influenza sera are seen as being above the cut-off value for H1, serum 16, 18 and 19. Sera 22 and 24 again exceed the cut off line for H3 however serum 17 this time does not. The 1:20 dilution was deemed ineffective at determining between influenza positive and negative sera as only 5 out of the 13 positive sera were detected.

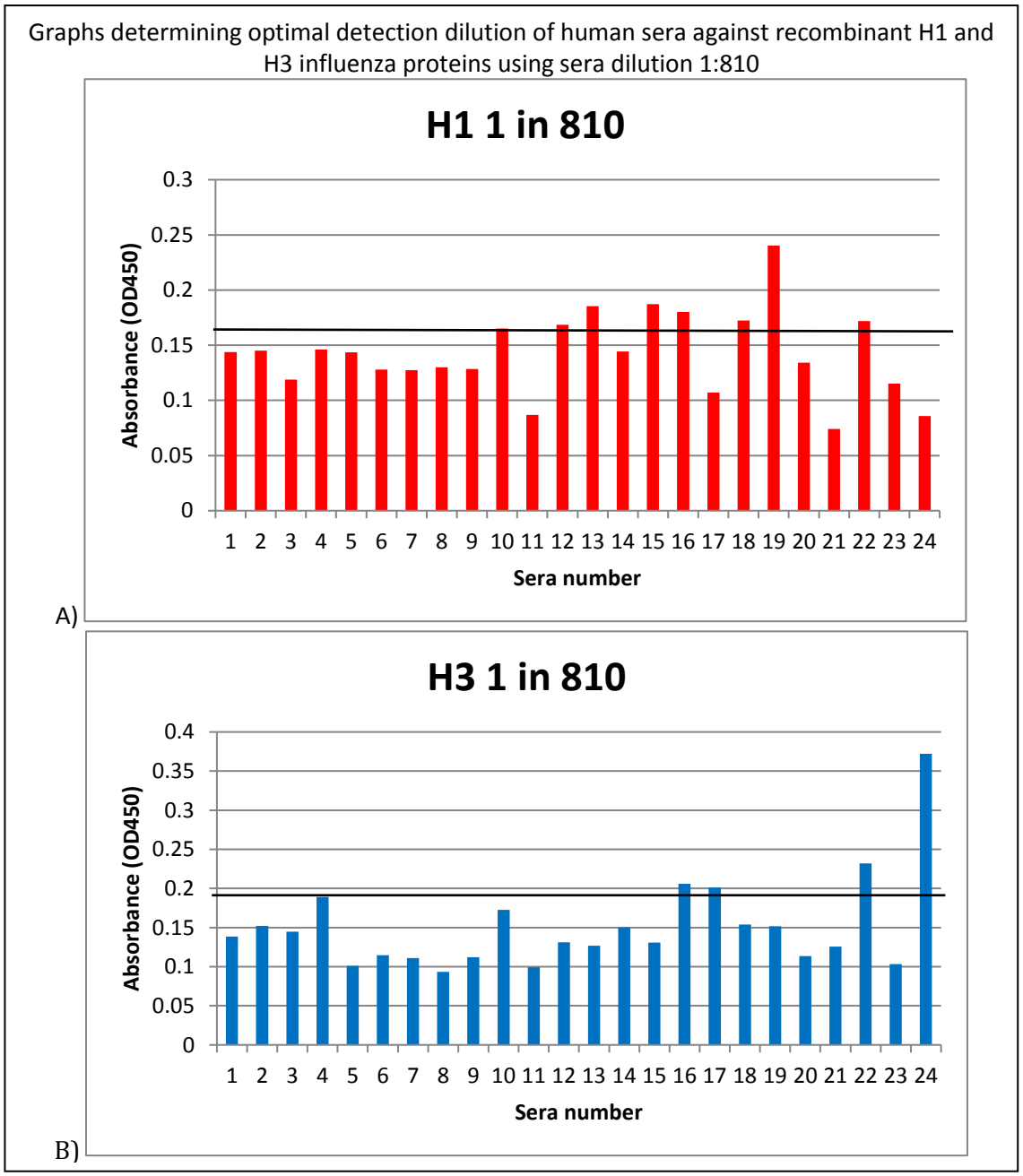


Figure 70: ELISA results obtained using a 1:810 sera dilution, conducted using negative and positive influenza sera and recombinant H1 and H3 proteins. Sera 1-11 deemed influenza negative and sera 12-24 deemed influenza positive. Figure a) results using recombinant H1 protein in red b) results using recombinant H3 protein in blue. Error bars not shown due to lack of repeats.

Figure 70 shows the sera at a 1:810 dilution; positive influenza sera start becoming more defined in the H1 screening with sera 12, 13, 15, 16, 18,19 and 23 all surpassing the cut-off line. Sera 16, 17, 22 and 24 exceed the cut-off line for H3. The 1:810 dilution shows a marked improvement on previous dilutions screened as 9 out of the 13 positive sera were detected. The only positive influenza sera that did not exceed the cut-off level for either H1 or H3 were 14, 20, 21 and 23.

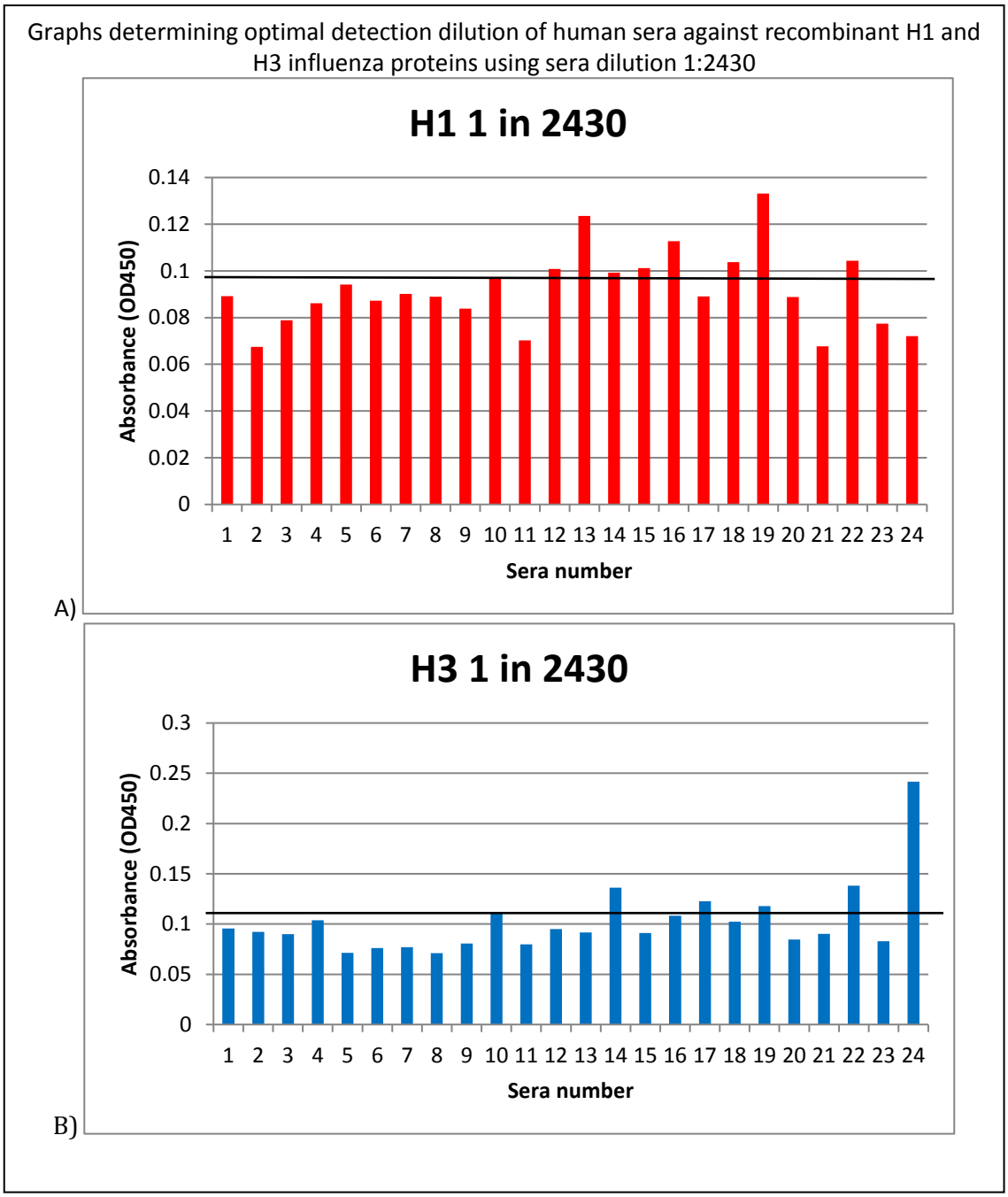


Figure 71: ELISA results obtained using a 1:2430 sera dilution, conducted using negative and positive influenza sera and recombinant H1 and H3 proteins. Sera 1-11 deemed influenza negative and sera 12-24 deemed influenza positive. Figure a) results using recombinant H1 protein in red b) results using recombinant H3 protein in blue. Error bars not shown due to lack of repeats.

When the sera is used to conduct the ELISA at a 1:2430 dilution, shown in Figure 71Figure 66, positive influenza sera are again detected in the H1 screening with sera 12, 13, 14, 15, 16, 18,19 and 22 all surpassing the cut-off line. Sera 14, 17, 19, 22 and 24 exceed the cut-off line for H3. The 1:2430 dilution detects 10 out of the 13 positive sera were detected. The only positive influenza sera that did not exceed the cut-off level for either H1 or H3 were 20, 21 and 23.

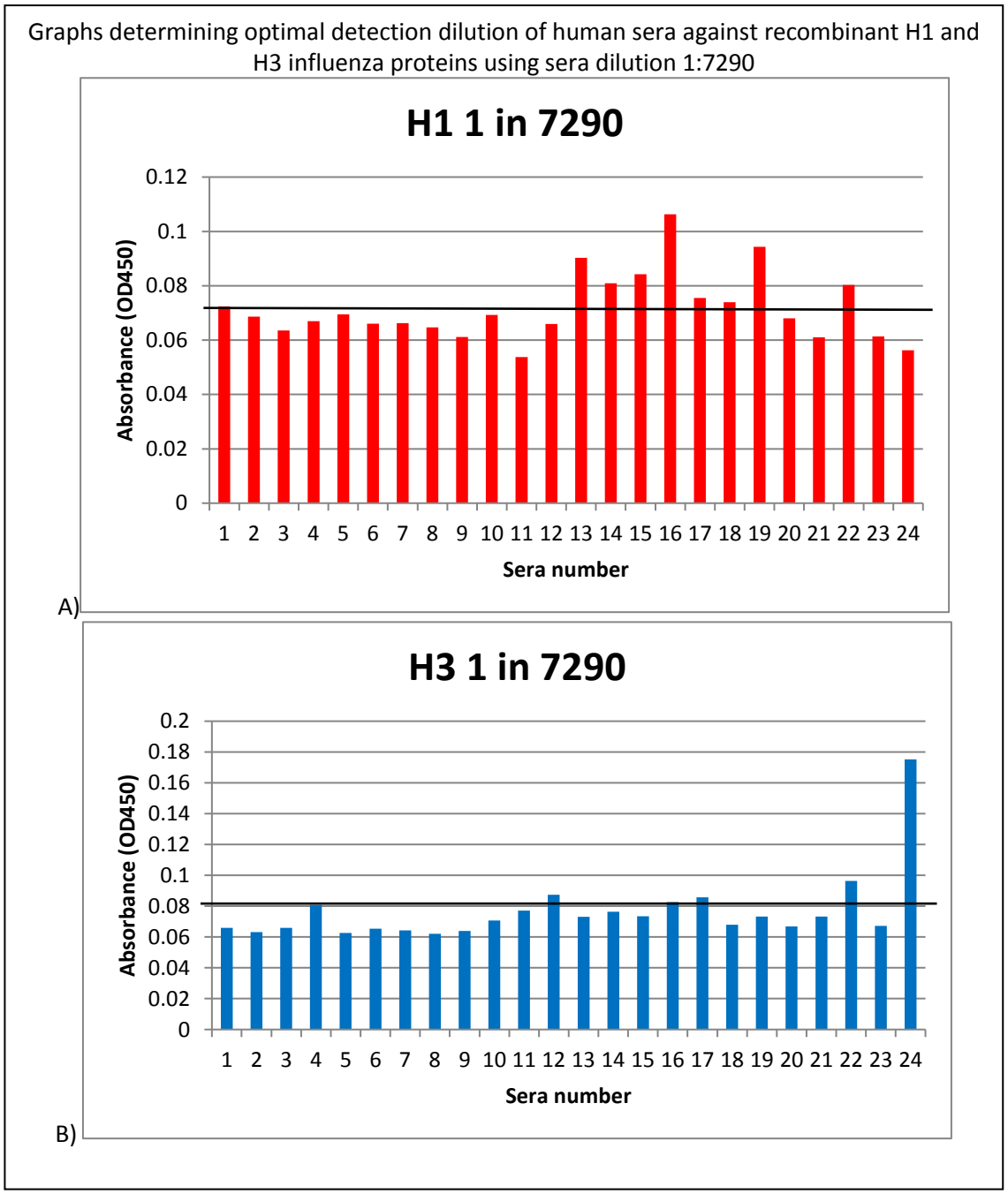


Figure 72: ELISA results obtained using a 1:7290 sera dilution, conducted using negative and positive influenza sera and recombinant H1 and H3 proteins. Sera 1-11 deemed influenza negative and sera 12-24 deemed influenza positive. Figure a) results using recombinant H1 protein in red b) results using recombinant H3 protein in blue. Error bars not shown due to lack of repeats.

When the sera is used to conduct the ELISA at a 1:7,290 dilution, shown in Figure 72, positive influenza sera deemed above the cut-off limit in the H1 screening were 13, 14, 15, 16, 17, 18,19 and 22. Sera 12, 17, 22 and 24 exceed the cut-off line for H3. The 1:7,290 dilution detected 10 out of the 13 positive sera, similar to the results seen for the previous dilution of 1:2,430. The only positive influenza sera that did not exceed the cut-off level for either H1 or H3 were 20, 21 and 23.

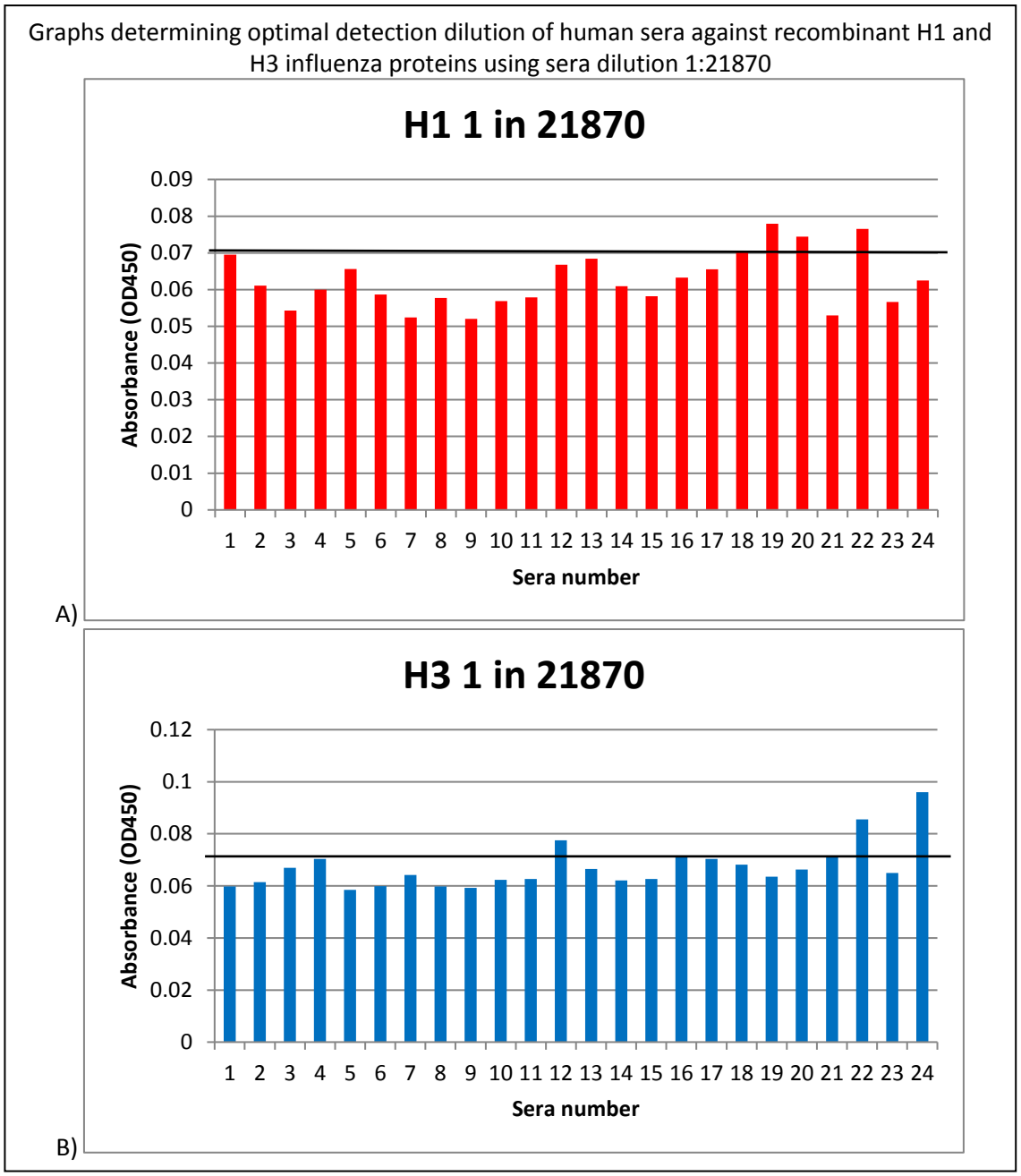


Figure 73: ELISA results obtained using a 1:21,870 sera dilution, conducted using negative and positive influenza sera and recombinant H1 and H3 proteins. Sera 1-11 deemed influenza negative and sera 12-24 deemed influenza positive. Figure a) results using recombinant H1 protein in red b) results using recombinant H3 protein in blue. Error bars not shown due to lack of repeats.

When the sera is used to conduct the ELISA at the highest dilution screened 1:21,870, shown in Figure 73, differentiation between positive and negative influenza serum reduces. The H1 screening detects only sera 19, 20 and 22 above the cut-off limit and only sera 12, 22 and 24 exceed the cut-off line for H3. The 1:21,870 dilution shows a clear reduction in efficiency on

previous dilutions screened as only 5 out of the 13 positive sera were detected missing out sera 13, 14, 15, 16, 17, 18, 21 and 23.

6.4.5.3 Summary of optimal ELISA protein dilution

The graphs determining optimal detection dilution of human sera against recombinant H1 and H3 influenza proteins using varying sera dilutions, shown in Figure 66-Figure 73, show that the dilutions with the fewest number of positive influenza sera missing were 1: 2,430 and 1: 7,290. If the purpose of this study was to truly use the recombinant H1 and H3 proteins in diagnostic testing then a more scientific cut-off value would need be to ascertained, for example the "mean + 3 standard deviation of negative control" could be used (Lardeux, Torrico et al. 2016) and the test's sensitivity and specificity could be ascertained.

Going forward, the same set of 24 sera was used for ELISA screen this time using the CoV recombinant nucleocapsid proteins. Positive values were assessed arbitrarily by eye.

6.4.6 Coronavirus protein ELISAs

6.4.6.1 By individual sera

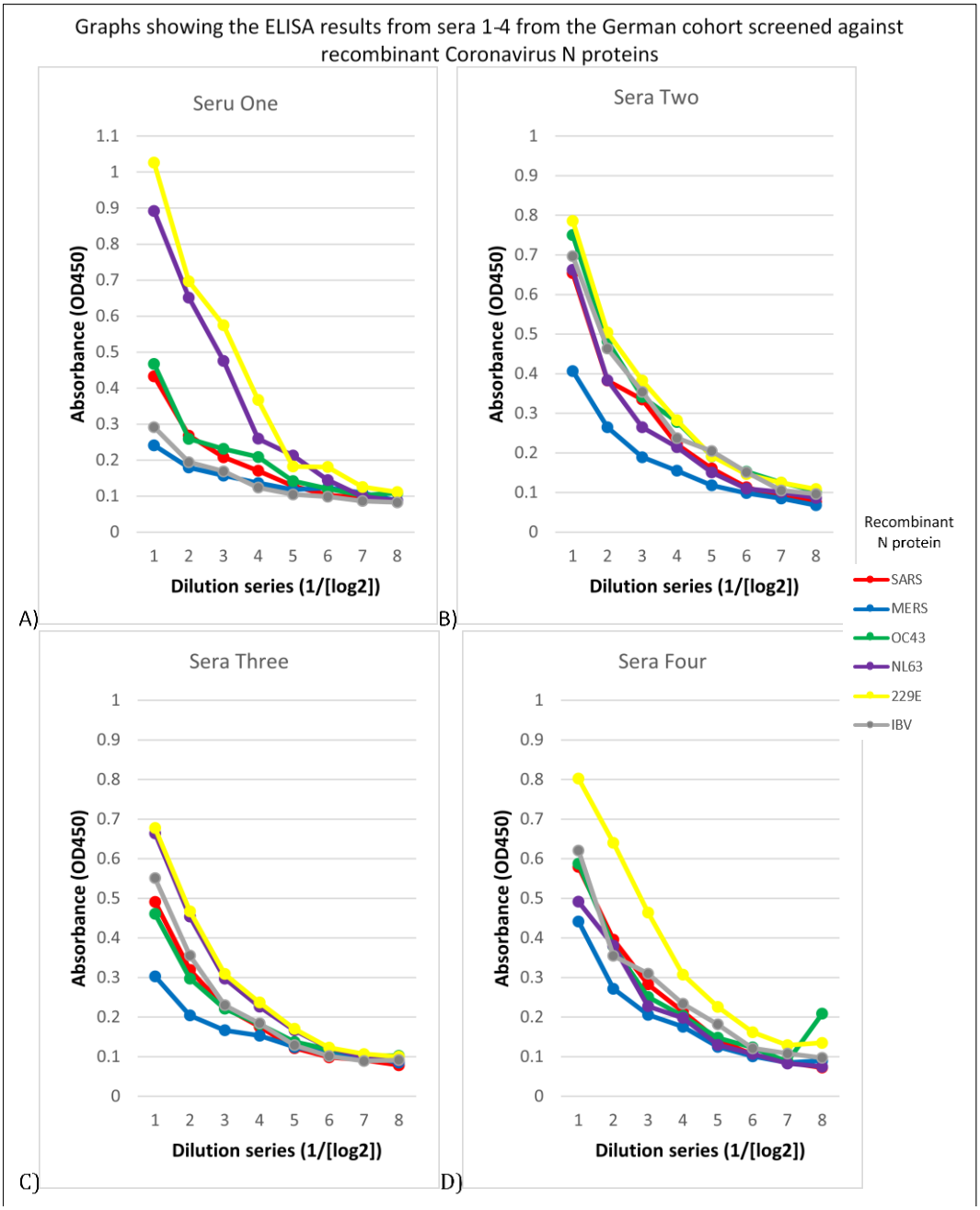


Figure 74: ELISA results showing 4 human sera against recombinant CoV proteins (colour coded; SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey). The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Figures a-d show sera 1-4 respectively. Error bars not shown due to lack of repeats.

Figure 74 shows sera 1 and 3 as having suspected 229E-CoV and NL63-CoV positive, sera 4 as suspected 229E-CoV positivity and sera 2 was not thought to have any Co-V positivity.

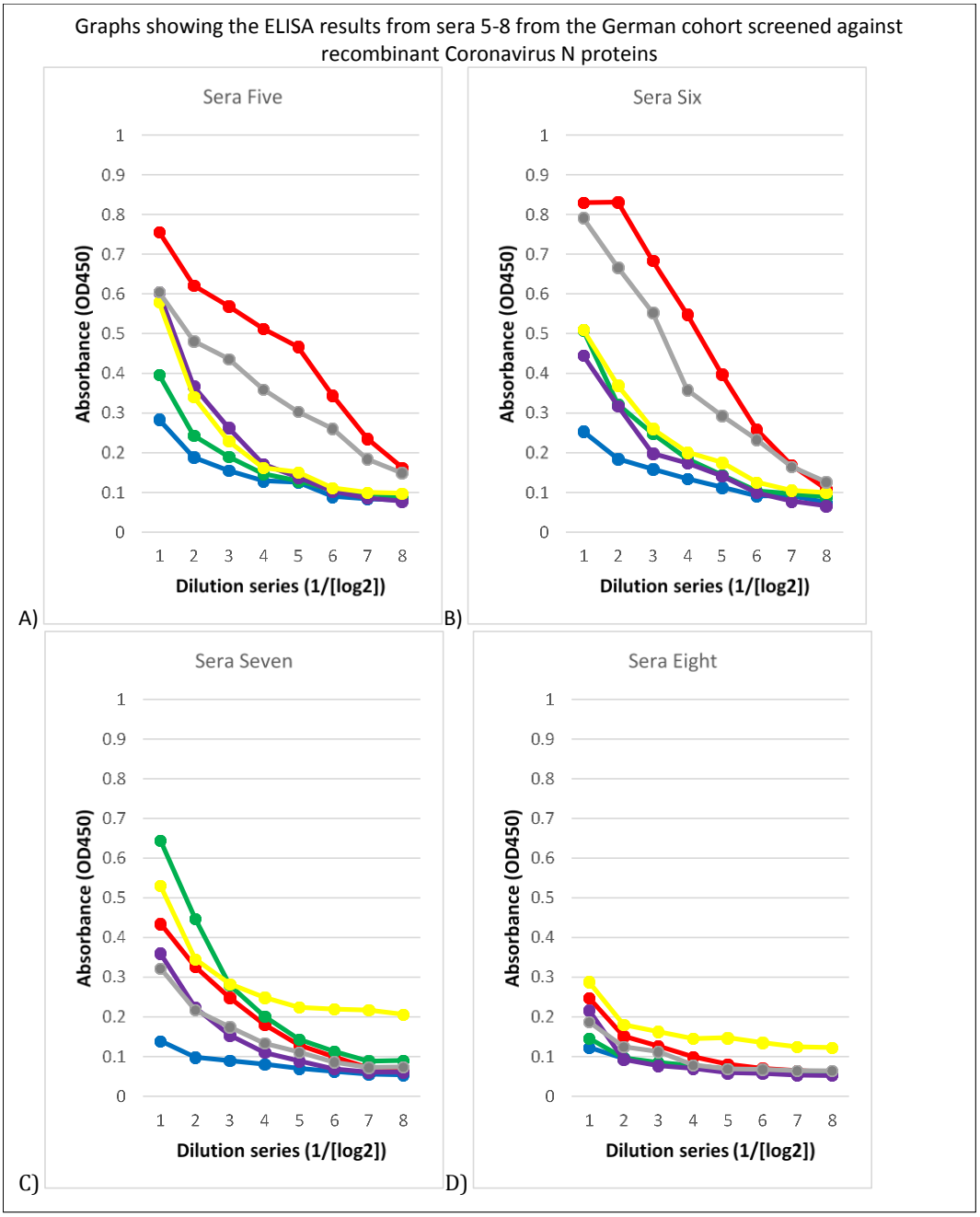


Figure 75: ELISA results showing 4 human sera against recombinant CoV proteins (SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey). The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Figures a-d show sera 5-8 respectively. Error bars not shown due to lack of repeats.

Figure 75 shows sera 5 and 6 as possible SARS-CoV reactivity although the high IBV levels also seen make this questionable, Sera 7 and 8 both show probably 229E-CoV reactivity with sera 7 additionally having OC43-CoV reactivity.

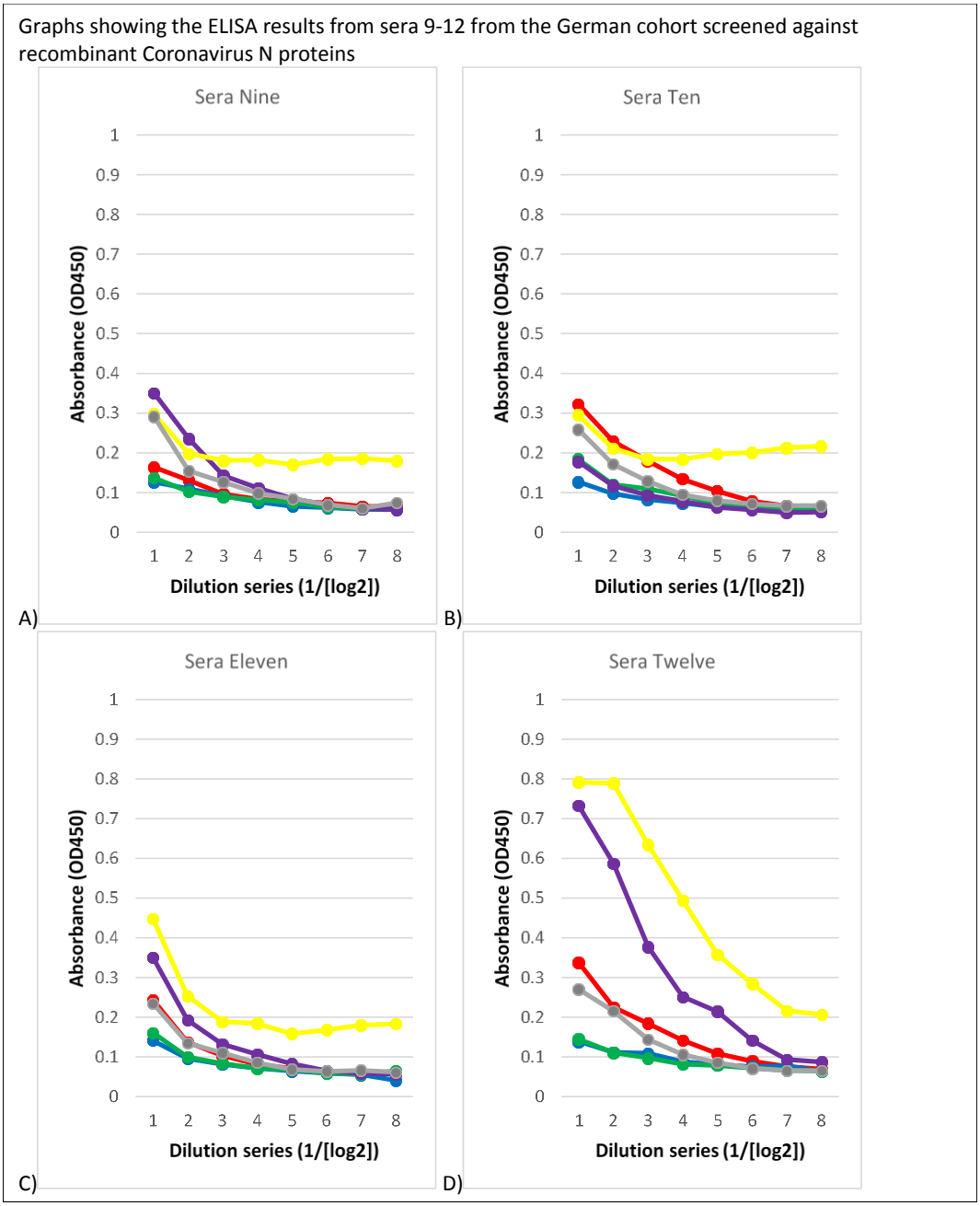


Figure 76: ELISA results showing 4 human sera against recombinant CoV proteins (SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey). The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Figures a-d show sera 9-12 respectively. Error bars not shown due to lack of repeats.

Figure 76 shows sera 9-12 as all having suspected reactivity to 229E-CoV, they also all show possible reactivity to NL63-CoV except for sera 10 which does not.

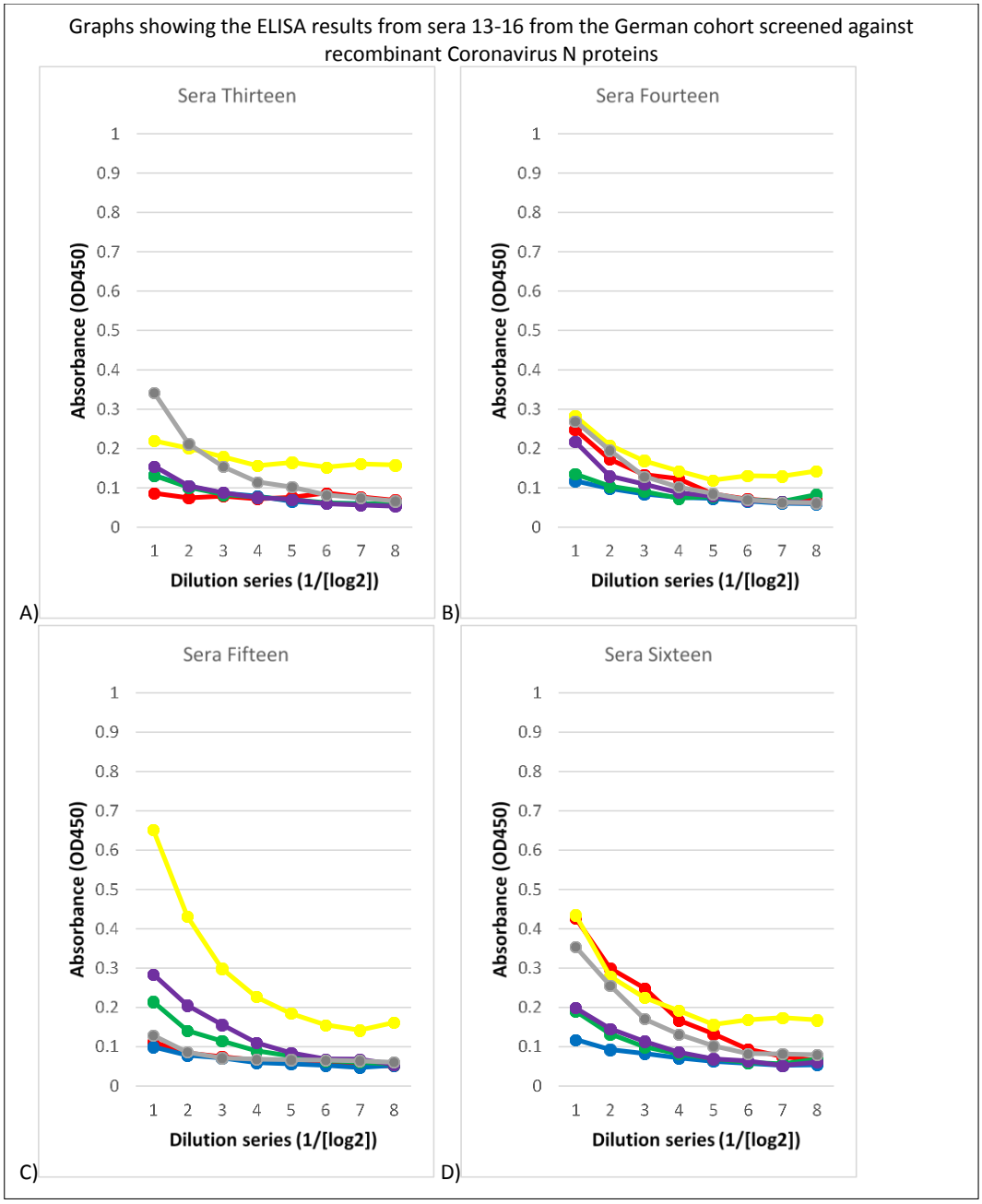


Figure 77: ELISA results showing 4 human sera against recombinant CoV proteins (SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey). The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Figures a-d show sera 13-16 respectively. Error bars not shown due to lack of repeats.

