An investigation into the cross-variant binding and neutralization potentials of patient-derived polyclonal antibodies against SARS-CoV-2 spike protein variants

Simon Cobley BSc (Hons)

Student ID: 14279214

Thesis submitted to the University of Nottingham
for the degree of *Master of Philosophy*

Original submission: October 2024

Resubmission: October 2025

Acknowledgments

I would like to express my gratitude to my supervisor Professor Alexander Tarr for his mentorship during my time as a PhD student in his research group. His patience, guidance, and unwavering support have been instrumental in shaping this work. His insights and encouragement have not only helped me overcome numerous challenges but also inspired me to complete my research. I am especially grateful for his encouragement and assistance during the pandemic, when significant demands were placed on him.

I would also like to sincerely thank Professor Patrick Jason Tighe for his insightful guidance and support during my research. His valuable advice and constructive feedback have greatly contributed to the progress and completion of this work.

I am deeply appreciative of Dr. Karen Robinson, whose kindness and support were instrumental in helping me overcome various challenges during my PhD journey. Her help in overcoming difficulties outside of research played a crucial role in a very difficult period.

I would like to extend my sincere thanks to Dr. Will Irving for his invaluable assistance with patient data.

To the staff at QMC Hospital, particularly those in Clinical Chemistry, I would like to extend thanks for their assistance, particularly while dealing with the pandemic response.

I would like to thank all study participants; without their contribution this research would have not been possible.

I would like to acknowledge the Medical Research Council for their financial support, which made this research possible.

My heartfelt thanks go to my friends, whose support and kindness were a source of strength for me, particularly Abram who could always give me a boost when needed.

Lastly, I would like to express my deepest gratitude to my partner, Vanisha Patel, for her unwavering support, especially during the most challenging times of this journey. Her patience, encouragement, and understanding were essential in helping me stay focused and motivated. I could not have completed this work without her constant belief in me.

Abstract

The novel betacoronavirus **SARS-CoV-2** emerged in December 2019 and was responsible for the global COVID-19 pandemic. It caused significant worldwide impact and presented a serious challenge to public health systems. The virus is closely related to other members of the *Coronaviridae* family, particularly **SARS-CoV-1**. SARS-CoV-2 causes a spectrum of disease ranging from asymptomatic or mild respiratory symptoms to severe respiratory illness, systemic infection, and death.

The **SARS-CoV-2** pandemic highlighted the importance of understanding both the general mechanisms underlying antibody responses to viral infection and the distinct immunological behaviours exhibited by specific viruses.

Successful viral clearance in humans is associated with a robust IgG antibody response. Disease severity and recovery are linked to the development of a neutralising antibody response, particularly targeting the spike glycoprotein.

This work aimed to characterise the humoral response to SARS-CoV-2 by analysing antibody binding and neutralisation activity against different spike protein variants, with a focus on how specific antibody profiles influence viral entry inhibition. Using sera from patients admitted to Queen's Medical Centre Hospital with severe COVID-19, combined with ELISA and pseudovirus neutralisation assays, we investigated variant-specific IgG responses and assessed neutralising potency across major variants of concern.

We found that many of the patients exhibited varying degrees of antibody-dependent enhancement across the three main variants of concern circulating at the time: Alpha, Beta, and Delta. These findings provide insight into how immunity evolves in response to SARS-CoV-2 exposure and vaccination. They contribute to the broader understanding of protective immunity, with implications for vaccine design and public health monitoring of emerging variants.

Table of Contents

List of Figures

Abbreviations

AA	amino acids
ACE2	Angiotensin converting enzyme II
ARDS	Acute respiratory distress syndrome
BCR	B cell receptor
BSA	Bovine serum albumin
CDMEM	Complete Dulbecco's Modified Eagle Medium
CFR	Case fatality ratio
CI	Confidence interval
CMV	Cytomegalovirus
CTD	C-terminal domain
CTDD	C terminal dimerization domain
CTT	C terminal tail
DC	Dendritic cell
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic Acid
E	Envelope glycoprotein
E. coli	*Escherichia coli*
EBV	Epstein Barr virus
ELISA	Enzyme Linked Immunosorbent Assay
EM	Effector memory T cells
ER	Endoplasmic reticulum
FBS	Foetal bovine serum
FCS	Foetal calf sera
GC	Germinal centre
GMT	Geometric mean titre
H2SO4	Sulfuric Acid
HCoV-229E	Human coronavirus 229E
HCoV-HKU1	Human coronavirus HKU1
HCoV-NL63	Human coronavirus NL63
HCoV-OC43	Human coronavirus OC43
HEK293T	Human embryonic kidney 293 cell line transfected with tsA1609 allele of Simian virus 40 large T antigen
HRP	Horseradish Peroxidase
IFN	Interferon
IFN-α	Interferon alpha
IFN-β	Interferon beta
IFNγ	Interferon gamma
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IgM	Immunoglobulin M

LB	Lysogeny Broth
LKR	linker region
M	Membrane glycoprotein
MERS	Middle East respiratory syndrome
MERS-CoV	Middle East respiratory syndrome coronavirus
MLV	Murine Leukaemia Virus
mRNA-1273	Nuclear factor kappa B
N	Nucleocapsid glycoprotein
NEB5α	New England Biolabs 5 alpha *E. coli*
NIBSC	National Institute for Biological Standards and Control
NIHR	National Institute of Health Research
NK	Natural killer T cells
NOE	Neutralisation of cell entry assay
NRBD	N terminal RNA binding domain
NSP	Non-structural protein
NTD	N-terminal domain
ORF	Open reading frame
ORF	open reading frame
PASC	Post-Acute Sequelae of SARS-CoV-2 infection
PBS	Phosphate buffered Saline
PBST	PBS
PEI	Polyethyleneimine
phCMV	Plasmid human Cytomegalovirus
pSARS2-A	pcDNA3.1+ SARS CoV-2 Alpha variant Spike glycoprotein encoding sequence
pSARS2-B	pcDNA3.1+ SARS CoV-2 Spike Beta variant glycoprotein encoding sequence
pSARS2-D	pcDNA3.1+ SARS CoV-2 Spike Delta glycoprotein encoding sequence
pSARS2-WT	pcDNA3.1+ SARS CoV-2 Wild Type Spike glycoprotein encoding sequence
RBD	Receptor binding domain
RLU	Relative light units
RNA	Ribonucleic acid
rpm	Revolutions per minute
RTC	Replication and transcription complex S Spike glycoprotein
SARS	Severe acute respiratory syndrome
SARS-CoV-1	Severe acute respiratory syndrome Coronavirus 1
SARS-CoV-2	Severe acute respiratory syndrome Coronavirus 2
SDS	Sodium Dodecyl Sulphate
TE	Terminal effector T cells
TGN	Trans-Golgi network
TMB	3,3′,5,5′-Tetramethylbenzidine
TMD	Transmembrane domain
TMPRSS2	Type 2 transmembrane protease
VERO E6	*Chlorocebus aethiops* kidney epithelial cell line

VOC	Variant of concern
WHO	World Health Organisation
WT	Wild type

Chapter 1. Introduction

1.1 Coronaviruses

Coronaviruses are a group of viruses that are enveloped, possessing an outer lipid membrane that surrounds a protein capsid. Their genome is large positive-sense single stranded RNA with a genome size ranging from 25-32 kilobases in length. Coronaviruses are a group of viruses belonging to the order Nidovirales and family Coronaviridae. The Coronaviridae family is split into four genera: alphacoronavirus, betacoronavirus, gammacoronavirus and deltacoronavirus (Payne, 2017a). All four genera are capable of infecting mammals, however gammacornoaviruses primarily infect avian species, with the only reported mammalian infections being within marine mammals. Deltacoronaviruses exhibit a broader host range and are known to infect avians and a broader range of mammals including pigs, ruminants and chiropterans (Woo et al., 2012). Alphacoronaviruses and betacoronaviruses infect primarily mammals (Cui et al., 2019). Included in these genera are viruses that infect humans and cause significant respiratory illness, ranging from a mild cold to severe disease. Coronaviruses typically cause mild to moderate respiratory illness in humans, however zoonotic strains have emerged which show significant pathogenicity (Forni et al., 2017; Payne, 2017b; Tiwari et al., 2020).

Two alphacoronavirus species are known to infect humans: HCoV-229E and HCoV-NL63. HCoV-229E is one of the causes of the common cold and typically symptoms include mild to moderate respiratory symptoms such as a sore throat, cough, runny nose and fever. HCoV-NL63 is another cause of the common cold with similar symptoms to HCoV-229E. Symptoms in vulnerable individuals, such the elderly, infants or immunocompromised, can be more severe and infections can progress to bronchitis or pneumonia, requiring intervention (van der Hoek et al., 2004). Between humans, transmission of alphacoronaviruses is usually through inhalation of respiratory droplets from a cough or sneeze however infection by fomite transmission via a mucosal membrane is not uncommon (Gaunt et al., 2010). Fomite transmission is possible due to the protective outer lipid membrane of coronaviruses which allow them to persist for some time on surfaces, dependent on conditions (Lim et al., 2016).

Similar to the alphacoronavirus genus, the betacoronavirus genus contain two species that are responsible for the common cold: HCoV-OC43 and HCoV-HKU1 (Woo et al., 2005). These two strains have broadly similar pathogenicity and transmission characteristics in humans. The betacoronavirus genus also contains the most clinically significant strains of coronavirus including MERS CoV (Middle East Respiratory Syndrome Coronavirus), SARS CoV-1 (Severe Acute Respiratory Syndrome Coronavirus 1) and SARS CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2). These strains are all zoonotic in origin and have a high human mortality rate compared to the other coronaviruses and are explored in more detail below (Lim et al., 2016; Woo et al., 2012).

1.2 MERS CoV

MERS CoV (Middle East Respiratory Syndrome Coronavirus) was first identified in Saudi Arabia in 2012 (Zaki et al., 2012) and likely originated in bats. The intermediary host species was dromedary camels with consumption of raw or undercooked camel products likely the route of initial human infection. Human to human transmission was primarily documented in healthcare settings (Z. Zhou et al., 2023). Like many other coronaviruses respiratory droplets and fomite transmission are the main routes of transmission. MERS was the name of the disease caused by MERS CoV infection. MERS is characterised by fever, cough and pneumonia. In more severe cases MERS can cause acute respiratory distress syndrome (ARDS), septic shock and in some cases multi-organ failure and death. Overall, the case fatality rate (CFR) for MERS was 36% globally (Azhar et al., 2019; Fisman et al., 2014).