Figure 77 shows serum 13 with suspected 229E-CoV reactivity although a questionable IBV reactivity, serum 15 as having possible OC43-CoV, NL63-CoV and 229E-CoV reactivity and sera 14 and 16 as having probably 229E-CoV reactivity.

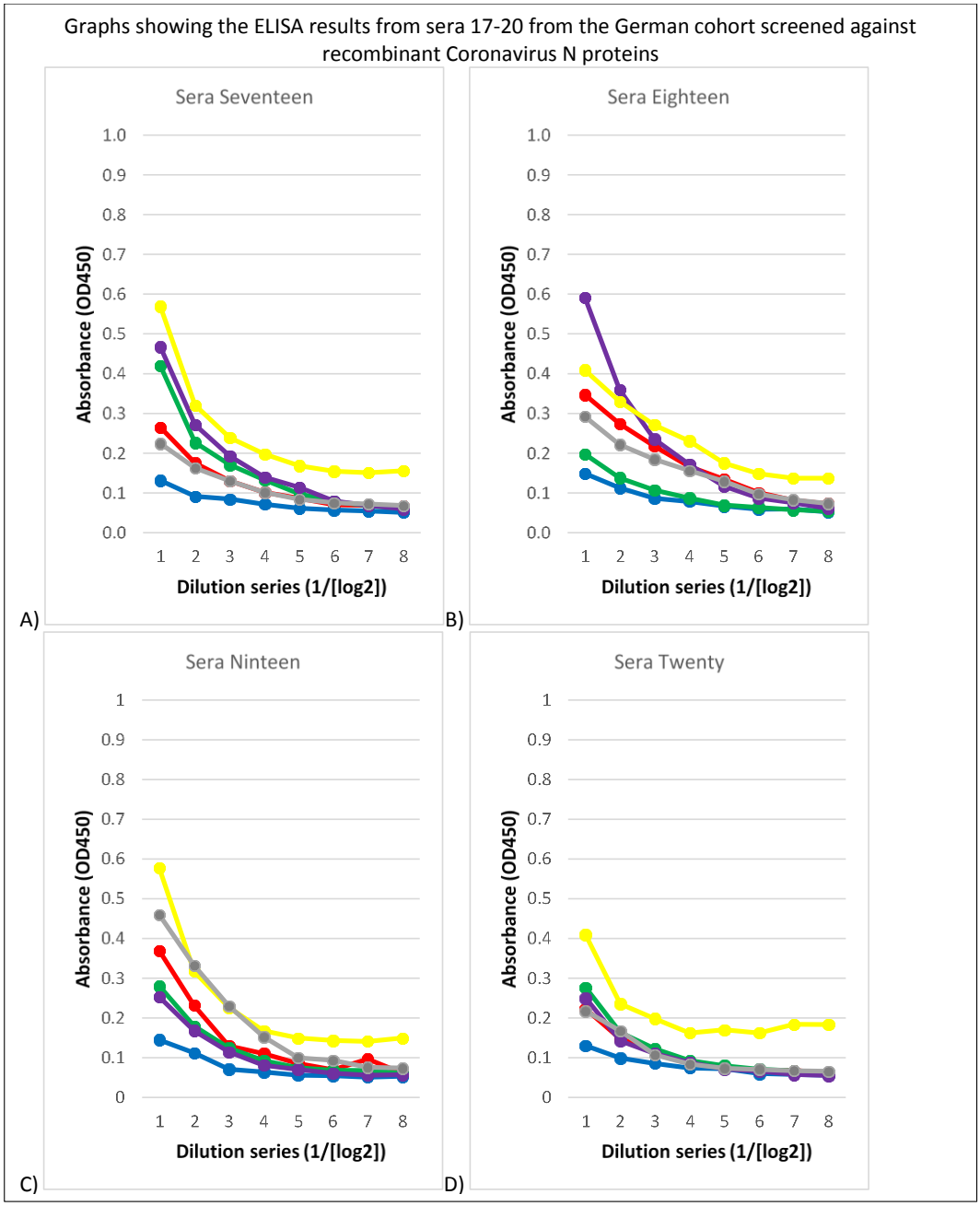


Figure 78: ELISA results showing 4 human sera against recombinant CoV proteins (SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey). The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Figures a-d show sera 16-20 respectively. Error bars not shown due to lack of repeats.

Figure 78 shows serum 17 as having possible OC43-CoV, NL63-CoV and 229E-CoV reactivity. Serum 18 as having NL63-CoV and 229E-CoV reactivity and sera 19 and 20 as having 229E-CoV reactivity.

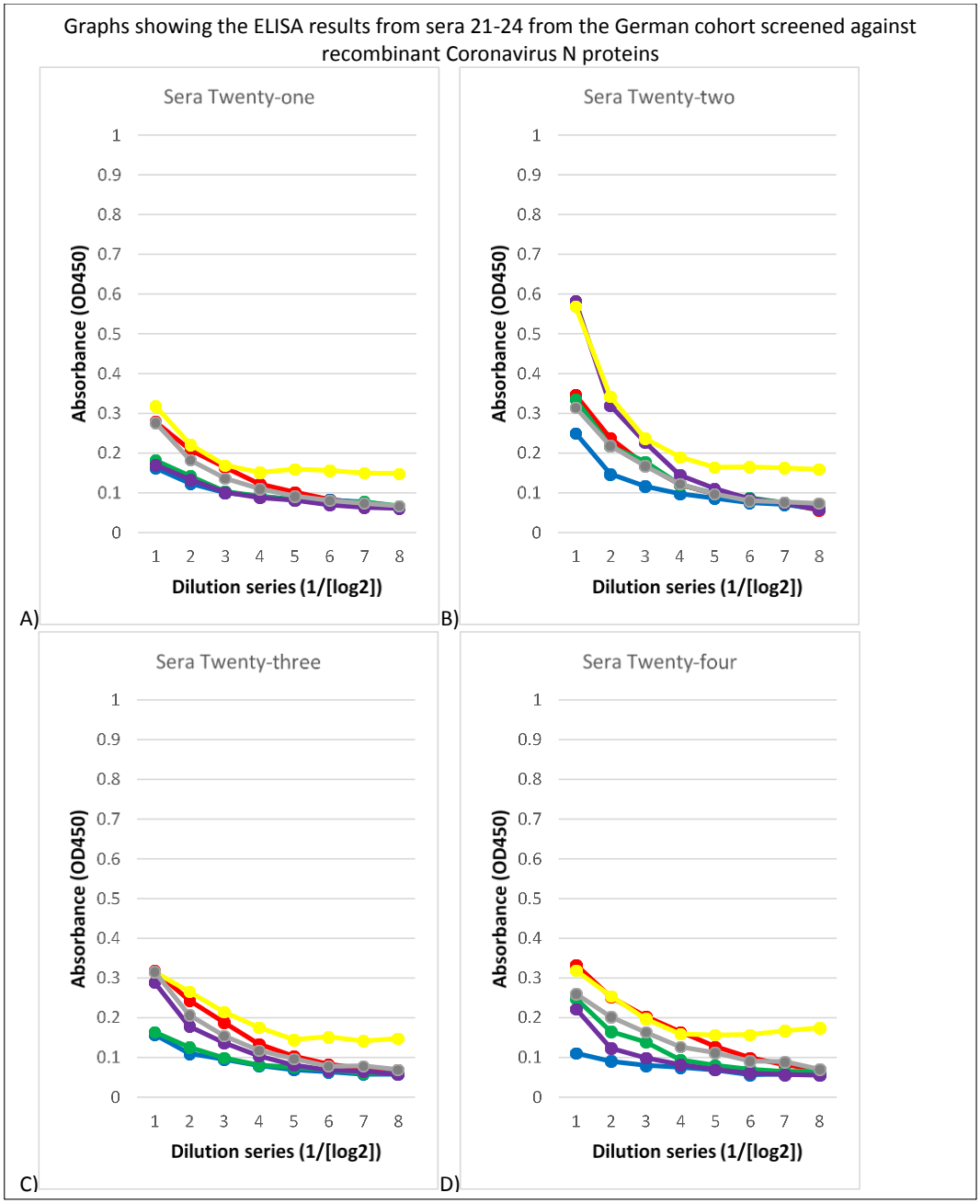


Figure 79: ELISA results showing 4 human sera against recombinant CoV proteins (SARS-CoV in red, MERS-CoV in blue, OC43-CoV in green, NL63-CoV in purple, 229E-CoV in yellow and IBV in grey). The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Figures a-d show sera 21-24 respectively. Error bars not shown due to lack of repeats.

Figure 79 shows sera 21-24 as all showing signs of 229E-CoV reactivity with serum 22 additionally showing NL63-CoV reactivity.

The following graphs show the ELISA results based on individual coronavirus recombinant N protein as opposed to sera.

6.4.6.2 By Coronavirus N protein

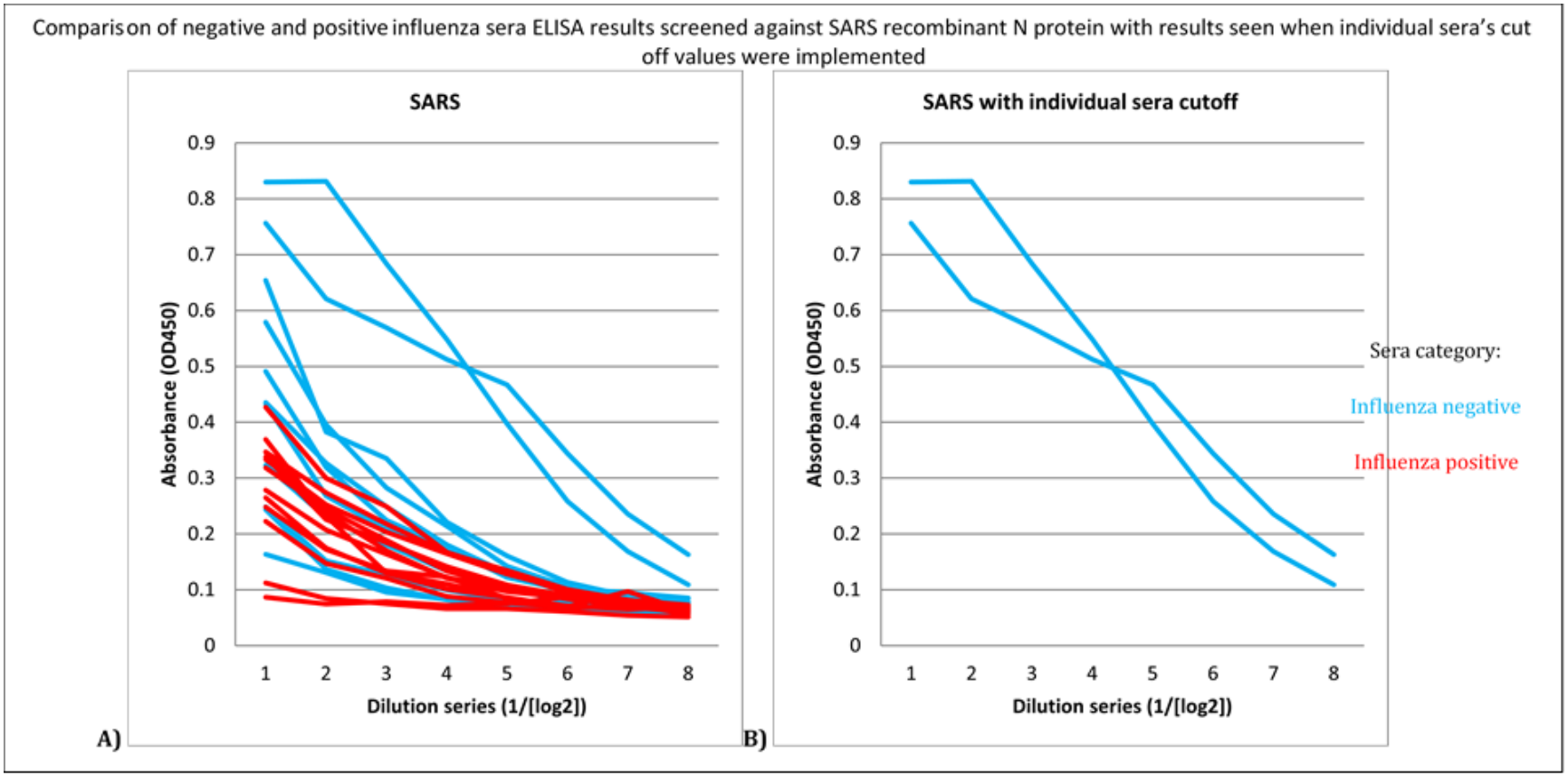


Figure 80: ELISA results showing the comparison of influenza positive vs influenza negative sera screened against recombinant SARS-CoV N protein, with results seen when individual sera's cut off values were implemented. Sera deemed influenza negative and positive are coloured blue and red respectively. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 80 shows the only two sera shown as having possible SARS-CoV reactivity were both from the influenza negative subgroup.

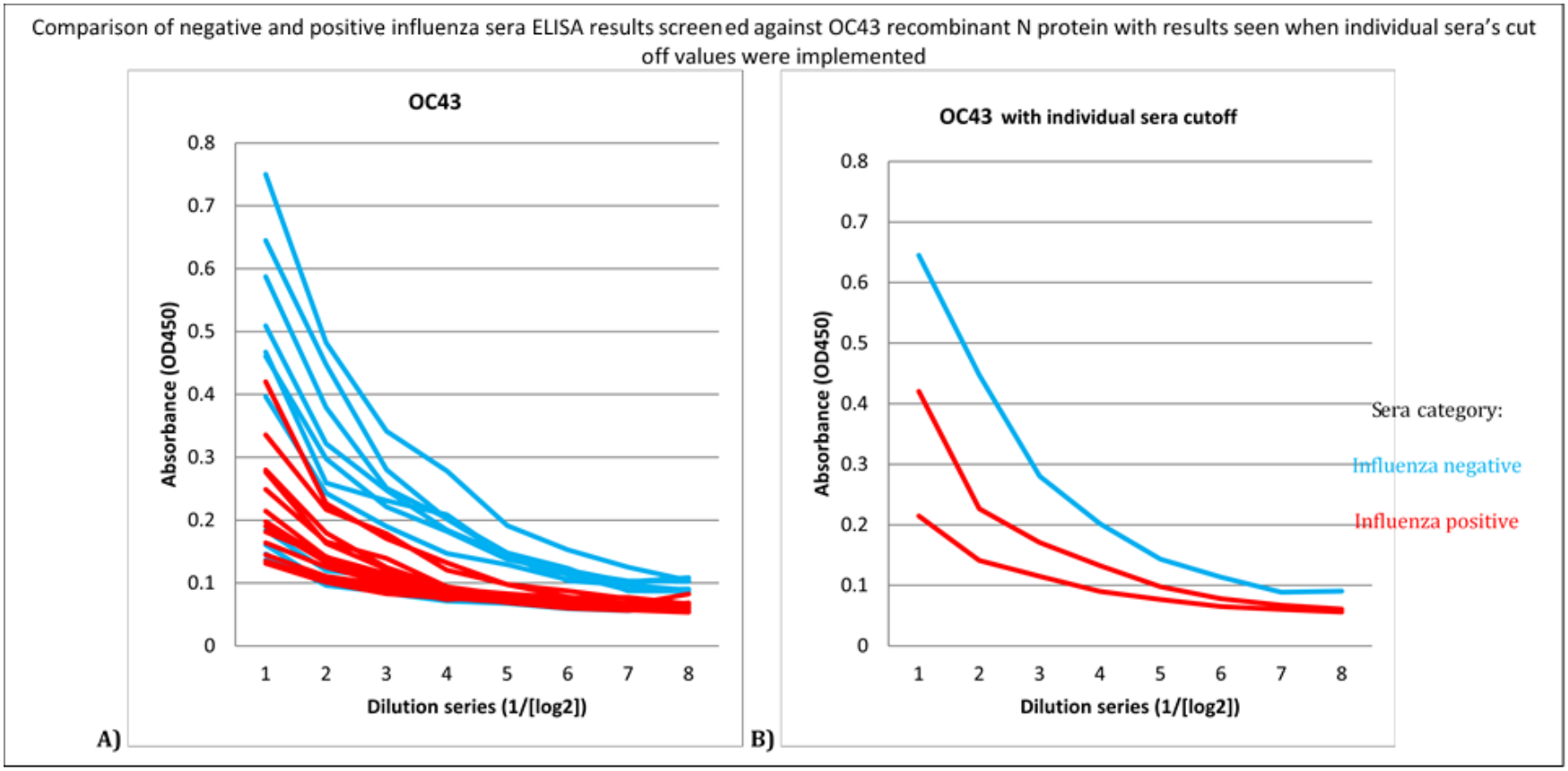


Figure 81: ELISA results showing the comparison of influenza positive vs influenza negative sera screened against recombinant OC43-CoV N protein, with results seen when individual sera's cut off values were implemented. Sera deemed influenza negative and positive are coloured blue and red respectively. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 81 shows that of the three sera deemed positive with cut off values two were influenza positive and one was influenza negative, more interestingly the graph without cut-off limits shown in figure A suggests that high OD450 readings seem to be seen in influenza negative sera overall.

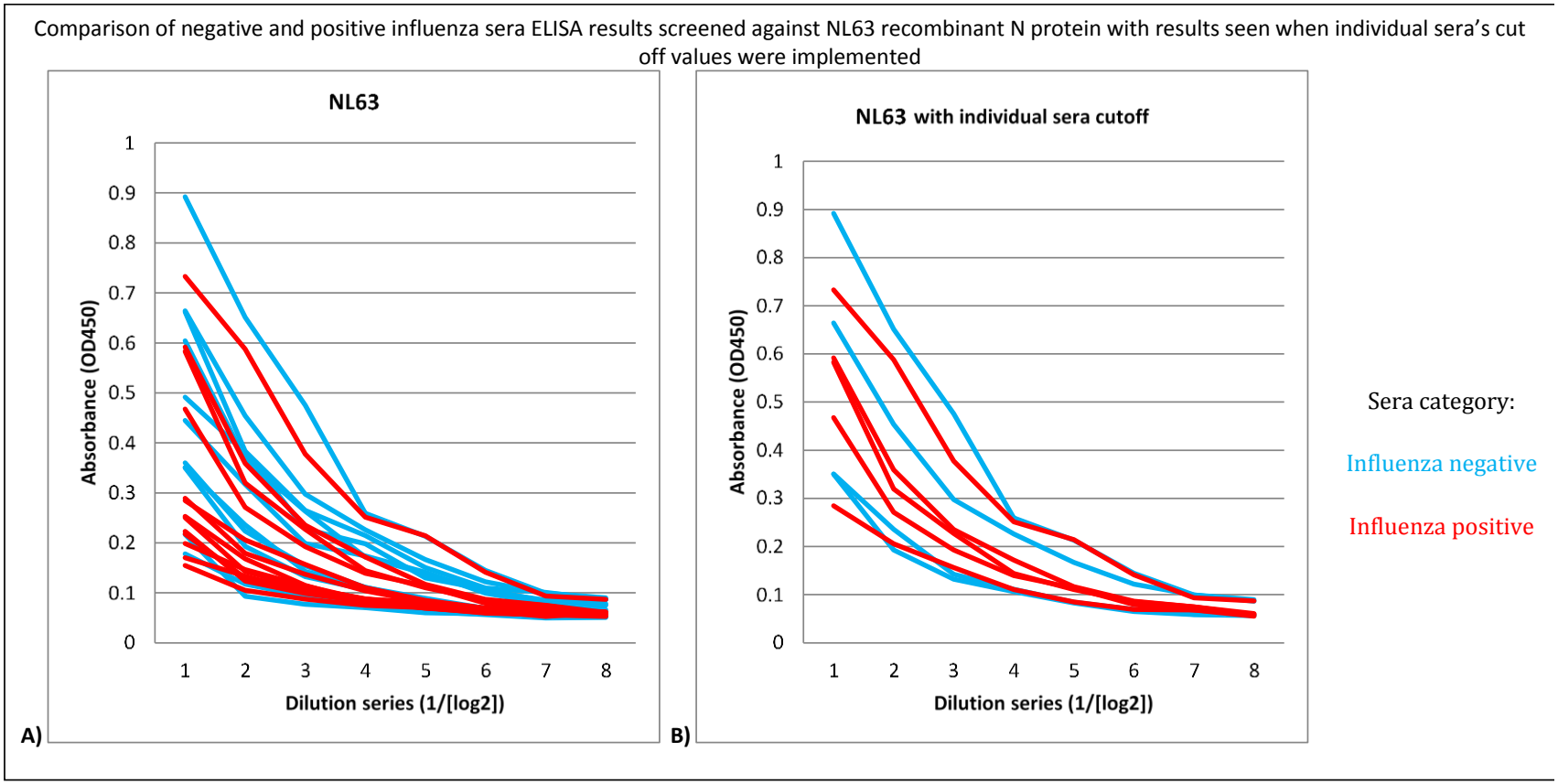


Figure 82: ELISA results showing the comparison of influenza positive vs influenza negative sera screened against recombinant NL63-CoV N protein, with results seen when individual sera's cut off values were implemented. Sera deemed influenza negative and positive are coloured blue and red respectively. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 82 shows there to be a fairly even split between influenza positive and influenza negative sera showing reactivity to recombinant NL63-CoV N protein once cut-off levels were in place.

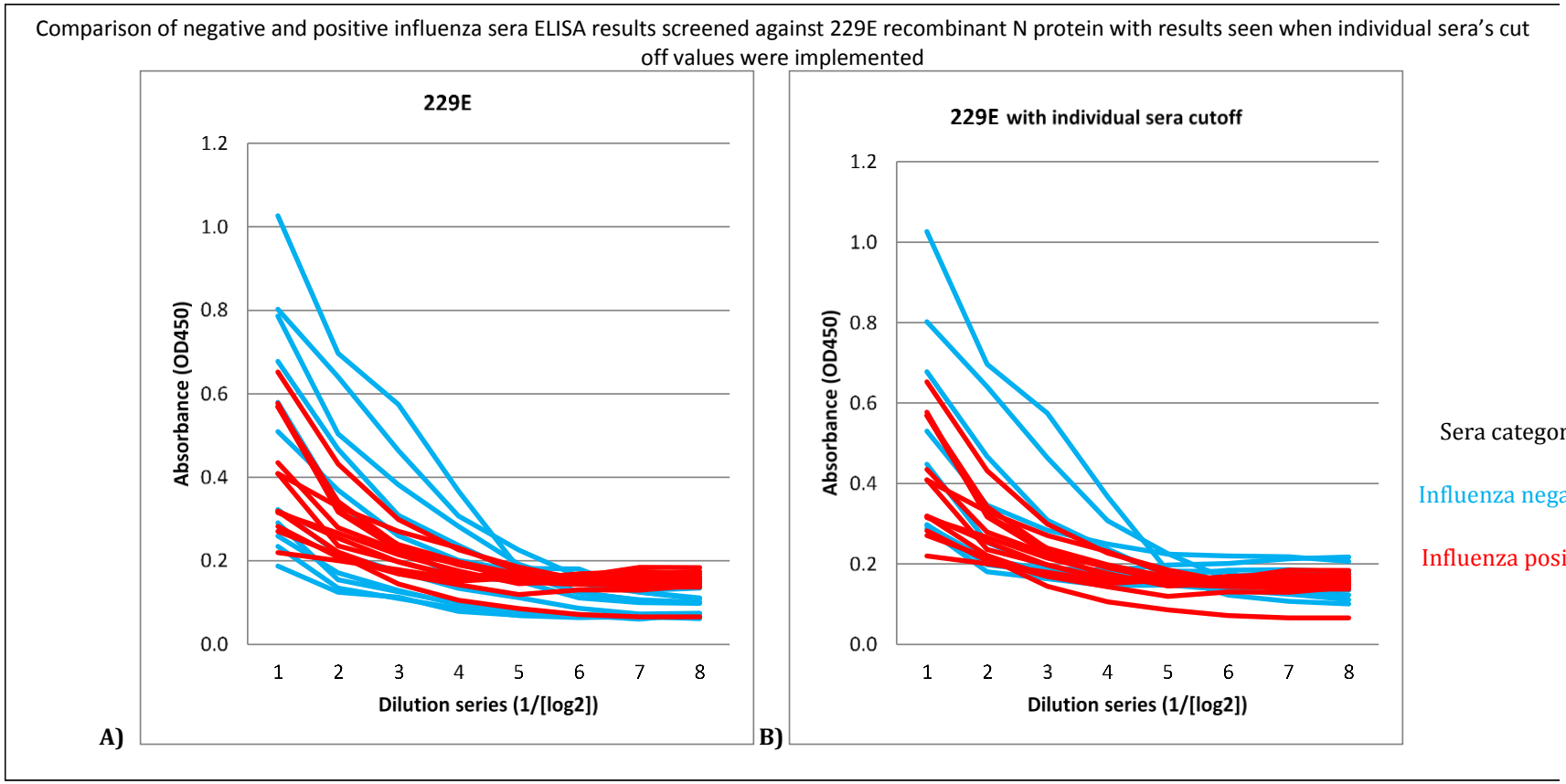


Figure 83: ELISA results showing the comparison of influenza positive vs influenza negative sera screened against recombinant 229E-CoV N protein, with results seen when individual sera's cut off values were implemented. Sera deemed influenza negative and positive are coloured blue and red respectively. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 83 shows there to be a fairly even split between influenza positive and influenza negative sera showing reactivity to recombinant 229E-CoV N protein once cut-off levels were in place.

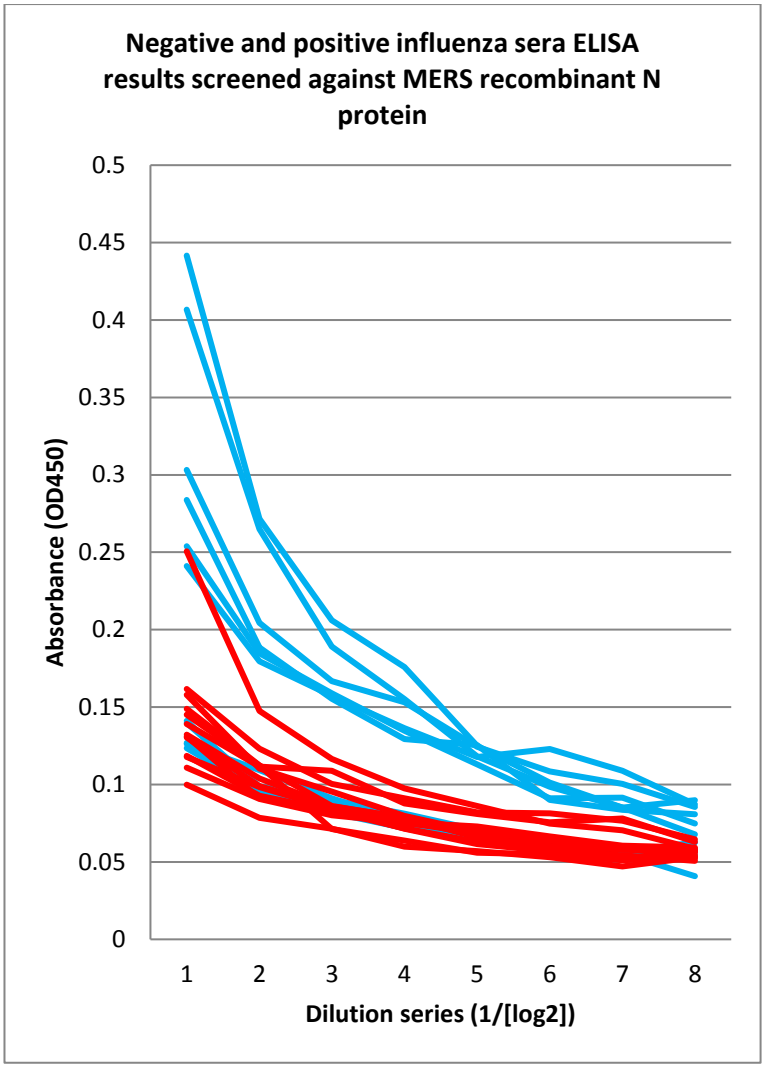


Figure 84: ELISA results showing the comparison of influenza positive vs influenza negative sera screened against recombinant MERS-CoV N protein. Sera deemed influenza negative and positive are coloured blue and red respectively. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 84 shows the colour coded results seen when influenza positive and negative sera was screened against recombinant MERS-CoV N proteins. Although it is highly unlikely that anybody screened would have true reactivity to MERS-CoV the data could be used as a negative control to work out cut-off values.

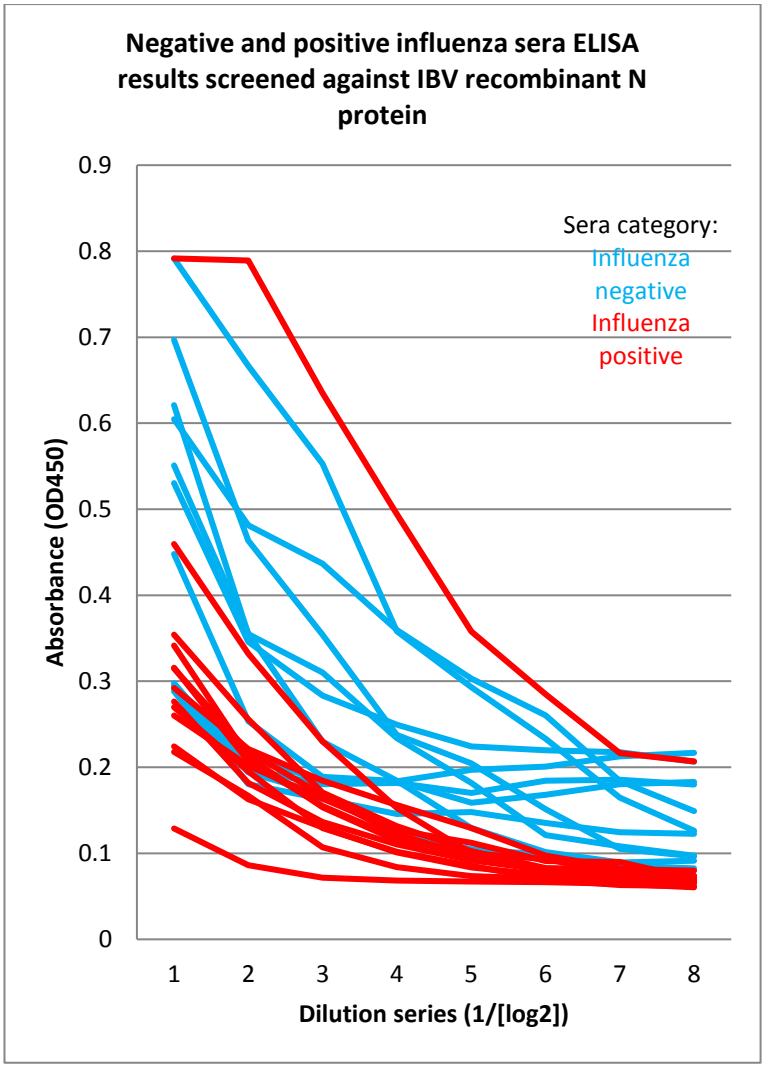


Figure 85: ELISA results showing the comparison of influenza positive vs influenza negative sera screened against recombinant IBV N protein. Sera deemed influenza negative and positive are coloured blue and red respectively. The dilution series started with the stock sera being diluted 1:50 and followed a two-fold dilution series. Error bars not shown due to lack of repeats.

Figure 85 shows the colour coded results seen when influenza positive and negative sera was screened against recombinant IBV N proteins. Similar to the results seen during screening with diabetic positive and negative sera the MERS-CoV results are actually of a lower magnitude OD450 than the IBV and may perhaps be a better data set to work out the cut-off level.

6.4.6.3 Statistical analysis of influenza positive/negative sera on Coronavirus N protein

Coronavirus N protein	Influenza negative	Influenza positive
SARS-CoV	2	0
MERS-CoV	0	0
OC43-CoV	1	0
NL63-CoV	1	3
229E-CoV	8	13
IBV	0	0

Table 18: The number of positive reactions to HCoV N proteins using ELISAs with both influenza positive and negative sera

Figures from Table 18 were used to create the contingency tables seen in Table 19 to allow Fisher's exact test to take place, similar to section 6.3.5.10.2.

6.4.6.3.1 2x2 contingency tables for each coronavirus recombinant N protein

SARS-CoV			
	Reactivity	No Reactivity	Total
Influenza positive	0	13	13
Influenza negative	2	9	11
Total	2	22	24
MERS-CoV and IBV			
	Reactivity	No Reactivity	Total
Influenza positive	0	13	13
Influenza negative	0	11	11
Total	0	24	24
NL63-CoV			
	Reactivity	No Reactivity	Total
Influenza positive	3	10	13
Influenza negative	1	10	11
Total	4	20	24
OC43-CoV			
	Reactivity	No Reactivity	Total
Influenza positive	0	13	13
Influenza negative	1	10	11
Total	1	23	24
229E-CoV			
	Reactivity	No Reactivity	Total
Influenza positive	13	0	13
Influenza negative	8	3	11
Total	20	4	24

Table 19: The six 2x2 contingency tables for each recombinant coronavirus N protein using suspected influenza positive and negative sera

For the SARS-CoV, IBV, MERS-CoV, NL63-CoV, OC43-CoV results displayed in Table 19, the following outcome was observed "*The two-tailed P value equals 1.0000, so the association between rows (groups) and columns (outcomes) is considered to be not statistically significant*". This allows the null hypothesis to be accepted and states that there is no association between influenza positive and negative status and coronavirus recombinant N protein reactivity seen in this study of a limited number of individual sera.

The smallest P value was seen in the case of 229E-CoV and was 0.0815, this time *"the association between rows (groups) and columns (outcomes) is considered to be not quite statistically significant*". Interestingly if the trend of 100% 229E-CoV reactivity in influenza positive and 73% 229E-CoV reactivity in influenza negative sera were found in a higher population/sample number such as 100 the P value would be 0.0001 and the result would be considered to be "extremely statistically significant". Showing the main downfall of this study to be the limited number of sera available to screen on.

It has been reported that as many as 20% of patients with influenza also have additional viral infections (Esper, Spahlinger et al. 2011). This can be somewhat explained by the times at which viruses circulate within the year; predominantly winter in the Northern hemisphere. Influenza has been identified with co-existing viruses (Peng, Zhao et al. 2009). Whether coinfection effects patient morbidity remains controversial in literature making the clinical relevance of respiratory virus copathogens unclear (Esper, Spahlinger et al. 2011). A 2011 study by Esper *et al* argued that certain copathogen pairings had higher clinical relevance than others. The group used multiplex RT-PCR to identify prominent viral copathogens in the H1N1 influenza pandemic; they used respiratory samples from 229 patients identified without influenza and 267 samples from patients with influenza and screened for the presence of 13 seasonal respiratory viruses, including coronaviruses. The study found that 13.1% of influenza samples were positive for 31 viral copathogens, the most prominent being rhinovirus (61.3%) followed by coronaviruses (16.1%).

At no point does the data from the study distinguish between NL62, OC42 and 229E-CoV, the three they conducted screening for.

Further work will be needed on a larger data set to test the null hypothesis and see if there is a link between influenza infection and coronavirus infection, specifically 229E-CoV.

Chapter 7: Summary and future work

7.1 Purpose of investigation

The emergence of SARS in 2003 highlighted an inadequacy in preparation for the first pandemic of the twenty-first century (de Wit, van Doremalen et al. 2016). Several months were needed before the causative agent was identified as SARS-CoV. Fortunately advances in next generation sequence technologies (NGST) paved the way for a quicker identification of MERS-CoV in 2012 (Zaki, van Boheemen et al. 2012). Most emerging diseases, such as SARS, MERS and Covid19 are zoonotic (Jones, Patel et al. 2008) indicating that viral cross-species transmission is a significant threat to human health (Chan, To et al. 2013, de Wit, van Doremalen et al. 2016). The high mortality rates shown by SARS, Covid19 and MERS demonstrate the importance of studying Coronaviridae as emerging human pathogens.

In 2015 it was suggested that a cluster of SARS-like CoV circulating in bat populations also show potential for human emergence (Menachery, Yount et al. 2015) (Drexler, Corman et al. 2014). It is also possible that there are additional CoVs circulating in other animal species that could go on to infect humans. A worldwide large-scale study screening samples from various wild animals would need to be conducted in order to investigate this further.

There is still a need for broad-spectrum vaccinations or therapeutic strategies for both current and future emerging Coronaviruses to be established (Omrani, Saad et al.). The development of animal models closer to the natural host targets is an important part of research aims moving forward. Although there are effective vaccine prototypes for both SARS-CoV and MERS-CoV available (Haagmans, van den Brand et al. 2016) (Roper and Rehm 2009, Honda-Okubo, Barnard et al. 2015) none are yet licensed, nor are there any licensed therapeutic drugs to treat either infection (Lu 2020). This could have disastrous consequences in the event of re-emergence.

In the case of MERS-CoV, vaccination of dromedary camels has been looked into (Haagmans, van den Brand et al. 2016) however there may be social complications by way of the reticence of camel owners to vaccinate their animals, although a dual vaccine against both MERS-CoV and camelpox virus has been suggested as a solution (Haagmans, van den Brand et al. 2016) in a similar design to a dual MERS-CoV and rabis vaccine (Wirblich, Coleman et al. 2017). Issues surrounding political and intellectual property rights play roles in the lack of vaccine currently on the market as pharmaceutical companies are unlikely to invest in development unless there is a current need (Simon, Claassen et al. 2005).

Alongside developments in vaccinations and therapeutic strategies, diagnostic techniques also require further study. Speed is a crucial factor when it comes to the identification of a new virus outbreak, as appropriate treatment regimen for patients and disease containment are most advantageous when implemented swiftly. This study focuses on the diagnostic aspect of Coronaviruses, both in terms of serosurveillance to test for current disease as well as retrospective infections.

The use of recombinant proteins as antigens may prove beneficial as they involve no biohazard risk, have the potential to be easily upscaled and the process can be readily automated for large-scale screening. This study had two research objectives; to create purified soluble N proteins for MERS-CoV, SARS-CoV, OC43-CoV, NL63-CoV, 229E-CoV and IBV, and to use these recombinant N proteins for sera testing, both of which have been met.

7.2 Results

7.2.1 A summary of the results and hypothesis

The purpose of this investigation was to ascertain if recombinant nucleocapsid proteins could be used as diagnostic antigens when screening sera suspected of having prior human coronavirus infection. Sequences were cloned into a commercial PTriEx1.1 vector, modified with a polyhistidine tag at the N terminal domain and a stop coding to prevent translation of the pre-existing polyhistine tag located at the C terminal domain of the vector, with the aim of producing soluble proteins. A variety of prokaryotic *E. coli* strains were investigated and SoluBL21 subsequently proved the most viable for all Coronaviruses recombinant nucleocapsid proteins produced. Proteolytic degradation was observed, most noticeably in the IBV sample, and so eukaryotic expression was also examined without a marked increase in success. As a result, a truncated version of the protein, omitting the central disordered region, was created and shown to elicit similar antigenic properties to the full-length construct. All six recombinant nucleocapsid proteins (SARS-CoV, MERS-CoV, NL63-CoV, OC43-CoV, 229E-CoV, IBV and ΔIBV) were produced and purified using IMAC and subsequently used to conduct a variety of ELISA screenings; showing great promise for their use as diagnostic antigens.

The individual hypotheses within this study were detailed in Chapter 6 and state that being a type one diabetic may predispose you to Coronavirus infection and secondly that showing signs of Coronavirus infection may correlate to an increased likelihood of an Influenza infection or vice versa. Due to the limited numbers of sera collected and screened neither hypothesis could be accepted, however the fact that 100% of influenza positive sera showed 229E-CoV reactivity remains an interesting find.

7.2.2 Existing research

7.2.2.1 Recombinant protein production

Although polyhistidine tags generally have no significant effect on the structure of the native protein (Carson, Johnson et al. 2007) a 2004 study compared four different vectors for expression of 20 human proteins with N- or C-terminal polyhistidine tags in *E. coli*, noting, among other things, the solubility of the target proteins (Woestenenk, Hammarstrom et al. 2004). The study showed that in general both N- and C-terminal polyhistine tags have a noticeable negative effect on protein solubility but found that the effect is target protein specific. Evidently in the case of the recombinant Coronavirus nucleocapsid proteins, the relocation from the C terminus to the N terminus elicited a positive effect in the proteins' solubility, an issue that had proved problematic in previous studies (McCrory 2009).

7.2.2.2 The use of recombinant Coronavirus nucleocapsids as diagnostic antigens

Similar studies using recombinant nucleocapsid proteins for diagnostic purposes have been implemented, some with an N terminus polyhistidine tag; screening SARS-CoV (Zuo, Mattern et al. 2005) and NL63-CoV (Zuwała, Golda et al. 2015), others with a C terminus polyhistidine tag including SARS-CoV (Haynes, Miao et al. 2007) and MERS-CoV (Chen, Chan et al. 2015).

This study is different to most in that it focuses on 5 HCoVs whereas predominantly studies chose to focus on one; mostly MERS-CoV or SARS-CoV. In examples whereby more than one Coronavirus is screened for RT-PCR appears the most common technique (Gaunt, Hardie et al. 2010) (Lau, Woo et al. 2006, Dominguez, Robinson et al. 2009, Mackay, Arden et al. 2012, Cabeca, Granato et al. 2013).

7.2.2.3 The benefits of truncated nucleocapsid proteins

Truncated proteins have similarly been examined and found to be beneficial. Blanchard *et al* divided the OC43-CoV N protein into three truncated proteins; N1 aa1-119, N2 aa120-332 and N3 aa333-448 with molecular weights of 17kDa, 25kDa and 17.5kDa respectively (Blanchard, Miao et al. 2011). The proteins contained a C terminal polyhistidine tag and were expressed in BL21

cells. 11 convalescent OC43-CoV sera were used to evaluate antibody response to the truncated proteins, the predominant response was found when using N3 (11 out of 11), followed by N2 (4 out of 11) then N1 (3 out of 11). As well as having the best response to OC43-CoV sera, N3 also had the least sequence homology to other HCoV N proteins. As a result, it was used to also screen sera from 229E-CoV and SARS-CoV positive patients to assess cross-reactivity. When the full-length protein was used 74% (20 out of 27) of 229E-CoV and 100% (20 out of 20) of the SARS-CoV sera showed cross reactivity. In contrast, when the truncated N3 protein was used, these figures reduced to 7.4% (2 out of 27) for 229E-CoV and 30% for SARS-CoV (6 out of 20). Similar studies have also been conducted using truncated SARS-CoV nucleocapsid proteins (Yu, Le et al. 2007, Lee, Lee et al. 2008). These studies highlight the importance that truncated proteins may provide not only in terms of additional structural stability but also in their potential to decrease cross-reactivity.

7.2.2.4 Screening

As previously mentioned, Coronaviruses that are potentially pathogenic towards humans may already be circulating in animal species, including bats, rodents and livestock. Viral surveillance studies are a crucial way of examining what strains exist in the environment that may be of concern (Coleman and Frieman 2014).

7.2.2.4.1 Bats

Small scale studies have highlighted the potential for a SARS-like virus, SHC014-CoV currently circulating in Chinese horseshoe bat populations, to emerge, as it was shown *in vitro* to be able to utilize ACE2 as a receptor and replicate in primary human airway cells as well as being shown to cause pathogenesis *in vivo* in mice (Menachery, Yount Jr et al. 2015). A novel MERS-like Coronavirus has also been reported in *Pipistrellus hesperidus* bats in Uganda giving further credence to MERS-CoV having a bat origin (Anthony, Gilardi et al. 2017). Evolution notwithstanding, the virus' spike protein differs sufficiently that it is not deemed as being a threat

to human health. Recent suspected zoonotic spread includes that of a new bat-HKU2-like porcine Coronavirus found in piglets with diarrhoea in China (Gong, Li et al. 2017).

7.2.2.4.2 Camels

Similar work has been successfully conducted in order to screen camels (Song, Ha et al. 2015). Due to their role in infection transmission in the case of MERS-CoV, metagenomic sequence analysis on nasopharyngeal swabs of 108 MERS-CoV positive camels were studied to ascertain if there are any additional, currently unidentified, viruses in camels' upper respiratory tracks, with a potential to infect humans (Li, Khalafalla et al. 2017). The study collected a total of 846.72 million high-quality reads from the samples, of which 2.88 million (0.34%) were related to viral sequences. 512.63 million (60.5%) and 50.87 million (6%) matched bacterial and eukaryotic sequences respectively. Sequences related to mammalian viruses from 13 genera in 10 viral families were identified, including Coronaviridae, Nairoviridae, Paramyxoviridae, Parvoviridae, Polyomaviridae, Papillomaviridae, Astroviridae, Picornaviridae, Poxviridae, and Genomoviridae. The study showed that although some viral sequence belonged to pre-known camel or human viruses, others were from potentially novel camel viruses with only limited sequence similarity to virus sequences found in GenBank. Five potentially novel virus species or strains were identified.

7.2.2.4.3 Other animals

Studies published in 2017 have also described serological evidence of Coronavirus infections in hamadryas baboons (*Papio hamadryas hamadryas)* in the Kingdom of Saudi Arabia (Olarinmoye, Olugasa et al. 2017). It is also thought that genetic recombination is responsible for the discovery of a novel canine respiratory Coronavirus (Lu, Wang et al. 2017)

7.3 Importance of the study

Few studies have been conducted on multiple Coronavirus-detection in immunosurveillance studies, with most focusing on the two emerging viruses of SARS-CoV and MERS-CoV. Although the screening studies provided no statistically significant results due to a low sample number, the principle remains that recombinant nucleocapsid proteins can be used in future

immunosurveillance studies to ascertain the seroprevalence of human Coronaviruses amongst populations.

7.3.1 Implications and practical applications of the study

The results of this study could provide the foundation for a functional ELISA screening kit for multiple Coronavirus. Current ELISA kits available seem to only specify a positive/negative result for Coronaviruses in general as opposed to giving a detailed account of which may be present. Pre-existing kits including a general Coronavirus ELISA kit to detect IgG available from my Biosource (MyBioSource 2020) and a sandwich ELISA kit to detect MERS-CoV specifically (MyBioSource 2020).

7.3.2 Limitations of the study

The study had several limitations, namely the limited number of sera collected and screened decreasing the statistical significance of any result.

7.3.2.1 KHU1-CoV omitted

One limitation of the study was the incomplete set of HCoVs used, as human Coronavirus KHU1-CoV (Drexler, Corman et al. 2014) was left out of this study due to lack of availability.

7.3.2.2 Cross-reactivity

Betacoronaviruses have been found to cross-react, which can cause false positives and limit the use of the procedure in diagnostics (Corman, Muller et al. 2012) (Woo, Lau et al. 2012) (Qiu, Shi et al. 2005) (Maache, Komurian-Pradel et al. 2006) (Che, Qiu et al. 2005). One way to compensate is to use an additional confirmation assay, such as a viral neutralization test (VNT), on any sera deemed positive by an ELISA. Although this is labour-intensive and would not be ideal in an outbreak situation.

7.3.2.3 Positive and negative sera

As discussed in the Chapter 6 and 7, positive and negative sera would have been beneficial in order to establish a scientific cut-off limit; although this may have been possible for the influenza screens

there was no such sera available for the Coronavirus screening and so the recombinant IBV data was utilised as a negative control instead.

7.3.2.4 Sample size

The small sample sizes screened hampered statistical analysis and did not enable any correlation for either hypothesis to be determined.

7.3.2.5 Analytical and statistical analysis

The purpose of this study was to test out the practicality of using recombinant nucleocapsid proteins as diagnostic antigens in serosurveillance as opposed to drawing conclusions on prevalence of Coronaviruses in the population screened and so statistical analysis was limited.

7.4 Recommendations for further research

Further work could be conducted in several areas in order to best improve this study, as detailed below.

7.4.1 ELISA optimisation and performance

Before the ELISA method could be expected to perform as a true diagnostic test it would need to be optimised and have the optimal N protein concentration, serum dilution and conjugate dilution assessed.

The main flaw with the ELISAs conducted in the study is the lack of data corresponding to their sensitivity (measure of the proportion of positive results correctly identified) and specificity (measure of the proportion of negatives that are correctly identified) due to a lack of positive and negative control serum for the Coronavirus examples. This makes it impossible to truly compare the test's validity against other methods, although the data would not be hard to generate with the necessary sera.

True positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) results could be looked at to evaluate the diagnostic test's performance.

Accuracy/sensitivity = $\frac{TP}{TP+FN}$ (To determine how good the assay picked up actual positive)

Precision/specificity = $\frac{TN}{TN+FP}$ (To determine how good the assay excluded actual negative)

Positive Predictive Value (PPV) = $\frac{TP}{TP+FP}$ (The probability of observed positive result)

Negative Predictive Value (NPV) = $\frac{TN}{TN+FN}$ (The probability of observed negative result)

Where: **TP** = the number of tests correctly assigned as positive.
TN = the number of tests correctly assigned as negative.
FP = the number of tests incorrectly assigned as positive.
FN = the number of tests incorrectly assigned negative

Ideally diagnostic tests would have 100% sensitivity and specificity however realistically this is almost impossible to achieve and so values over 90% are deemed to have high credibility in differential diagnosis (Parikh, Mathai et al. 2008).

In the case of the influenza sera screened against recombinant H1 and H3 proteins, sensitivity and specificity could be examined, however the cut-off point was merely set by the highest negative value, adding bias into equations as FP will automatically equal 0. That being said, using the current cut-off level and the optimum sera dilution of 1:2430 or 1:7290, only 10 out of the 13 samples were detected as TP leaving 3 FN and 11 TN, giving the test a specificity of 1 and sensitivity of 0.77.

A second use of positive and negative sera sets would be in validating assays. A positive control mean (PC) and negative control mean (NC) could be ascertained and OD boundaries set accordingly i.e. assay is valid when PC–NC > x and NC ≤ y, the figures would be arbitrarily set by the test conductor but could remain constant throughout subsequent assays.

7.4.2 Statistical analysis

"There is no single statistic that can adequately represent the agreement between a diagnostic test and a reference standard" (Okeh and Okoro 2012). The cut-off points generated in this study were done crudely and could be improved upon again with the use of known negative and positive sera. Statistical analysis could also be improved with the use of statistical software and, perhaps more importantly, a larger data set. A variety of measures can be used to summarise and quantify the

diagnostic accuracy of a test including amongst others; sensitivity, specificity, ROC curve, likelihood ratio (LR) for positive/negative test, odds ratio (OR), error rates and confidence interval. Perhaps the one most likely to be beneficial in a study such as this is the ROC curve which could be created by plotting the true positive rate (TPR), equivalent to sensitivity, against the false positive rate (FPR), equal to 1- specificity, at various threshold settings in order to analyse cost (false positives) vs benefit (true positives) when it comes to diagnostic decision making. A diagonal line can be used to divide the ROC space whereby sensitivity=specificity and points above the diagonal represent results that are better than random and points below the line represent results worse than random. The perfect classification would be a test with no false negatives (100% sensitivity) and no false positives (100% specificity) this would like on the (0,1) point of the ROC space graph (Powers 2011).

7.4.3 Further investigation into protein structure

The truncated IBV protein provided similar ELISA results to that of the full-length protein, it would be interesting to apply the same truncation to the remaining recombinant nucleocapsid proteins to see if they too become more structurally stable without compromising their antigenicity.

7.4.4 Other methodologies

Once the sensitivity and specificity of the ELISA method using the recombinant Coronavirus nucleocapsid proteins is established using a larger sample set of both known positive and negative sera; the method can be compared with other diagnostic tests such as PCR, RT-PCR and IFA. Newer methods could also be compared against including a MassARRAY matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) which has already been used to screen Alphavoronavirus and Betacoronavirus in bat, rodent and human samples (Xiu, Zhang et al. 2017). The results of the Xiu *et al* study showed good agreement with the results of metagenomic analysis or PCR-sequencing and could also be used to provide phylogenetic

evidence about unknown CoVs. Quantitative polymerase chain reaction (qPCR) has been shown to be a sensitive detection method to detect HCoV sensitivity and has the additional benefit of being quantitative (Vijgen, Keyaerts et al. 2005, Gaunt, Hardie et al. 2010). The downside to qPCR-based methods is that most target a single gene of a virus. Although a few studies have developed methods to target different genes of MERS-CoV such as *upE, ORF1b and ORF1a* (Corman, Eckerle et al. 2012, Corman, Muller et al. 2012) as well as the nucleocapsid gene (Lu, Whitaker et al. 2014). The MALDI-TOF method targets two or four genes and showed that using multiple targets was beneficial to avoid false negative results (Xiu, Zhang et al. 2017). NGST have been widely used in pathogen detection (Wu, Yang et al. 2016) as they do not require prior knowledge of targets which is beneficial in the case of CoVs as they have a high frequency of mutation and recombination. However, NGST often come with a high cost, long turn-around time and require complex bioinforamatic analysis which can limit the number of samples that can be sequenced and the overall practicality of the method. Other methods such as MALDI-TOF could be used initially and then positive results further investigated using NGST.

7.4.5 Large scale human screening

Future larger studies with statistical analyses would be of interest. Sera could be examined both from cohorts suspected of having prior Coronavirus exposure, people from the Middle East who may have been exposed to MERS-CoV for example, as well as cohorts not suspected of having any exposure above that of common Coronaviruses. Statistical analysis could be conducted to see to what extent age, gender, season, socio-economic status, location and other factors, effect likelihood of each Coronavirus infection. One such study by Liljander, Meyer et al. focused on the detection of MERS-CoV antibodies in 1122 human sera collected as part of a household survey conducted during 2013–2014 in 2 eastern counties of Kenya, Garissa and Tana River (Liljander, Meyer et al. 2016). Where possible gender, age and occupation was recorded. The households of nearly all participants kept or owned livestock, including goats, sheep, cattle, and donkeys; although camels are not commonly kept, they are widespread in the region. The study used a

commercial anti-MERS-CoV recombinant ELISA, based on the viral spike protein. The study identified 16 samples as having positive results. Although the number of samples tested was approximately one tenth of the number of samples tested during a similar Saudi Arabia study conducted by Muller, Meyer *et al.*, the proportion of seropositive specimens was shown to be similar (Müller, Meyer et al. 2015). Prior to the study autochthonous human MERS-CoV infections have not been recorded in Africa, resulting in the hypothesis that there is a difference in transmission between Africa and the Arabian Peninsula and raising doubts as to the involvement of camels in the disease spread. The study showed the presence of previously unrecorded human MERS-CoV infections in Kenya; the apparent lack of reported cases may therefore be down to either a lack of a well-developed public health system and lack of diagnoses or perhaps a less virulent strain. The recombinant nucleocapsid proteins produced in this study could be similarly used to conduct large scale human sera screens.

7.4.6 Animal screening

Following on from the paper highlighting the presence of CoV in European hedgehogs (*Erinaceus europaeus)* (Corman, Kallies et al. 2014) the recombinant nucleocapsid proteins produced in this study could be used to conduct ELISAs on animal sera too. Feral sera samples could be sourced from the Royal Society for the Protection of Animals (RSPCA) and/or Animal Health and Veterinary Laboratory Agency (AHVLA); who have a large selection of animal and bat sera as a result of their rabies screening program.

7.4.7 Production of a functional screening kit

As well as using the recombinant nucleocapsid proteins to produce ELISA kits to conduct serosurveillance, other styles of kits could also be manufactured such as an immunochromatographic test which is beneficial in that in can be used for diagnosis in remote areas where laboratory facilities are not available (Guan, Chen et al. 2004). Such kit has been produced using recombinant MERS-CoV nucleocapsid protein with a relative sensitivity and

specificity of 93.90% and 100%, respectively, compared to that of the UpE and Orf1A RT-PCR (Song, Ha et al. 2015) shown in Figure 86 .

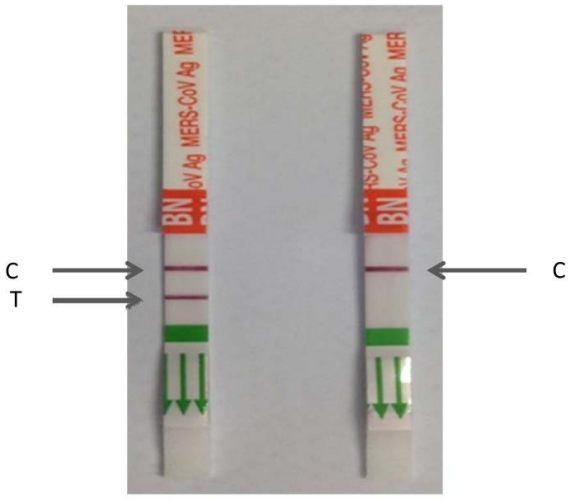


Figure 86: Two test strips showing the immunochromatographic screening of sera to MERS-CoV. C= control line; T= test line. A) shows a positive result and B) a negative result. Image taken from (Song, Ha et al. 2015).

Other potential tests that could be produced as a result of this study include an antigen spot test (AST). This has been used previously to detect bovine Coronavirus (BCV) in bovine faecal samples (Gaber and Kapil 1999). The Gaber and Kapil study compared the sensitivity and specificity of AST for the detection of BCV antigen to those of a sandwich ELISA and a haemagglutination assay (HA). 347 field samples were collected and screened using all three methods. 94.2% were deemed positive using AST, 91.4% by ELISA, and 86.7% by HA. The sensitivity of the AST was determined to be 100% however the test had a low specificity of 67% as a result of a lower limit of detection (10^4 viral particles per mL in a 10% faecal suspension). As mentioned in section 1.2.10.2 such kits have been manufactured for SARS-CoV-2 detection.

7.5 Summary

7.5.1 Pre-diagnosis

Even before diagnostic methods come in to play healthcare providers play a crucial role in disease containment. During a six-month period (December 2015–May 2016) a series of unannounced "mystery patient drills" were carried out to assess New York City Emergency Departments' (EDs) abilities to identify and respond to patients with communicable diseases of public health concern (Foote, Styles et al. 2017). The aim of the study was to highlight significant health care system vulnerabilities and stress the importance of rapid recognition and isolation of patients with potentially severe infectious diseases. The drill scenarios presented a patient reporting signs or symptoms and having a travel history consistent with possible measles or MERS. The expectation was that once the patient was identified as being at high risk for having a communicable disease with a potential for respiratory transmission, they would be asked to wear a mask and be placed into an airborne infection isolation room. The evaluators used a variety of infection control performance measures. The exercise was considered successful/passed if the patient was given a mask and isolated from other patients and staff members, and unsuccessful whereby waiting time exceeded 30 minutes before triage which was deemed as a fail.

In total ninety-five drills (53 measles and 42 MERS) were conducted in 49 EDs with patients masked and isolated in 78% of drills. Median time from entry to masking was 1.5 minutes (range = 0–47 minutes) and from entry to isolation was 8.5 minutes (range = 1–57). Overall, 80% patients were asked about recent fevers, and 85% were asked about recent travel. In total 88% patients were given a mask; 85% patients in the measles scenarios and 93% patients in the MERS scenarios. Nineteen (39%) of 49 hospitals failed at least one drill.

Other infection control practices found that only 36% of staff members performed personal hand hygiene. In the 76 (80%) drills that resulted in the patient being isolated, precaution signage was posted outside the patient's airborne isolation room of 53 (70%), and staff members used recommended PPE when entering these rooms in 56 (74%) drills.

The results clearly show that hospitals varied in their ability to identify potentially infectious patients and implement recommended infection control measures in a timely manner. Drill findings were used to inform hospital improvement planning to more rapidly and consistently identify and isolate patients with a potentially highly infectious disease. Studies evaluating hospital emergency plans are important, as waiting areas have been shown to facilitate the transmission of infections, to patients and health care workers, leading to spread within hospitals and surrounding communities (McDonald, Simor et al. 2004, Adini, Goldberg et al. 2008) (Maltezou and Wicker 2013). The findings were used to set a performance goals of 1 minute from entry to masking and 10 minutes from entry to isolation. The overall median time from entry to isolation achieved in this study (8.5 minutes) was comparable to times achieved in an earlier Ebola drill analysis (9 minutes) (Foote, Daver et al. 2017).

7.5.2 Diagnosis

Clearly diagnostic testing remains a crucial element of disease monitoring and potential containment. This study has successfully shown that recombinant nucleocapsids proteins can be used as diagnostic antigens in regard to human coronavirus cases and refers to a variety of ways in which they could be commercialised into functional screening kits. Further work could be undertaken to see if truncating the N proteins could improve specificity and sensitivity.

7.5.3 Current situation and concluding remarks

The emergence of novel pathogens is a global cause for concern. A team at University College London has developed a computer-modelling program that uses data about societal impact on the environment, including climate and changes in animal habitat, to predict where future zoonotic pandemics might emerge (Hassell, Newbold et al. 2021). This approach successfully predicted the origins, but not the specific timing, of the Ebola outbreak (Gibb, Redding et al. 2020). As mentioned in the prologue the timing of the submission of this thesis coincides with an outbreak of SARS-CoV-2. Whilst the pandemic was somewhat predictable given the nature of coronaviruses, globally the world seemed ill-prepared for disease treatment, diagnosis or containment.

Although RT-PCR is the standard method for diagnosing a SARS-CoV-2 infection high false negative rates have been reported (Li, Yi et al. 2020) (Guo, Ren et al. 2020). Scientists around the world are searching to find a diagnostic kit with a high enough sensitivity and specificity to be of use in clinical test to help with disease control. An accurate and rapid test could allow large number of infected and asymptomatic carriers to be identified and help prevent virus transmission as well as allowing for timely treatment of patients. Whilst figures for SARS-CoV-2 infections and death seem large (at the time of writing, 17th August 2021, WHO reports Covid-19's case numbers are estimated to be 207,784,507 resulting in 4,370,424 deaths) the fatality rate of SARS-CoV-2 remains lower than both that of MERS-CoV and SARS-CoV.

It remains possible, if not probable, that novel coronaviruses may emerge in future, potentially leading to further pandemics. The ability to produce a rapid, sensitive, specific, cost-effective point of care test for the current and potential future epidemics and pandemics, whether coronavirus or otherwise, remains of high important.

References

(ECDC), E. C. f. D. P. a. C. (2020). Options for the use of rapid antigen tests for COVID-19 in the EU/EEA and the UK. Stockholm, ECDC.

A., D. C., A. C. Ghani, G. M. Leung, A. J. Hedley, C. Fraser, S. Riley, L. J. Abu-Raddad, L. M. Ho, T. Q. Thach, P. Chau, K. P. Chan, T. H. Lam, L. Y. Tse, T. Tsang, S. H. Liu, J. H. Kong, E. M. Lau, N. M. Ferguson and R. M. Anderson (2003). "Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong." Lancet **361**(9371): 1761-1766.

Adams, S. F. (1926). "The Seasonal Variation in the Onset of Acute Diabetes." Archives of Internal Medicine **37**(6): 861.

Adedeji, A. O., K. Singh, A. Kassim, C. M. Coleman, R. Elliott, S. R. Weiss, M. B. Frieman and S. G. Sarafianos (2014). "Evaluation of SSYA10-001 as a replication inhibitor of severe acute respiratory syndrome, mouse hepatitis, and Middle East respiratory syndrome coronaviruses." Antimicrob Agents Chemother **58**(8): 4894-4898.

Adini, B., A. Goldberg, R. Cohen and Y. Bar-Dayan (2008). "Relationship between standards of procedures for pandemic flu and level of hospital performance in simulated drills." Ann Emerg Med **52**(3): 223-229.

Adney, D. R., H. Bielefeldt-Ohmann, A. E. Hartwig and R. A. Bowen (2016). "Infection, Replication, and Transmission of Middle East Respiratory Syndrome Coronavirus in Alpacas." Emerg Infect Dis **22**(6): 1031-1037.

Adney, D. R., L. Wang, N. van Doremalen, W. Shi, Y. Zhang, W. P. Kong, M. R. Miller, T. Bushmaker, D. Scott, E. de Wit, K. Modjarrad, N. Petrovsky, B. S. Graham, R. A. Bowen and V. J. Munster (2019). "Efficacy of an Adjuvanted Middle East Respiratory Syndrome Coronavirus Spike Protein Vaccine in Dromedary Camels and Alpacas." Viruses **11**(3).

Agency, E. M. (2016). "Public summary of opinion on orphan designation. Interferon alfa-n3 for the treatment of Middle East respiratory syndrome." from http://www.ema.europa.eu/docs/en_GB/document_library/Orphan_designation/2016/01/WC500199406.pdf.

Agnihothram, S., R. Gopal, B. L. Yount, Jr., E. F. Donaldson, V. D. Menachery, R. L. Graham, T. D. Scobey, L. E. Gralinski, M. R. Denison, M. Zambon and R. S. Baric (2014). "Evaluation of serologic and antigenic relationships between middle eastern respiratory syndrome coronavirus and other coronaviruses to develop vaccine platforms for the rapid response to emerging coronaviruses." J Infect Dis **209**(7): 995-1006.

Al-Tawfiq, J. A., H. Momattin, J. Dib and Z. A. Memish (2014). "Ribavirin and interferon therapy in patients infected with the Middle East respiratory syndrome coronavirus: an observational study." Int J Infect Dis **20**: 42-46.

Alexandersen, S., A. Chamings and T. R. Bhatta (2020). "SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication." Nat Commun **11**(1): 6059.

Almazan, F., M. L. DeDiego, I. Sola, S. Zuniga, J. L. Nieto-Torres, S. Marquez-Jurado, G. Andres and L. Enjuanes (2013). "Engineering a replication-competent, propagation-defective Middle East respiratory syndrome coronavirus as a vaccine candidate." mBio **4**(5): e00650-00613.

Alraddadi, B. M., J. T. Watson, A. Almarashi, G. R. Abedi, A. Turkistani, M. Sadran, A. Housa, M. A. Almazroa, N. Alraihan, A. Banjar, E. Albalawi, H. Alhindi, A. J. Choudhry, J. G. Meiman, M. Paczkowski, A. Curns, A. Mounts, D. R. Feikin, N. Marano, D. L. Swerdlow, S. I. Gerber, R. Hajjeh and T. A. Madani (2016). "Risk Factors for Primary Middle East Respiratory Syndrome Coronavirus Illness in Humans, Saudi Arabia, 2014." Emerg Infect Dis **22**(1): 49-55.

Andersen, K. R., N. C. Leksa and T. U. Schwartz (2013). "Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification." Proteins **81**(11): 1857-1861.

Anderson, E. J., N. G. Rouphael, A. T. Widge, L. A. Jackson, P. C. Roberts, M. Makhene, J. D. Chappell, M. R. Denison, L. J. Stevens, A. J. Pruijssers, A. B. McDermott, B. Flach, B. C. Lin, N. A. Doria-Rose, S. O'Dell, S. D. Schmidt, K. S. Corbett, P. A. Swanson, 2nd, M. Padilla, K. M. Neuzil, H. Bennett, B. Leav, M. Makowski, J. Albert, K. Cross, V. V. Edara, K. Floyd, M. S. Suthar, D. R. Martinez, R. Baric, W. Buchanan, C. J. Luke, V. K. Phadke, C. A. Rostad, J. E. Ledgerwood, B. S. Graham, J. H. Beigel and R. N. A. S. G. m (2020). "Safety and Immunogenicity of SARS-CoV-2 mRNA-1273 Vaccine in Older Adults." N Engl J Med **383**(25): 2427-2438.

Anthony, S. J., K. Gilardi, V. D. Menachery, T. Goldstein, B. Ssebide, R. Mbabazi, I. Navarrete-Macias, E. Liang, H. Wells, A. Hicks, A. Petrosov, D. K. Byarugaba, K. Debbink, K. H. Dinnon, T. Scobey, S. H. Randell, B. L. Yount, M. Cranfield, C. K. Johnson, R. S. Baric, W. I. Lipkin and J. A. Mazet (2017). "Further Evidence for Bats as the Evolutionary Source of Middle East Respiratory Syndrome Coronavirus." mBio **8**(2): e00373-00317.

Arabi, Y. M., A. A. Arifi, H. H. Balkhy, H. Najm, A. S. Aldawood, A. Ghabashi, H. Hawa, A. Alothman, A. Khaldi and B. Al Raiy (2014). "Clinical course and outcomes of critically ill patients with Middle East respiratory syndrome coronavirus infection." Ann Intern Med **160**(6): 389-397.
Arbeitskreis Blut, U. (2009). "Influenza Virus." Transfus Med Hemother **36**(1): 32-39.
Baden, L. R., H. M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Rouphael, C. B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B. S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, T. Zaks and C. S. Group (2021). "Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine." N Engl J Med **384**(5): 403-416.
Banatvala, J. E., J. Bryant, G. Schernthaner, M. Borkenstein, E. Schober, D. Brown, L. M. De Silva, M. A. Menser and M. Silink (1985). "Coxsackie B, mumps, rubella, and cytomegalovirus specific IgM responses in patients with juvenile-onset insulin-dependent diabetes mellitus in Britain, Austria, and Australia." Lancet **1**(8443): 1409-1412.
Barkley, M. D., A. D. Riggs, A. Jobe and S. Burgeois (1975). "Interaction of effecting ligands with lac repressor and repressor-operator complex." Biochemistry **14**(8): 1700-1712.
Barros, M. M., S. C. Cartagena and F. L. Bavestrello (2005). "[Prevention of community-acquired pneumonia in adults]." Rev Chilena Infectol **22 Suppl 1**: s67-74.
Bell, C. E. and M. Lewis (2000). "A closer view of the conformation of the Lac repressor bound to operator." Nat Struct Biol **7**(3): 209-214.
Berg, J. M., J. L. Tymoczko and L. Stryer (2012). Biochemistry/Jeremy M. Berg, John L. Tymoczko, Lubert Stryer; with Gregory J. Gatto, Jr, New York: WH Freeman.
Berry, M., T. L. Manasse, Y. J. Tan and B. C. Fielding (2012). "Characterisation of human coronavirus-NL63 nucleocapsid protein." African Journal of Biotechnology **11**(75): 13962-13968.
Blanchard, E. G., C. Miao, T. E. Haupt, L. J. Anderson and L. M. Haynes (2011). "Development of a recombinant truncated nucleocapsid protein based immunoassay for detection of antibodies against human coronavirus OC43." J Virol Methods **177**(1): 100-106.
Blowey, R., C. Gray, J. Griffiths and J. Rowe (2017). "Bovine TB in the pilot badger cull zone in Gloucestershire." Vet Rec **181**(9): 243.
Booth, C. M., L. M. Matukas, G. A. Tomlinson, A. R. Rachlis, D. B. Rose, H. A. Dwosh, S. L. Walmsley, T. Mazzulli, M. Avendano, P. Derkach, I. E. Ephtimios, I. Kitai, B. D. Mederski, S. B. Shadowitz, W. L. Gold, L. A. Hawryluck, E. Rea, J. S. Chenkin, D. W. Cescon, S. M. Poutanen and A. S. Detsky (2003). "Clinical features and short-term outcomes of 144 patients with SARS in the greater Toronto area." JAMA **289**(21): 2801-2809.
Borchers, A. T., R. Uibo and M. E. Gershwin (2010). "The geoepidemiology of type 1 diabetes." Autoimmun Rev **9**(5): A355-365.
Breslin, J. J., L. G. Smith and J. S. Guy (2001). "Baculovirus expression of turkey coronavirus nucleocapsid protein." Avian Dis **45**(1): 136-143.
Cabeca, T. K., C. Granato and N. Bellei (2013). "Epidemiological and clinical features of human coronavirus infections among different subsets of patients." Influenza Other Respir Viruses **7**(6): 1040-1047.
Cano, M., P. Iglesias, G. Perez and J. J. Diez (2010). "[Influenza A virus (H1N1) infection as a cause of severe diabetic ketoacidosis in type 1 diabetes]." Endocrinol Nutr **57**(1): 37-38.
Carson, M., D. H. Johnson, H. McDonald, C. Brouillette and L. J. Delucas (2007). "His-tag impact on structure." Acta Crystallogr D Biol Crystallogr **63**(Pt 3): 295-301.
Carstensen, B. P. M. L. E. and M. Hills (2013). Epi: A package for statistical analysis in epidemiology. R package version 1.1.49.
Cavanagh, D. (2007). "Coronavirus avian infectious bronchitis virus." Vet Res **38**(2): 281-297.
Chan, J. F., K. H. Chan, G. K. Choi, K. K. To, H. Tse, J. P. Cai, M. L. Yeung, V. C. Cheng, H. Chen, X. Y. Che, S. K. Lau, P. C. Woo and K. Y. Yuen (2013). "Differential cell line susceptibility to the emerging novel human betacoronavirus 2c EMC/2012: implications for disease pathogenesis and clinical manifestation." J Infect Dis **207**(11): 1743-1752.
Chan, J. F., K. K. To, H. Tse, D. Y. Jin and K. Y. Yuen (2013). "Interspecies transmission and emergence of novel viruses: lessons from bats and birds." Trends Microbiol **21**(10): 544-555.
Chan, J. F., Y. Yao, M. L. Yeung, W. Deng, L. Bao, L. Jia, F. Li, C. Xiao, H. Gao, P. Yu, J. P. Cai, H. Chu, J. Zhou, H. Chen, C. Qin and K. Y. Yuen (2015). "Treatment With Lopinavir/Ritonavir or Interferon-beta1b Improves Outcome of MERS-CoV Infection in a Nonhuman Primate Model of Common Marmoset." J Infect Dis **212**(12): 1904-1913.
Chan, J. W., C. K. Ng, Y. H. Chan, T. Y. Mok, S. Lee, S. Y. Chu, W. L. Law, M. P. Lee and P. C. Li (2003). "Short term outcome and risk factors for adverse clinical outcomes in adults with severe acute respiratory syndrome (SARS)." Thorax **58**(8): 686-689.