1.3 SARS CoV-1

SARS CoV-1 likely originated in bats and the first human infection were in Guandon province (Wu et al., 2020), China in 2002. Similar to MERS CoV an intermediary species was likely the origin of human infection for SARS CoV-1, with civets proposed as the likely intermediary host. Human to human transmission was primarily through respiratory droplets and fomite transmission (Cui et al., 2019). The disease caused by SARS CoV-1, named SARS (severe acute respiratory syndrome), was characterised by fever, muscle pain and fatigue on initial infection. As the disease progressed the development of severe respiratory symptoms can occur including pneumonia and ARDS. The reported CFR for SARS was approximately 9.6% globally with a CFR in excess of 50% in vulnerable individuals (Donnelly et al., 2003).

1.4 SARS CoV-2

SARS CoV-2 was the virus responsible for the global pandemic starting in 2019. Originating in Wuhan, China, it was first reported to the World Health Organisation (WHO) on 31st of December 2019 as an unknown viral infection (Zhu et al., 2020). The virus was identified and named SARS CoV-2 within days of the report, with the first reported human death following on 13th January 2020 (Figure 2.). The disease caused by SARS CoV-2 was named COVID-19 by the WHO. The virus rapidly spread internationally and was declared a global pandemic by the WHO 10 weeks after the first report (Figure 3-6). Among the coronaviruses, SARS CoV-2 has a particularly high transmission rate, spreading through respiratory droplets and fomite transmission. Human to human transmission occurred in presymptomatic and asymptomatic individuals soon after infection, significantly contributing to the rapid global spread of the virus. As of September 2024, globally over 776 million cases of COVID-19 (Figure 1.)(World Health Organization 2023 data.who.int, n.d.-a) and in excess of 7 million deaths have been reported (World Health Organization 2023 data.who.int, n.d.-a).

Infection with SARS CoV-2 can be asymptomatic (You et al., 2024) or cause mild symptoms similar to the common cold. More severe cases of COVID-19 result in symptoms including shortness of breath, fatigue, muscle pain and loss of taste or smell. In the most severe cases symptoms can progress to pneumonia, ARDS, multi-organ failure and death (Cevik et al., 2020). As the pandemic progressed mutations in the viral genome resulted in the emergence of new variants of SARS CoV-2. These variants caused varying symptoms, and had different CFR and degrees of transmissibility. Measures were taken to limit the spread of the virus. These included: local and national lockdowns, isolation of infected individuals and workplace adjustments. Other measures taken included: emergency rapid deployment of new vaccines, legislation mandating face masks and social distancing, local and international travel restrictions, and widespread testing initiatives. The trade-offs between the effectiveness of these public health interventions and the resulting economic damage and health consequences remain a contentious issue in public discourse (Plümper & Neumayer, 2022).

The CFR for COVID-19 was initially 2-3%, reducing to around 1% with the introduction of non-sterilising vaccines and improvement in both diagnosis and treatments as the pandemic progressed (Cao et al., 2020; Horita & Fukumoto, 2023). The CFR varied globally with availability and quality of healthcare being a contributing factor. With age being a significant risk factor the severity of symptoms varied by age group, with children generally experiencing mild symptoms or being asymptomatic. The elderly (60+) generally had higher rates of severe disease and a higher mortality rate, with pre-existing conditions being a contributing factor. In adults COVID-19 symptoms ranged from mild to severe. Despite this trend severe disease and death was reported for all age groups. CFR for vulnerable individuals was estimated at 5-15%. COVID-19 CFR was challenging to calculate due to underreporting of cases, especially in mild or asymptomatic cases. The variants of SARS CoV-2 which exhibited different symptoms, transmission rates and mortality rates also contributed to difficulty in establishing CFR.

Post-Acute Sequelae of SARS-CoV-2 infection (PASC), also known as long Covid, is a condition that arose in many individuals after the acute phase of a SARS-CoV-2 infection. Development of PASC is independent of the severity of symptoms exhibited during the acute stage of infection, including asymptomatic cases. Symptoms were diverse including: fatigue, cognitive impairment, shortness of breath and loss of taste or smell. Neurological, cardiovascular and gastrointestinal issues were also reported as PASC symptoms. The pathophysiology of PASC is still unknown however it is hypothesised to be impacted by factors such as persistent viral reservoirs, immune dysregulation, and microvascular damage (Jiao et al., 2024).
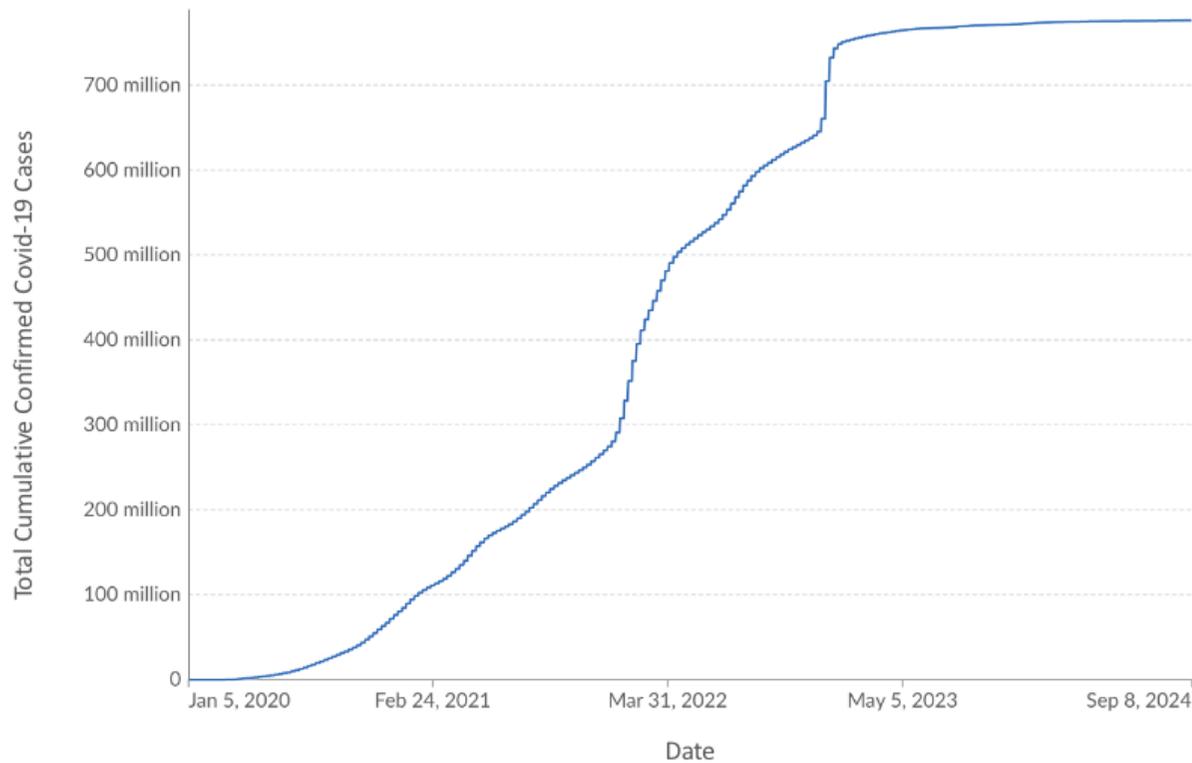
Figure 1. Total worldwide cumulative reported COVID-19 cases. Data from WHO, 2024.
This figure illustrates the cumulative global total number of cases of confirmed SARS-CoV-2 infections from December 2019 to 2024. The sharp increases correspond to major pandemic waves driven by variants of concern (VOCs) such as Alpha, Delta and Omicron. The steep rises in early 2020 and again around mid-2021 highlight periods of high viral transmissibility and limited immunity within the global population to the circulating dominant variant. The gradual slowing of the curve in later years reflects increasing vaccine coverage and post-infection immunity. Graph adapted from (Mathieu et al., 2020).
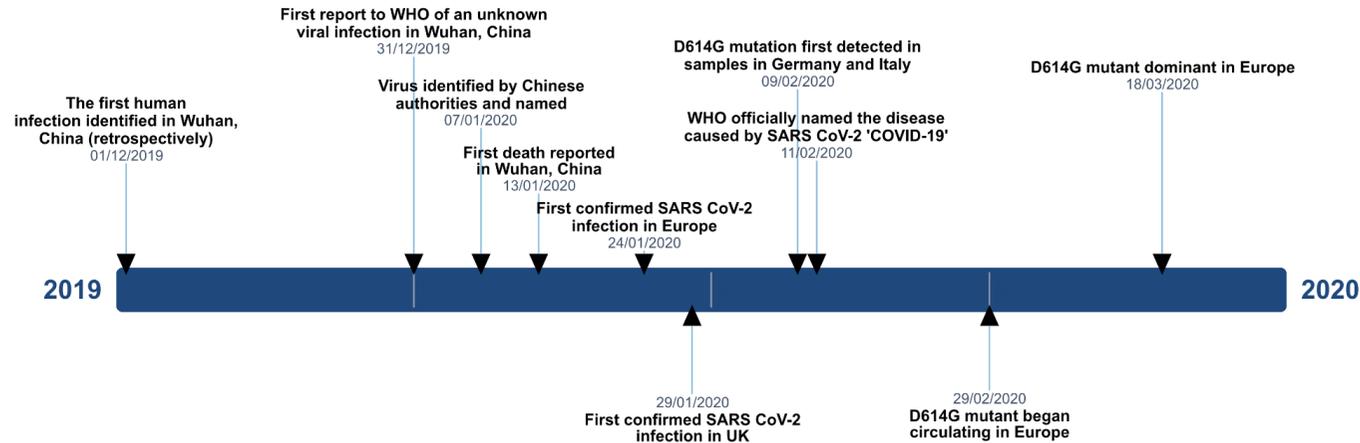
Figure 2. Timeline outlining the emergence of SARS CoV-2 and its global spread between December 2019 to November 2020.
This figure shows the early phase of the pandemic, beginning with the first identified case in Wuhan, China, and culminating in global spread late in 2020. Key milestones include the Public Health Emergency of International Concern announced by (January 2020) and later a pandemic (March 2020). The critical period between December 2019 and March 2020 is the time period during which uncontrolled transmission allowed SARS-CoV-2 to spread across multiple continents, laying the groundwork for later variant emergence. The timeline illustrates how rapidly the situation evolved from emergence to spread into Europe and the rapid emergence of mutants.
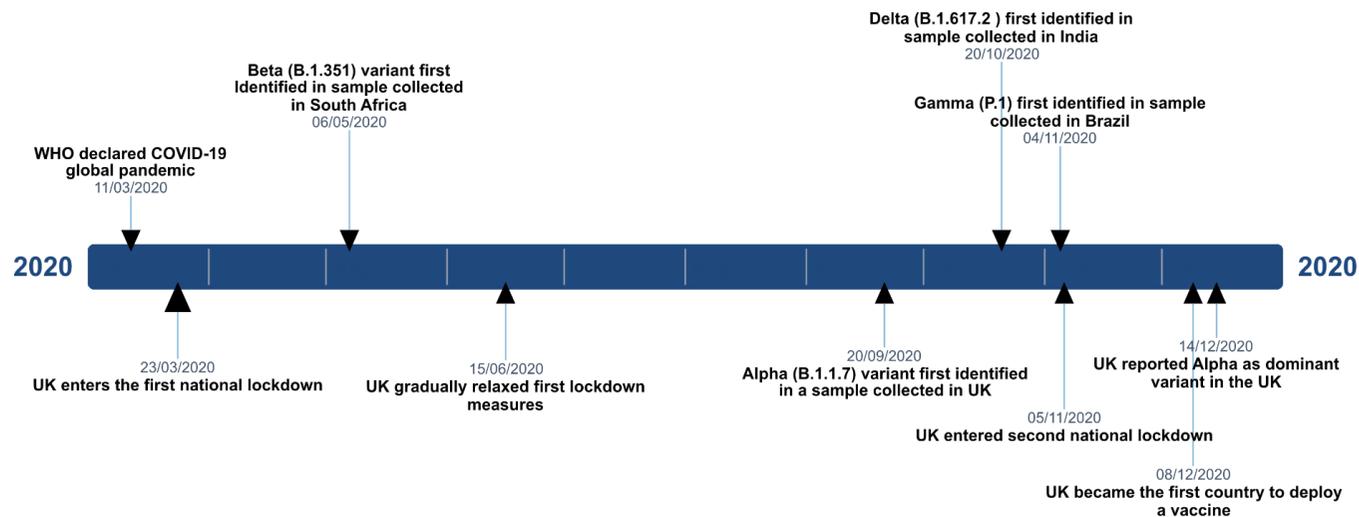
Figure 3. Timeline outlining the emergence of SARS CoV-2 and its global spread between November 2020 to December 2020.
This figure shows the appearance of the first major variants of concern: Alpha (B.1.1.7) in the United Kingdom, Beta (B.1.351) in South Africa and Gamma (P.1) in Brazil. This short interval marks the transition from the original wild-type strain to variant driven transmission. The increase in case numbers despite ongoing public health measures highlights how mutations within the spike protein enhanced viral fitness, enabling SARS-CoV-2 to adapt more effectively to infect Humans