Chang, C. K., C. M. Chen, M. H. Chiang, Y. L. Hsu and T. H. Huang (2013). "Transient oligomerization of the SARS-CoV N protein--implication for virus ribonucleoprotein packaging." PLoS One **8**(5): e65045.
Chang, C. K., Y. L. Hsu, Y. H. Chang, F. A. Chao, M. C. Wu, Y. S. Huang, C. K. Hu and T. H. Huang (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, C. K., S. C. Sue, T. H. Yu, C. M. Hsieh, C. K. Tsai, Y. C. Chiang, S. J. Lee, H. H. Hsiao, W. J. Wu, C. F. Chang and T. H. Huang (2005). "The dimer interface of the SARS coronavirus nucleocapsid protein adapts a porcine respiratory and reproductive syndrome virus-like structure." FEBS Lett **579**(25): 5663-5668.
Chang, C. K., S. C. Sue, T. H. Yu, C. M. Hsieh, C. K. Tsai, Y. C. Chiang, S. J. Lee, H. H. Hsiao, W. J. Wu, W. L. Chang, C. H. Lin and T. H. Huang (2006). "Modular organization of SARS coronavirus nucleocapsid protein." J Biomed Sci **13**(1): 59-72.
Chang, C. Y., J. Y. Peng, Y. H. Cheng, Y. C. Chang, Y. T. Wu, P. S. Tsai, H. Y. Chiou, C. R. Jeng and H. W. Chang (2019). "Development and comparison of enzyme-linked immunosorbent assays based on recombinant trimeric full-length and truncated spike proteins for detecting antibodies against porcine epidemic diarrhea virus." BMC Vet Res **15**(1): 421.
Channappanavar, R., L. Lu, S. Xia, L. Du, D. K. Meyerholz, S. Perlman and S. Jiang (2015). "Protective Effect of Intranasal Regimens Containing Peptidic Middle East Respiratory Syndrome Coronavirus Fusion Inhibitor Against MERS-CoV Infection." J Infect Dis **212**(12): 1894-1903.
Chartrand, C., M. M. Leeflang, J. Minion, T. Brewer and M. Pai (2012). "Accuracy of rapid influenza diagnostic tests: a meta-analysis." Ann Intern Med **156**(7): 500-511.
Chatila, W. M., E. A. Hoffman, J. Gaughan, G. B. Robinswood and G. J. Criner (2006). "Advanced emphysema in African-American and white patients: do differences exist?" CHEST Journal **130**(1): 108-118.
Che, X. Y., L. W. Qiu, Z. Y. Liao, Y. D. Wang, K. Wen, Y. X. Pan, W. Hao, Y. B. Mei, V. C. C. Cheng and K. Y. Yuen (2005). "Antigenic Cross-Reactivity between Severe Acute Respiratory Syndrome-Associated Coronavirus and Human Coronaviruses 229E and OC43." Journal of infectious diseases **191**(12): 2033-2037.
Chen, C. Y., C. K. Chang, Y. W. Chang, S. C. Sue, H. I. Bai, L. Riang, C. D. Hsiao and T. H. Huang (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.
Chen, N., M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, J. Xia, T. Yu, X. Zhang and L. Zhang (2020). "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study." Lancet **395**(10223): 507-513.
Chen, R. and E. C. Holmes (2006). "Avian influenza virus exhibits rapid evolutionary dynamics." Mol Biol Evol **23**(12): 2336-2341.
Chen, Y., K. H. Chan, Y. Kang, H. Chen, H. K. Luk, R. W. Poon, J. F. Chan, K. Y. Yuen, N. Xia, S. K. Lau and P. C. Woo (2015). "A sensitive and specific antigen detection assay for Middle East respiratory syndrome coronavirus." Emerg Microbes Infect **4**(4): e26.
Chen, Y. R., S. Zhong, Z. Fei, Y. Hashimoto, J. Z. Xiang, S. Zhang and G. W. Blissard (2013). "The transcriptome of the baculovirus Autographa californica multiple nucleopolyhedrovirus in Trichoplusia ni cells." J Virol **87**(11): 6391-6405.
Cheng, V. C., K. K. To, H. Tse, I. F. Hung and K. Y. Yuen (2012). "Two years after pandemic influenza A/2009/H1N1: what have we learned?" Clin Microbiol Rev **25**(2): 223-263.
Cho, K. O., A. E. Hoet, S. C. Loerch, T. E. Wittum and L. J. Saif (2001). "Evaluation of concurrent shedding of bovine coronavirus via the respiratory tract and enteric route in feedlot cattle." Am J Vet Res **62**(9): 1436-1441.
Choi, H. K., G. Lu, S. Lee, G. Wengler and M. G. Rossmann (1997). "Structure of Semliki Forest virus core protein." Proteins **27**(3): 345-359.
Choi, H. K., L. Tong, W. Minor, P. Dumas, U. Boege, M. G. Rossmann and G. Wengler (1991). "Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion." Nature **354**(6348): 37-43.
Chu, D. K., L. L. Poon, M. M. Gomaa, M. M. Shehata, R. A. Perera, D. Abu Zeid, A. S. El Rifay, L. Y. Siu, Y. Guan, R. J. Webby, M. A. Ali, M. Peiris and G. Kayali (2014). "MERS coronaviruses in dromedary camels, Egypt." Emerg Infect Dis **20**(6): 1049-1053.
Clem, A. and S. Galwankar (2009). "Seasonal influenza: waiting for the next pandemic." J Glob Infect Dis **1**(1): 51-56.
Clements, G. B., D. N. Galbraith and K. W. Taylor (1995). "Coxsackie B virus infection and onset of childhood diabetes." Lancet **346**(8969): 221-223.
ClinicalTrials.gov. (2014). "Anti-MERS-COV Convalescent Plasma Therapy." from https://clinicaltrials.gov/ct2/show/NCT02190799.

Coleman, C. M. and M. B. Frieman (2014). "Coronaviruses: important emerging human pathogens." J Virol **88**(10): 5209-5212.
Coleman, C. M., J. M. Sisk, R. M. Mingo, E. A. Nelson, J. M. White and M. B. Frieman (2016). "Abelson Kinase Inhibitors Are Potent Inhibitors of Severe Acute Respiratory Syndrome Coronavirus and Middle East Respiratory Syndrome Coronavirus Fusion." J Virol **90**(19): 8924-8933.
Cong, Y., M. Ulasli, H. Schepers, M. Mauthe, P. V'Kovski, F. Kriegenburg, V. Thiel, C. A. M. de Haan and F. Reggiori (2020). "Nucleocapsid Protein Recruitment to Replication-Transcription Complexes Plays a Crucial Role in Coronaviral Life Cycle." J Virol **94**(4).
Conrad, B., R. N. Weissmahr, J. Boni, R. Arcari, J. Schupbach and B. Mach (1997). "A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes." Cell **90**(2): 303-313.
Consortium, I. S. A. R. a. E. I. (2017). "Severe acute respiratory infection data tools." from https://isaric.tghn.org/protocols/severe-acute-respiratory-infection-data-tools/.
Coppieters, K. T., T. Boettler and M. von Herrath (2012). "Virus infections in type 1 diabetes." Cold Spring Harb Perspect Med **2**(1): a007682.
Coppieters, K. T. and M. G. von Herrath (2009). "Histopathology of type 1 diabetes: old paradigms and new insights." Rev Diabet Stud **6**(2): 85-96.
Corman, V. M., I. Eckerle, T. Bleicker, A. Zaki, O. Landt, M. Eschbach-Bludau, S. van Boheemen, R. Gopal, M. Ballhause, T. M. Bestebroer, D. Muth, M. A. Muller, J. F. Drexler, M. Zambon, A. D. Osterhaus, R. M. Fouchier and C. Drosten (2012). "Detection of a novel human coronavirus by real-time reverse-transcription polymerase chain reaction." Euro Surveill **17**(39).
Corman, V. M., N. L. Ithete, L. R. Richards, M. C. Schoeman, W. Preiser, C. Drosten and J. F. Drexler (2014). "Rooting the phylogenetic tree of middle East respiratory syndrome coronavirus by characterization of a conspecific virus from an African bat." J Virol **88**(19): 11297-11303.
Corman, V. M., M. A. Muller, U. Costabel, J. Timm, T. Binger, B. Meyer, P. Kreher, E. Lattwein, M. Eschbach-Bludau, A. Nitsche, T. Bleicker, O. Landt, B. Schweiger, J. F. Drexler, A. D. Osterhaus, B. L. Haagmans, U. Dittmer, F. Bonin, T. Wolff and C. Drosten (2012). "Assays for laboratory confirmation of novel human coronavirus (hCoV-EMC) infections." Euro Surveill **17**(49).
Corti, D., J. Zhao, M. Pedotti, L. Simonelli, S. Agnihothram, C. Fett, B. Fernandez-Rodriguez, M. Foglierini, G. Agatic, F. Vanzetta, R. Gopal, C. J. Langrish, N. A. Barrett, F. Sallusto, R. S. Baric, L. Varani, M. Zambon, S. Perlman and A. Lanzavecchia (2015). "Prophylactic and postexposure efficacy of a potent human monoclonal antibody against MERS coronavirus." Proc Natl Acad Sci U S A **112**(33): 10473-10478.
Corum, J., D. Grady, S.-L. Wee and C. Zimmer (2020). "Coronavirus vaccine tracker." The New York Times **5**.
Couch, R. B. (2000). "Prevention and treatment of influenza." N Engl J Med **343**(24): 1778-1787.
Craighead, J. E. and M. F. McLane (1968). "Diabetes mellitus: induction in mice by encephalomyocarditis virus." Science **162**(3856): 913-914.
Crameri, G., P. A. Durr, R. Klein, A. Foord, M. Yu, S. Riddell, J. Haining, D. Johnson, M. G. Hemida, J. Barr, M. Peiris, D. Middleton and L. F. Wang (2016). "Experimental Infection and Response to Rechallenge of Alpacas with Middle East Respiratory Syndrome Coronavirus." Emerg Infect Dis **22**(6): 1071-1074.
Daber, R., S. Stayrook, A. Rosenberg and M. Lewis (2007). "Structural analysis of lac repressor bound to allosteric effectors." J Mol Biol **370**(4): 609-619.
de Wilde, A. H., D. Jochmans, C. C. Posthuma, J. C. Zevenhoven-Dobbe, S. van Nieuwkoop, T. M. Bestebroer, B. G. van den Hoogen, J. Neyts and E. J. Snijder (2014). "Screening of an FDA-approved compound library identifies four small-molecule inhibitors of Middle East respiratory syndrome coronavirus replication in cell culture." Antimicrob Agents Chemother **58**(8): 4875-4884.
de Wit, E., N. van Doremalen, D. Falzarano and V. J. Munster (2016). "SARS and MERS: recent insights into emerging coronaviruses." Nat Rev Microbiol **14**(8): 523-534.
Diaz-Valencia, P. A., P. Bougneres and A. J. Valleron (2015). "Global epidemiology of type 1 diabetes in young adults and adults: a systematic review." BMC Public Health **15**: 255.
Doan, D. N. and T. Dokland (2003). "Structure of the nucleocapsid protein of porcine reproductive and respiratory syndrome virus." Structure **11**(11): 1445-1451.
Dodge, J. (2014). "Camels emit dangerous MERS virus, CSU confirms." Retrieved 26/04/2017, from http://source.colostate.edu/camels-emit-mers/.
Dokland, T., M. Walsh, J. M. Mackenzie, A. A. Khromykh, K. H. Ee and S. Wang (2004). "West Nile virus core protein; tetramer structure and ribbon formation." Structure **12**(7): 1157-1163.
Dominguez, S. R., C. C. Robinson and K. V. Holmes (2009). "Detection of four human coronaviruses in respiratory infections in children: a one-year study in Colorado." J Med Virol **81**(9): 1597-1604.
Donaldson, L. J., P. D. Rutter, B. M. Ellis, F. E. Greaves, O. T. Mytton, R. G. Pebody and I. E. Yardley (2009). "Mortality from pandemic A/H1N1 2009 influenza in England: public health surveillance study." BMJ **339**: b5213.

Dormitzer, P., W. Andrews, P. Rinella, D. ROSA and G. Palladino (2013). Assay for influenza virus hemagglutinins, Google Patents.
Doron, N. (2015). "Bacterial Strains for Protein Expression." Retrieved 26/08/2019, from http://wolfson.huji.ac.il/expression/bac-strains-prot-exp.html.
Drexler, J. F., V. M. Corman and C. Drosten (2014). "Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS." Antiviral Res **101**: 45-56.
Du, L., G. Zhao, Y. Yang, H. Qiu, L. Wang, Z. Kou, X. Tao, H. Yu, S. Sun, C. T. Tseng, S. Jiang, F. Li and Y. Zhou (2014). "A conformation-dependent neutralizing monoclonal antibody specifically targeting receptor-binding domain in Middle East respiratory syndrome coronavirus spike protein." J Virol **88**(12): 7045-7053.
Dyall, J., C. M. Coleman, B. J. Hart, T. Venkataraman, M. R. Holbrook, J. Kindrachuk, R. F. Johnson, G. G. Olinger, Jr., P. B. Jahrling, M. Laidlaw, L. M. Johansen, C. M. Lear-Rooney, P. J. Glass, L. E. Hensley and M. B. Frieman (2014). "Repurposing of clinically developed drugs for treatment of Middle East respiratory syndrome coronavirus infection." Antimicrob Agents Chemother **58**(8): 4885-4893.
Dyson, M. R., S. P. Shadbolt, K. J. Vincent, R. L. Perera and J. McCafferty (2004). "Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression." BMC Biotechnol **4**: 32.
Eckerle, I., V. M. Corman, M. A. Muller, M. Lenk, R. G. Ulrich and C. Drosten (2014). "Replicative Capacity of MERS Coronavirus in Livestock Cell Lines." Emerg Infect Dis **20**(2): 276-279.
Erles, K., C. Toomey, H. W. Brooks and J. Brownlie (2003). "Detection of a group 2 coronavirus in dogs with canine infectious respiratory disease." Virology **310**(2): 216-223.
Esper, F. P., T. Spahlinger and L. Zhou (2011). "Rate and influence of respiratory virus co-infection on pandemic (H1N1) influenza disease." J Infect **63**(4): 260-266.
Falzarano, D., E. de Wit, A. L. Rasmussen, F. Feldmann, A. Okumura, D. P. Scott, D. Brining, T. Bushmaker, C. Martellaro, L. Baseler, A. G. Benecke, M. G. Katze, V. J. Munster and H. Feldmann (2013). "Treatment with interferon-alpha2b and ribavirin improves outcome in MERS-CoV-infected rhesus macaques." Nat Med **19**(10): 1313-1317.
Fan, H., A. Ooi, Y. W. Tan, S. Wang, S. Fang, D. X. Liu and J. Lescar (2005). "The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal structure of its N-terminal domain and multimerization properties." Structure **13**(12): 1859-1868.
Farag, E. A., C. B. Reusken, B. L. Haagmans, K. A. Mohran, V. Stalin Raj, S. D. Pas, J. Voermans, S. L. Smits, G. J. Godeke, M. M. Al-Hajri, F. H. Alhajri, H. E. Al-Romaihi, H. Ghobashy, M. M. El-Maghraby, A. M. El-Sayed, M. H. Al Thani, S. Al-Marri and M. P. Koopmans (2015). "High proportion of MERS-CoV shedding dromedaries at slaughterhouse with a potential epidemiological link to human cases, Qatar 2014." Infect Ecol Epidemiol **5**: 28305.
Fauci, A. S. (2006). "Seasonal and pandemic influenza preparedness: science and countermeasures." J Infect Dis **194 Suppl 2**: S73-76.
Ferrer, M., T. N. Chernikova, M. M. Yakimov, P. N. Golyshin and K. N. Timmis (2003). "Chaperonins govern growth of Escherichia coli at low temperatures." Nat Biotechnol **21**(11): 1266-1267.
Fields, B. N., D. M. Knipe and P. M. Howley (2007). Fields virology. Philadelphia, Wolters Kluwer Health/Lippincott Williams & Wilkins.
Fisher, R. A. (1922). "On the Interpretation of χ 2 from Contingency Tables, and the Calculation of P." Journal of the Royal Statistical Society **85**(1): 87.
Flyvbjerg, A. (2010). "Diabetic angiopathy, the complement system and the tumor necrosis factor superfamily." Nat Rev Endocrinol **6**(2): 94-101.
Folegatti, P. M., M. Bittaye, A. Flaxman, F. R. Lopez, D. Bellamy, A. Kupke, C. Mair, R. Makinson, J. Sheridan, C. Rohde, S. Halwe, Y. Jeong, Y. S. Park, J. O. Kim, M. Song, A. Boyd, N. Tran, D. Silman, I. Poulton, M. Datoo, J. Marshall, Y. Themistocleous, A. Lawrie, R. Roberts, E. Berrie, S. Becker, T. Lambe, A. Hill, K. Ewer and S. Gilbert (2020). "Safety and immunogenicity of a candidate Middle East respiratory syndrome coronavirus viral-vectored vaccine: a dose-escalation, open-label, non-randomised, uncontrolled, phase 1 trial." Lancet Infect Dis **20**(7): 816-826.
Foote, M., R. Daver and C. Quinn (2017). "Using "Mystery Patient" Drills to Assess Hospital Ebola Preparedness in New York City, 2014-2015." Health Secur **15**(5): 500-508.
Foote, M. M. K., T. S. Styles and C. L. Quinn (2017). "Assessment of Hospital Emergency Department Response to Potentially Infectious Diseases Using Unannounced Mystery Patient Drills - New York City, 2016." MMWR Morb Mortal Wkly Rep **66**(36): 945-949.
Fouchier, R. A., V. Munster, A. Wallensten, T. M. Bestebroer, S. Herfst, D. Smith, G. F. Rimmelzwaan, B. Olsen and A. D. Osterhaus (2005). "Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls." J Virol **79**(5): 2814-2822.

Francis, D. M. and R. Page (2010). "Strategies to optimize protein expression in E. coli." Curr Protoc Protein Sci **Chapter 5**: Unit 5 24 21-29.
Frey, A., J. Di Canzio and D. Zurakowski (1998). "A statistically defined endpoint titer determination method for immunoassays." J Immunol Methods **221**(1-2): 35-41.
Friedman, A. M., T. O. Fischmann and T. A. Steitz (1995). "Crystal structure of lac repressor core tetramer and its implications for DNA looping." Science **268**(5218): 1721-1727.
Fuchtenbusch, M., A. Irnstetter, G. Jager and A. G. Ziegler (2001). "No evidence for an association of coxsackie virus infections during pregnancy and early childhood with development of islet autoantibodies in offspring of mothers or fathers with type 1 diabetes." J Autoimmun **17**(4): 333-340.
Gaber, F. and S. Kapil (1999). "Development of an antigen spot test for detection of coronavirus in bovine fecal samples." Clin Diagn Lab Immunol **6**(4): 542-544.
Gale, E. A. (2008). "Congenital rubella: citation virus or viral cause of type 1 diabetes?" Diabetologia **51**(9): 1559-1566.
Gamble, D. R., M. L. Kinsley, M. G. FitzGerald, R. Bolton and K. W. Taylor (1969). "Viral antibodies in diabetes mellitus." Br Med J **3**(5671): 627-630.
Gaunt, E. R., A. Hardie, E. C. Claas, P. Simmonds and K. E. Templeton (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.
Geerlings, S. E. and A. I. Hoepelman (1999). "Immune dysfunction in patients with diabetes mellitus (DM)." FEMS Immunol Med Microbiol **26**(3-4): 259-265.
Gensheimer, K. F., M. I. Meltzer, A. S. Postema and R. A. Strikas (2003). "Influenza pandemic preparedness." Emerg Infect Dis **9**(12): 1645-1648.
Gibb, R., D. W. Redding, K. Q. Chin, C. A. Donnelly, T. M. Blackburn, T. Newbold and K. E. Jones (2020). "Zoonotic host diversity increases in human-dominated ecosystems." Nature **584**(7821): 398-402.
Gilbert, W. and B. Muller-Hill (1966). "Isolation of the lac repressor." Proc Natl Acad Sci U S A **56**(6): 1891-1898.
Gomaa, M. H., D. Yoo, D. Ojkic and J. R. Barta (2008). "Seroprevalence of turkey coronavirus in North American turkeys determined by a newly developed enzyme-linked immunosorbent assay based on recombinant antigen." Clin Vaccine Immunol **15**(12): 1839-1844.
Gong, L., J. Li, Q. Zhou, Z. Xu, L. Chen, Y. Zhang, C. Xue, Z. Wen and Y. Cao (2017). "A New Bat-HKU2-like Coronavirus in Swine, China, 2017." Emerg Infect Dis **23**(9).
Gorse, G. J., G. B. Patel, J. N. Vitale and T. Z. O'Connor (2010). "Prevalence of antibodies to four human coronaviruses is lower in nasal secretions than in serum." Clin Vaccine Immunol **17**(12): 1875-1880.
Goto, A., Y. Takahashi, M. Kishimoto, Y. Nakajima, K. Nakanishi, H. Kajio and M. Noda (2008). "A case of fulminant type 1 diabetes associated with significant elevation of mumps titers." Endocr J **55**(3): 561-564.
Gov.UK (2021). UK government announces partnership with British rapid test manufacturers.
Graham, R. L. and R. S. Baric (2010). "Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission." J Virol **84**(7): 3134-3146.
Graham, R. L., E. F. Donaldson and R. S. Baric (2013). "A decade after SARS: strategies for controlling emerging coronaviruses." Nat Rev Microbiol **11**(12): 836-848.
Gramberg, T., H. Hofmann, P. Moller, P. F. Lalor, A. Marzi, M. Geier, M. Krumbiegel, T. Winkler, F. Kirchhoff, D. H. Adams, S. Becker, J. Munch and S. Pohlmann (2005). "LSECtin interacts with filovirus glycoproteins and the spike protein of SARS coronavirus." Virology **340**(2): 224-236.
GraphPad. (2017). "QuickCalcs: Analyze a 2x2 contingency table." Retrieved 28/09/2017, from https://www.graphpad.com/quickcalcs/contingency1.cfm.
Graves, P. M., H. A. Rotbart, W. A. Nix, M. A. Pallansch, H. A. Erlich, J. M. Norris, M. Hoffman, G. S. Eisenbarth and M. Rewers (2003). "Prospective study of enteroviral infections and development of beta-cell autoimmunity. Diabetes autoimmunity study in the young (DAISY)." Diabetes Res Clin Pract **59**(1): 51-61.
Greffex. (2014). "MERS." from http://www.greffex.com/vaccines/mers/.
Guan, M., H. Y. Chen, S. Y. Foo, Y. J. Tan, P. Y. Goh and S. H. Wee (2004). "Recombinant protein-based enzyme-linked immunosorbent assay and immunochromatographic tests for detection of immunoglobulin G antibodies to severe acute respiratory syndrome (SARS) coronavirus in SARS patients." Clin Diagn Lab Immunol **11**(2): 287-291.
Guan, Y., B. J. Zheng, Y. Q. He, X. L. Liu, Z. X. Zhuang, C. L. Cheung, S. W. Luo, P. H. Li, L. J. Zhang, Y. J. Guan, K. M. Butt, K. L. Wong, K. W. Chan, W. Lim, K. F. Shortridge, K. Y. Yuen, J. S. Peiris and L. L. Poon (2003). "Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China." Science **302**(5643): 276-278.
Guarino, L. A. (1990). "Identification of a viral gene encoding a ubiquitin-like protein." Proc Natl Acad Sci U S A **87**(1): 409-413.

Guberski, D. L., V. A. Thomas, W. R. Shek, A. A. Like, E. S. Handler, A. A. Rossini, J. E. Wallace and R. M. Welsh (1991). "Induction of type I diabetes by Kilham's rat virus in diabetes-resistant BB/Wor rats." Science **254**(5034): 1010-1013.
Gui, M., X. Liu, D. Guo, Z. Zhang, C. C. Yin, Y. Chen and Y. Xiang (2017). "Electron microscopy studies of the coronavirus ribonucleoprotein complex." Protein Cell **8**(3): 219-224.
Guo, L., L. Ren, S. Yang, M. Xiao, Chang, F. Yang, C. S. Dela Cruz, Y. Wang, C. Wu, Y. Xiao, L. Zhang, L. Han, S. Dang, Y. Xu, Q. W. Yang, S. Y. Xu, H. D. Zhu, Y. C. Xu, Q. Jin, L. Sharma, L. Wang and J. Wang (2020). "Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19)." Clin Infect Dis **71**(15): 778-785.
Guy, J. S., L. G. Smith, J. J. Breslin and S. Pakpinyo (2002). "Development of a competitive enzyme-linked immunosorbent assay for detection of turkey coronavirus antibodies." Avian Dis **46**(2): 334-341.
Haagmans, B. L., J. M. van den Brand, L. B. Provacia, V. S. Raj, K. J. Stittelaar, S. Getu, L. de Waal, T. M. Bestebroer, G. van Amerongen, G. M. Verjans, R. A. Fouchier, S. L. Smits, T. Kuiken and A. D. Osterhaus (2015). "Asymptomatic Middle East respiratory syndrome coronavirus infection in rabbits." J Virol **89**(11): 6131-6135.
Haagmans, B. L., J. M. van den Brand, V. S. Raj, A. Volz, P. Wohlsein, S. L. Smits, D. Schipper, T. M. Bestebroer, N. Okba, R. Fux, A. Bensaid, D. Solanes Foz, T. Kuiken, W. Baumgartner, J. Segales, G. Sutter and A. D. Osterhaus (2016). "An orthopoxvirus-based vaccine reduces virus excretion after MERS-CoV infection in dromedary camels." Science **351**(6268): 77-81.
Hao, X. Y., Q. Lv, F. D. Li, Y. F. Xu and H. Gao (2019). "The characteristics of hDPP4 transgenic mice subjected to aerosol MERS coronavirus infection via an animal nose-only exposure device." Animal Model Exp Med **2**(4): 269-281.
Harries, A. D., Y. Lin, S. Satyanarayana, K. Lonnroth, L. Li, N. Wilson, L. S. Chauhan, R. Zachariah, M. A. Baker, C. Y. Jeon, M. B. Murray, D. Maher, I. C. Bygbjerg, D. A. Enarson, N. E. Billo and A. Kapur (2011). "The looming epidemic of diabetes-associated tuberculosis: learning lessons from HIV-associated tuberculosis." Int J Tuberc Lung Dis **15**(11): 1436-1444, i.
Harrison, A. G., T. Lin and P. Wang (2020). "Mechanisms of SARS-CoV-2 Transmission and Pathogenesis." Trends Immunol **41**(12): 1100-1115.
Hassell, J. M., T. Newbold, A. P. Dobson, Y. M. Linton, L. H. V. Franklinos, D. Zimmerman and K. M. Pagenkopp Lohan (2021). "Towards an ecosystem model of infectious disease." Nat Ecol Evol **5**(7): 907-918.
Hayden, F. G. (2006). "Antivirals for influenza: historical perspectives and lessons learned." Antiviral Res **71**(2-3): 372-378.
Hayden, F. G. and A. T. Pavia (2006). "Antiviral management of seasonal and pandemic influenza." J Infect Dis **194 Suppl 2**: S119-126.
Hayden, F. G. and N. Shindo (2019). "Influenza virus polymerase inhibitors in clinical development." Curr Opin Infect Dis **32**(2): 176-186.
Haynes, L. M., C. Miao, J. L. Harcourt, J. M. Montgomery, M. Q. Le, S. A. Dryga, K. I. Kamrud, B. Rivers, G. J. Babcock, J. B. Oliver, J. A. Comer, M. Reynolds, T. M. Uyeki, D. Bausch, T. Ksiazek, W. Thomas, H. Alterson, J. Smith, D. M. Ambrosino and L. J. Anderson (2007). "Recombinant protein-based assays for detection of antibodies to severe acute respiratory syndrome coronavirus spike and nucleocapsid proteins." Clin Vaccine Immunol **14**(3): 331-333.
He, R., F. Dobie, M. Ballantine, A. Leeson, Y. Li, N. Bastien, T. Cutts, A. Andonov, J. Cao, T. F. Booth, F. A. Plummer, S. Tyler, L. Baker and X. Li (2004). "Analysis of multimerization of the SARS coronavirus nucleocapsid protein." Biochem Biophys Res Commun **316**(2): 476-483.
He, Y. X., H. Lu, P. Siddiqui, Y. S. Zhou and S. B. Jiang (2005). "Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies." Journal of Immunology **174**(8): 4908-4915.
Heath, P. T., E. P. Galiza, D. N. Baxter, M. Boffito, D. Browne, F. Burns, D. R. Chadwick, R. Clark, C. Cosgrove, J. Galloway, A. L. Goodman, A. Heer, A. Higham, S. Iyengar, A. Jamal, C. Jeanes, P. A. Kalra, C. Kyriakidou, D. F. McAuley, A. Meyrick, A. M. Minassian, J. Minton, P. Moore, I. Munsoor, H. Nicholls, O. Osanlou, J. Packham, C. H. Pretswell, A. San Francisco Ramos, D. Saralaya, R. P. Sheridan, R. Smith, R. L. Soiza, P. A. Swift, E. C. Thomson, J. Turner, M. E. Viljoen, G. Albert, I. Cho, F. Dubovsky, G. Glenn, J. Rivers, A. Robertson, K. Smith, S. Toback and V. S. G. nCo (2021). "Safety and Efficacy of NVX-CoV2373 Covid-19 Vaccine." N Engl J Med.
Hegewald, M. J. and R. O. Crapo (2007). "Socioeconomic status and lung function." Chest **132**(5): 1608-1614.
Hemida, M. G., D. K. Chu, L. L. Poon, R. A. Perera, M. A. Alhammadi, H. Y. Ng, L. Y. Siu, Y. Guan, A. Alnaeem and M. Peiris (2014). "MERS coronavirus in dromedary camel herd, Saudi Arabia." Emerg Infect Dis **20**(7): 1231-1234.

Hemida, M. G., R. A. Perera, P. Wang, M. A. Alhammadi, L. Y. Siu, M. Li, L. L. Poon, L. Saif, A. Alnaeem and M. Peiris (2013). "Middle East Respiratory Syndrome (MERS) coronavirus seroprevalence in domestic livestock in Saudi Arabia, 2010 to 2013." Euro Surveill **18**(50): 20659.
Heymann, D. L. and G. Rodier (2004). "Global surveillance, national surveillance, and SARS." Emerg Infect Dis **10**(2): 173-175.
Hober, D. and F. Sane (2010). "Enteroviral pathogenesis of type 1 diabetes." Discov Med **10**(51): 151-160.
Hofmann, H., K. Pyrc, L. van der Hoek, M. Geier, B. Berkhout and S. Pohlmann (2005). "Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry." Proc Natl Acad Sci U S A **102**(22): 7988-7993.
Holm, L. and C. Sander (1998). "Touring protein fold space with Dali/FSSP." Nucleic Acids Res **26**(1): 316-319.
Holzworth, J. (1963). "Some important disorders of cats." The Cornell Veterinarian **53**: 157-160.
Honda-Okubo, Y., D. Barnard, C. H. Ong, B. H. Peng, C. T. Tseng and N. Petrovsky (2015). "Severe acute respiratory syndrome-associated coronavirus vaccines formulated with delta inulin adjuvants provide enhanced protection while ameliorating lung eosinophilic immunopathology." J Virol **89**(6): 2995-3007.
Honeyman, M. C., B. S. Coulson, N. L. Stone, S. A. Gellert, P. N. Goldwater, C. E. Steele, J. J. Couper, B. D. Tait, P. G. Colman and L. C. Harrison (2000). "Association between rotavirus infection and pancreatic islet autoimmunity in children at risk of developing type 1 diabetes." Diabetes **49**(8): 1319-1324.
Hu, B., X. Ge, L. F. Wang and Z. Shi (2015). "Bat origin of human coronaviruses." Virol J **12**: 221.
Huang, C. Y., Y. L. Hsu, W. L. Chiang and M. H. Hou (2009). "Elucidation of the stability and functional regions of the human coronavirus OC43 nucleocapsid protein." Protein Sci **18**(11): 2209-2218.
Huang, X., W. Dong, A. Milewska, A. Golda, Y. Qi, Q. K. Zhu, W. A. Marasco, R. S. Baric, A. C. Sims, K. Pyrc, W. Li and J. Sui (2015). "Human Coronavirus HKU1 Spike Protein Uses O-Acetylated Sialic Acid as an Attachment Receptor Determinant and Employs Hemagglutinin-Esterase Protein as a Receptor-Destroying Enzyme." J Virol **89**(14): 7202-7213.
Hurst, K. R., C. A. Koetzner and P. S. Masters (2009). "Identification of in vivo-interacting domains of the murine coronavirus nucleocapsid protein." J Virol **83**(14): 7221-7234.
Huynh, J., S. Li, B. Yount, A. Smith, L. Sturges, J. C. Olsen, J. Nagel, J. B. Johnson, S. Agnihothram, J. E. Gates, M. B. Frieman, R. S. Baric and E. F. Donaldson (2012). "Evidence supporting a zoonotic origin of human coronavirus strain NL63." J Virol **86**(23): 12816-12825.
Ishida, T. and K. Kinoshita (2007). "PrDOS: prediction of disordered protein regions from amino acid sequence." Nucleic Acids Res **35**(Web Server issue): W460-464.
Israr, M. Z., D. Bernieh, A. Salzano, S. Cassambai, Y. Yazaki and T. Suzuki (2020). "Matrix-assisted laser desorption ionisation (MALDI) mass spectrometry (MS): basics and clinical applications." Clin Chem Lab Med **58**(6): 883-896.
Jacob, F. and J. Monod (1961). "Genetic regulatory mechanisms in the synthesis of proteins." J Mol Biol **3**: 318-356.
Jayaram, H., H. Fan, B. R. Bowman, A. Ooi, J. Jayaram, E. W. Collisson, J. Lescar and B. V. Prasad (2006). "X-ray structures of the N- and C-terminal domains of a coronavirus nucleocapsid protein: implications for nucleocapsid formation." J Virol **80**(13): 6612-6620.
Jensen, T. H., J. H. Andersen, C. K. Hjulsager, M. Chriel and M. F. Bertelsen (2017). "Evaluation of a Commercial Competitive Enzyme-Linked Immunosorbent Assay for Detection of Avian Influenza Virus Subtype H5 Antibodies in Zoo Birds." J Zoo Wildl Med **48**(3): 882-885.
Jiang, L., N. Wang, T. Zuo, X. Shi, K. M. Poon, Y. Wu, F. Gao, D. Li, R. Wang, J. Guo, L. Fu, K. Y. Yuen, B. J. Zheng, X. Wang and L. Zhang (2014). "Potent neutralization of MERS-CoV by human neutralizing monoclonal antibodies to the viral spike glycoprotein." Sci Transl Med **6**(234): 234ra259.
Jiang, X., K. Oohira, Y. Iwasaki, H. Nakano, S. Ichihara and T. Yamane (2002). "Reduction of protein degradation by use of protease-deficient mutants in cell-free protein synthesis system of Escherichia coli." J Biosci Bioeng **93**(2): 151-156.
Jiang, Y., J. Li, Y. Teng, H. Sun, G. Tian, L. He, P. Li, Y. Chen, Y. Guo, J. Li, G. Zhao, Y. Zhou and S. Sun (2019). "Complement Receptor C5aR1 Inhibition Reduces Pyroptosis in hDPP4-Transgenic Mice Infected with MERS-CoV." Viruses **11**(1).
Jobe, A. and S. Bourgeois (1972). "lac Repressor-operator interaction. VI. The natural inducer of the lac operon." J Mol Biol **69**(3): 397-408.
Jones, I. and Y. Morikawa (1996). "Baculovirus vectors for expression in insect cells." Curr Opin Biotechnol **7**(5): 512-516.
Jones, K. E., N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman and P. Daszak (2008). "Global trends in emerging infectious diseases." Nature **451**(7181): 990-993.