WHO declared Gamma variant of concern
11/01/2021

2020 | 2021

19/12/2020
UK imposed stricter lockdown measures in London

20/12/2020
UK imposed international travel bans

18/12/2020
WHO declared Alpha variant of concern

04/01/2021
UK entered third national lockdown

26 Dec
Delta dominant variant in UK
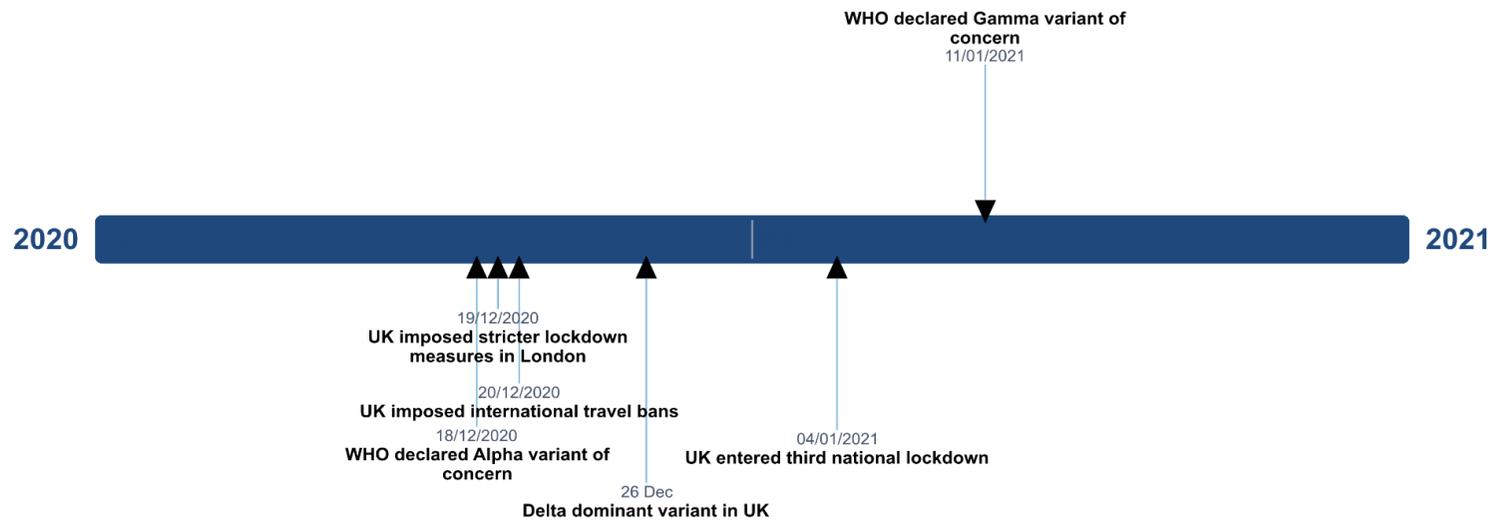
Figure 4. Timeline outlining the emergence of SARS CoV-2 and its global spread between December 2020 to November 2021.
This figure depicts the dominance of the Alpha variant in early 2021 and its subsequent replacement by Delta (B.1.617.2) from mid-2021 onwards. The global expansion of Delta was associated with higher viral loads and increased transmissibility This drove large waves of infection, even in partially immunised populations.
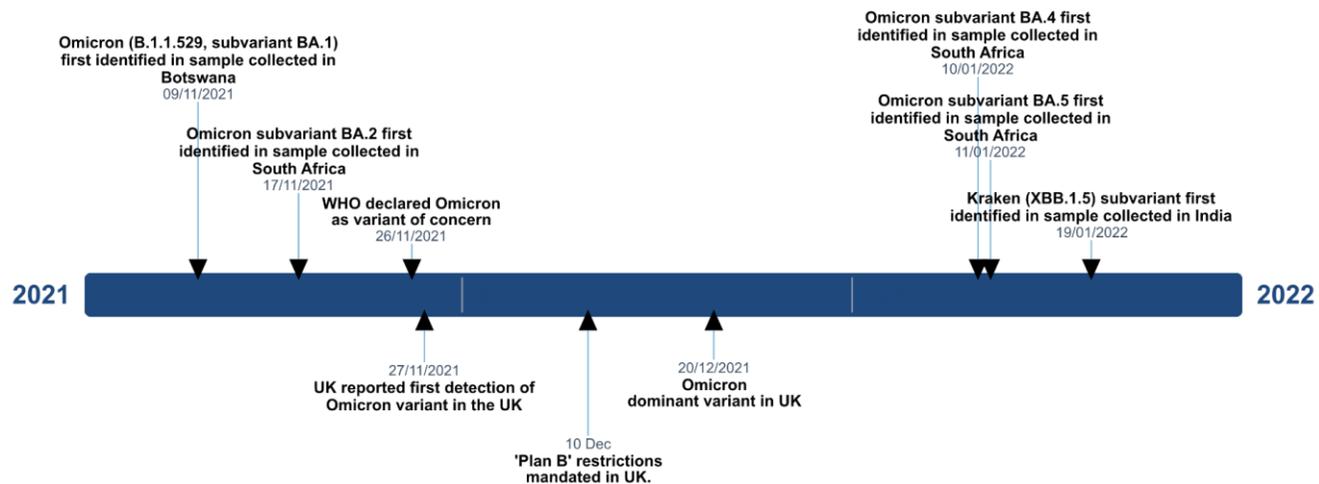
Figure 5. **Timeline outlining the emergence of SARS CoV-2 and its global spread between November 2021 to January 2022.**
This figure depicts the rapid emergence and global dominance of the Omicron (B.1.1.529) variant. Omicron carried over thirty mutations within the spike gene, many concentrated in the receptor-binding domain, leading to substantial immune evasion. The sharp rise in case numbers in late 2021 demonstrates Omicron's transmission advantage and its capacity to evade pre-existing neutralising antibodies. This period highlights the virus's ability to adapt under strong immune selection pressure. It also highlights the ongoing need to monitor evolution of the spike protein to maintain vaccine efficacy.

2022 ─── **2023**

27/01/2022
End of 'Plan B' restrictions in UK

04/03/2022
BA.2 subvariant dominant in UK

28/06/2022
Kraken subvariant first
Identified in sample collected in UK

20/06/2022
BA.4 BA.5 subvariants dominant in uk

27/01/2023
Kraken subvariant dominant in UK

05/09/2023
Bivalent vaccines deployed in UK
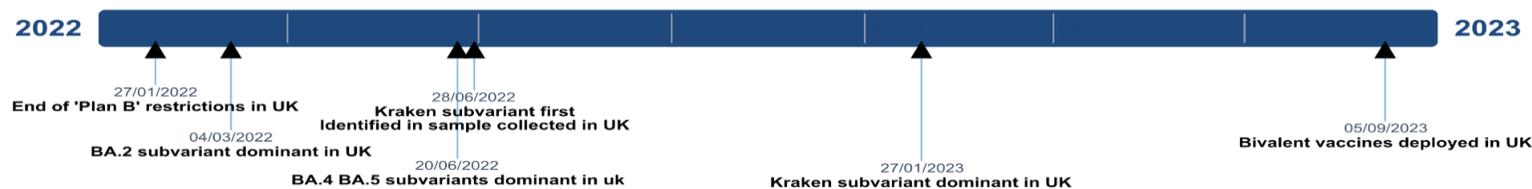
Figure 6. Timeline outlining the emergence of SARS CoV-2 and its global spread between January 2022 to September 2023.
This figure shows the transition from pandemic expansion to endemic persistence, characterised by the diversification of Omicron into multiple sub-lineages, including BA.1, BA.2, BA.4/5 and XBB. The gradual plateau in cumulative case numbers indicates a new equilibrium between viral adaptation and widespread immunity. The emergence of successive sub-variants demonstrates ongoing antigenic drift driven by immune pressure. These developments provide a rationale for investigating cross-variant antibody binding and neutralisation.

1.4.1 Genomic Organization and Protein Functionality

SARS CoV-2, like all coronaviruses, is an enveloped virus with a positive-sense single stranded RNA genome. The SARS CoV-2 genome is approximately 29.9 kilobases in length (Harcourt et al., 2020; P. Zhou et al., 2020) containing a number of open reading frames (ORFs) which encode for 29 proteins in total (Figure 7.). These proteins are: 4 major structural proteins, 16 non-structural proteins (NSPs) and 8 accessory proteins. The four structural proteins: spike (S), envelope (E), membrane (M), and nucleocapsid (N) are encoded by ORF2, ORF4, ORF5 and ORF9 respectively. The NSPs are encoded by ORF1 which is split into ORF1a and ORF1b. All the accessory proteins are named for the ORF from which they are encoded, for example ORF3a encodes ORF3a protein. The accessory proteins are encoded by ORFS 3a, 3b, 6, 7a, 7b, 8, 9b and 9c (Finkel et al., 2021; Kim et al., 2020; Redondo et al., 2021). The accessory proteins are not essential for successful cell entry of the assembled virion. However, they are involved in suppression of the hosts immune response, virus release from infected cells, and virion assembly. ORF10 is an additional ORF that is thought to encode a redundant protein that is not translated in normal conditions and is not beneficial to the viral life cycle (Redondo et al., 2021; Wang et al., 2020).
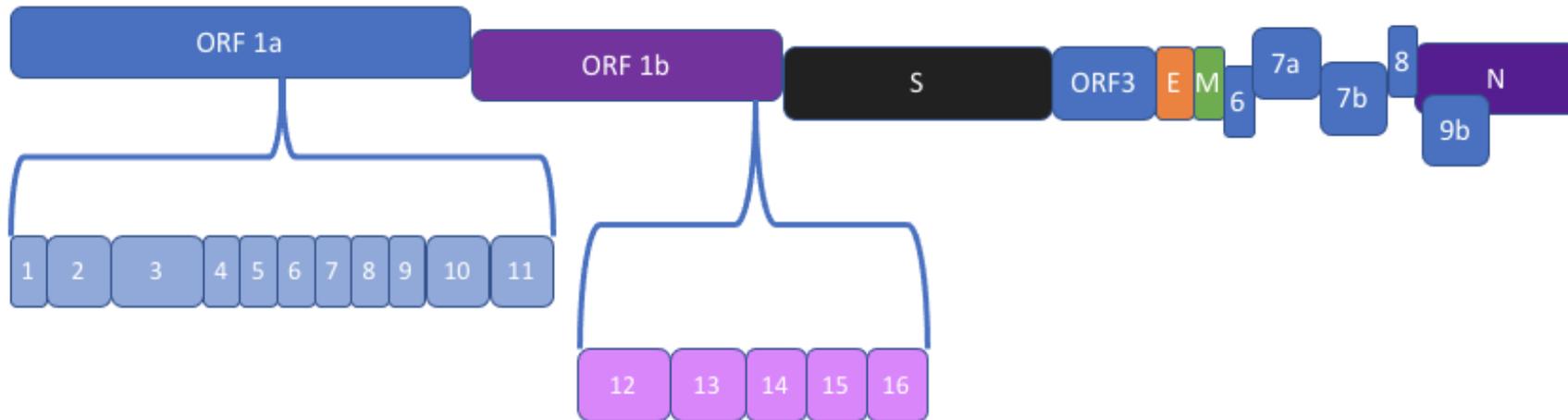
Figure 7. SARS CoV-2 Viral genome structure.

In this figure open reading frame 1a (ORF1a) open reading frame 1b (ORF1b) and open reading frame 3 (ORF3) non structural proteins (NSPs) are labelled with their respective number. NSPs 1-11 are encoded in ORF1A, ORF1B encodes for NSPs 12-16 and genes encoding for structural proteins are labelled with their corresponding letter: spike (S), envelope (E), membrane (M) and nucleocapsid (N). Adapted from (Zhang et al., 2021).

1.4.2 Structural proteins

The envelope of SARS CoV-2 consists of a phospholipid bilayer embedded with viral structural proteins. The viral genome is protected within the envelope as a nucleocapsid complex. Some NSPs and host cell proteins may also be present in the lumen of the viral envelope. There are four essential structural proteins of the virion (Figure 8.). One of the four major structural proteins N binds to the single stranded RNA genome, forming a ribonuclease complex. The other three major structural proteins of the virus, E, M and S, are embedded in the phospholipid bilayer (which is derived from the host cell during replication) of the virion (Wang et al., 2020).
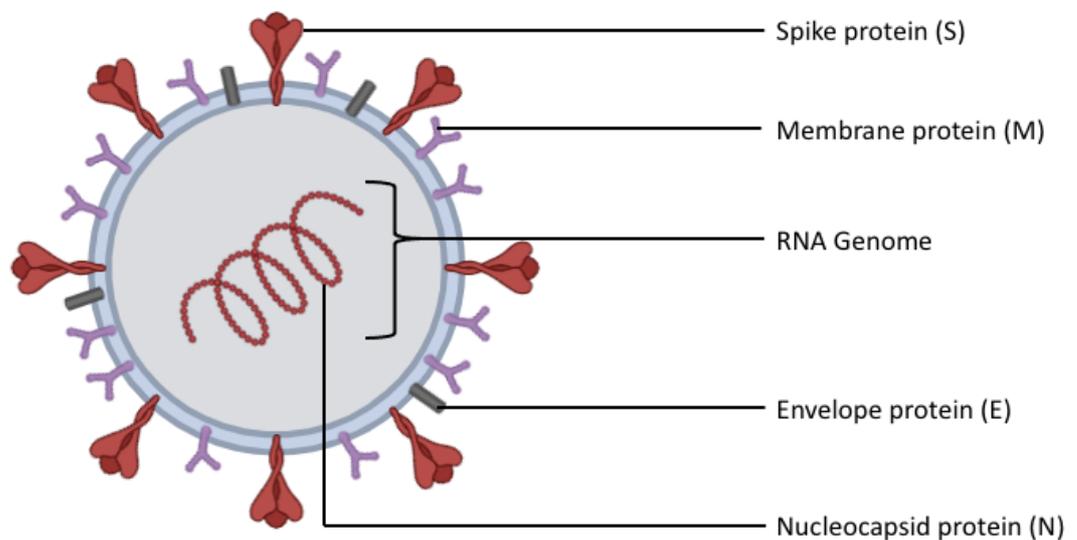
Figure 8. SARS CoV-2 virion structure
This figure shows a simplified diagram of the SARS CoV-2 virion with structural proteins. The virion is roughly spherical and derived from host cell membrane. The viral envelope is studded with spike proteins, which give the coronavirus their characteristic appearance and name. Membrane proteins are embedded in the viral envelope and they interact with the nucleocapsid and envelope proteins in assembly of the virion. Nucleocapsid proteins are show in complex with the RNA genome. Nucleocapsid proteins compact and help localise the genome in the virion during budding from the host cell.

1.4.3 Nucleocapsid protein

The N protein is approximately 419 amino acids (AA) long and contains four structural domains: N terminal RNA binding domain (NRBD), linker region, C terminal dimerization domain (CTDD) and the C terminal tail (CTT). The N terminal RNA binding domain is approximately 135 amino acid long. (Chang et al., 2006). Structurally the NRBD 23 adopts a compact, folded structure. The viral RNA interaction with N occurs in a negatively charged cleft within the NRBD. Beta sheet

elements of the NRBD containing positively charged AA interact with the negatively charged RNA phosphate backbone, binding it in place. The linker region domain (LKR) provides structural flexibility to the protein, allowing for conformational change as the protein binds to RNA. The LKR is an approximately 142 AA long region with a number of serine residues that can be phosphorylated modulating the protein's interactions. The CTDD is a domain approximately 115 AA long that is facilitates dimerization and oligomerization of the N protein. Dimerization allows the protein complex to bind to RNA at two points with the two NRBD of each unit. Oligomerization of the protein dimers is essential in efficient and compact genome packaging. The N dimers oligomerize to form a protein scaffold structure which compacts the bound RNA during virion assembly. The smallest of the domains CTT at 54 AA long and is thought to be a stabilising element of the protein. Once the virion infects a cell, the N protein also interacts with host proteins to promote efficient viral RNA translation (Marra et al., 2003; McBride et al., 2014). Among these proteins are G3BP1 and G3BP2, core stress granule proteins. Under normal conditions these proteins cause translational shutdown during cellular stress however N can bind to and inhibit these proteins, preventing efficient translational shutdown.

1.4.4 SARS CoV-2 Envelope protein

The E protein is the smallest of the four structural proteins at ~75 AA long. It is a multifunctional protein playing a role in viral assembly, virus budding, host cell interactions and in pathogenesis. The three functional domains of the E protein are: the N terminal domain (NTD), the transmembrane domain (TMD) and the C terminal domain (CTD) (Marra et al., 2003). The TMD is a ~25 AA long hydrophobic domain, organised into alpha helices that embed within membranes of the host's intracellular organelles. The protein can form a pentameric complex via interactions between five E protein's TMDs, forming a viroporin complex. These viroporins, with a hydrophilic core, act as ion channels facilitating transport of Na+ and K+ ions. This disruption of a host cells ionic balance helps facilitate viral egress. E protein viroporins are situated in host intracellular membranes such as the ERGIC. Their ion channel activity promotes potassium efflux from the cytoplasm, which activates the NLRP3 inflammasome. This activation leads to inflammatory signalling and the release of IL-1β, a cytokine linked to disease severity in COVID-19. https://pubmed.ncbi.nlm.nih.gov/26331680/

The NTD is a short hydrophilic domain that acts to anchor the protein within membranes. With the TMD embedded in the membrane the NTD is exposed to the luminal side of intracellular organelles. The hydrophilic nature of the NTD aids in correct orientation of the E protein in membranes (Cao et al., 2021). The CTD is a ~36 AA long hydrophilic domain of the protein, when the E protein is embedded in an intracellular organelle membrane the CTD is present on the cytosolic side of the membrane. A PDZ binding motif in the CTD is capable of binding to proteins involved in cell signalling pathways and cell structure, allowing the virus to modulate host

immune signalling and disrupt cell architecture. During virion assembly The CTD interacts with the M protein to facilitate membrane curvature, essential for budding of the virion (Breitinger et al., 2022; Santos-Mendoza, 2023).