Jones, R. (2010). "Europe: history, current situation and control measures for infectious bronchitis." Revista Brasileira de Ciência Avícola **12**: 125-128.
Joshi, N., G. M. Caputo, M. R. Weitekamp and A. W. Karchmer (1999). "Infections in patients with diabetes mellitus." N Engl J Med **341**(25): 1906-1912.
Kaddoura, M., M. AlIbrahim, G. Hijazi, N. Soudani, A. Audi, H. Alkalamouni, S. Haddad, A. Eid and H. Zaraket (2020). "COVID-19 Therapeutic Options Under Investigation." Front Pharmacol **11**: 1196.
Kandeil, A., M. Gomaa, A. Nageh, M. M. Shehata, A. E. Kayed, J. S. M. Sabir, A. Abiadh, J. Jrijer, Z. Amr, M. A. Said, D. K. Byarugaba, F. Wabwire-Mangen, T. Tugume, N. S. Mohamed, R. Attar, S. M. Hassan, S. A. Linjawi, Y. Moatassim, O. Kutkat, S. Mahmoud, O. Bagato, N. M. A. Shama, R. El-Shesheny, A. Mostafa, R. A. Perera, D. K. Chu, N. Hassan, B. Elsokary, A. Saad, H. Sobhy, I. El Masry, P. P. McKenzie, R. J. Webby, M. Peiris, Y. J. Makonnen, M. A. Ali and G. Kayali (2019). "Middle East Respiratory Syndrome Coronavirus (MERS-CoV) in Dromedary Camels in Africa and Middle East." Viruses **11**(8).
Kasuga, A., R. Harada and T. Saruta (1996). "Insulin-dependent diabetes mellitus associated with parvovirus B19 infection." Ann Intern Med **125**(8): 700-701.
Katsarou, A., S. Gudbjornsdottir, A. Rawshani, D. Dabelea, E. Bonifacio, B. J. Anderson, L. M. Jacobsen, D. A. Schatz and A. Lernmark (2017). "Type 1 diabetes mellitus." Nat Rev Dis Primers **3**: 17016.
Katze, M. G., Y. He and M. Gale, Jr. (2002). "Viruses and interferon: a fight for supremacy." Nat Rev Immunol **2**(9): 675-687.
Kaufman, D. L., M. G. Erlander, M. Clare-Salzler, M. A. Atkinson, N. K. Maclaren and A. J. Tobin (1992). "Autoimmunity to two forms of glutamate decarboxylase in insulin-dependent diabetes mellitus." J Clin Invest **89**(1): 283-292.
Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes and A. Drummond (2012). "Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data." Bioinformatics **28**(12): 1647-1649.
Keech, C., G. Albert, I. Cho, A. Robertson, P. Reed, S. Neal, J. S. Plested, M. Zhu, S. Cloney-Clark, H. Zhou, G. Smith, N. Patel, M. B. Frieman, R. E. Haupt, J. Logue, M. McGrath, S. Weston, P. A. Piedra, C. Desai, K. Callahan, M. Lewis, P. Price-Abbott, N. Formica, V. Shinde, L. Fries, J. D. Lickliter, P. Griffin, B. Wilkinson and G. M. Glenn (2020). "Phase 1-2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine." N Engl J Med **383**(24): 2320-2332.
Khalafalla, A. I., X. Lu, A. I. Al-Mubarak, A. H. Dalab, K. A. Al-Busadah and D. D. Erdman (2015). "MERS-CoV in Upper Respiratory Tract and Lungs of Dromedary Camels, Saudi Arabia, 2013-2014." Emerg Infect Dis **21**(7): 1153-1158.
Kilianski, A., A. M. Mielech, X. Deng and S. C. Baker (2013). "Assessing activity and inhibition of Middle East respiratory syndrome coronavirus papain-like and 3C-like proteases using luciferase-based biosensors." J Virol **87**(21): 11955-11962.
Kim, E., K. Okada, T. Kenniston, V. S. Raj, M. M. AlHajri, E. A. Farag, F. AlHajri, A. D. Osterhaus, B. L. Haagmans and A. Gambotto (2014). "Immunogenicity of an adenoviral-based Middle East Respiratory Syndrome coronavirus vaccine in BALB/c mice." Vaccine **32**(45): 5975-5982.
Kim, H. Y. (2017). "Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test." Restor Dent Endod **42**(2): 152-155.
Kitts, P. A. and R. D. Possee (1993). "A method for producing recombinant baculovirus expression vectors at high frequency." Biotechniques **14**(5): 810-817.
Koyama, T., D. Platt and L. Parida (2020). "Variant analysis of SARS-CoV-2 genomes." Bull World Health Organ **98**(7): 495-504.
Krammer, F. (2020). "SARS-CoV-2 vaccines in development." Nature **586**(7830): 516-527.
Kunkel, F. and G. Herrler (2003). "Structural and functional analysis of the surface protein of human coronavirus OC4." Virology **195**(1): 195-202.
Kuo, L. and P. S. Masters (2002). "Genetic evidence for a structural interaction between the carboxy termini of the membrane and nucleocapsid proteins of mouse hepatitis virus." J Virol **76**(10): 4987-4999.
Kupferschmidt, K. (2015). "Infectious diseases. MERS surges again, but pandemic jitters ease." Science **347**(6228): 1296-1297.
Labella, A. M. and S. E. Merel (2013). "Influenza." Med Clin North Am **97**(4): 621-645, x.
Lachance, C., N. Arbour, N. R. Cashman and P. J. Talbot (1998). "Involvement of aminopeptidase N (CD13) in infection of human neural cells by human coronavirus 229E." J Virol **72**(8): 6511-6519.
Lai, M. M. and D. Cavanagh (1997). "The molecular biology of coronaviruses." Adv Virus Res **48**: 1-100.
Lardeux, F., G. Torrico and C. Aliaga (2016). "Calculation of the ELISA's cut-off based on the change-point analysis method for detection of Trypanosoma cruzi infection in Bolivian dogs in the absence of controls." Mem Inst Oswaldo Cruz **111**(8): 501-504.

Lau, S. K., R. W. Poon, B. H. Wong, M. Wang, Y. Huang, H. Xu, R. Guo, K. S. Li, K. Gao, K. H. Chan, B. J. Zheng, P. C. Woo and K. Y. Yuen (2010). "Coexistence of different genotypes in the same bat and serological characterization of Rousettus bat coronavirus HKU9 belonging to a novel Betacoronavirus subgroup." J Virol **84**(21): 11385-11394.
Lau, S. K., P. C. Woo, K. S. Li, Y. Huang, H. W. Tsoi, B. H. Wong, S. S. Wong, S. Y. Leung, K. H. Chan and K. Y. Yuen (2005). "Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats." Proc Natl Acad Sci U S A **102**(39): 14040-14045.
Lau, S. K., P. C. Woo, B. H. Wong, H. W. Tsoi, G. K. Woo, R. W. Poon, K. H. Chan, W. I. Wei, J. S. Peiris and K. Y. Yuen (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.
Lau, S. K., P. C. Woo, C. C. Yip, H. Tse, H. W. Tsoi, V. C. Cheng, P. Lee, B. S. Tang, C. H. Cheung, R. A. Lee, L. Y. So, Y. L. Lau, K. H. Chan and K. Y. Yuen (2006). "Coronavirus HKU1 and other coronavirus infections in Hong Kong." J Clin Microbiol **44**(6): 2063-2071.
Lee, D. H. and A. L. Goldberg (1998). "Proteasome inhibitors: valuable new tools for cell biologists." Trends Cell Biol **8**(10): 397-403.
Lee, H. K., B. H. Lee, N. K. Dutta, S. H. Seok, M. W. Baek, H. Y. Lee, D. J. Kim, Y. R. Na, K. J. Noh, S. H. Park, H. Kariwa, M. Nakauchi, Q. Mai le, S. J. Heo and J. H. Park (2008). "Detection of antibodies against SARS-Coronavirus using recombinant truncated nucleocapsid proteins by ELISA." J Microbiol Biotechnol **18**(10): 1717-1721.
Lee, N., D. Hui, A. Wu, P. Chan, P. Cameron, G. M. Joynt, A. Ahuja, M. Y. Yung, C. B. Leung, K. F. To, S. F. Lui, C. C. Szeto, S. Chung and J. J. Sung (2003). "A major outbreak of severe acute respiratory syndrome in Hong Kong." N Engl J Med **348**(20): 1986-1994.
Lelesius, R., A. Karpovaite, R. Mickiene, T. Drevinskas, N. Tiso, O. Ragazinskiene, L. Kubiliene, A. Maruska and A. Salomskas (2019). "In vitro antiviral activity of fifteen plant extracts against avian infectious bronchitis virus." BMC Vet Res **15**(1): 178.
Lewis, M., G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan and P. Lu (1996). "Crystal structure of the lactose operon repressor and its complexes with DNA and inducer." Science **271**(5253): 1247-1254.
Li, W., M. J. Moore, N. Vasilieva, J. Sui, S. K. Wong, M. A. Berne, M. Somasundaran, J. L. Sullivan, K. Luzuriaga, T. C. Greenough, H. Choe and M. Farzan (2003). "Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus." Nature **426**(6965): 450-454.
Li, Y., A. I. Khalafalla, C. R. Paden, M. F. Yusof, Y. M. Eltahir, Z. M. Al Hammadi, Y. Tao, K. Queen, F. A. Hosani, S. I. Gerber, A. J. Hall, S. Al Muhairi and S. Tong (2017). "Identification of diverse viruses in upper respiratory samples in dromedary camels from United Arab Emirates." PLoS One **12**(9): e0184718.
Li, Z., Y. Yi, X. Luo, N. Xiong, Y. Liu, S. Li, R. Sun, Y. Wang, B. Hu, W. Chen, Y. Zhang, J. Wang, B. Huang, Y. Lin, J. Yang, W. Cai, X. Wang, J. Cheng, Z. Chen, K. Sun, W. Pan, Z. Zhan, L. Chen and F. Ye (2020). "Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis." J Med Virol **92**(9): 1518-1524.
LifeTechnologies (2011). "Guide to Baculovirus Expression Vector Systems (BEVS) and Insect Cell Culture Techniques." Invitrogen Life Technologies: Manual, Instructions.
Liljander, A., B. Meyer, J. Jores, M. A. Muller, E. Lattwein, I. Njeru, B. Bett, C. Drosten and V. M. Corman (2016). "MERS-CoV Antibodies in Humans, Africa, 2013-2014." Emerg Infect Dis **22**(6): 1086-1089.
Liu, S., X. Zhang, Y. Wang, C. Li, Z. Han, Y. Shao, H. Li and X. Kong (2009). "Molecular characterization and pathogenicity of infectious bronchitis coronaviruses: complicated evolution and epidemiology in china caused by cocirculation of multiple types of infectious bronchitis coronaviruses." Intervirology **52**(4): 223-234.
Lo, Y. S., S. Y. Lin, S. M. Wang, C. T. Wang, Y. L. Chiu, T. H. Huang and M. H. Hou (2013). "Oligomerization of the carboxyl terminal domain of the human coronavirus 229E nucleocapsid protein." FEBS Lett **587**(2): 120-127.
Logunov, D. Y., I. V. Dolzhikova, O. V. Zubkova, A. I. Tukhvatulin, D. V. Shcheblyakov, A. S. Dzharullaeva, D. M. Grousova, A. S. Erokhova, A. V. Kovyrshina, A. G. Botikov, F. M. Izhaeva, O. Popova, T. A. Ozharovskaya, I. B. Esmagambetov, I. A. Favorskaya, D. I. Zrelkin, D. V. Voronina, D. N. Shcherbinin, A. S. Semikhin, Y. V. Simakova, E. A. Tokarskaya, N. L. Lubenets, D. A. Egorova, M. M. Shmarov, N. A. Nikitenko, L. F. Morozova, E. A. Smolyarchuk, E. V. Kryukov, V. F. Babira, S. V. Borisevich, B. S. Naroditsky and A. L. Gintsburg (2020). "Safety and immunogenicity of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine in two formulations: two open, non-randomised phase 1/2 studies from Russia." Lancet **396**(10255): 887-897.
López-Ratón, M., M. X. Rodríguez-Álvarez, C. C. Suárez and F. G. Sampedro (2014). "OptimalCutpoints: AnRPackage for Selecting Optimal Cutpoints in Diagnostic Tests." Journal of Statistical Software **61**(8): 36.

Low, P., F. J. Doherty, E. Fellinger, M. Sass, R. J. Mayer and L. Laszlo (1995). "Related organelles of the endosome-lysosome system contain a different repertoire of ubiquitinated proteins in Sf9 insect cells." FEBS Lett **368**(1): 125-131.
Lu, H. (2020). "Drug treatment options for the 2019-new coronavirus (2019-nCoV)." Biosci Trends **14**(1): 69-71.
Lu, S., Y. Wang, Y. Chen, B. Wu, K. Qin, J. Zhao, Y. Lou and W. Tan (2017). "Discovery of a novel canine respiratory coronavirus support genetic recombination among betacoronavirus1." Virus Res **237**: 7-13.
Lu, W., S. D. Chapple, O. Lissini and I. M. Jones (2002). "Characterization of a truncated soluble form of the baculovirus (AcMNPV) major envelope protein Gp64." Protein Expr Purif **24**(2): 196-201.
Lu, X., B. Whitaker, S. K. Sakthivel, S. Kamili, L. E. Rose, L. Lowe, E. Mohareb, E. M. Elassal, T. Al-sanouri, A. Haddadin and D. D. Erdman (2014). "Real-time reverse transcription-PCR assay panel for Middle East respiratory syndrome coronavirus." J Clin Microbiol **52**(1): 67-75.
Luckow, V. A., S. C. Lee, G. F. Barry and P. O. Olins (1993). "Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli." J Virol **67**(8): 4566-4579.
Luke, T., H. Wu, J. Zhao, R. Channappanavar, C. M. Coleman, J. A. Jiao, H. Matsushita, Y. Liu, E. N. Postnikova, B. L. Ork, G. Glenn, D. Flyer, G. Defang, K. Raviprakash, T. Kochel, J. Wang, W. Nie, G. Smith, L. E. Hensley, G. G. Olinger, J. H. Kuhn, M. R. Holbrook, R. F. Johnson, S. Perlman, E. Sullivan and M. B. Frieman (2016). "Human polyclonal immunoglobulin G from transchromosomic bovines inhibits MERS-CoV in vivo." Sci Transl Med **8**(326): 326ra321.
Luo, H., J. Chen, K. Chen, X. Shen and H. Jiang (2006). "Carboxyl terminus of severe acute respiratory syndrome coronavirus nucleocapsid protein: self-association analysis and nucleic acid binding characterization." Biochemistry **45**(39): 11827-11835.
Luo, H., D. Wu, C. Shen, K. Chen, X. Shen and H. Jiang (2006). "Severe acute respiratory syndrome coronavirus membrane protein interacts with nucleocapsid protein mostly through their carboxyl termini by electrostatic attraction." Int J Biochem Cell Biol **38**(4): 589-599.
Luo, H., F. Ye, K. Chen, X. Shen and H. Jiang (2005). "SR-rich motif plays a pivotal role in recombinant SARS coronavirus nucleocapsid protein multimerization." Biochemistry **44**(46): 15351-15358.
Luo, J. and C. Xiong (2012). "DiagTest3Grp: An R Package for Analyzing Diagnostic Tests with Three Ordinal Groups." J Stat Softw **51**(3): 1-24.
Lynn, D. E. (2007). "Lepidopteran insect cell line isolation from insect tissue." Methods Mol Biol **388**: 139-154.
Maahs, D. M., N. A. West, J. M. Lawrence and E. J. Mayer-Davis (2010). "Epidemiology of type 1 diabetes." Endocrinol Metab Clin North Am **39**(3): 481-497.
Mackay, I. M., K. E. Arden, D. J. Speicher, N. T. O'Neil, P. K. McErlean, R. M. Greer, M. D. Nissen and T. P. Sloots (2012). "Co-circulation of four human coronaviruses (HCoVs) in Queensland children with acute respiratory tract illnesses in 2004." Viruses **4**(4): 637-653.
Maltezou, H. C. and S. Wicker (2013). "Measles in health-care settings." Am J Infect Control **41**(7): 661-663.
Manual, I. (2011). "Guide to Baculovirus Expression Vector Systems (BEVS) and Insect Cell Culture Techniques." Invitrogen Life Technologies.
Maranga, L., A. S. Coroadinha and M. J. Carrondo (2002). "Insect cell culture medium supplementation with fetal bovine serum and bovine serum albumin: effects on baculovirus adsorption and infection kinetics." Biotechnol Prog **18**(4): 855-861.
Martensen, P. M. and J. Justesen (2001). "Specific inhibitors prevent proteolytic degradation of recombinant proteins expressed in High Five cells." Biotechniques **30**(4): 782-784, 786, 788 passim.
Martin, K. (2020). "A Deep Dive Into Induction with IPTG." Retrieved 29/03/2020, from https://www.goldbio.com/articles/article/a-deep-dive-into-iptg-induction.
Masihi, K. N. (2001). "Fighting infection using immunomodulatory agents." Expert Opin Biol Ther **1**(4): 641-653.
McBride, R., M. van Zyl and B. C. Fielding (2014). "The coronavirus nucleocapsid is a multifunctional protein." Viruses **6**(8): 2991-3018.
McCrory, S. A. (2009). SARS coronavirus : the nucleocapsid protein and the human immune response / Sarah Ann McCrory. PhD, University of Reading.
McDonald, J. H. (2014). Handbook of Biological Statistics. Baltimore, Maryland, Sparky House Publishing.
McDonald, L. C., A. E. Simor, I. J. Su, S. Maloney, M. Ofner, K. T. Chen, J. F. Lando, A. McGeer, M. L. Lee and D. B. Jernigan (2004). "SARS in healthcare facilities, Toronto and Taiwan." Emerg Infect Dis **10**(5): 777-781.
McPherson, C., R. Chubet, K. Holtz, Y. Honda-Okubo, D. Barnard, M. Cox and N. Petrovsky (2016). "Development of a SARS Coronavirus Vaccine from Recombinant Spike Protein Plus Delta Inulin Adjuvant." Methods Mol Biol **1403**: 269-284.

Menachery, V. D., A. J. Eisfeld, A. Schafer, L. Josset, A. C. Sims, S. Proll, S. Fan, C. Li, G. Neumann, S. C. Tilton, J. Chang, L. E. Gralinski, C. Long, R. Green, C. M. Williams, J. Weiss, M. M. Matzke, B. J. Webb-Robertson, A. A. Schepmoes, A. K. Shukla, T. O. Metz, R. D. Smith, K. M. Waters, M. G. Katze, Y. Kawaoka and R. S. Baric (2014). "Pathogenic influenza viruses and coronaviruses utilize similar and contrasting approaches to control interferon-stimulated gene responses." mBio **5**(3): e01174-01114.
Menachery, V. D., B. L. Yount, Jr., K. Debbink, S. Agnihothram, L. E. Gralinski, J. A. Plante, R. L. Graham, T. Scobey, X. Y. Ge, E. F. Donaldson, S. H. Randell, A. Lanzavecchia, W. A. Marasco, Z. L. Shi and R. S. Baric (2015). "A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence." Nat Med **21**(12): 1508-1513.
Menachery, V. D., B. L. Yount Jr, K. Debbink, S. Agnihothram, L. E. Gralinski, J. A. Plante, R. L. Graham, T. Scobey, X.-Y. Ge, E. F. Donaldson, S. H. Randell, A. Lanzavecchia, W. A. Marasco, Z.-L. Shi and R. S. Baric (2015). "A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence." Nat Med **advance online publication**.
Middleton, D. (2014). "Hendra virus." Vet Clin North Am Equine Pract **30**(3): 579-589.
Mills, C. E., J. M. Robins and M. Lipsitch (2004). "Transmissibility of 1918 pandemic influenza." Nature **432**(7019): 904-906.
Miroux, B. and J. E. Walker (1996). "Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels." J Mol Biol **260**(3): 289-298.
Monto, A. S., J. L. McKimm-Breschkin, C. Macken, A. W. Hampson, A. Hay, A. Klimov, M. Tashiro, R. G. Webster, M. Aymard, F. G. Hayden and M. Zambon (2006). "Detection of influenza viruses resistant to neuraminidase inhibitors in global surveillance during the first 3 years of their use." Antimicrob Agents Chemother **50**(7): 2395-2402.
Monto, A. S., D. P. Robinson, M. L. Herlocher, J. M. Hinson, Jr., M. J. Elliott and A. Crisp (1999). "Zanamivir in the prevention of influenza among healthy adults: a randomized controlled trial." JAMA **282**(1): 31-35.
Mu, F., D. Niu, J. Mu, B. He, W. Han, B. Fan, S. Huang, Y. Qiu, B. You and W. Chen (2008). "The expression and antigenicity of a truncated spike-nucleocapsid fusion protein of severe acute respiratory syndrome-associated coronavirus." BMC Microbiol **8**: 207.
Mukhopadhyay, S., R. J. Kuhn and M. G. Rossmann (2005). "A structural perspective of the flavivirus life cycle." Nat Rev Microbiol **3**(1): 13-22.
Muller, L. M., K. J. Gorter, E. Hak, W. L. Goudzwaard, F. G. Schellevis, A. I. Hoepelman and G. E. Rutten (2005). "Increased risk of common infections in patients with type 1 and type 2 diabetes mellitus." Clin Infect Dis **41**(3): 281-288.
Müller, M. A., B. Meyer, V. M. Corman, M. Al-Masri, A. Turkestani, D. Ritz, A. Sieberg, S. Aldabbagh, B.-J. Bosch and E. Lattwein (2015). "Presence of Middle East respiratory syndrome coronavirus antibodies in Saudi Arabia: a nationwide, cross-sectional, serological study." The Lancet Infectious Diseases **15**(5): 559-564.
Müller, M. A. R., V.S.; Muth, D.; Meyer, B.; Kallies, S.; Smits, S. L.; Wollny, R.; Bestebroer, T. M.; Specht, S.; Sulimana, T.; Zimmermannd, K.; Bingera, T.; Eckerlea, I.; Tschapkae, M.; Zakif, A. M.; Osterhausb, A. D. M. E.; Fouchierb, R. A. M.; Haagmansb, B. L.; Drostena, C. (2012). "Human coronavirus EMC does not require the SARS-coronavirus receptor and maintains broad replicative capability in mammalian cell lines." American Society for Microbiology(3): 6.
Mulligan, M. J., K. E. Lyke, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, K. Neuzil, V. Raabe, R. Bailey, K. A. Swanson, P. Li, K. Koury, W. Kalina, D. Cooper, C. Fontes-Garfias, P. Y. Shi, O. Tureci, K. R. Tompkins, E. E. Walsh, R. Frenck, A. R. Falsey, P. R. Dormitzer, W. C. Gruber, U. Sahin and K. U. Jansen (2020). "Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults." Nature **586**(7830): 589-593.
Muthumani, K., D. Falzarano, E. L. Reuschel, C. Tingey, S. Flingai, D. O. Villarreal, M. Wise, A. Patel, A. Izmirly, A. Aljuaid, A. M. Seliga, G. Soule, M. Morrow, K. A. Kraynyak, A. S. Khan, D. P. Scott, F. Feldmann, R. LaCasse, K. Meade-White, A. Okumura, K. E. Ugen, N. Y. Sardesai, J. J. Kim, G. Kobinger, H. Feldmann and D. B. Weiner (2015). "A synthetic consensus anti-spike protein DNA vaccine induces protective immunity against Middle East respiratory syndrome coronavirus in nonhuman primates." Sci Transl Med **7**(301): 301ra132.
MyBioSource. (2020). HCoV-HKU-IgG elisa kit :: Human Coronavirus IgG (HCoV-HKU-IgG) ELISA Kit Retrieved 06/04/2020, from https://www.mybiosource.com/human-elisa-kits/coronavirus-igg-hcov-hku-igg/9301037.
MyBioSource. (2020). MERS-COV elisa kit :: Human Middle East Respiratory Syndrome Coronavirus (MERS-COV) ELISA Kit Retrieved 06/04/2020, from https://www.mybiosource.com/prods/ELISA-Kit/Human/Middle-East-Respiratory-Syndrome-Coronavirus-MERS-COV/MERS-COV/datasheet.php?products_id=109244.
Naesens, L., A. Stevaert and E. Vanderlinden (2016). "Antiviral therapies on the horizon for influenza." Curr Opin Pharmacol **30**: 106-115.

Najafi, H., A. Ghalyanchi Langeroudi, M. Hashemzadeh, V. Karimi, O. Madadgar, R. Khaltabadi Farahani, S. A. Ghafouri, H. Maghsoudloo, P. Seifouri and A. Madhi (2017). "Pathogenicity study of Iranian genotype of avian infectious bronchitis virus (IR-1)." Vet Res Forum **8**(1): 35-41.
Nelson, G. W., S. A. Stohlman and S. M. Tahara (2000). "High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA." J Gen Virol **81**(Pt 1): 181-188.
Neumann, G. and Y. Kawaoka (2015). "Transmission of influenza A viruses." Virology **479-480**: 234-246.
Ng, D. L., F. Al Hosani, M. K. Keating, S. I. Gerber, T. L. Jones, M. G. Metcalfe, S. Tong, Y. Tao, N. N. Alami, L. M. Haynes, M. A. Mutei, L. Abdel-Wareth, T. M. Uyeki, D. L. Swerdlow, M. Barakat and S. R. Zaki (2016). "Clinicopathologic, Immunohistochemical, and Ultrastructural Findings of a Fatal Case of Middle East Respiratory Syndrome Coronavirus Infection in the United Arab Emirates, April 2014." Am J Pathol **186**(3): 652-658.
Ng, W. F., K. F. To, W. W. Lam, T. K. Ng and K. C. Lee (2006). "The comparative pathology of severe acute respiratory syndrome and avian influenza A subtype H5N1--a review." Hum Pathol **37**(4): 381-390.
Nikolai, L. A., C. G. Meyer, P. G. Kremsner and T. P. Velavan (2020). "Asymptomatic SARS Coronavirus 2 infection: Invisible yet invincible." Int J Infect Dis **100**: 112-116.
Nuorti, J. P., J. C. Butler, M. M. Farley, L. H. Harrison, A. McGeer, M. S. Kolczak and R. F. Breiman (2000). "Cigarette smoking and invasive pneumococcal disease. Active Bacterial Core Surveillance Team." N Engl J Med **342**(10): 681-689.
Ohnuma, K., B. L. Haagmans, R. Hatano, V. S. Raj, H. Mou, S. Iwata, N. H. Dang, B. J. Bosch and C. Morimoto (2013). "Inhibition of Middle East respiratory syndrome coronavirus infection by anti-CD26 monoclonal antibody." J Virol **87**(24): 13892-13899.
Oikarinen, S., M. Martiskainen, S. Tauriainen, H. Huhtala, J. Ilonen, R. Veijola, O. Simell, M. Knip and H. Hyoty (2011). "Enterovirus RNA in blood is linked to the development of type 1 diabetes." Diabetes **60**(1): 276-279.
Okeh, U. M. and C. N. Okoro (2012). "Evaluating Measures of Indicators of Diagnostic Test Performance: Fundamental Meanings and Formulars." J Biomet Biostat **3**(1).
Olarinmoye, A. O., B. O. Olugasa, H. Niphuis, R. V. Herwijnen, E. Verschoor, A. Boug, O. O. Ishola, H. Buitendijk, Z. Fagrouch and K. Al-Hezaimi (2017). "Serological evidence of coronavirus infections in native hamadryas baboons (Papio hamadryas hamadryas) of the Kingdom of Saudi Arabia." Epidemiol Infect **145**(10): 2030-2037.
Omrani, A. S., M. M. Saad, K. Baig, A. Bahloul, M. Abdul-Matin, A. Y. Alaidaroos, G. A. Almakhlafi, M. M. Albarrak, Z. A. Memish and A. M. Albarrak (2014). "Ribavirin and interferon alfa-2a for severe Middle East respiratory syndrome coronavirus infection: a retrospective cohort study." Lancet Infect Dis **14**(11): 1090-1095.
Osterhaus, A. D., G. F. Rimmelzwaan, B. E. Martina, T. M. Bestebroer and R. A. Fouchier (2000). "Influenza B virus in seals." Science **288**(5468): 1051-1053.
Pak, C. Y., H. M. Eun, R. G. McArthur and J. W. Yoon (1988). "Association of cytomegalovirus infection with autoimmune type 1 diabetes." Lancet **2**(8601): 1-4.
Pan, A. A., G. B. Rosenberg, M. K. Hurley, G. J. Schock, V. P. Chu and A. Aiyappa (1992). "Clinical evaluation of an EIA for the sensitive and specific detection of serum antibody to Trypanosoma cruzi (Chagas' disease)." J Infect Dis **165**(3): 585-588.
Parikh, R., A. Mathai, S. Parikh, G. C. Sekhar and R. Thomas (2008). "Understanding and using sensitivity, specificity and predictive values." Indian journal of ophthalmology **56**(1): 45.
Pascal, K. E., C. M. Coleman, A. O. Mujica, V. Kamat, A. Badithe, J. Fairhurst, C. Hunt, J. Strein, A. Berrebi, J. M. Sisk, K. L. Matthews, R. Babb, G. Chen, K. M. Lai, T. T. Huang, W. Olson, G. D. Yancopoulos, N. Stahl, M. B. Frieman and C. A. Kyratsous (2015). "Pre- and postexposure efficacy of fully human antibodies against Spike protein in a novel humanized mouse model of MERS-CoV infection." Proc Natl Acad Sci U S A **112**(28): 8738-8743.
Patterson, K. D. and G. F. Pyle (1991). "The geography and mortality of the 1918 influenza pandemic." Bull Hist Med **65**(1): 4-21.
Peiris, J. S., C. M. Chu, V. C. Cheng, K. S. Chan, I. F. Hung, L. L. Poon, K. I. Law, B. S. Tang, T. Y. Hon, C. S. Chan, K. H. Chan, J. S. Ng, B. J. Zheng, W. L. Ng, R. W. Lai, Y. Guan, K. Y. Yuen and H. U. S. S. Group (2003). "Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study." Lancet **361**(9371): 1767-1772.
Peiris, J. S., Y. Guan and K. Y. Yuen (2004). "Severe acute respiratory syndrome." Nat Med **10**(12 Suppl): S88-97.
Peleg, A. Y., T. Weerarathna, J. S. McCarthy and T. M. Davis (2007). "Common infections in diabetes: pathogenesis, management and relationship to glycaemic control." Diabetes Metab Res Rev **23**(1): 3-13.