1.4.5   Membrane protein

The most abundant structural viral protein, M is a ~222 AA long protein with three functional domains: NTD, TMD and CTD. When the M protein is embedded in a host cell membrane the NTD is exposed to the luminal side of intracellular organelles. The NTD is a ~17 AA long hydrophilic domain that aids in correct orientation of the M protein the membrane. The TMD is a ~81 AA long domain that folds into three alpha helices. Once assembled and folded the TMD is hydrophobic and integrates into the phospholipid membrane. The TMD also facilitates a number of protein-protein interaction essential for virion assembly and budding. Oligomerization of M proteins, mediated by the TMD, is essential in the formation of the viral envelope. The M protein TMD interacts with the TMD of the S protein to ensure correct incorporation of S into the membrane. Shaping of the viral membrane and budding is achieved partly through interaction of hydrophobic interfaces of the TMDs of M and E. The CTD of the M protein is the largest domain at ~101 AA long. The M protein CTD is hydrophilic and is present on the luminal side of the assembled virion where it interacts with the CTD of N to ensure correct localisation of the N-RNA complex. This interaction is key in assembly of functional infective virions (Siu et al., 2008; Wang et al., 2020).

1.4.6   Spike protein

 The S protein is a large glycoprotein (~1,273 AA) present on the virion membrane as a homotrimer that plays an essential role in infection of host cells. The S protein mediates cell entry via interaction with host cell proteins (Tortorici & Veesler, 2019). ACE2 (Angiotensin-Converting Enzyme 2) is the primary receptor for SARS CoV-2. ACE2 is expressed in the epithelial cells of the respiratory tract where it takes part in SARS CoV-2 cell entry (Figure 10.). TMPRSS2 (Transmembrane Protease Serine 2) is a key host protein involved in the complex multistage process of cell entry. The protease cleaves the S2' site, activating the spike for membrane fusion.

 The S protein can be divided into two subunits S1 and S2, each with distinct functions (Figure 9.). The S1 subunit is the most distal part of the protein from the virion membrane and contains the NTD and receptor binding domain (RBD) (Marra et al., 2003). The S2 subunit mediates fusion between the viral envelope and host cell membrane. It contains three major functional domains: TMD, fusion peptide (FP), heptad repeat 1 (HR1) and heptad repeat 2 (HR2) (Xu et al., 2004). The TMD is a ~22 AA long domain that is hydrophobic and holds the S protein firmly in the viral envelope during the membrane fusion process. The TMD forms an alpha helical structure which maintains its orientation in the viral envelope. The trimeric

structurer of the protein is also stabilised by interactions between the TMDs of the S monomers. The FP is a hydrophobic sequence ~20 AA long that plays a key role in membrane fusion. The S protein also contains two important cleavage sites: the S1/S2 cleavage site and S2' cleavage site. These two sites are cleaved by host proteases and cleavage at these sites is essential for S-receptor binding and for membrane fusion. The S1/S2 site is located between S1 and S2 and the S2' site is within the S2 subunit. The heptad repeats, HR1 and HR2, within S2 subunit, are essential in fusing the viral envelope with the host cell membrane (Walls et al., 2020).
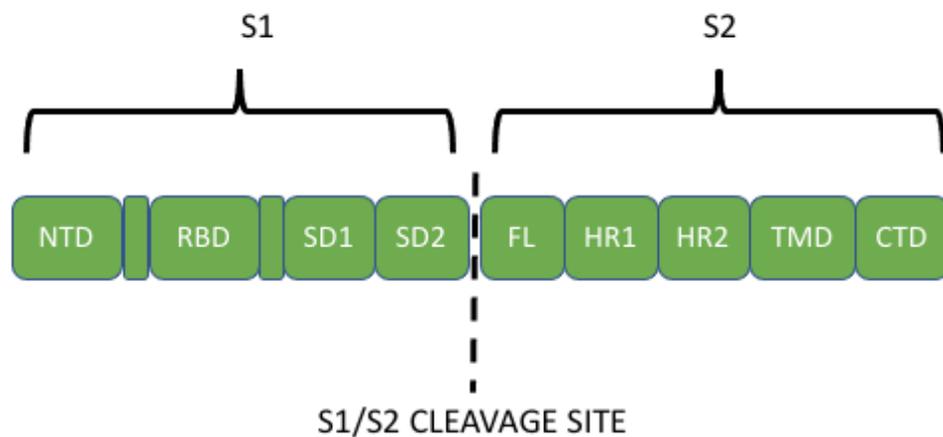
Figure 9. SARS CoV-2 Spike protein functional domains and order.
SARS CoV-2 Spike protein functional domains.

This figure illustrates the overall primary structure of the spike glycoprotein. The S1 region contains the N terminal domain (NTD), receptor binding domain (RBD), the SD1 and SD2 subdomains. The S1 region, responsible for ACE2 binding, is located at the N terminal section of the spike protein. The S2 region contains the fusion peptide or fusion loop (FL) domain, heptad repeat 1 (HR1), heptad repeat 2 (HR2) , the transmembrane domain (TMD) and the C terminal domain (CTD). The S2 region is responsible for membrane fusion in the viral entry process. This figure highlights the modular organisation of the spike protein. (Zhang et al., 2021).
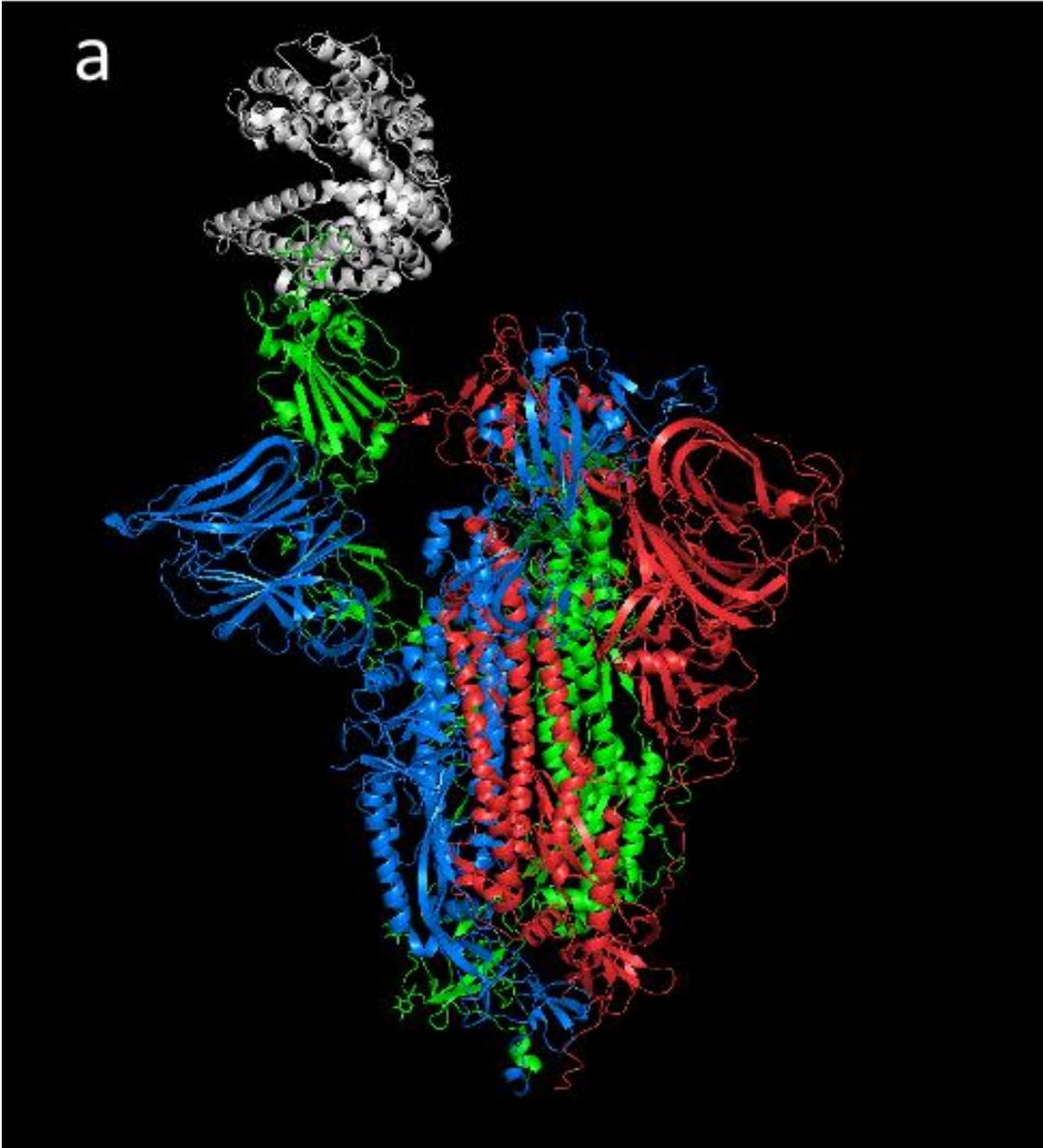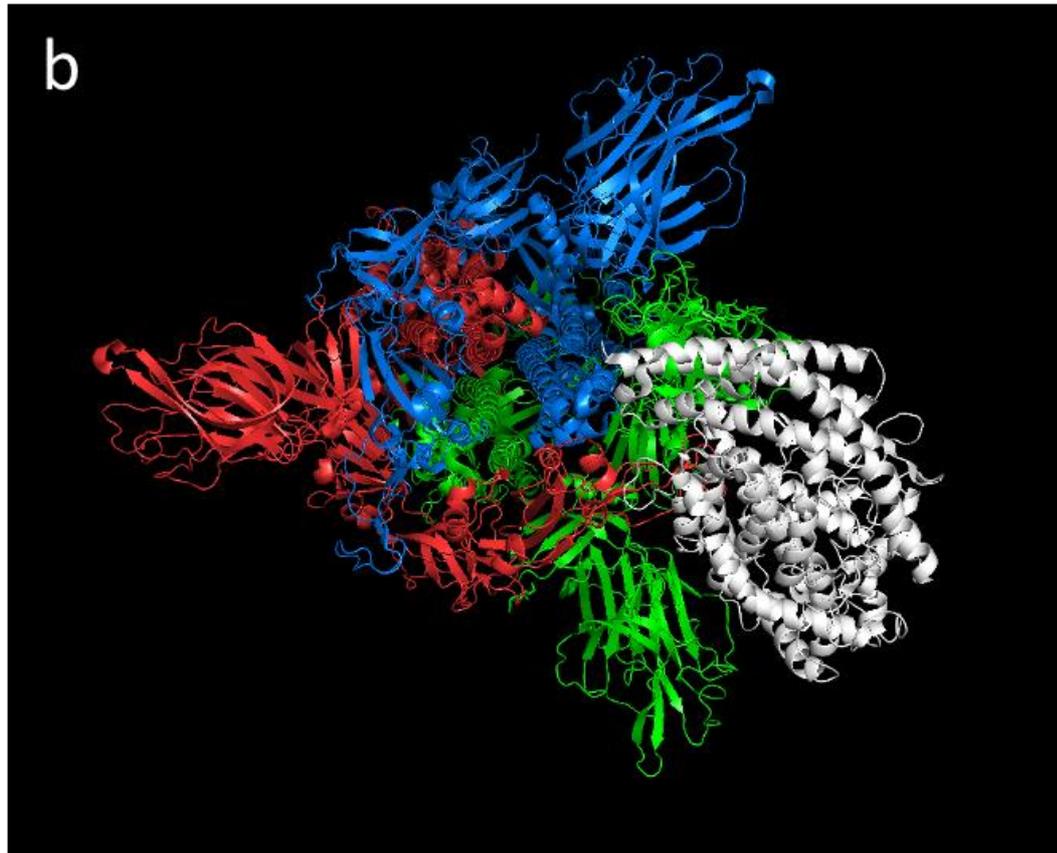
Figure 10. Homotrimeric SARS CoV-2 spike protein tertiary structure, bound to ACE2.