Peng, D., D. Zhao, J. Liu, X. Wang, K. Yang, H. Xicheng, Y. Li and F. Wang (2009). "Multipathogen infections in hospitalized children with acute respiratory infections." Virol J **6**: 155.
Peng, T. Y., K. R. Lee and W. Y. Tarn (2008). "Phosphorylation of the arginine/serine dipeptide-rich motif of the severe acute respiratory syndrome coronavirus nucleocapsid protein modulates its multimerization, translation inhibitory activity and cellular localization." FEBS J **275**(16): 4152-4163.
Pfefferle, S., S. Oppong, J. F. Drexler, F. Gloza-Rausch, A. Ipsen, A. Seebens, M. A. Muller, A. Annan, P. Vallo, Y. Adu-Sarkodie, T. F. Kruppa and C. Drosten (2009). "Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in bats, Ghana." Emerg Infect Dis **15**(9): 1377-1384.
Pijlman, G. P., J. E. van Schijndel and J. M. Vlak (2003). "Spontaneous excision of BAC vector sequences from bacmid-derived baculovirus expression vectors upon passage in insect cells." J Gen Virol **84**(Pt 10): 2669-2678.
Polack, F. P., S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. Perez Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, Jr., L. L. Hammitt, O. Tureci, H. Nell, A. Schaefer, S. Unal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Sahin, K. U. Jansen, W. C. Gruber and C. C. T. Group (2020). "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine." N Engl J Med **383**(27): 2603-2615.
Porowinska, D., M. Wujak, K. Roszek and M. Komoszynski (2013). "[Prokaryotic expression systems]." Postepy Hig Med Dosw (Online) **67**: 119-129.
Post, N., D. Eddy, C. Huntley, M. C. I. van Schalkwyk, M. Shrotri, D. Leeman, S. Rigby, S. V. Williams, W. H. Bermingham, P. Kellam, J. Maher, A. M. Shields, G. Amirthalingam, S. J. Peacock and S. A. Ismail (2020). "Antibody response to SARS-CoV-2 infection in humans: A systematic review." PLoS One **15**(12): e0244126.
Powers, D. M. W. (2011). "Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation." Journal of Machine Learning Technologies **2**(1): 37-63.
Prompetchara, E., C. Ketloy and T. Palaga (2020). "Immune responses in COVID-19 and potential vaccines: Lessons learned from SARS and MERS epidemic." Asian Pac J Allergy Immunol **38**(1): 1-9.
PublicHealthEngland. (2019). "Annual National Flu programme letter 2019 to 2020." Retrieved 06/04/2020, from https://www.england.nhs.uk/wp-content/uploads/2019/03/annual-national-flu-programme-2019-to-2020-1.pdf.
Raj, V. S., H. Mou, S. L. Smits, D. H. Dekkers, M. A. Muller, R. Dijkman, D. Muth, J. A. Demmers, A. Zaki, R. A. Fouchier, V. Thiel, C. Drosten, P. J. Rottier, A. D. Osterhaus, B. J. Bosch and B. L. Haagmans (2013). "Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC." Nature **495**(7440): 251-254.
Ramasamy, M. N., A. M. Minassian, K. J. Ewer, A. L. Flaxman, P. M. Folegatti, D. R. Owens, M. Voysey, P. K. Aley, B. Angus, G. Babbage, S. Belij-Rammerstorfer, L. Berry, S. Bibi, M. Bittaye, K. Cathie, H. Chappell, S. Charlton, P. Cicconi, E. A. Clutterbuck, R. Colin-Jones, C. Dold, K. R. W. Emary, S. Fedosyuk, M. Fuskova, D. Gbesemete, C. Green, B. Hallis, M. M. Hou, D. Jenkin, C. C. D. Joe, E. J. Kelly, S. Kerridge, A. M. Lawrie, A. Lelliott, M. N. Lwin, R. Makinson, N. G. Marchevsky, Y. Mujadidi, A. P. S. Munro, M. Pacurar, E. Plested, J. Rand, T. Rawlinson, S. Rhead, H. Robinson, A. J. Ritchie, A. L. Ross-Russell, S. Saich, N. Singh, C. C. Smith, M. D. Snape, R. Song, R. Tarrant, Y. Themistocleous, K. M. Thomas, T. L. Villafana, S. C. Warren, M. E. E. Watson, A. D. Douglas, A. V. S. Hill, T. Lambe, S. C. Gilbert, S. N. Faust, A. J. Pollard and C. V. T. G. Oxford (2021). "Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial." Lancet **396**(10267): 1979-1993.
Redondo, M. J., J. Jeffrey, P. R. Fain, G. S. Eisenbarth and T. Orban (2008). "Concordance for islet autoimmunity among monozygotic twins." N Engl J Med **359**(26): 2849-2850.
Reilly, L. M. and L. A. Guarino (1996). "The viral ubiquitin gene of Autographa californica nuclear polyhedrosis virus is not essential for viral replication." Virology **218**(1): 243-247.
Restrepo, B. I., A. J. Camerlin, M. H. Rahbar, W. Wang, M. A. Restrepo, I. Zarate, F. Mora-Guzman, J. G. Crespo-Solis, J. Briggs, J. B. McCormick and S. P. Fisher-Hoch (2011). "Cross-sectional assessment reveals high diabetes prevalence among newly-diagnosed tuberculosis cases." Bull World Health Organ **89**(5): 352-359.
Rhee, C., S. Kanjilal, M. Baker and M. Klompas (2021). "Duration of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infectivity: When Is It Safe to Discontinue Isolation?" Clin Infect Dis **72**(8): 1467-1474.
Rikhtegaran Tehrani, Z., S. Saadat, E. Saleh, X. Ouyang, N. Constantine, A. L. DeVico, A. D. Harris, G. K. Lewis, S. Kottilil and M. M. Sajadi (2020). "Performance of nucleocapsid and spike-based SARS-CoV-2 serologic assays." PLoS One **15**(11): e0237828.
Rivals, I., L. Personnaz, L. Taing and M. C. Potier (2007). "Enrichment or depletion of a GO category within a class of genes: which test?" Bioinformatics **23**(4): 401-407.

Robin, X., N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J. C. Sanchez and M. Muller (2011). "pROC: an open-source package for R and S+ to analyze and compare ROC curves." BMC Bioinformatics **12**: 77.
Roper, R. L. and K. E. Rehm (2009). "SARS vaccines: where are we?" Expert Rev Vaccines **8**(7): 887-898.
Rosano, G. L. and E. A. Ceccarelli (2014). "Recombinant protein expression in Escherichia coli: advances and challenges." Front Microbiol **5**: 172.
Sastre, P., R. Dijkman, A. Camunas, T. Ruiz, M. F. Jebbink, L. van der Hoek, C. Vela and P. Rueda (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.
Sehr, P., K. Zumbach and M. Pawlita (2001). "A generic capture ELISA for recombinant proteins fused to glutathione S-transferase: validation for HPV serology." J Immunol Methods **253**(1-2): 153-162.
Severance, E. G., I. Bossis, F. B. Dickerson, C. R. Stallings, A. E. Origoni, A. Sullens, R. H. Yolken and R. P. Viscidi (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.
Sezonov, G., D. Joseleau-Petit and R. D'Ari (2007). "Escherichia coli physiology in Luria-Bertani broth." J Bacteriol **189**(23): 8746-8749.
Shahwan, K., M. Hesse, A. K. Mork, G. Herrler and C. Winter (2013). "Sialic acid binding properties of soluble coronavirus spike (S1) proteins: differences between infectious bronchitis virus and transmissible gastroenteritis virus." Viruses **5**(8): 1924-1933.
Shao-Hua, C., S. Hong-Liang and L. Zuo-Hu (1998). "Effect of temperature oscillation on insect cell growth and baculovirus replication." Appl Environ Microbiol **64**(6): 2237-2239.
Sharma, O., A. A. Sultan, H. Ding and C. R. Triggle (2020). "A Review of the Progress and Challenges of Developing a Vaccine for COVID-19." Front Immunol **11**: 585354.
Shi, X. and D. L. Jarvis (2007). "Protein N-glycosylation in the baculovirus-insect cell system." Curr Drug Targets **8**(10): 1116-1125.
Shi, Y., Y. Yi, P. Li, T. Kuang, L. Li, M. Dong, Q. Ma and C. Cao (2003). "Diagnosis of severe acute respiratory syndrome (SARS) by detection of SARS coronavirus nucleocapsid antibodies in an antigen-capturing enzyme-linked immunosorbent assay." J Clin Microbiol **41**(12): 5781-5782.
Shultz, D. (2014). "Cows with human chromosomes enlisted to fight hantavirus." from http://www.sciencemag.org/news/2014/11/cows-human-chromosomes-enlisted-fight-hantavirus.
Simon, J. H., E. Claassen, C. E. Correa and A. D. Osterhaus (2005). "Managing severe acute respiratory syndrome (SARS) intellectual property rights: the possible role of patent pooling." Bull World Health Organ **83**(9): 707-710.
Singh, A., V. Upadhyay, A. K. Upadhyay, S. M. Singh and A. K. Panda (2015). "Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process." Microb Cell Fact **14**: 41.
Smith, G. E., M. D. Summers and M. J. Fraser (1983). "Production of human beta interferon in insect cells infected with a baculovirus expression vector." Mol Cell Biol **3**(12): 2156-2165.
Song, D., G. Ha, W. Serhan, Y. Eltahir, M. Yusof, F. Hashem, E. Elsayed, B. Marzoug, A. Abdelazim and S. Al Muhairi (2015). "Development and validation of a rapid immunochromatographic assay for detection of Middle East respiratory syndrome coronavirus antigen in dromedary camels." J Clin Microbiol **53**(4): 1178-1182.
Song, F., R. Fux, L. B. Provacia, A. Volz, M. Eickmann, S. Becker, A. D. Osterhaus, B. L. Haagmans and G. Sutter (2013). "Middle East respiratory syndrome coronavirus spike protein delivered by modified vaccinia virus Ankara efficiently induces virus-neutralizing antibodies." J Virol **87**(21): 11950-11954.
Stano, N. M. and S. S. Patel (2004). "T7 lysozyme represses T7 RNA polymerase transcription by destabilizing the open complex during initiation." J Biol Chem **279**(16): 16136-16143.
Steele, K. H., B. J. Stone, K. M. Franklin, A. Fath-Goodin, X. Zhang, H. Jiang, B. A. Webb and C. Geisler (2017). "Improving the baculovirus expression vector system with vankyrin-enhanced technology." Biotechnol Prog **33**(6): 1496-1507.
Stene, L. C., S. Oikarinen, H. Hyoty, K. J. Barriga, J. M. Norris, G. Klingensmith, J. C. Hutton, H. A. Erlich, G. S. Eisenbarth and M. Rewers (2010). "Enterovirus infection and progression from islet autoimmunity to type 1 diabetes: the Diabetes and Autoimmunity Study in the Young (DAISY)." Diabetes **59**(12): 3174-3180.
Stevaert, A. and L. Naesens (2016). "The Influenza Virus Polymerase Complex: An Update on Its Structure, Functions, and Significance for Antiviral Drug Design." Med Res Rev **36**(6): 1127-1173.
Stockman, L. J., R. Bellamy and P. Garner (2006). "SARS: systematic review of treatment effects." PLoS Med **3**(9): e343.
Studier, F. W. and B. A. Moffatt (1986). "Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes." J Mol Biol **189**(1): 113-130.

Studier, F. W., A. H. Rosenberg, J. J. Dunn and J. W. Dubendorff (1990). "Use of T7 RNA polymerase to direct expression of cloned genes." Methods Enzymol **185**: 60-89.
Su, M., C. Li, D. Guo, S. Wei, X. Wang, Y. Geng, S. Yao, J. Gao, E. Wang, X. Zhao, Z. Wang, J. Wang, R. Wu, L. Feng and D. Sun (2016). "A recombinant nucleocapsid protein-based indirect enzyme-linked immunosorbent assay to detect antibodies against porcine deltacoronavirus." J Vet Med Sci **78**(4): 601-606.
Surjit, M., R. Kumar, R. N. Mishra, M. K. Reddy, V. T. Chow and S. K. Lal (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., B. Liu, P. Kumar, V. T. Chow and S. K. Lal (2004). "The nucleocapsid protein of the SARS coronavirus is capable of self-association through a C-terminal 209 amino acid interaction domain." Biochem Biophys Res Commun **317**(4): 1030-1036.
Tan, Y. W., S. Fang, H. Fan, J. Lescar and D. X. Liu (2006). "Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells." Nucleic Acids Res **34**(17): 4816-4825.
Tang, T. K., M. P. Wu, S. T. Chen, M. H. Hou, M. H. Hong, F. M. Pan, H. M. Yu, J. H. Chen, C. W. Yao and A. H. Wang (2005). "Biochemical and immunological studies of nucleocapsid proteins of severe acute respiratory syndrome and 229E human coronaviruses." Proteomics **5**(4): 925-937.
Tang, X. C., S. S. Agnihothram, Y. Jiao, J. Stanhope, R. L. Graham, E. C. Peterson, Y. Avnir, A. S. Tallarico, J. Sheehan, Q. Zhu, R. S. Baric and W. A. Marasco (2014). "Identification of human neutralizing antibodies against MERS-CoV and their role in virus adaptive evolution." Proc Natl Acad Sci U S A **111**(19): E2018-2026.
Taubenberger, J. K. and D. M. Morens (2008). "The pathology of influenza virus infections." Annu Rev Pathol **3**: 499-522.
Timothy M. Uyeki, K. J. E., George Korch, Michael O'Hara, Michael Wathen, Jean Hu-Primmer, Sally Hojvat, Erik J. Stemmy, and Armen Donabedian (2016). "Development of Medical Countermeasures to Middle East Respiratory Syndrome Coronavirus." Emerging infectious diseases **22**(7).
Tong, S., X. Zhu, Y. Li, M. Shi, J. Zhang, M. Bourgeois, H. Yang, X. Chen, S. Recuenco, J. Gomez, L. M. Chen, A. Johnson, Y. Tao, C. Dreyfus, W. Yu, R. McBride, P. J. Carney, A. T. Gilbert, J. Chang, Z. Guo, C. T. Davis, J. C. Paulson, J. Stevens, C. E. Rupprecht, E. C. Holmes, I. A. Wilson and R. O. Donis (2013). "New world bats harbor diverse influenza A viruses." PLoS Pathog **9**(10): e1003657.
Trials.gov., C. (2016). "Phase I, open label dose ranging safety study of GLS-5300 in healthy volunteers." from https://clinicaltrials.gov/ct2/show/NCT02670187?term=mers-cov&rank=7.
Triggle, C. R., D. Bansal, H. Ding, M. M. Islam, E. Farag, H. A. Hadi and A. A. Sultan (2021). "A Comprehensive Review of Viral Characteristics, Transmission, Pathophysiology, Immune Response, and Management of SARS-CoV-2 and COVID-19 as a Basis for Controlling the Pandemic." Front Immunol **12**: 631139.
Tseng, C. T., E. Sbrana, N. Iwata-Yoshikawa, P. C. Newman, T. Garron, R. L. Atmar, C. J. Peters and R. B. Couch (2012). "Immunization with SARS coronavirus vaccines leads to pulmonary immunopathology on challenge with the SARS virus." PLoS One **7**(4): e35421.
Uyeki, T. M., H. H. Bernstein, J. S. Bradley, J. A. Englund, T. M. File, A. M. Fry, S. Gravenstein, F. G. Hayden, S. A. Harper, J. M. Hirshon, M. G. Ison, B. L. Johnston, S. L. Knight, A. McGeer, L. E. Riley, C. R. Wolfe, P. E. Alexander and A. T. Pavia (2019). "Clinical Practice Guidelines by the Infectious Diseases Society of America: 2018 Update on Diagnosis, Treatment, Chemoprophylaxis, and Institutional Outbreak Management of Seasonal Influenzaa." Clin Infect Dis **68**(6): e1-e47.
van Doremalen, N., T. Bushmaker, D. H. Morris, M. G. Holbrook, A. Gamble, B. N. Williamson, A. Tamin, J. L. Harcourt, N. J. Thornburg, S. I. Gerber, J. O. Lloyd-Smith, E. de Wit and V. J. Munster (2020). "Aerosol and Surface Stability of SARS-CoV-2 as Compared with SARS-CoV-1." N Engl J Med **382**(16): 1564-1567.
van Doremalen, N., K. L. Miazgowicz, S. Milne-Price, T. Bushmaker, S. Robertson, D. Scott, J. Kinne, J. S. McLellan, J. Zhu and V. J. Munster (2014). "Host species restriction of Middle East respiratory syndrome coronavirus through its receptor, dipeptidyl peptidase 4." J Virol **88**(16): 9220-9232.
van Doremalen, N. and V. J. Munster (2015). "Animal models of Middle East respiratory syndrome coronavirus infection." Antiviral Res **122**: 28-38.
Vardakas, K. Z., Siempos, II and M. E. Falagas (2007). "Diabetes mellitus as a risk factor for nosocomial pneumonia and associated mortality." Diabet Med **24**(10): 1168-1171.
Vaughn, J. L., R. H. Goodwin, G. J. Tompkins and P. McCawley (1977). "The establishment of two cell lines from the insect Spodoptera frugiperda (Lepidoptera; Noctuidae)." In Vitro **13**(4): 213-217.
Velavan, T. P. and C. G. Meyer (2020). "The COVID-19 epidemic." Trop Med Int Health **25**(3): 278-280.
Vijgen, L., E. Keyaerts, E. Moes, P. Maes, G. Duson and M. Van Ranst (2005). "Development of one-step, real-time, quantitative reverse transcriptase PCR assays for absolute quantitation of human coronaviruses OC43 and 229E." J Clin Microbiol **43**(11): 5452-5456.

Volz, A., A. Kupke, F. Song, S. Jany, R. Fux, H. Shams-Eldin, J. Schmidt, C. Becker, M. Eickmann, S. Becker and G. Sutter (2015). "Protective Efficacy of Recombinant Modified Vaccinia Virus Ankara Delivering Middle East Respiratory Syndrome Coronavirus Spike Glycoprotein." J Virol **89**(16): 8651-8656.
Voysey, M., S. A. C. Clemens, S. A. Madhi, L. Y. Weckx, P. M. Folegatti, P. K. Aley, B. Angus, V. L. Baillie, S. L. Barnabas, Q. E. Bhorat, S. Bibi, C. Briner, P. Cicconi, A. M. Collins, R. Colin-Jones, C. L. Cutland, T. C. Darton, K. Dheda, C. J. A. Duncan, K. R. W. Emary, K. J. Ewer, L. Fairlie, S. N. Faust, S. Feng, D. M. Ferreira, A. Finn, A. L. Goodman, C. M. Green, C. A. Green, P. T. Heath, C. Hill, H. Hill, I. Hirsch, S. H. C. Hodgson, A. Izu, S. Jackson, D. Jenkin, C. C. D. Joe, S. Kerridge, A. Koen, G. Kwatra, R. Lazarus, A. M. Lawrie, A. Lelliott, V. Libri, P. J. Lillie, R. Mallory, A. V. A. Mendes, E. P. Milan, A. M. Minassian, A. McGregor, H. Morrison, Y. F. Mujadidi, A. Nana, P. J. O'Reilly, S. D. Padayachee, A. Pittella, E. Plested, K. M. Pollock, M. N. Ramasamy, S. Rhead, A. V. Schwarzbold, N. Singh, A. Smith, R. Song, M. D. Snape, E. Sprinz, R. K. Sutherland, R. Tarrant, E. C. Thomson, M. E. Torok, M. Toshner, D. P. J. Turner, J. Vekemans, T. L. Villafana, M. E. E. Watson, C. J. Williams, A. D. Douglas, A. V. S. Hill, T. Lambe, S. C. Gilbert, A. J. Pollard and C. V. T. G. Oxford (2021). "Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK." Lancet **397**(10269): 99-111.
Wade-Jardetzky, N., R. P. Bray, W. W. Conover, O. Jardetzky, N. Geisler and K. Weber (1979). "Differential mobility of the N-terminal headpiece in the lac-repressor protein." J Mol Biol **128**(2): 259-264.
Wan, X. F. (2012). "Lessons from emergence of A/goose/Guangdong/1996-like H5N1 highly pathogenic avian influenza viruses and recent influenza surveillance efforts in southern China." Zoonoses Public Health **59 Suppl 2**(0 2): 32-42.
Wang, C., C. Hua, S. Xia, W. Li, L. Lu and S. Jiang (2019). "Combining a Fusion Inhibitory Peptide Targeting the MERS-CoV S2 Protein HR1 Domain and a Neutralizing Antibody Specific for the S1 Protein Receptor-Binding Domain (RBD) Showed Potent Synergism against Pseudotyped MERS-CoV with or without Mutations in RBD." Viruses **11**(1).
Wang, C., X. Zheng, W. Gai, Y. Zhao, H. Wang, H. Wang, N. Feng, H. Chi, B. Qiu, N. Li, T. Wang, Y. Gao, S. Yang and X. Xia (2017). "MERS-CoV virus-like particles produced in insect cells induce specific humoural and cellular imminity in rhesus macaques." Oncotarget **8**(8): 12686-12694.
Wang, L., W. Shi, M. G. Joyce, K. Modjarrad, Y. Zhang, K. Leung, C. R. Lees, T. Zhou, H. M. Yassine, M. Kanekiyo, Z. Y. Yang, X. Chen, M. M. Becker, M. Freeman, L. Vogel, J. C. Johnson, G. Olinger, J. P. Todd, U. Bagci, J. Solomon, D. J. Mollura, L. Hensley, P. Jahrling, M. R. Denison, S. S. Rao, K. Subbarao, P. D. Kwong, J. R. Mascola, W. P. Kong and B. S. Graham (2015). "Evaluation of candidate vaccine approaches for MERS-CoV." Nat Commun **6**: 7712.
Wang, Z., Y. Jiyuan, C. Su, Q. Xinyuan, T. Lijie and L. Yijing (2015). "Development of an antigen capture enzyme-linked immunosorbent assay for virus detection based on porcine epidemic diarrhea virus monoclonal antibodies." Viral Immunol **28**(3): 184-189.
Warren, T. K., J. Wells, R. G. Panchal, K. S. Stuthman, N. L. Garza, S. A. Van Tongeren, L. Dong, C. J. Retterer, B. P. Eaton, G. Pegoraro, S. Honnold, S. Bantia, P. Kotian, X. Chen, B. R. Taubenheim, L. S. Welch, D. M. Minning, Y. S. Babu, W. P. Sheridan and S. Bavari (2014). "Protection against filovirus diseases by a novel broad-spectrum nucleoside analogue BCX4430." Nature **508**(7496): 402-405.
Webster, R. G. (2002). "The importance of animal influenza for human disease." Vaccine **20 Suppl 2**: S16-20.
Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers and Y. Kawaoka (1992). "Evolution and ecology of influenza A viruses." Microbiol Rev **56**(1): 152-179.
Weiss, J. N. (1997). "The Hill equation revisited: uses and misuses." FASEB J **11**(11): 835-841.
WHO (2017). "Middle East respiratory syndrome coronavirus (MERS-CoV)."
WHO. (2017). "Middle East respiratory syndrome coronavirus (MERS-CoV) :MERS-CoV in Republic of Korea at a glance." Retrieved 25th April 2017, from http://www.wpro.who.int/outbreaks_emergencies/wpro_coronavirus/en/.
WHO. (2019). "Recommended composition of influenza virus vaccines for use in the 2019-2020 northern hemisphere influenza season." Retrieved 06/04/2020, from https://www.who.int/influenza/vaccines/virus/recommendations/2019_20_north/en/.
WHO. (2020, December 2020). "MERS situation update, December 2020." Retrieved 19th June 2021, from http://www.emro.who.int/health-topics/mers-cov/mers-outbreaks.html.
WHO. (2020). "Severe acute respiratory syndrome." Retrieved 19/06/2021, 2021, from http://www.emro.who.int/health-topics/severe-acute-respiratory-syndrome/.
WHO. (2021). "COVID-19 Weekly Epidemiological Update " Edition 44. from https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---15-june-2021.
WHO. (2021). "Tracking SARS-CoV-2 variants." Retrieved 18/08/2021, from https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/.

WHO. (2021). "WHO Coronavirus (COVID-19) Dashboard." Retrieved 18/08/2021, from https://covid19.who.int/.
Wibowo, M. H., S. Tarigan, Sumarningsih, S. Artanto, R. Indriani, D. Anggoro, C. P. Putra, S. Idris, T. Untari, W. Asmara, C. R. Tabbu and J. Ignjatovic (2017). "Use of M2e ELISAs for longitudinal surveillance of commercial poultry in Indonesia vaccinated against highly pathogenic avian influenza." J Virol Methods **249**: 181-188.
Widagdo, W., N. M. A. Okba, M. Richard, D. de Meulder, T. M. Bestebroer, P. Lexmond, E. Farag, M. Al-Hajri, K. J. Stittelaar, L. de Waal, G. van Amerongen, J. M. A. van den Brand, B. L. Haagmans and S. Herfst (2019). "Lack of Middle East Respiratory Syndrome Coronavirus Transmission in Rabbits." Viruses **11**(4).
Wirblich, C., C. M. Coleman, D. Kurup, T. S. Abraham, J. G. Bernbaum, P. B. Jahrling, L. E. Hensley, R. F. Johnson, M. B. Frieman and M. J. Schnell (2017). "One-Health: a Safe, Efficient, Dual-Use Vaccine for Humans and Animals against Middle East Respiratory Syndrome Coronavirus and Rabies Virus." J Virol **91**(2).
Wise, A. G., M. Kiupel and R. K. Maes (2006). "Molecular characterization of a novel coronavirus associated with epizootic catarrhal enteritis (ECE) in ferrets." Virology **349**(1): 164-174.
Woestenenk, E. A., M. Hammarstrom, S. van den Berg, T. Hard and H. Berglund (2004). "His tag effect on solubility of human proteins produced in Escherichia coli: a comparison between four expression vectors." J Struct Funct Genomics **5**(3): 217-229.
Woo, P. C., S. K. Lau, C. S. Lam, C. C. Lau, A. K. Tsang, J. H. Lau, R. Bai, J. L. Teng, C. C. Tsang, M. Wang, B. J. Zheng, K. H. Chan and K. Y. Yuen (2012). "Discovery of seven novel Mammalian and avian coronaviruses in the genus deltacoronavirus supports bat coronaviruses as the gene source of alphacoronavirus and betacoronavirus and avian coronaviruses as the gene source of gammacoronavirus and deltacoronavirus." J Virol **86**(7): 3995-4008.
Wu, Z., L. Yang, X. Ren, G. He, J. Zhang, J. Yang, Z. Qian, J. Dong, L. Sun, Y. Zhu, J. Du, F. Yang, S. Zhang and Q. Jin (2016). "Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases." ISME J **10**(3): 609-620.
Wycoff, K., J. Maclean, A. Belle, L. Yu, Y. Tran, C. Roy and F. Hayden (2015). "Anti-infective immunoadhesins from plants." Plant Biotechnol J **13**(8): 1078-1093.
Xia, S., K. Duan, Y. Zhang, D. Zhao, H. Zhang, Z. Xie, X. Li, C. Peng, Y. Zhang, W. Zhang, Y. Yang, W. Chen, X. Gao, W. You, X. Wang, Z. Wang, Z. Shi, Y. Wang, X. Yang, L. Zhang, L. Huang, Q. Wang, J. Lu, Y. Yang, J. Guo, W. Zhou, X. Wan, C. Wu, W. Wang, S. Huang, J. Du, Z. Meng, A. Pan, Z. Yuan, S. Shen, W. Guo and X. Yang (2020). "Effect of an Inactivated Vaccine Against SARS-CoV-2 on Safety and Immunogenicity Outcomes: Interim Analysis of 2 Randomized Clinical Trials." JAMA **324**(10): 951-960.
Xia, S., Q. Lan, J. Pu, C. Wang, Z. Liu, W. Xu, Q. Wang, H. Liu, S. Jiang and L. Lu (2019). "Potent MERS-CoV Fusion Inhibitory Peptides Identified from HR2 Domain in Spike Protein of Bat Coronavirus HKU4." Viruses **11**(1).
Xiu, L., C. Zhang, Z. Wu and J. Peng (2017). "Establishment and Application of a Universal Coronavirus Screening Method Using MALDI-TOF Mass Spectrometry." Front Microbiol **8**: 1510.
Xu, P. W., X. Wu, H. N. Wang, B. C. Ma, M. D. Ding and X. Yang (2016). "Assembly and immunogenicity of baculovirus-derived infectious bronchitis virus-like particles carrying membrane, envelope and the recombinant spike proteins." Biotechnol Lett **38**(2): 299-304.
Yamaji, H. (2011). Production of antibody in insect cells. Antibody Expression and Production, Springer**:** 53-76.
Yamamoto, M., S. Matsuyama, X. Li, M. Takeda, Y. Kawaguchi, J. I. Inoue and Z. Matsuda (2016). "Identification of Nafamostat as a Potent Inhibitor of Middle East Respiratory Syndrome Coronavirus S Protein-Mediated Membrane Fusion Using the Split-Protein-Based Cell-Cell Fusion Assay." Antimicrob Agents Chemother **60**(11): 6532-6539.
Yang, J. K., Y. Feng, M. Y. Yuan, S. Y. Yuan, H. J. Fu, B. Y. Wu, G. Z. Sun, G. R. Yang, X. L. Zhang, L. Wang, X. Xu, X. P. Xu and J. C. Chan (2006). "Plasma glucose levels and diabetes are independent predictors for mortality and morbidity in patients with SARS." Diabet Med **23**(6): 623-628.
Yeung, W. C., W. D. Rawlinson and M. E. Craig (2011). "Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies." BMJ **342**: d35.
Ying, T., L. Du, T. W. Ju, P. Prabakaran, C. C. Lau, L. Lu, Q. Liu, L. Wang, Y. Feng, Y. Wang, B. J. Zheng, K. Y. Yuen, S. Jiang and D. S. Dimitrov (2014). "Exceptionally potent neutralization of Middle East respiratory syndrome coronavirus by human monoclonal antibodies." J Virol **88**(14): 7796-7805.
You, J., B. K. Dove, L. Enjuanes, M. L. DeDiego, E. Alvarez, G. Howell, P. Heinen, M. Zambon and J. A. Hiscox (2005). "Subcellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein." J Gen Virol **86**(Pt 12): 3303-3310.
Yu, F., M. Q. Le, S. Inoue, F. Hasebe, C. Parquet Mdel, S. Morikawa and K. Morita (2007). "Recombinant truncated nucleocapsid protein as antigen in a novel immunoglobulin M capture enzyme-linked

immunosorbent assay for diagnosis of severe acute respiratory syndrome coronavirus infection." Clin Vaccine Immunol **14**(2): 146-149.
Yu, F., M. Q. Le, S. Inoue, H. T. Thai, F. Hasebe, M. Del Carmen Parquet and K. Morita (2005). "Evaluation of inapparent nosocomial severe acute respiratory syndrome coronavirus infection in Vietnam by use of highly specific recombinant truncated nucleocapsid protein-based enzyme-linked immunosorbent assay." Clin Diagn Lab Immunol **12**(7): 848-854.
Yu, I. M., C. L. Gustafson, J. Diao, J. W. Burgner, 2nd, Z. Li, J. Zhang and J. Chen (2005). "Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain." J Biol Chem **280**(24): 23280-23286.
Yu, I. M., M. L. Oldham, J. Zhang and J. Chen (2006). "Crystal structure of the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein dimerization domain reveals evolutionary linkage between corona- and arteriviridae." J Biol Chem **281**(25): 17134-17139.
Zaki, A. M., S. van Boheemen, T. M. Bestebroer, A. D. Osterhaus and R. A. Fouchier (2012). "Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia." N Engl J Med **367**(19): 1814-1820.
Zhang, Q., H. Zhang, K. Huang, Y. Yang, X. Hui, J. Gao, X. He, C. Li, W. Gong and Y. Zhang (2020). "SARS-CoV-2 neutralizing serum antibodies in cats: a serological investigation." BioRxiv.
Zhang, R., Y. Li, A. L. Zhang, Y. Wang and M. J. Molina (2020). "Identifying airborne transmission as the dominant route for the spread of COVID-19." Proc Natl Acad Sci U S A **117**(26): 14857-14863.
Zhang, W., L. Hou, J. Song, S. Zhang, Y. Li, J. Li, L. Sun, W. Fan and W. Liu (2017). "[Establishment of a high sensitive indirect ELISA for detecting specific antibodies against H9 subtype avian influenza virus]." Sheng Wu Gong Cheng Xue Bao **33**(8): 1253-1264.
Zhang, Y., J. Corver, P. R. Chipman, W. Zhang, S. V. Pletnev, D. Sedlak, T. S. Baker, J. H. Strauss, R. J. Kuhn and M. G. Rossmann (2003). "Structures of immature flavivirus particles." EMBO J **22**(11): 2604-2613.
Zhao, G., L. He, S. Sun, H. Qiu, W. Tai, J. Chen, J. Li, Y. Chen, Y. Guo, Y. Wang, J. Shang, K. Ji, R. Fan, E. Du, S. Jiang, F. Li, L. Du and Y. Zhou (2018). "A Novel Nanobody Targeting Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Receptor-Binding Domain Has Potent Cross-Neutralizing Activity and Protective Efficacy against MERS-CoV." J Virol **92**(18).
Zhao, Y., D. A. Chapman and I. M. Jones (2003). "Improving baculovirus recombination." Nucleic Acids Res **31**(2): E6-6.
Zhou, Y., Y. Yang, J. Huang, S. Jiang and L. Du (2019). "Advances in MERS-CoV Vaccines and Therapeutics Based on the Receptor-Binding Domain." Viruses **11**(1).
Zhu, F. C., X. H. Guan, Y. H. Li, J. Y. Huang, T. Jiang, L. H. Hou, J. X. Li, B. F. Yang, L. Wang, W. J. Wang, S. P. Wu, Z. Wang, X. H. Wu, J. J. Xu, Z. Zhang, S. Y. Jia, B. S. Wang, Y. Hu, J. J. Liu, J. Zhang, X. A. Qian, Q. Li, H. X. Pan, H. D. Jiang, P. Deng, J. B. Gou, X. W. Wang, X. H. Wang and W. Chen (2020). "Immunogenicity and safety of a recombinant adenovirus type-5-vectored COVID-19 vaccine in healthy adults aged 18 years or older: a randomised, double-blind, placebo-controlled, phase 2 trial." Lancet **396**(10249): 479-488.
Zitzmann, J., G. Sprick, T. Weidner, C. Schreiber and P. Czermak (2017). Process Optimization for Recombinant Protein Expression in Insect Cells**:** 43-98.
Zlotnick, A. (2005). "Theoretical aspects of virus capsid assembly." J Mol Recognit **18**(6): 479-490.
Zuo, X., M. R. Mattern, R. Tan, S. Li, J. Hall, D. E. Sterner, J. Shoo, H. Tran, P. Lim, S. G. Sarafianos, L. Kazi, S. Navas-Martin, S. R. Weiss and T. R. Butt (2005). "Expression and purification of SARS coronavirus proteins using SUMO-fusions." Protein Expr Purif **42**(1): 100-110.
Zuwała, K., A. Golda, W. Kabala, M. Burmistrz, M. Zdzalik, P. Nowak, S. Kedracka-Krok, M. Zarebski, J. Dobrucki, D. Florek, S. Zeglen, J. Wojarski, J. Potempa, G. Dubin and K. Pyrc (2015). "The nucleocapsid protein of human coronavirus NL63." PLoS one **10**(2).

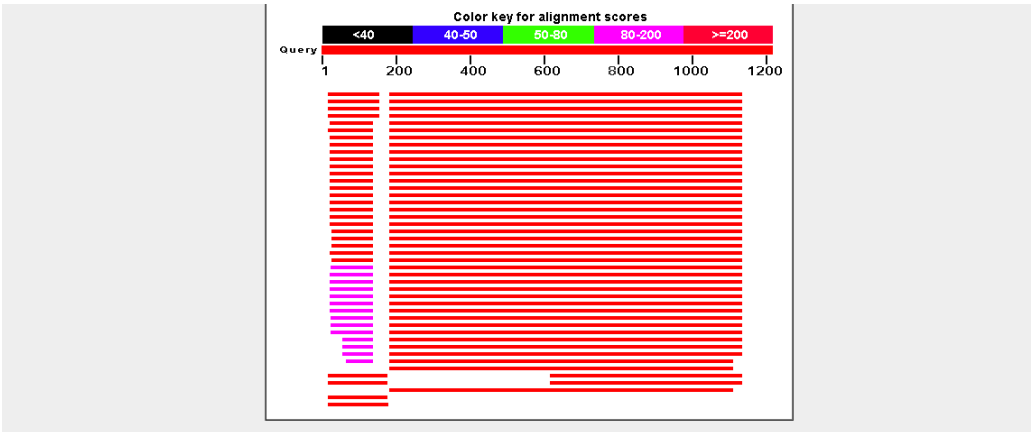
Appendix one: A table showing the percentages of each symptom displayed during four SARS cohort studies

SARS Clinical symptoms at presentation (in %)				
	(Lee, Hui et al. 2003) n=138	(Peiris, Chu et al. 2003) n=50	(A., Ghani et al. 2003) n > 1250	(Booth, Matukas et al. 2003). n=144
Fever	100	100	94	99
Chills or rigors	73	74	65	28
Cough	57	62	50	69
Myalgia	61	54	51	49
Malaise	n/a	50	64	31
Runny nose	23	24	25	2
Sore throat	23.	20	23	12
Shortness of breath	n/a	20	31	n/a
Diarrhoea	20	10	27	24
Headache	56	20	50	35

Appendix two: BLAST results

Sequencing data for 229E-CoV samples

```
                                                                                                    100
                                                                                                    |
229EReference  ATGGGCAGCAGCCATCATCATCATCATCACGCTACAGTCAAATGGGCTGATGCATCTGAACCACAACGTGGTCGTCAGGGTAGAATACCTTATTCTCTTT
229EA          ....................................................................................................
229EB          ....................................................................................................

                                                                                                    200
                                                                                                    |
229EReference  ATAGCCCTTTGCTTGTTGATAGTGAACAACCTTGGAAGGTGATACCTCGTAATTTGGTACCCATCAACAAGAAAGACAAAAATAAGCTTATAGGCTATTG
229EA          ....................................................................................................
229EB          ....................................................................................................

                                                                                                    300
                                                                                                    |
229EReference  GAATGTTCAAAAACGTTTCAGAACTAGAAAGGGCAAACGGGTGGATTTGTCACCCAAGTTACATTTTTATTATCTTGGCACAGGACCCCATAAAGCTGCA
229EA          ..........................................................C.G.....................................A....
229EB          ..........................................................C.G.....................................A....

                                                                                                    400
                                                                                                    |
229EReference  AAATTTAGAGAGCCGTGTTGAAGGTGTTGTCTGGGTTGCTGTTGATGGTGCTAAAACTGAACCTACAGGTTACGGTGTTAGGCGCAAGAATTCAGAACCAG
229EA          ..........................C.........................................................................
229EB          ..........................C.........................................................................

                                                                                                    500
                                                                                                    |
229EReference  AGATACCACACTTCAATCAAAAGCTCCCAAATGGTGTTACTGTTGTTGAAGAACCTGACTCCCGTGCTCCTTCCCGTTCTCAGTCAAGGTCTCAGAGTCG
229EA          ..........................................................G.........G.....G.........
229EB          ..........................................................G.........G.....G.........

                                                                                                    600
                                                                                                    |
229EReference  CGGTCGTGGTGAATCCAAATCTCAATCTCGGAATCCTTCAAGTGACAGAAACCATAACAGTCAGGATGACATCATGAAGGCAGTCGCTGCGGCTCTTAAA
229EA          ..................C.......A.....................................................T...............
229EB          ..................C.......A.....................................................T...............

                                                                                                    700
                                                                                                    |
229EReference  TCTTTAGGTTTTGACAAGCCTCAGGAAAAAGACAAAAAGTCAGCGAAAACGGGTACTCCTAAGCCTTCTCGTAATCAGAGTCCTTCTTCTTTTCAATCTG
229EA          ..............................T.....................................G......C....A..T
229EB          ..............................T.......T.............................G......C....A..T

                                                                                                    800
                                                                                                    |
229EReference  CTGCCAAGATTCTTGCTCGTTCTCAGAGTTCTGAAACAAAAGAACAAAAGCATGAAATGCAAAAGCCCACGGTGGAAAAGACAGCCTAACGATGATGTGAC
229EA          .........G.....................................................................T..........
229EB          .........G.....................................................................T..........

                                                                                                    850
                                                                                                    |
229EReference  ATCTAATGTCACACAATGTTTTGGCCCCAGAGACCTTGACCACAACTTTG
229EA          ...................................................
229EB          ...................................................
```



Descriptions

Sequences producing significant alignments:

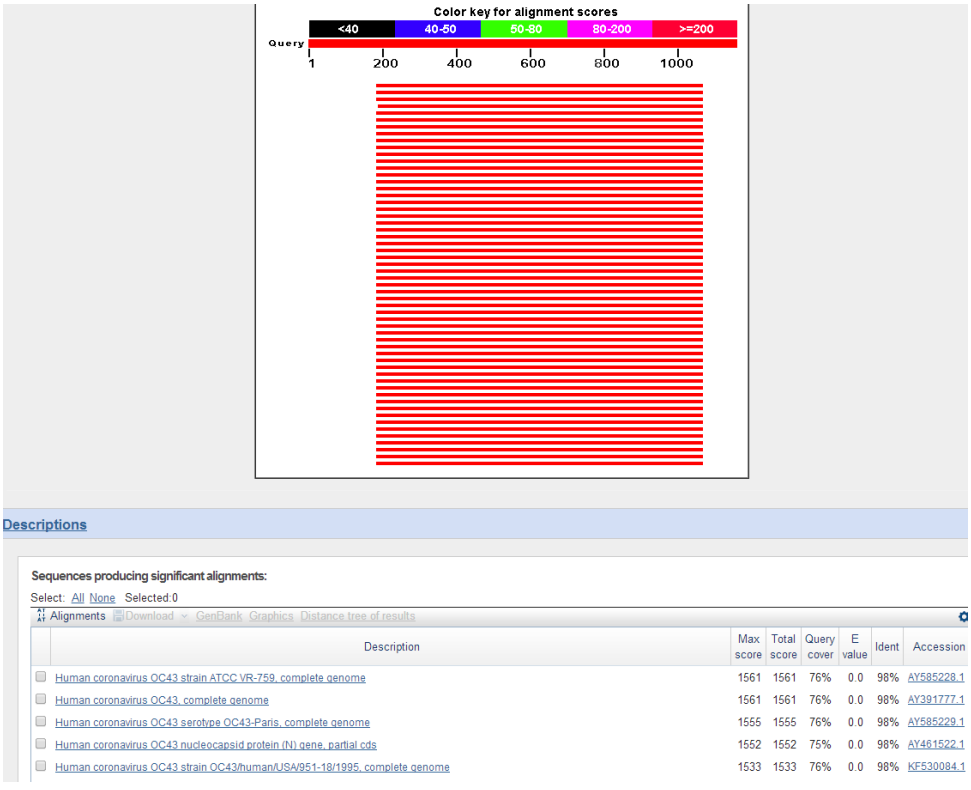
Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

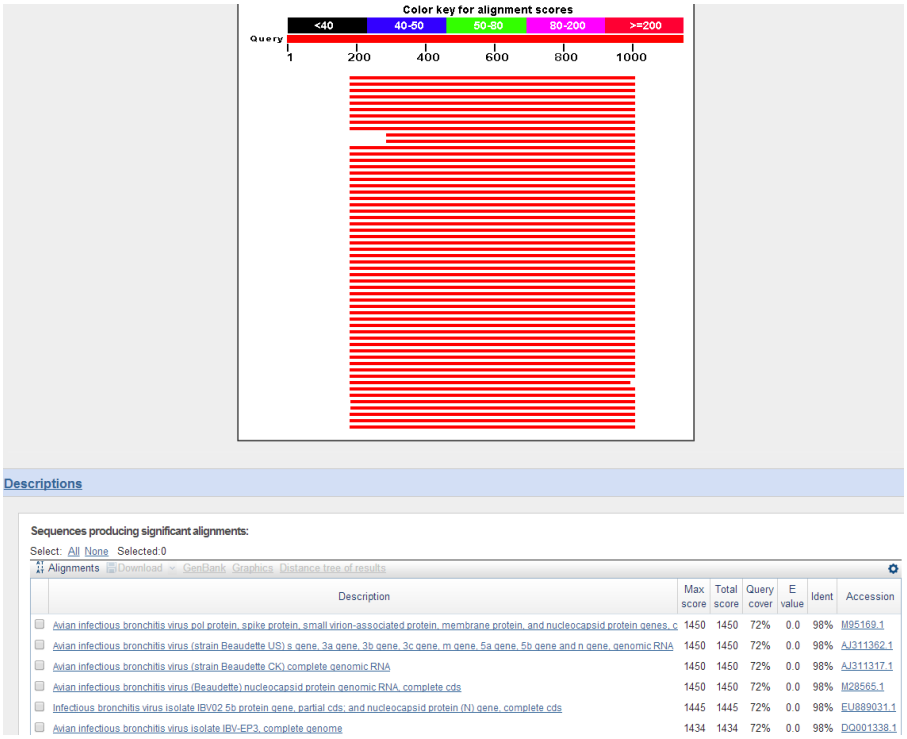
Description	Max score	Total score	Query cover	E value	Ident	Accession
Human coronavirus 229E clone K22-p11, partial genome	1659	1659	78%	0.0	98%	KF293666.1
Human coronavirus 229E clone mock-p11, partial genome	1659	1659	78%	0.0	98%	KF293665.1
Human coronavirus 229E clone p0, partial genome	1659	1659	78%	0.0	98%	KF293664.1
Human coronavirus 229E strain ATCC VR-740 nucleoprotein (N) gene, complete cds	1659	1659	78%	0.0	98%	DQ243939.1
Human coronavirus 229E, complete genome	1659	1659	78%	0.0	98%	AF304460.1

Sequencing data for OC43-CoV samples

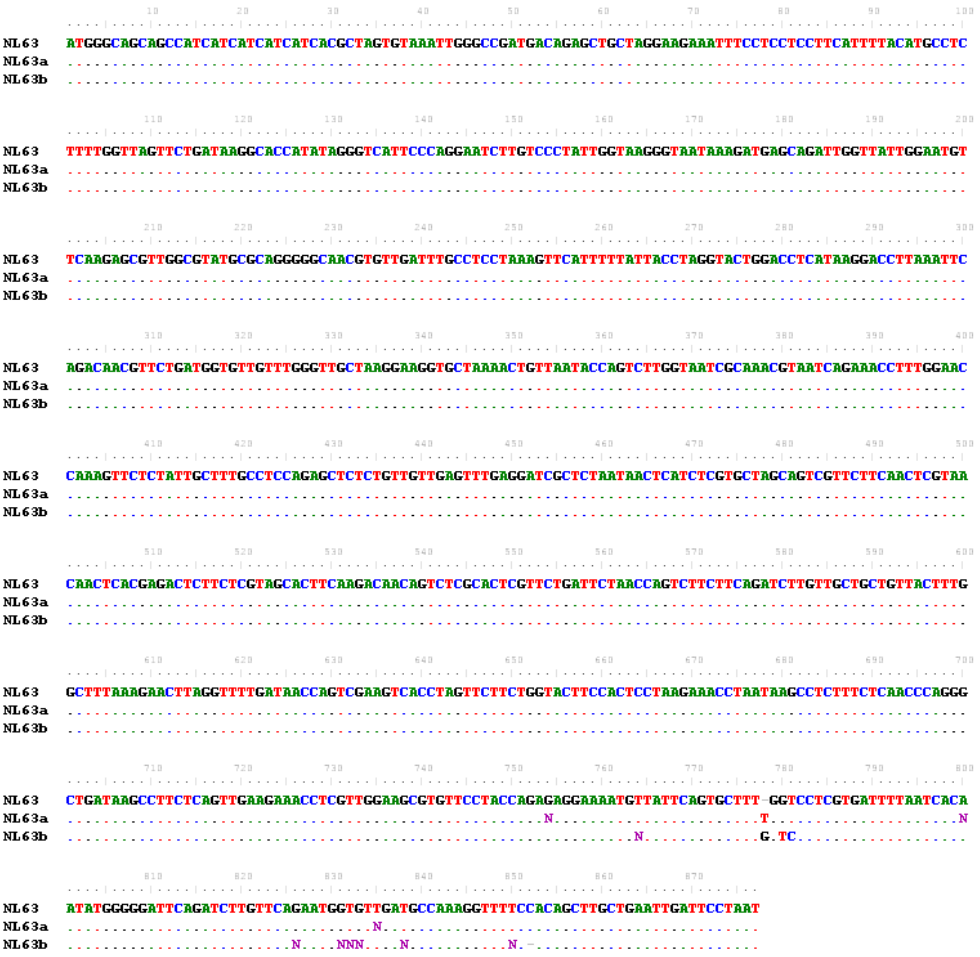
```
                10        20        30        40        50        60        70        80        90        100
                |         |         |         |         |         |         |         |         |         |
OC43ref         ATGGGCAGCAGCCATCATCATCATCATCACTCTTTTACTCCTGGTAAGCAATCCAGTAGTAGAGCGTCCTCTGGAAATCGGTCTGTTAATGGCATCCTCA
OC43a           ...................................................................................T....G..............
OC43b           ...................................................................................T....G..............

                110       120       130       140       150       160       170       180       190       200
                |         |         |         |         |         |         |         |         |         |
OC43ref         AGTGGGCCGATCAGTCCGATCAGTTTTAGAAATGTTCAAACCAGGGGTAGAAGAGCTCAACCCAAGCAAACTGCTACTTCTCAGCAACCATCAGGAGGGAA
OC43a           ..................C................G.....................................C.....................
OC43b           ..................C................G.....................................C.....................

                210       220       230       240       250       260       270       280       290       300
                |         |         |         |         |         |         |         |         |         |
OC43ref         TGTTGTACCCTACTATTCTTGGTTCTCTGGAATTACTCAGTTTCAAAAGGGAAAGGAGTTTGAGTTTGCAGAAGGACAAGGTGTGCCTATTGCACCAGGA
OC43a           ...................................................................T.............C.................
OC43b           ...................................................................T.............C.................

                310       320       330       340       350       360       370       380       390       400
                |         |         |         |         |         |         |         |         |         |
OC43ref         GTCCCAGCTACTGAAGCTAAGGGGTACTGGTACAGACACAACAGACGTTCTTTTAAAACAGCCGATGGCAACCAGCGTCAACTGCTGCCACGATGGTATT
OC43a           ..................................................................................................
OC43b           ..................................................................................................

                410       420       430       440       450       460       470       480       490       500
                |         |         |         |         |         |         |         |         |         |
OC43ref         TTTACTATCTGGGAACAGGACCGCATGCTAAAGACCAGTACGGCACCGATATTGACGGAGTCTACTGGGTCGCTAGCAACCAGGCTGATGTCAATACCCC
OC43a           ..................................................................................................
OC43b           ..................................................................................................

                510       520       530       540       550       560       570       580       590       600
                |         |         |         |         |         |         |         |         |         |
OC43ref         GGCTGACATTGTCGATCGGGACCCAAGTAGCGATGAGGCTATTCCGACTAGGTTTCCGCCTGGCACGGTACTCCCTCAGGGTTACTATATTGAAGGCTCA
OC43a           ..................................................................................................
OC43b           ..................................................................................................

                610       620       630       640       650       660       670       680       690       700
                |         |         |         |         |         |         |         |         |         |
OC43ref         GGAAGGTCTGCTCCTAATTCCAGATCTACTTCGCGCACATCCAGTAGAGCCTCTAGTGCAGGATCGCGTAGTAGAGCCAATTCTGGCAATAGAACCCCTA
OC43a           ..........................................C.......................................................
OC43b           ..........................................C.......................................................

                710       720       730       740       750       760       770       780       790       800
                |         |         |         |         |         |         |         |         |         |
OC43ref         CCTCTGGTGTAACACCTGACATGGCTGATCAAATTGCTAGTCTTGTTCTGGCAAAACTTGGCAAGGATGCCACTAAACCTAAGCAAGTAACTAAGCATAC
OC43a           ..................................................................................C...............
OC43b           ..................................................................................C...............

                810       820       830       840
                |         |         |         |
OC43ref         TGCCAAAGAAGTCAGACAGAAAATTTTGAATAAACCCCGCCAGA
OC43a           ..................................G...........
OC43b           ..................................G...........
```

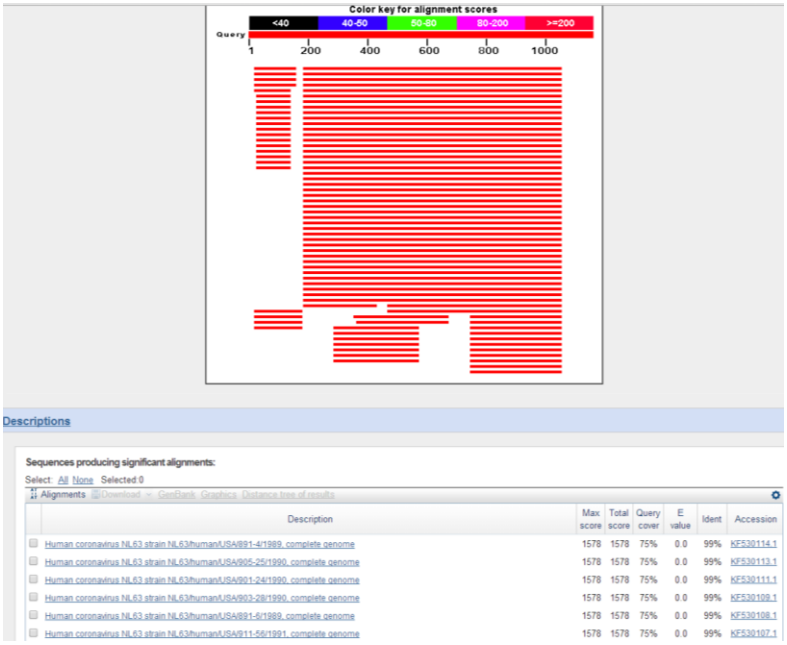


Sequencing data for IBV samples

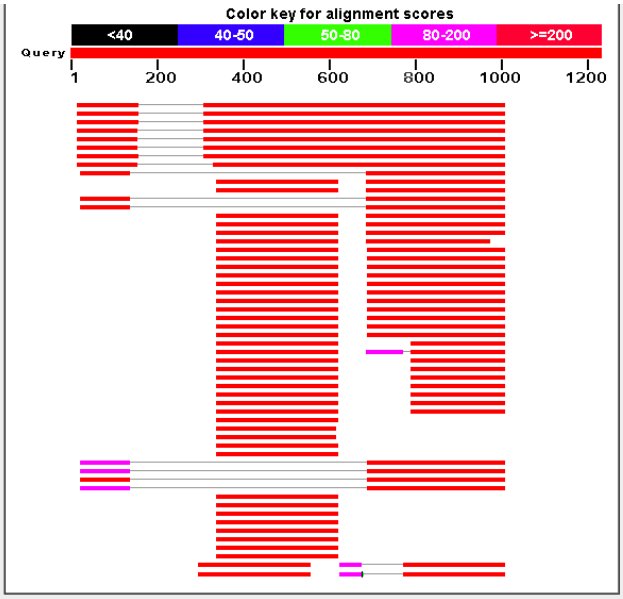


Sequencing data for NL63-CoV samples





Blast search SARS (sequence not shown)



Descriptions

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
Expression vector pOPINVL, complete sequence	1253	1507	68%	0.0	99%	EU733646.1
Expression vector pOPINVH, complete sequence	1253	1507	68%	0.0	99%	EU733645.1
Expression vector pOPING-ET, complete sequence	1253	1507	68%	0.0	99%	EU733644.1
Expression vector pOPINE, complete sequence	1253	1507	68%	0.0	99%	EF372397.1
Expression vector pOPINM, complete sequence	1253	1507	68%	0.0	99%	EF372396.1

Appendix three: A table showing potential SARS-CoV-2 therapeutic agents sourced from (Kaddoura, AlIbrahim et al. 2020)

Compound/drug	**Mechanism of action**	***In vitro* studies**	**Randomized clinical trials**	
			Ongoing	**Complete**
Virus targeting agents				
Remdesivir	Inhibits RNA-dependent RNA polymerase	Yes EC50 = 0.77 - 23.15 μM Vero E6 cells	Yes	Yes
Lopinavir/Ritonavir (Kaletra)	Protease inhibitor with a CYP3A4 inhibitory activity	Yes EC50 = 26.63 μM Vero E6 cells	Yes	Yes
Favipiravir	Inhibits RNA-dependent RNA polymerase	No	Yes	Yes
Ribavirin	Blocks viral RNA synthesis and viral mRNA capping	No	Yes	Yes

Famotidine	Histamine-2 (H2) receptor antagonist $3CL^{pro}$ targeting	No	Yes	No
EIED 2801	Impairs viral replication by incorporating into the genome of the newly formed virions	Yes EC50 = 0.3 μM Vero E6	Yes	No
Oseltamivir	Inhibits neuraminidase enzyme	No	Yes	No
Sofosbuvir	Inhibits RNA-dependent RNA polymerase	No	No	No
Penciclovir	Inhibits viral DNA polymerase	Yes EC50 = 95.96 μM Vero E6	No	No
Azvudine	Inhibiting nucleoside reverse-transcriptase	No	Yes	No
Triazavirin	Inhibits RNA synthesis	No	Yes	No
ACE2 decoy receptor	ACE2 antagonist	Yes 6–100 μg/ml	Yes	No
Host-targeting agents				

Azithromycin	Stimulates the interferon pathway Interferes with virus internalization	Yes	Yes	No
Ivermectin	Impairs nuclear import by interacting with importin (IMP) α/β1 heterodimer	Yes EC50 = 2 µM Vero-hSLAM cells	Yes	No
Nafamostat and Camostat	Inhibits fusion-activation of the virus through inhibition of the host protease (TMPRSS2)	Yes EC50 = 22.50 µM Vero E6 cells	Yes	No
Teicoplanin	Suppresses the entry by blocking the activity of cathepsin L in the late endosome/lysosome	Yes EC50 = 1.66 µM Vero E6 cells	No	No
Nitazoxanide	Blocks viral entry and replication Inhibits the production of pro-inflammatory cytokines	Yes EC50 = 2.12 µM Vero E6 cells	Yes	No
Drugs with mixed action				
Umifenovir	-Inhibits membrane fusion through interacting with the viral glycoproteins - Elevate endosomal pH	Yes EC50 = 4.11 µM Vero E6 cells	Yes	No

Chloroquine phosphate and hydroxychloroquine	-Hinders the auto-immune response -Impairing ACE2 terminal glycosylation-Increasing the endosomal pH	Yes CQ EC50 = 1.13–5.74 μM HCQ EC50 = 0.72 μM Vero E6 cells	Yes	Yes
Adjunctive therapy				
Immunomodulatory agents				
Fingolimod	Targets sphingosine-1-phosphate (S1P) receptors and alters the signaling of the S1P pathway	No	Yes	No
Thymosin α1	Triggers lymphocyte maturation Enhances T cell activation	No	Yes	No
Tocilizumab	Recombinant anti-human interleukin-6 receptor (IL-6R) monoclonal antibody	No	Yes	No
Bevacizumab	Humanized monoclonal antibody against the angiogenic vascular endothelial growth factor (VEGF)	No	Yes	No
Colchicine	Down-regulates multiple inflammatory pathways through tubulin disruption	No	Yes	Yes

	Inhibits microtubule-dependent chemotaxis of neutrophils, generation of leukotrienes and cytokines, phagocytosis, and the (TNF-α)-induced NF-κB pathway			
Methylprednisolone	Anti-inflammatory properties at high doses	Not applicable	Yes	No
Dexamethasone	Anti-inflammatory properties	Not applicable	Yes	Yes
Convalescent plasma	Provides passive immunization	Not applicable	Yes	No

Appendix four: Summary of COVID-19 therapeutics that have completed clinical trials, sourced from (Kaddoura, Allbrahim et al. 2020)

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
Remdesivir	Intravenous	200 mg on day 1 followed by 100 mg on days 2–10 in single daily infusions	Randomized, double-blind, placebo-controlled, multicenter trial	Moderate	Aim: To evaluate the safety and efficacy of remdesivir in hospitalized adults infected with SARS-CoV-2. Key findings: No significant difference in improvement time. Adverse effects: Same as placebo.
Remdesivir	Intravenous	200 mg on the first day followed by a 100-mg once daily maintenance dose for up to a 10 days	Multicenter, adaptive, randomized, double-blind, placebo-controlled trial	Strong	Aim: To evaluate the safety and efficacy of remdesivir in hospitalized adults infected with SARS-CoV-2. Key findings: Patients receiving remdesivir recovered faster than those treated with placebo

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
					(median recovery time of 11 days and 15 days, respectively). The risk of death by 14 days was less in the remdesivir group compared with the placebo one; 7.1% and 11.9% respectively. Adverse effects: Same as placebo.
Lopinavir/ritonavir	oral	400 mg/100 mg twice daily for 14 days	Randomized, controlled, open-label trial	Moderate	Aim: To study the efficacy of lopinavir/ritonavir in hospitalized adult patients with severe COVID-19. Key findings: No clinical benefit. Adverse effects: Greater than placebo primarily gastrointestinal side effects.

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
Favipiravir or umifenovir	Oral	Favipiravir 1,600 mg twice daily 1st day then 600 mg twice daily for 10 days vs. umifenovir 200 mg three times daily for 10 days	Prospective, randomized, controlled, open-label.	Moderate	Aim: To evaluate the clinical efficacy and safety of favipiravir versus umifenovir as a treatment for COVID-19. Key findings: No improvement in clinical recovery at day 7. Improved the time to relief for pyrexia and cough compared to umifenovir. Adverse effects: raised serum uric acid was more frequently observed in favipiravir group.
Favipiravir	Oral	Favipiravir 1,600 mg twice daily 1st day then 600 mg twice daily for 14 days + 5mIU of IFN-α twice daily vs. lopinavir/ritonavir 400	Open-label nonrandomized-comparative controlled study	Weak	Aim: To examine the efficacy of favipiravir versus lopinavir/ritonavir for the treatment of COVID-19. Key Findings: FPV showed better therapeutic responses than LPV/RTV.

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
		mg/100 mg twice daily + 5 mIU of IFN-α twice daily for 14 days			Adverse effects: Generally mild but less common in the favipiravir treated group.
Ribavirin	Oral	400mg twice daily for 14 days	Prospective, randomized, controlled, open-label trial	Moderate	Aim: To evaluate the safety and efficacy of IFN-ß-1b, lopinavir/ritonavir, and ribavirin combination. Key findings: Recovery was accelerated, viral load was suppressed, hospitalization was shortened and mortality was reduced after the combination of lopinavir/ritonavir, ribavirin, and IFN-ß-1b compared with to lopinavir/ritonavir alone (control). Adverse effects: Same as placebo.

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
Umifenovir or lopinavir/ritinovir	Oral	Lopinavir 200 mg plus ritonavir 500mg twice daily for 7–14 days vs. umifenovir 200 mg three times daily for 7–14 days	Open-label randomized controlled trial	Moderate	Aim: To explore the efficacy and safety of lopinavir/ritonavir or umifenovir monotherapy for the treatment of patients hospitalized with mild/moderate COVID-19. Key findings: lopinavir/ritonavir or umifenovir monotherapy offered minimal added benefit compared to standard of care. Adverse events: Greater than control with diarrhea being most common.
HCQ/AZM	Oral	HCQ 200 mg three times daily for 10 days Azithromycin 500 mg on day	Open-label non-randomized clinical trial	Weak	Aim: To investigate the efficacy of HCQ in COVID-19 patients and the role of adding AZM Key findings: Significant reduction in viral load in patients receiving HCQ alone. 100% recovery in patients receiving a combination of AZM and

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
		1 followed by 250 mg daily for four consecutive days			HCQ. Adverse effects: Not described.
HCQ	Oral	800 mg as a first dose, followed by 600 mg after 6 to 8 h, then 600 mg daily for 4 days	Randomized, double-blind, placebo-controlled trial	Strong	Aim: To assess HCQ as post-exposure (within 4 days of exposure) prophylaxis for COVID-19 Key findings: HCQ did not prevent laboratory confirmed infection or COVID-19 like illness compared to placebo Adverse effects: Greater than placebo but not serious
HCQ	Oral	800 mg on the first day, followed by 400 mg once daily for 6 days	Multicenter, open label, randomized controlled trail	Moderate	Aim: To evaluate the efficacy of early administration of HCQ in non-hospitalized adults with mild COVID-19 Key findings: No significant difference in viral

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
					load reduction, risk of hospitalization, and clinical recovery compared to standard care. Adverse effects: Same as placebo
HCQ	Oral	800 mg as a first dose followed by 600 mg after 6 to 8 h, then 600 mg daily for 4 days	Randomized, double blinded, placebo controlled trial	Strong	Aim: To assess the efficacy of HCQ in decreasing the disease severity in adult outpatients with early, mild COVID-19 Key findings: No significant decrease in the severity of symptoms compared to placebo Adverse effects: Greater than placebo- none were serious
Colchicine	Oral	1.5-mg loading dose followed by 0.5 mg after 60 min	Prospective, open-label, randomized clinical trial	Moderate	Aim: To evaluate the effect of treatment with colchicine on cardiac and inflammatory biomarkers and clinical outcomes in COVID-19

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
		and maintenance doses of 0.5 mg twice daily for 3 weeks			hospitalized patients Key findings: Improved time to clinical deterioration. Adverse effects: Similar to the control group except for diarrhea being more frequent with colchicine.
Dexamethasone	Oral or intravenous	6 mg once daily for 10 days	Randomized, controlled, open-label, adaptive, platform trial	Moderate	Aim: To evaluate dexamethasone in hospitalized COVID-19 patients. Key findings: Dexamethasone reduced mortality among those receiving invasive mechanical ventilation or oxygen but not among milder cases. Adverse effects: Not described.

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
Convalescent Plasma	Intravenous	200 – 500 mL	Non-randomized	Weak	Aim: To investigate the safety of convalescent plasma treatment in hospitalized COVID-19 patients Key findings: Transfusion of convalescent plasma is safe. Adverse effects: Frequency of <1% of all transfusions. Include transfusion-associated circulatory overload (TACO), transfusion-related acute lung injury (TRALI), severe allergic transfusion reaction, and death.
Convalescent Plasma	Intravenous	4 to 13 ml/kg	Open-label, multicenter, randomized clinical trial	Moderate	Aim: To evaluate the efficacy and adverse effects of convalescent plasma therapy for patients with COVID-19. Key findings: No statistically significant

Compound	Route	Dosing regimen	Study design	Evidence type[#]	Summary
					difference in clinical improvement within 28 days. Negative viral PCR conversion rate was significantly higher in the convalescent plasma group. Adverse effects: Reported in two patients. Included chills and rashes in one patient and shortness of breath, cyanosis, and severe dyspnea within 6 h of transfusion in another one.

Appendix five: Mass spectrometry results

MALDI mass fingerprint of tryptic digest on Micromass MALDImicroMX. Raw data, centred spectrum, externally calibrated
Spectrum then lockmass-calibrated, de-isotoped and the resulting
values sent to Mascot search engine; these may differ slightly from the mass values shown in this spectrum

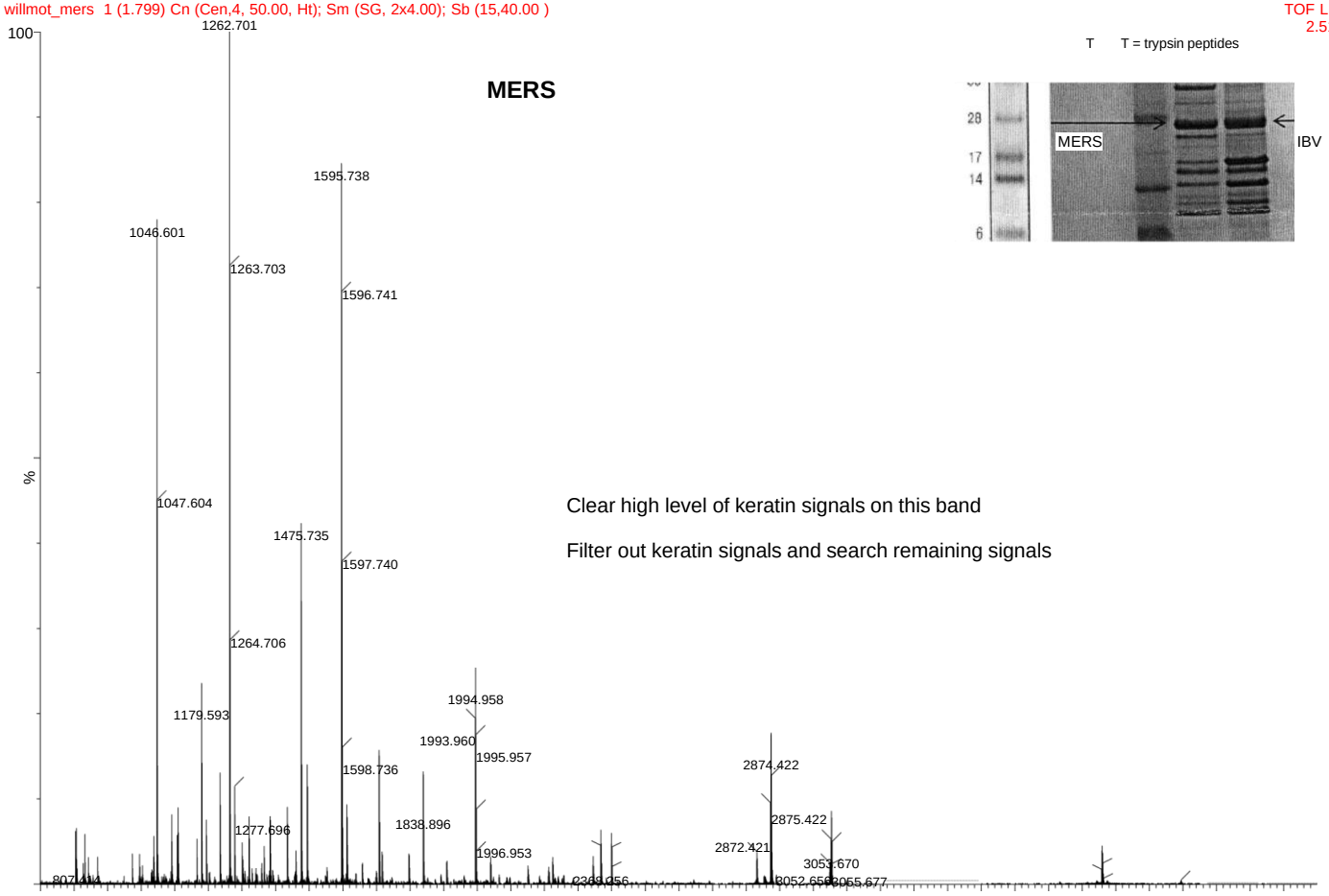
C18zt desalt
CHCA matrix 120pwr, 2080pv
external calibration
int lockmass calib to 2211.104 tryptic peptide

03-Jul-2015

0.00000000

willmot_mers 1 (1.799) Cn (Cen,4, 50.00, Ht); Sm (SG, 2x4.00); Sb (15,40.00)

TOF LD+
2.51e4



T T = trypsin peptides

Clear high level of keratin signals on this band

Filter out keratin signals and search remaining signals

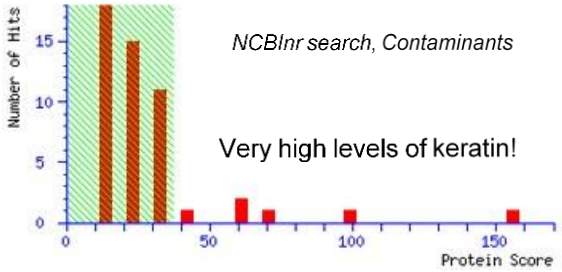
```
Search title     : Willmot MERS psdmx
MS data file     : willmot_mers_psdmx.pkl
Database         : contaminants 20090624 (262 sequences; 133770 residues)
Timestamp        : 6 Jul 2015 at 11:47:03 GMT

Top Score        : 156 for 00009865.2, Tax_Id=9606 Gene_Symbol=KRT10 Keratin, type I cytoskeletal 10
```

http://www.matrixscience.com/cgi/master_results.pl?file=.%2Fdata%2F20150706%2FFTTtiaHnO.dat&REPTYPE=protein&_sigthreshold=0.05&REPORT=AUTO&percolate=0&_minpeplen=7&_server_mudpit_switch=0.000000001&_ignoreionsscorebelow=0&_showsubsets=0&_showpopups=TRUE&_sortunassigned=scoredown&_requireboldred=0&sessionID=guest_guestsession

Mascot Score Histogram

MERS psdmx analysis

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 37 are significant (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



NCBInr search, Contaminants

Components 1-3

Very high levels of keratin!