This figure shows the three spike monomers (red, green and blue) bound to the ACE2 cell surface receptor (grey). Two of the monomers (red and blue) are in the closed conformation and the other (green) is in the open conformation, exposing its RBD, and binding to ACE2. Figure **a** shows this protein from the side on perspective, where the virus membrane would be below the structure and **b** show the spike trimer and ACE2 from top down perspective where the virus membrane would be behind the structure. (Berman et al., 2003)

1.4.7 Molecular dynamics of SARS CoV-2 cell entry

The entry of the SARS CoV-2 virion to a host cell is a complex multistage process. The spike protein uses ACE2, a surface protein present mainly on the cell surface of two cell types of the respiratory tract. ACE2 is also expressed by many other cells throughout the body, which likely explains systemic infections and non-respiratory symptoms. On initial infection ciliated epithelial cells of the upper respiratory tract are infected. Once the infection progresses Type II alveolar cells deeper in the lungs are infected. Initial exposure occurs in the upper respiratory tract where inhaled virions contact ciliated epithelial cells. Once replication starts, virions are released into the airways where they are inhaled deeper in the lungs where Type II alveolar cells may be infected.

For successful infection the spike glycoprotein must undergo a series of ordered conformational changes. Prior to release from the host cell, the spike protein undergoes priming by host proteases, an essential step in virion maturation. The initial cleavage at the S1/S2 site is mediated by the protease furin within the Golgi apparatus and trans-Golgi network. A second cleavage at the S2' site, performed by TMPRSS2 or endosomal cathepsins, occurs during subsequent infection of a new host cell to activate membrane fusion. Host furin, present in the TGN cleaves the S1/S2 28 cleavage site, once this cleavage occurs the covalent bonds between S1 and S2 are severed, and they are then held together by non-covalent interactions. This cleavage of the protein acts to destabilise the individual spike monomers conformation, making conformational change more likely to occur. The first cleavage also exposes the second cleavage site S2'. Once S has been primed the mature virion buds from the infected cell and is free to infect another host cell. The RBD of S of the mature virion undergoes dynamic conformational change between "up" where the RBD is accessible and able to bond with the host receptor and "down" where the RBD is inaccessible. This switching may aid in preventing host antibodies from binding to the RBD while in the down formation (Jackson et al., 2022; Qing et al., 2021).

1.4.7.1 Direct fusion pathway

The first step in viral entry is an RBD in the up formation binding to a host cell ACE2 receptor's peptidase domain. Once bound to ACE2 S becomes locked in the open conformation it maintains the bond, keeping the virus in close proximity. The host protease, Transmembrane serine protease 2 (TMPRSS2) on the cell membrane cleaves the S2' cleavage site progressing the process of viral entry. On cleavage of S2' site the S1 subunit dissociates from S and the S2 subunit progresses membrane fusion. The FP of S2 is exposed by the previous step and it makes contact with the membrane where its arrangement of polar and charged AA guide the FP to the cell membrane. The FP partially inserts into the cell membrane and causes a disruption in the phospholipid bilayer, aiding membrane fusion. This disruption causes a thinning and convex curving of the cell membrane, weaking it. The FP anchors the virus to the cell membrane while S2 undergoes conformation change. HR1 of S2, in a helical formation, extends towards the cell and embeds itself in the cell membrane. Once this occurs the HR2 of S2 undergoes conformation change, folding towards HR1 and forming a six-helix bundle structure with HR1. This folding pulls the viral and cell membranes closer, and the six-helix bundle stabilises the conformation. Membrane fusion occurs at this point, creating a continuous membrane between the viral envelope and host cell in a process called hemifusion. Destabilisation of the cell membrane occurs resulting in the formation of a fusion pore. Initially the fusion pore is small but as membrane fusion continues the pore increases in diameter until it is large enough for the viral RNA to pass into the host cell (Jackson et al., 2022; Li et al., 2023).

1.4.7.2 Endosomal pathway

An alternative route through which viral RNA can enter the cell is the endosomal pathway. After the RBD binds to ACE2 on the extracellular side of the membrane clathrin proteins coat the corresponding cytosolic side of the cell membrane. The membrane is deformed in a concave manner and a clathrin-coated pit forms. The membrane invaginates and the virus is endocytosed by the host cell into an early endosome. The endosome is transported deeper into the cell and the endosome matures and the endosomal environment becomes acidic, which activates host cathepsin proteases. The cathepsins perform the same function as TMPRSS2 in the previously described process and the same process occurs to achieve membrane fusion, fusion pore formation and release of the viral RNA into the cytosol (Jackson et al., 2022; Shang et al., 2020; Walls et al., 2020).

1.4.7.3 Viral replication and innate immune sensing

Once viral RNA is released into the cytoplasm, replication begins immediately. The positive-sense RNA genome serves as mRNA for translation of ORF1a and ORF1b into polyproteins that are cleaved into non-structural proteins (NSPs). These NSPs form the replication–transcription complex (RTC) on double-membrane vesicles derived from the endoplasmic reticulum, where genomic and subgenomic RNAs are synthesised. The resulting double-stranded RNA intermediates are sensed by cytosolic receptors such as MDA5 and RIG-I, activating interferon signalling and antiviral responses (V'Kovski et al., 2021; Xue et al., 2021).

1.4.8 SARS CoV-2 Variants and critical mutations

Over the course of the pandemic a number of variants of concern (VOC) have emerged (Figure 11.). A VOC is a viral variant designated by the WHO that has evolved significant phenotypic changes in transmission, pathogenicity or immune evasion. Due to the short replication time and large population size viruses can evolve relatively quickly. As the viruses replicate, they accumulate synonymous and non-synonymous mutations in the viral genome. These mutations can affect viral fitness and cause phenotypic changes in the mutated virus. The RNA dependent polymerase of SARS CoV-2 lacks a proofreading subunit. The lack of this subunit means that mutations that are introduced are replicated in virus progeny. In addition to mutations, viruses can undergo recombination. Recombination can occur in any infection however in a coinfection of two or more variants a recombination event may result in a new variant or subvariant, with modified phenotypes (Carabelli et al., 2023; World Health Organization 2023 data.who.int, n.d.-b).
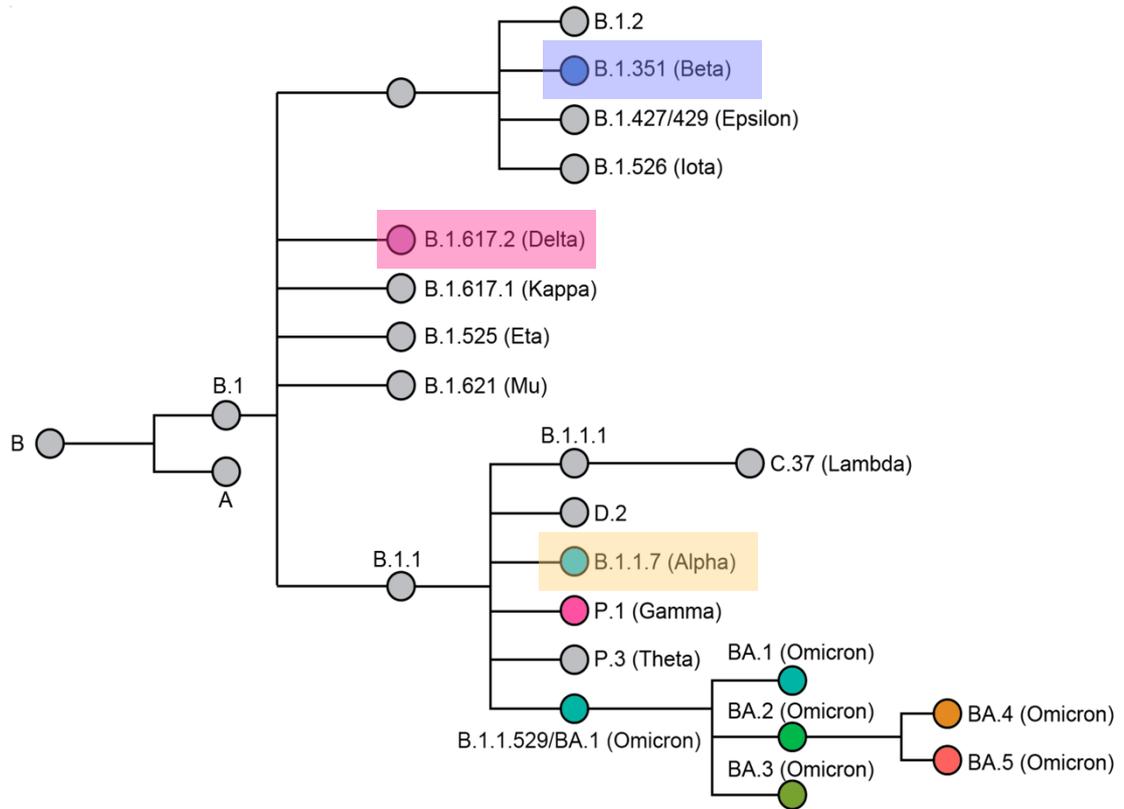
Figure 11. **Phylogenetic tree of SARS CoV-2 variants occurring between 2020-2022** (Wiegand et al., 2022).

Phylogenetic tree of SARS CoV-2 variants (grey) and variants of concern (coloured). Variants used in this study are highlighted in yellow (Alpha), blue (Beta) and pink (Delta). Each variant is labelled with its PANGO lineage code

1.4.8.1 Alpha (B.1.1.7)

The Alpha variant was the first variant of concern (VOC) declared by the WHO. It was distinguished by its increased infectivity and transmission over the wild type. It was first detected in the UK on 20th September 2020 and by 14th December 2020 it had become the dominate variant in the UK. This is because it has novel mutations in the spike protein causing an increased binding affinity between spike and the ACE2 receptor, compared to the wild type. It has substitution mutations in the receptor-binding domain of N501Y, P681H, and deletion mutations at 69-70 and 144. It also contains some nucleocapsid mutations, which are thought to contribute to the increase transmission rate. The N501Y mutation increases binding affinity to ACE2 receptors, enhancing transmissibility. The mutation near the S1/S2 cleavage site, P681H, potentially enhanced the virus's ability to enter cells. 69-70del in S has been associated with enhanced immune evasion. Finally, the D614G mutation enhances stability of the S protein, increasing transmissibility (Bloom et al., 2023; Khetran & Mustafa, 2024; Wiegand et al., 2022).

1.4.8.2 Beta (B.1.351)

The beta variant was first detected in South Africa in May 2020. It was similarly associated with increased rates of infection. It contained the N501Y mutation as found in alpha which increased its ACE2 binding. Alongside this key mutation it contains 2 others in its spike regions: K471N and E484K. Beta however lacked the deletion mutation at 69-70 which is found in other variants. The K417N and E484K mutation in the S1 subunit reduced neutralization by antibodies, aiding in immune evasion. The D614G mutation, also present in alpha, had the same affect previously described (Andre et al., 2023; Bloom et al., 2023; Khetran & Mustafa, 2024; Wiegand et al., 2022).

1.4.8.3 Gamma (P.1)

First reported in Brazil late 2020, this variant was found to evade previous immunity developed to the alpha and beta strains. It was also found to produce reinfections and was the most transmissible VOC. It has the second highest number of mutations in the spike protein of all the VOCs at 12 mutations, these include the substitution mutations: L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I, and V1176F. The mutations L18F, T20N, P26S, D138Y and R190S in the S NTD are thought to have reduced binding affinity of antibodies raised by previous infection with WT, alpha and beta variants. The K417T mutation in the S1 subunit reduced antibody neutralization, aiding in immune evasion. The H655Y mutation near the S1/S2 cleavage site enhanced the cleavage by host furin, enhancing infectivity. The two mutations near the TMD T1027I and V1176F could have affected the conformation changes undergone in cell entry, enhancing infectivity (Andre et al., 2023; Bloom et al., 2023; Khetran & Mustafa, 2024).

1.4.8.4 Delta (B.1.617.2)

The Delta variant had become the main variant across the globe by mid-2022 and although it causes similar symptoms to alpha, people infected suffered more severe symptoms resulting in more hospitalizations across most countries. Delta had over 20 new mutations compared to alpha, half which affected the spike protein (Andre et al., 2023). The delta variant had 8 substitution mutations: T19R, G142D, R158G, L452R, T478K, D614G, P681R and D950N in its spike protein and 2 deletion mutations. It also contained 2 mutations that were associated with high viral load. The mutations T19R, G14D and R158G were in the NTD of the S protein. The T19R mutation was associated with a minor modification to S processing and folding. All three mutation were associated with reduced efficacy of binding of antibodies raised by previous infection with other VOC. Mutations to the RBD, L452R and T478K both enhanced the binding affinity to ACE2. P681R near the S1/S2 cleavage site enhanced proteolytic cleavage, possibly enhancing viral fusion and leading to increased transmissibility. D950N in the HR1 of S2 enhanced the conformational capabilities of

the region, improving efficiency of membrane fusion, increasing transmissibility (Bloom et al., 2023; Khetran & Mustafa, 2024).

1.4.8.5 Omicron (B.1.617.2)

This variant was an interesting VOC as it contained over 30 mutations connected with the spike region. It was first reported in South Africa with symptoms similar to that of alpha, beta and delta. It has 15 mutations in the receptor-binding domain of spike, some of the substitution mutations include: S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, N501Y, and Y505H, some of these are also found in the previous variants discussed. The omicron variant had a combination of mutations in 3 areas: improved ACE2 binding from the mutations Q498R and N501Y, within the S1/S2 cleavage site: H655Y, N679K, and P681H, improving proteolytic cleavage and mutations within the N protein associated with higher viral load: P203K and G204R. It was found that although omicron contained more mutations it had a lower CFR (Andre et al., 2023; Bloom et al., 2023; Khetran & Mustafa, 2024; Wiegand et al., 2022).

1.5 Immune response to viral infection

The first line of defence to a viral infection is the innate immune system. It consists of a broad range of defences including barriers, pathogen detection, inflammation and immune cell recruitment. The innate immune response is not dependent on immune recognition by lymphocytes and can therefore act on pathogens which the body has not yet been exposed. As an infection progresses the adaptive immune response works to combat and clear the infection, and the immunological memory helps protect against future infections (Parija, 2023).

1.5.1 Innate immune response

Once the virus has gained entry into a mucosal membrane, usually the nasal epithelial cells of the upper respiratory tract, it will encounter a number of defences. First among these are mucosal barriers which trap the pathogen. Mucus consists of mucins, large glycoproteins, produced by goblet cells. It acts as a physical barrier and contains antimicrobial peptides which impair the virus. Antimicrobial peptides such as defensins and cathelicidins can disrupt viral envelopes and inhibit replication, and surfactant proteins A and D in the alveolar space can bind viral glycoproteins, enhancing clearance by immune cells. These innate molecules are an important first line of defence before the adaptive immune system is activated. (Ejemel et al., 2020; Macpherson et al., 2011; Parija, 2023).

1.5.2 Adaptive immune response

Adaptive immunity to an acute viral infection generally takes ~7 days to initiate with the peak response being achieved in ~14 days, in many cases (Mueller & Rouse, 2008). The

adaptive immune response to SARS CoV-2 can be broken down into two main mechanisms: B cell mediated, and T-cell mediated. In the B cell mediated response a viral antigen is recognised by a B cell which then secretes antibodies to neutralise the virus. In the T-cell mediated response T cells recognise and destroy infected cells and assist B cells in their function.

Within the mucus, immunoglobulin A (IgA) antibodies are also present. IgA antibodies are produced by plasma cells in mucosa-associated lymphoid tissue and are secreted into the mucus as IgA dimers. In a SARS-CoV-2 infection, the IgA can bind to the S protein, inhibiting its binding to ACE2 and thereby inhibiting viral infection. The dimeric structure of IgA allows it to bind two virus particles with the antigen binding site of each monomer. Cilia in the respiratory tract act to remove the IgA-bound pathogen-laden mucus (Ejemel et al., 2020; Macpherson et al., 2011; Parija, 2023).

In the adaptive immune response, the primary role of B cells is to recognise viral epitopes, differentiating into plasma cells that produce specific neutralising antibodies to clear the infection. B cells are produced in the bone marrow and migrate to lymphoid tissues including mucosal associated lymphoid tissues. B cell activation occurs when a specialised membrane bound protein, a B cell receptor (BCR), binds to an antigen. There are two pathways for B cells to become activated and begin producing antibodies: T cell-dependent (TCD) and T cell-independent (TCI). Activation via the TCI pathway occurs when antigens bind to the BCRs, multiple BCRs must bind to one antigen to start off the signalling cascades that cause B cells to proliferate, differentiate and produce Immunoglobulin M (IgM) antibodies. In this response IgM antibodies are the first type of antibody produced to an infection, they are short lived but able to provide limited protection until more specialised antibodies can be produced. For the TCD pathway, antigens bound to multiple BCRs are endocytosed, processed into smaller peptides and presented on the B cell surface by major histocompatibility complex (MHC) molecules. These peptide derived antigens presented on the MHC molecules can interact with a subtype of T cells, which then trigger B cells to proliferate, switch classes, undergo somatic hypermutation and produce high affinity antibodies (Kaduskar et al., 2022; Sette & Crotty, 2021).

1.5.3 T cell activation

Before a T cell can take part in the TCD pathway, to activate B cells, it must first be primed by antigen presenting cells, usually dendritic cells. This is because T cells exist in a naive state until they encounter antigen derived peptides via other cells. During an infection dendritic cells present peptides from antigens they have endocytosed on MHC molecules and T cells can bind to peptides displayed this way through their membrane bound T cell receptors (TCR), this binding is a key activation step. Activation of a T cell is completed through the release of co-stimulatory molecules

by the dendritic cell. Detection of these co-stimulatory molecules and cytokines, along with TCR binding to an MHC presented peptide, trigger the differentiation of the naive T cell into a specialised T cell. dentritic cells can activate two different classes of T cells by displaying the peptide on two different MHC molecules. To activate CD8+ T cells (cytotoxic T cells) the peptide is displayed on MHC class I molecules and to activate CD4+ T cells (T helper cells) the peptide is displayed on MHC class II molecules. In the TCD activation of B cells a subset of the CD4+ T cells, called T follicular helper (Tfh) cells, bind to the peptides displayed on MHC class II molecules through their TCRs. This interaction can only take place with a B cell that is displaying the same peptide that was used to prime the T cell (Tangye et al., 2000).

Once this interaction takes place the B cell will undergo proliferation, creating a large number of B cells specialised for the antigen which triggered the process. The population of activated B cells will also undergo somatic hypermutation where the genes encoding the BCR are mutated. This creates a population of B cells with BCRs that have variable binding to the antigen derived peptide used in the process. B cells with BCRs with high binding affinity to the antigen are preferentially selected for, increasing the specificity of BCR binding in the population of activated B cells. Subtypes of T cells, T helper 1 (Th1) cells, secrete cytokines which promote class switching in the activated B cells (Moss, 2022; Parija, 2023).