	Accession	Mass	Score	Description
1.	00009865.2	59703	156	Tax_Id=9606 Gene_Symbol=KRT10 Keratin, type I cytoskeletal 10
2.	00220327.3	66149	96	Tax_Id=9606 Gene_Symbol=KRT1 Keratin, type II cytoskeletal 1
3.	00021304.1	66110	69	Tax_Id=9606 Gene_Symbol=KRT2 Keratin, type II cytoskeletal 2 epidermal
4.	00300725.7	60293	60	Tax_Id=9606 Gene_Symbol=KRT6A Keratin, type II cytoskeletal 6A

00009865.2 Mass: 59703 Score: 156 Expect: 6.6e-14 Matches: 32
Tax Id=9606 Gene Symbol=KRT10 Keratin, type I cytoskeletal 10

00220327.3 Mass: 66149 Score: 94 Expect: 1e-07 Matches: 20
Tax Id=9606 Gene Symbol=KRT1 Keratin, type II cytoskeletal 1

00021304.1 Mass: 66110 Score: 82 Expect: 1.5e-06 Matches: 18
Tax_Id=9606 Gene_Symbol=KRT2 Keratin, type II cytoskeletal 2 epidermal

Match keratin signals and search remaining signals for additional component(s) →

Search title : Willmot MERS psdmx
Database : NCBInr 20150627 (68461039 sequences; 24550204531 residues)
Taxonomy : Escherichia coli (1469464 sequences)
Timestamp : 6 Jul 2015 at 11:50:44 GMT

Top Score : 156 for gi|585370710, peptidylprolyl isomerase, partial [Escherichia coli]

http://www.matrixscience.com/cgi/master_results.pl?file=..%2Fdata%2F20150706%2FFTTtiaHtT.dat&REPTYPE=protein&_sigthreshold=0.05&REPORT=AUTO&_minpeplen=7&_server_mudpit_switch=99999999&_ignoreionsscorebelow=0&_showsubsets=0&_showpopups=TRUE&_sortunassigned=scoredown&_requireboldred=0&_prefertaxonomy=0&sessionID=guest_guestsession

Mascot Score Histogram

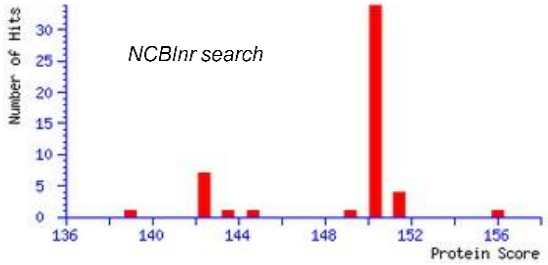
MERS psdmx analysis

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 74 are significant (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

peptidylprolyl isomerase, partial [Escherichia coli]

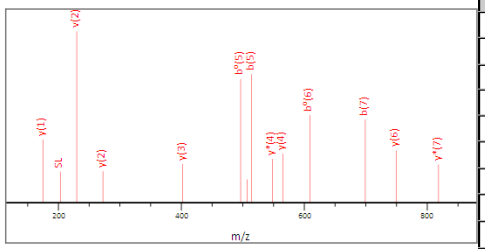
```
  1 MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS LISGLETALE
 51 GHEVGDKFDV AVGANDAYGQ YDENLVQRVP KDVFMGVDEL QVGMRFLAET
101 DQGPVPVEIT AVEDDHVVVD GNHMLAGQNL KFNVEVVAIR E
```

RMS error 19 ppm

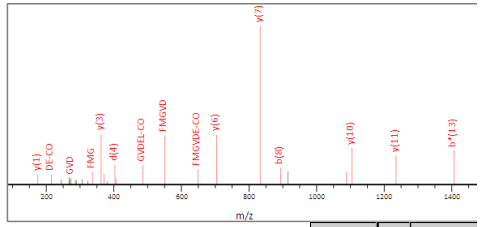


NCBInr search

Component 4



b	Seq.	y
116.0342	D	
229.1183	L	1147.6834
328.1867	V	1034.5993
427.2551	V	935.5309
514.2871	S	836.4625
627.3712	L	*749.4305*
698.4083	A	636.3464
861.4716	Y	*565.3093*
989.5302	Q	*402.2459*
1088.5986	V	*274.1874*
	R	*175.1190*



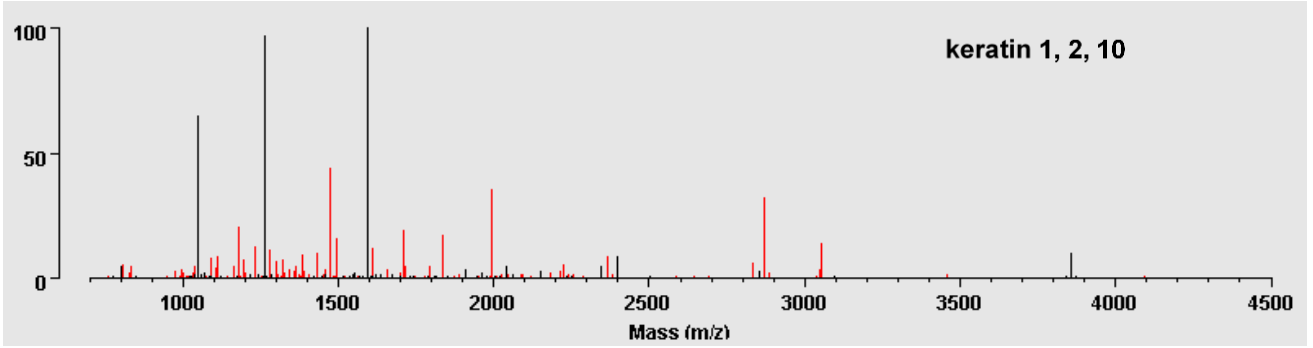
b	Seq.	y
116.0342	D	
215.1026	V	1480.7287
362.1710	F	1381.6603
493.2115	M	*1234.5919*
550.2330	G	*1103.5514*
649.3014	V	1046.5299
764.3284	D	947.4615
893.3709	E	*832.4346*
1006.4550	L	*703.3920*
1134.5136	Q	590.3079
1233.5820	V	462.2493
1290.6035	G	*363.1809*
1421.6440	M	*306.1594*
	R	*175.1190*

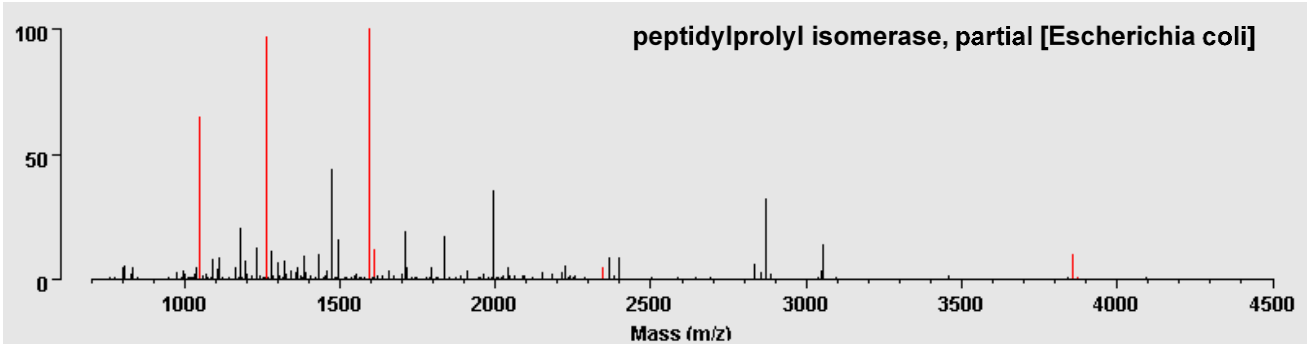
psdmx fragment match

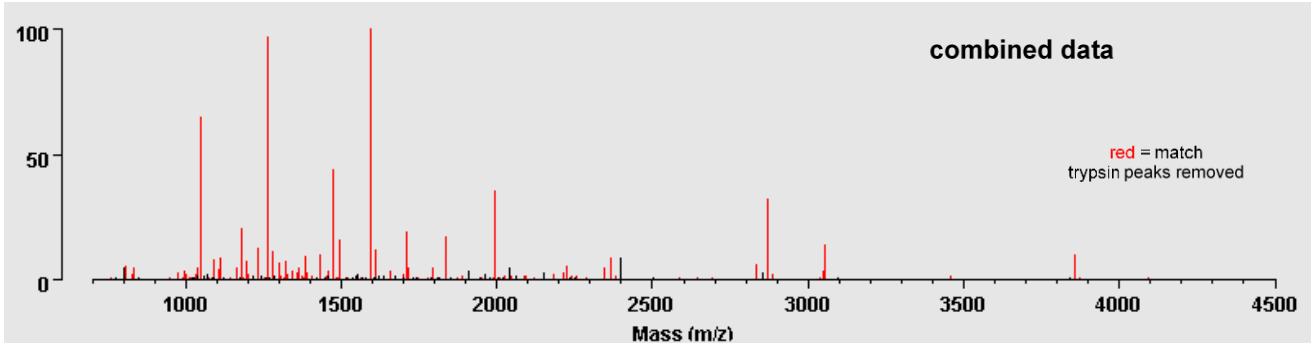
1. gi|585370710 Mass: 15265 Score: 156 Expect: 3.7e-10 Matches: 6
peptidylprolyl isomerase, partial [Escherichia coli]

Observed	Mr(expt)	Mr(calc)	ppm	Start	End	Miss	Ions	Peptide
1046.5991	1045.5918	1045.5920	-0.19	132	140	0	13	K.FNVEVVAIR.E
1262.6893	1261.6820	1261.7030	-16.65	6	16	0	47	K.DLVVSLAYQVR.T
1595.7361	1594.7288	1594.7484	-12.25	82	95	0	60	K.DVFMGVDELQVGMR.F
2344.0700	2343.0627	2343.0768	-5.99	58	78	0	---	K.FDVAVGANDAYGQYDENLVQR.V
3857.9920	3856.9847	3856.8942	23.5	96	131	0	---	R.FLAETDQGPVPVEITAVEDDHVVVDGNHMLAGQNLK.F
3874.0340	3873.0267	3872.8891	35.5	96	131	0	---	R.FLAETDQGPVPVEITAVEDDHVVVDGNHMLAGQNLK.F + Oxidation (M)

MERS - relative signals







MALDI mass fingerprint of tryptic digest on Micromass MALDImicroMX. Raw data, centred spectrum, externally calibrated
Spectrum then lockmass-calibrated, de-isotoped and the resulting
values sent to Mascot search engine; these may differ slightly from the mass values shown in this spectrum

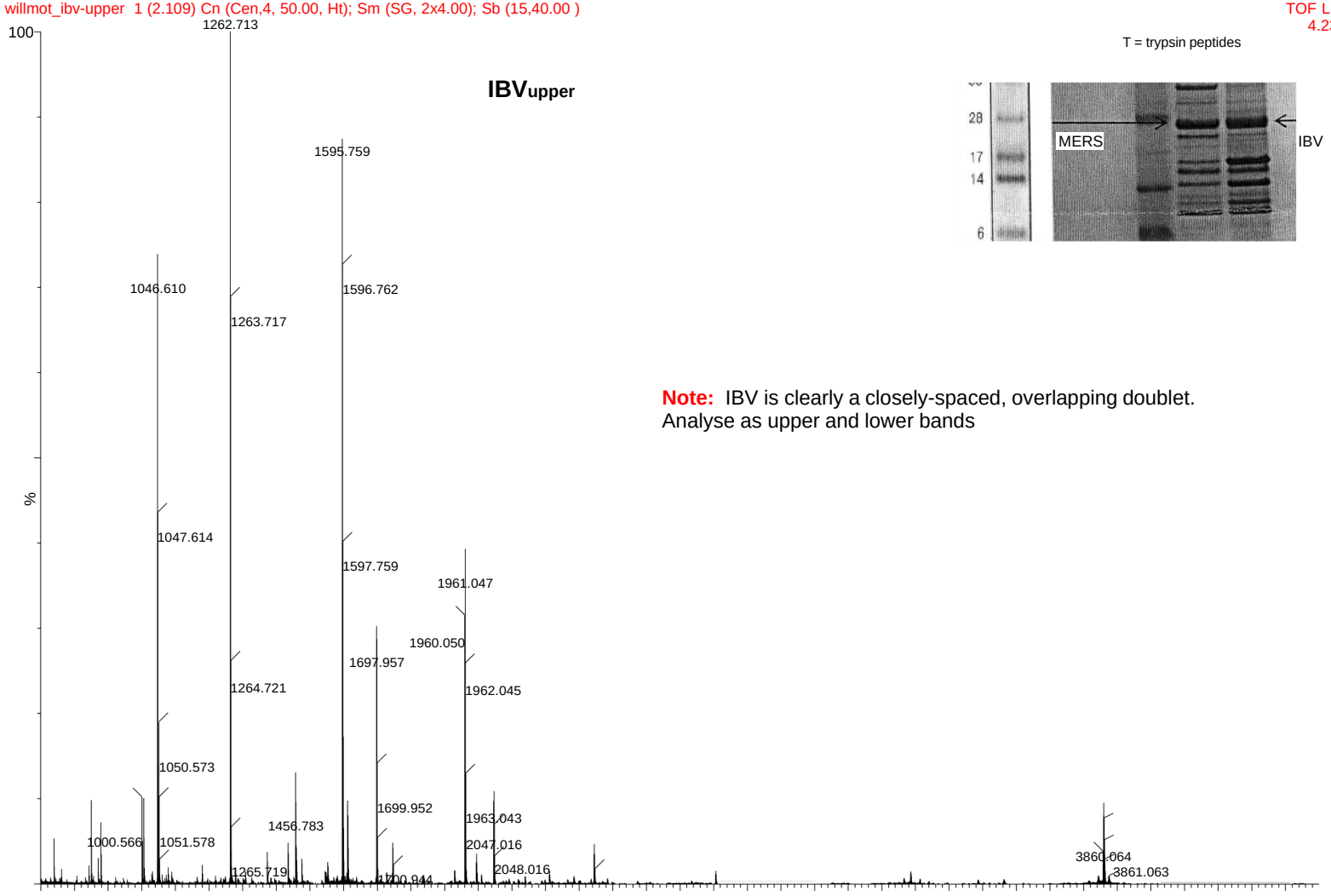
C18zt desalt
CHCA matrix 120pwr, 2080pv
external calibration
int lockmass calib to 2211.104 tryptic peptide

0.00000000

03-Jul-2015

willmot_ibv-upper 1 (2.109) Cn (Cen,4, 50.00, Ht); Sm (SG, 2x4.00); Sb (15,40.00)

TOF LD+
4.23e4



Note: IBV is clearly a closely-spaced, overlapping doublet.
Analyse as upper and lower bands

Search title : Willmot IBV upper
Database : NCBInr 20150627 (68461039 sequences; 24550204531 residues)
Taxonomy : Escherichia coli (1469464 sequences)
Timestamp : 6 Jul 2015 at 10:54:45 GMT
Top Score : 180 for Mixture 1, gi|607508696 + gi|585370710

http://www.matrixscience.com/cgi/master_results.pl?file=..%2Fdata%2F20150706%2FFTTtiaTmm.dat;sessionID=guest_guestsession

	NCBInr	Decoy
Protein hits above identity threshold	7	0
Highest scoring protein hit	113	61

IBVupper

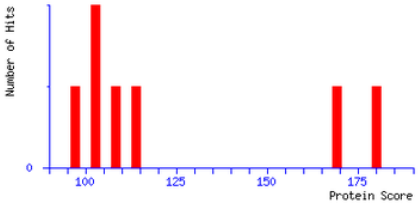
30S ribosomal protein S4, partial [Escherichia coli]

```
  1 GVLERQFRNY YKEAARLKGN TGENLLALLE GRLDNVVYRM GFGATRAEAR
 51 QLVSHKAIMV NGRVVNIASY QVSPNDVVSI REKAKKQSRV KAALELAEQR
101 EKPTWLEVDA GKMEGTFKRK PERSDLSADI NEHLIVELYS K
```

RMS error 15 ppm

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 74 are significant (p<0.05).



peptidylprolyl isomerase, partial [Escherichia coli]

```
  1 MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS LISGLETALE
 51 GHEVGDKFDV AVGANDAYGQ YDENLVQRVP KDVFMGVDEL QVGMRFLAET
101 DQGPVPVEIT AVEDDHVVVD GNHMLAGQNL KFNVEVVAIR E
```

RMS error 17 ppm

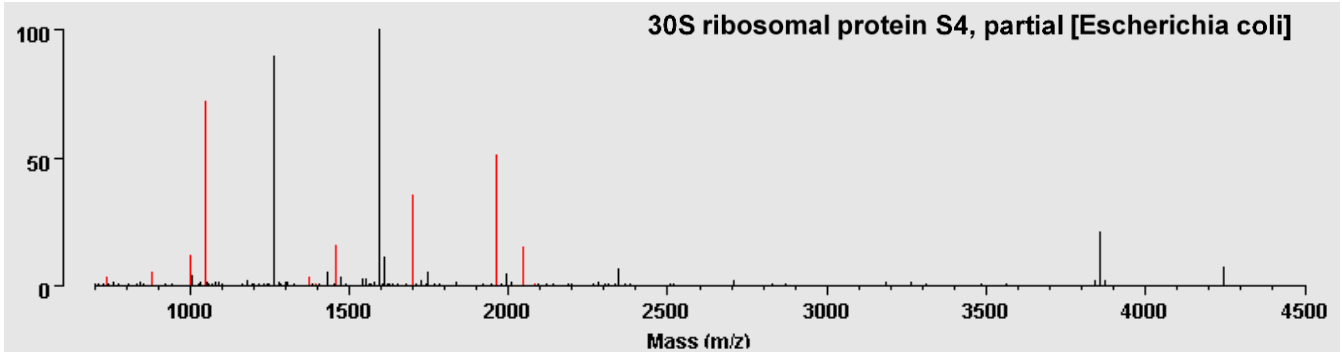
1. Mixture 1 Total score: 180 Expect: 1.5e-12 Matches: 16

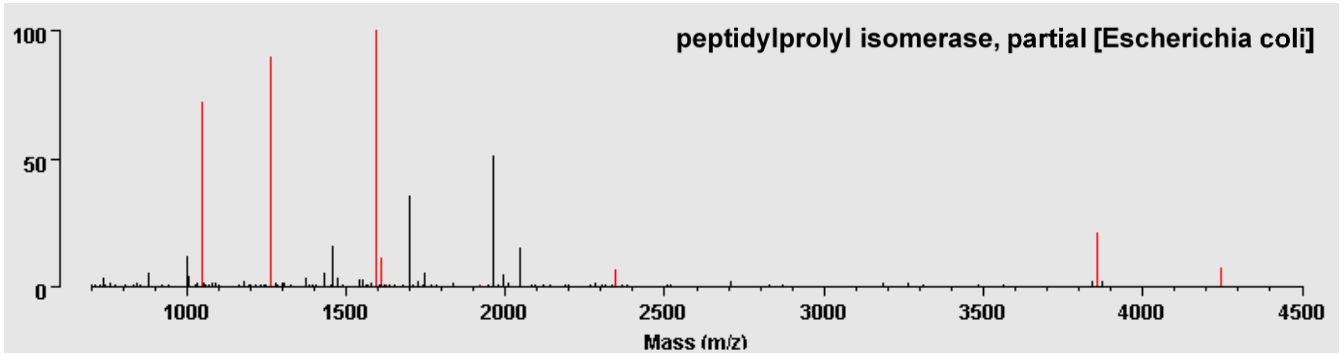
Components: 1. gi|607508696 30S ribosomal protein S4, partial [Escherichia coli O157:H7 str. K2324]
2. gi|585370710 peptidylprolyl isomerase, partial [Escherichia coli]

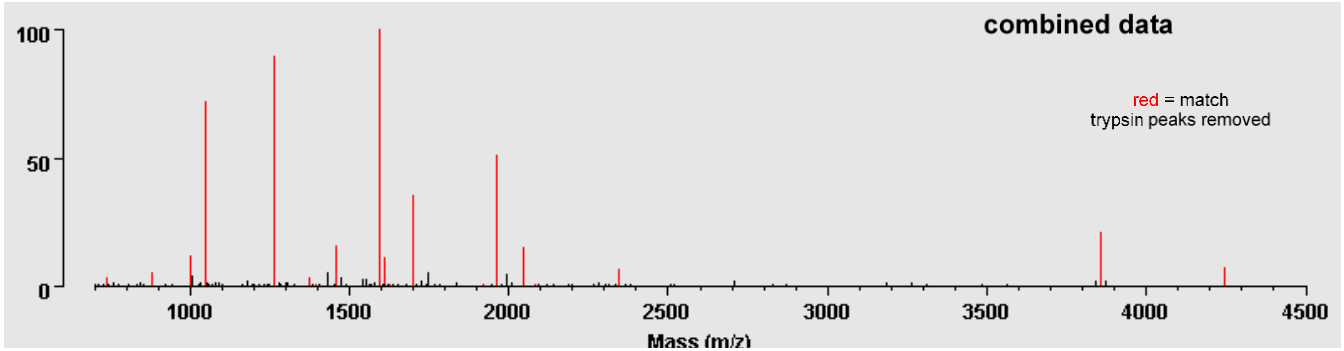
Observed	Mr(expt)	Mr(calc)	ppm	Start	End	Miss	Comp	Peptide
739.3760	738.3687	738.3483	27.7	40	46	0	1	R.MGFGATR.A
878.4860	877.4787	877.4658	14.8	33	39	0	1	R.LDNVVYR.M
1000.5650	999.5577	999.5349	22.9	92	100	0	1	K.AALELAEQR.E
1046.6090	1045.6017	1045.5920	9.28	132	140	0	2	K.FNVEVVAIR.E
1262.7110	1261.7037	1261.7030	0.55	6	16	0	2	K.DLVVSLAYQVR.T
1372.6910	1371.6837	1371.7034	-14.36	101	112	1	1	R.EKPTWLEVDAGK.M
1456.7780	1455.7707	1455.7681	1.78	19	32	0	1	K.GNTGENLLALLEGR.L
1595.7550	1594.7477	1594.7484	-0.40	82	95	0	2	K.DVFMGVDELQVGMR.F
1611.7350	1610.7277	1610.7433	-9.66	82	95	0	2	K.DVFMGVDELQVGMR.F + Oxidation (M)
1697.9510	1696.9437	1696.9471	-2.02	17	32	1	1	R.LKGNTGENLLALLEGR.L
1960.0400	1959.0327	1959.0425	-5.01	64	81	0	1	R.VVNIASYQVSPNDVVSIR.E
2046.0070	2044.9997	2045.0317	-15.62	124	141	0	1	R.SDLSADINEHLIVELYSK.-
2344.0990	2343.0917	2343.0768	6.39	58	78	0	2	K.FDVAVGANDAYGQYDENLVQR.V
3858.0070	3856.9997	3856.8942	27.4	96	131	0	2	R.FLAETDQGPVPVEITAVEDDHVVVDGNHMLAGQNLK.F
4243.0000	4241.9927	4242.0605	-15.97	17	57	0	2	R.TEDGVLVDESPVSAPLDYLHGHGSLISGLETALEGHEVGDK.F
4243.2160	4242.2087	4242.0605	35.0	17	57	0	2	R.TEDGVLVDESPVSAPLDYLHGHGSLISGLETALEGHEVGDK.F

No match to: 1006.5900, 1031.6010, 1052.5980, 1078.5930, 1088.6100, 1179.6040, 1278.7120, 1307.6820, 1434.6940, 1475.7420, 1544.7690, 1552.7710, 1727.8550, 1745.9280, 1835.9520, 1993.9840, 2705.2050, 3842.1430, 3874.0570

IBV$_{upper}$ - relative signals







MALDI mass fingerprint of tryptic digest on Micromass MALDImicroMX. Raw data, centred spectrum, externally calibrated
Spectrum then lockmass-calibrated, de-isotoped and the resulting
values sent to Mascot search engine; these may differ slightly from the mass values shown in this spectrum

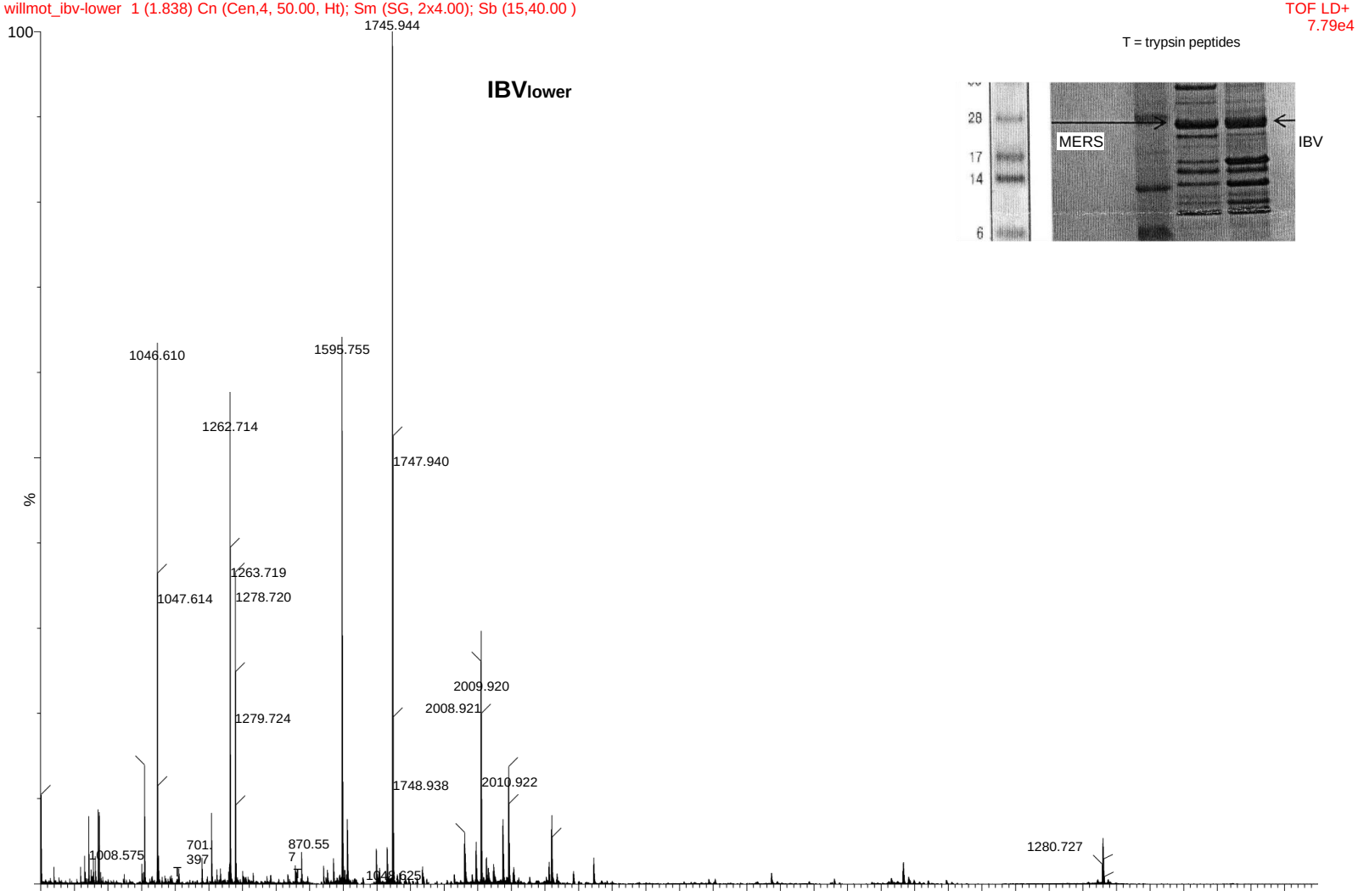
C18zt desalt
CHCA matrix 120pwr, 2080pv
external calibration
int lockmass calib to 2211.104 tryptic peptide

0.00000000

03-Jul-2015

willmot_ibv-lower 1 (1.838) Cn (Cen,4, 50.00, Ht); Sm (SG, 2x4.00); Sb (15,40.00)

TOF LD+
7.79e4



T = trypsin peptides

Search title : Willmot IBV lower
Database : NCBInr 20150627 (68461039 sequences; 24550204531 residues)
Taxonomy : Escherichia coli (1469464 sequences)
Timestamp : 6 Jul 2015 at 11:14:03 GMT
Top Score : 180 for gi|447062357, MULTISPECIES: GTP cyclohydrolase 1 [Enterobacteriaceae]

http://www.matrixscience.com/cgi/master_results.pl?file=..%2Fdata%2F20150706%2FFTTtiaTnh.dat;sessionID=guest_guestsession

	NCBInr	Decoy
Protein hits above identity threshold	1	0
Highest scoring protein hit	180	55

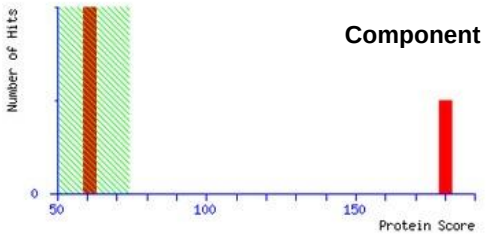
IBVlower

```
  1 MPSLSKEAAL VHEALVARGL ETPLRPFVHE MDNETRKSLI AGHMTEIMQL
 51 LNLDLADDSL METPHRIAKM YVDEIFSGLD YANFPKITLI ENKMKVDEMV
101 TVRDITLTST CEHHFVTIDG KATVAYIPKD SVIGLSKINR IVQFFAQRPQ
151 VQERLTQQIL IALQTLLGTN NVAVSIDAVH YCVKARGIRD ATSATTTTSL
201 GGLFKSSQNT RHEFLRAVRH HN
```

RMS error 14 ppm

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 74 are significant (p<0.05).



Component 1

1. gi|447062357 Mass: 24929 Score: 180 Expect: 1.5e-12 Matches: 15
MULTISPECIES: GTP cyclohydrolase 1 [Enterobacteriaceae]

Observed	Mr(expt)	Mr(calc)	ppm	Start	End	Miss	Peptide
701.3960	700.3887	700.3656	33.0	212	216	0	R.HEFLR.A
818.4770	817.4697	817.4545	18.6	130	137	0	K.DSVIGLSK.I
830.5090	829.5017	829.4909	13.1	87	93	0	K.ITLIENK.M
862.5110	861.5037	861.4960	8.98	122	129	0	K.ATVAYIPK.D
948.4970	947.4897	947.4746	15.9	96	103	0	K.VDEMVTVR.D
1008.5750	1007.5677	1007.5553	12.4	141	148	0	R.IVQFFAQR.P
1207.6250	1206.6177	1206.6101	6.35	94	103	1	K.MKVDEMVTVR.D
1278.7180	1277.7107	1277.7092	1.23	7	18	0	K.EAALVHEALVAR.G
1570.7700	1569.7627	1569.7886	-16.51	190	205	0	R.DATSATTTTSLGGLFK.S
1745.9380	1744.9307	1744.9373	-3.76	141	154	1	R.IVQFFAQRPQVQER.L
2008.9130	2007.9057	2007.9288	-11.49	70	86	0	K.MYVDEIFSGLDYANFPK.I
2024.9060	2023.8987	2023.9237	-12.35	70	86	0	K.MYVDEIFSGLDYANFPK.I + Oxidation (M)
2073.9830	2072.9757	2072.9837	-3.86	104	121	0	R.DITLTSTCEHHFVTIDGK.A
2091.0190	2090.0117	2090.0215	-4.66	19	36	1	R.GLETPLRPFVHEMDNETR.K
3264.6780	3263.6707	3263.5988	22.0	38	66	0	K.SLIAGHMTEIMQLLNLDLADDSLMETPHR.I

No match to: 878.4880, 1000.5580, 1046.6090, 1179.5990, 1262.7120, 1300.6180, 1456.7690, 1475.7490, 1540.8010, 1588.8640, 1595.7500, 1603.7680, 1611.7350, 1697.9430, 1707.7820, 1730.1660, 1835.0060, 1960.0400, 1993.9720, 2045.9800, 2106.0140, 2219.1170, 2344.0930, 2872.4350, 3857.9960

Re-search unmatched for additional component(s) →

Search title : Willmot IBV lower
Database : NCBInr 20150627 (68461039 sequences; 24550204531 residues)
Taxonomy : Escherichia coli (1469464 sequences)
Timestamp : 6 Jul 2015 at 11:15:44 GMT
Top Score : 100 for Mixture 1, gi|585370710 + gi|378021461

http://www.matrixscience.com/cgi/master_results.pl?file=..%2Fdata%2F20150706%2FFTTtiaTtL.dat;sessionID=guest_guestsession

	NCBInr	Decoy
Protein hits above identity threshold	3	0
Highest scoring protein hit	79	61

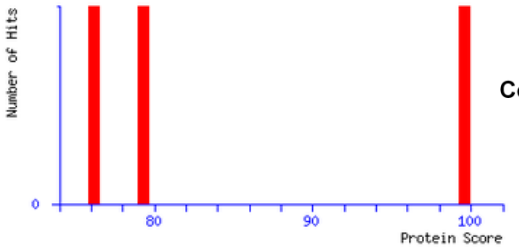
IBVlower

peptidylprolyl isomerase, partial [Escherichia coli]

```
  1 MKVAKDLVVS LAYQVRTEDG VLVDESPVSA PLDYLHGHGS LISGLETALE
 51 GHEVGDKFDV AVGANDAYGQ YDENLVQRVP KDVFMGVDEL QVGMRFLAET
101 DQGPVPVEIT AVEDDHVVVD GNHMLAGQNL KFNVEVVAIR E
                                              RMS error 11 ppm
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 74 are significant (p<0.05).



Components 2 & 3

ribosomal protein S4 [Escherichia coli DEC7C]

```
  1 MGPKLKLSRR EGTDLFLKSG VRAIDTKCKI EQALGQHGAR KPRLSDYGVQ
 51 LREKQKVRRI YGVLERQFRN YYKEAARLKG NTGENLLALL EGRLDNVVYR
101 MGFGATRAEA RQLVSHKAIM VNGRVVNIAS YQVSPNDVVS IREKAKKQSR
151 VKAALELAEQ REKPTWLEVD AGKMEGTFKR KPERSDLSAD INEHLIVELY
201 SK
                                              RMS error 16 ppm
```

1. Mixture 1 Total score: 100 Expect: 0.00016 Matches: 13
Components: 1. gi|585370710 peptidylprolyl isomerase, partial [Escherichia coli]
2. gi|378021461 ribosomal protein S4 [Escherichia coli DEC7C]

Observed	Mr(expt)	Mr(calc)	ppm	Start	End	Miss	Comp	Peptide
878.4880	877.4807	877.4658	17.1	94	100	0	2	R.LDNVVYR.M
1000.5580	999.5507	999.5349	15.9	153	161	0	2	K.AALELAEQR.E
1046.6090	1045.6017	1045.5920	9.28	132	140	0	1	K.FNVEVVAIR.E
1179.5990	1178.5917	1178.6156	-20.26	30	40	0	2	K.IEQALGQHGAR.K
1262.7120	1261.7047	1261.7030	1.34	6	16	0	1	K.DLVVSLAYQVR.T
1456.7690	1455.7617	1455.7681	-4.40	80	93	0	2	K.GNTGENLLALLEGR.L
1595.7500	1594.7427	1594.7484	-3.54	82	95	0	1	K.DVFMGVDELQVGMR.F
1611.7350	1610.7277	1610.7433	-9.66	82	95	0	1	K.DVFMGVDELQVGMR.F + Oxidation (M)
1697.9430	1696.9357	1696.9471	-6.73	78	93	1	2	R.LKGNTGENLLALLEGR.L
1960.0400	1959.0327	1959.0425	-5.01	125	142	0	2	R.VVNIASYQVSPNDVVSIR.E
2045.9800	2044.9727	2045.0317	-28.82	185	202	0	2	R.SDLSADINEHLIVELYSK.-
2344.0930	2343.0857	2343.0768	3.83	58	78	0	1	K.FDVAVGANDAYGQYDENLVQR.V
3857.9960	3856.9887	3856.8942	24.5	96	131	0	1	R.FLAETDQGPVPVEITAVEDDHVVVDGNHMLAGQNLK.F

No match to: 1300.6180, 1475.7490, 1540.8010, 1588.8640, 1603.7680, 1707.7820, 1730.1660, 1835.0060, 1993.9720, 2106.0140, 2219.1170, 2872.4350

No further matches found

IBV$_{lower}$ - relative signals