1.5.4   IgG antibodies

Class switching causes the B cell to switch from the production of IgM to producing other isotypes of antibodies, including Immunoglobulin G (IgG) antibodies. Class Switch Recombination (CSR) retains the same variable region of the antibody, so antigen binding remains the same, but the constant region of the antibody changes, switching the antibody isotype. There are a number of benefits associated with switching to IgG antibodies, a key one being that IgG antibodies have a much longer half-life than IgM antibodies, so they provide longer term immunity.

Antibodies are secreted molecules that help mediate the immune response. They are composed of four protein chains, two heavy and two light. The heavy chains used for each antibody produced by a single B cell are identical and so are both light chains. They have two distinct regions: fragment antigen binding (Fab) and fragment crystallisable (Fc) (Figure 12.). The Fab region confers antigenic binding specificity, and the Fc region interacts with Fc receptors on the surface of immune cells. In IgG, the heavy chain consists of the globular domains variable heavy (VH), constant heavy (CH) 1, CH2 and CH3. The light chains consist of V light (L) and CL globular domains. The Fab of the assembled antibody is made of VH and CH1 of the heavy chain and Variable light (VL) and constant light (CL) if the light chain. The Fc consists of the C2 and C3 of the heavy chains. The Fc and Fab are connected with the hinge region of both heavy chains (Janeway, Travers, Walport, & Shlomchik, 2001; Kaduskar et al., 2022).
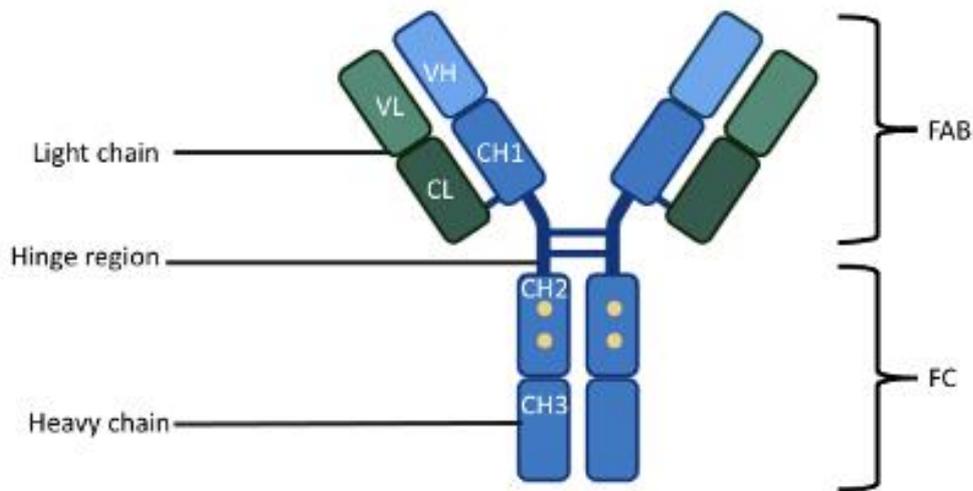
Figure 12. IgG antibody structure
The outer part of the fragment antigen binding (FAB) region is composed of the light chain. The inner part of the FAB and entirety of the fraction crystallisable (FC) is composed of the heavy chain. The FC is composed of the heavy chain regions Constant region 2 (CH2) and constant region 3 (CH3). The inner ppart of the FAB is composed of the heavy chain variable region (VH) and constant region 1 (CH1). The outer part of the FAB is composed of the light chain variable region (VL) and constant region (CL). The hinge region provides flexibility in the FAB, enabling epitope binding.

1.5.5 IgG Antibody diversity

Diversity in antibody binding is achieved through the recombination of genes encoding for polypeptides that make up the Fab section of the antibody. Recombination of different gene segments: variable (V), diversity (D) and joining (J) antibody is used to modify antigen binding of produced antibodies. The heavy chain Fab uses V, D and J and the light chain Fab uses V and J. Upon activation by a T cell, the B cell will switch antibody isotypes by using a different selection of VDJ genes. In the case of IgG heavy chain, it uses the Cγ region which has 38 to 46 V gene segments, ~23 D and ~6 J and only one of each V, D and J genes are selected. For the IgG light chain the B cell will use genes from only one of two available loci κ or λ. The κ locus has ~40 V and 5 J and the λ locus has ~30 V and ~4 J. Once the locus has been selected the B cell will use only one of each V and J for the light chain. This VDJ recombination increases the diversity of antibody binding in the antibodies produced by the activated B cell population. Somatic hypermutation triggered in the B cell also leads to further binding diversity of antibodies by enhancing mutation in genes encoding the V region of the antibodies. Diversity is further achieved by addition and deletion

mutations at the sections of the genes coding for the junctions between V, D and J (Janeway, Travers, Walport, & Shlomchik, 2001; Leder, 1982).

1.5.6 Antibody dependent enhancement

While most of the antibodies produced will, to a varying degree, bind and neutralise the virus it is possible for non-neutralising or sub-neutralising antibodies to enhance the infection, known as antibody-dependent enhancement (ADE). ADE in immune cells occurs through two pathways: Fc receptor mediated and complement mediated. In Fc receptor mediated ADE an antibody binds to the virus and the FC region of the antibody binds to a FC receptor of an immune cell. This binding triggers the immune cell to endocytosis the virus, this allows the virus to not only escape destruction by the immune cell but also infect it. In complement mediated ADE, the antibody-virus complex activates the complement system leading to opsonization of the complex which is then taken up by cells via complement receptors. Once inside the cell the virus escapes destruction and infects it. ADE can increase the viral load by infecting more cells and it also complicates vaccine development as antibodies raised through immunisation may enhance infection.

Antibody dependent enhancement can also enhance the infection of non-immune cells, particularly endothelial and epithelial cells. ADE without FC receptor or complement mediation can occur through antibody mediated enhanced membrane fusion or enhancement binding of the antibody-virus complex with cell surface receptors. Non-immune cell ADE has been reported in the case of infection with zika virus, Ebola virus and HIV-1. ADE has not been definitively proven in SARs CoV-2 infections; in some severe SARS CoV-1 infections IgG response has been linked with disease severity (Lee et al., 2006). In the case of ADE of SARS CoV-1 it was found that ADE was mediated by diluted antibodies against spike proteins in infectivity assays (Wen et al., 2020).

1.6 SARS CoV-2 vaccines

Pfizer-BioNTech, Moderna, and AstraZeneca each developed mRNA-based vaccines for SARS CoV-2: BNT162b2, mRNA-1273 and AZD1222 (ChAdOx1-S) respectively. These vaccines used a delivery method for SARS CoV-2 S encoding mRNA to be translated into proteins in the cell which, through the adaptive immune response, raised long term protection against the virus. The Pfizer-BioNTech vaccine, BNT162b2, and the Moderna mRNA-1273 vaccine both used lipid nanoparticles to deliver mRNA. The AstraZeneca AZD1222 vaccine used a non-replicating chimpanzee adenovirus as the vector to deliver mRNA. As the virus evolved the protection offered by monovalent vaccines became less efficacious and bivalent vaccines were developed. Bivalent vaccines generally utilize the original WT strain and newer variants such as Delta or Omicron (Chakraborty et al., 2023; Lapuente et al., 2023).

1.7   Research question and aims

The research question this study set out to address is: how does the binding and neutralization capacity of patient-derived polyclonal antibodies vary across different SARS-CoV-2 spike protein variants, and how do mutations in these proteins affect antibody efficacy?

Antibody binding ELISA

The primary aim for this experiment was to evaluate the binding ability of polyclonal antibodies present in patient sera samples to SARS CoV-2 spike proteins. These experiments were designed to assess the extent to which patient derived antibodies differentially recognised and interacted with spike protein of different SARS CoV-2 variants. The results of these experiments should contribute to the understanding of immune response variability and cross reactivity of antibodies of those with severe COVID-19. The degree of antibody binding was quantified by measuring the antibodies bound to immobilised SARS CoV-2 spike proteins.

Neutralisation of binding ELISA

The aim for these experiments were to assess the capacity for patient derived polyclonal antibodies to neutralise the interaction between SARS CoV-2 spike protein and its primary target receptor ACE2. The degree of inhibition of binding was quantified by measuring the reduction of ACE2-spike interaction following incubation of spike with patient antibodies.

Neutralisation of cell entry assay

The aim for these experiments were to evaluate the capacity for patient derived polyclonal antibodies to neutralise the entry into a cell of SARS CoV-2 spike pseudotypes. This study intended to elucidate the potential of these antibodies to block viral entry and assess their protective capacity. By comparing the effectiveness of these antibodies at neutralising entry of pseudotypes displaying different SARS CoV-2 spike variants, the effectiveness of cross-reactive antibodies in providing protection against different variants can be elucidated.

LEXSY expression system

The aim of these experiments was to utilise the LEXSY expression system for the production of SARS CoV-2 spike proteins, with human like post translational modifications. These proteins would then be used in applications such as antibody binding ELISA and neutralisation of binding ELISA to further study the binding characteristics of patient derived polyclonal antibodies. A second aim was to use this system to produce SARS CoV-2 spike proteins with single mutations and

combinations of mutations to understand how antibody binding is affected by individual mutations and combinations of mutations in our patient sera samples.

Novel computational approach to mutagenesis

The primary aim of this approach was to explore potential novel mutations in the spike protein of SARS CoV-2 that could affect the binding of antibodies. The design of this approach was to use synonymous sequences as initial templates for error prone PCR. A novel program was developed to generate synonymous sequences. These sequences would be used in combination with the LEXSY system to express mutant proteins for investigation in antibody binding ELISA and neutralisation of binding ELISA to further study the binding characteristics of patient derived polyclonal antibodies.

Thesis outline

Chapters 1 & 2 – Provide a background on SARS-CoV-2 biology and immune responses. They also describe experimental procedures and analytical methods used. Chapter 3 – Assays and Results:

- Presents the ELISA results measuring total and variant-specific IgG binding to SARS-CoV-2 spike proteins. It also compares antibody binding strength across Alpha, Beta, and Delta variants and explores associations with disease severity and immune status.
- Reports pseudovirus assay results assessing the serum neutralising activity against different spike variants. It evaluates how antibody binding correlates with functional neutralisation and identifies differences in potency between variants.
- Examines evidence of antibody-dependent enhancement in sera from patients with severe COVID-19 and assesses how enhancement varies across variants.

Chapter 4 – Discussion and future work: Integrates all findings, interprets their immunological significance, and discusses study limitations.

Chapter 2. Materials and Methods

2.1   Materials

2.1.1   Serum samples

Patient cohort

Samples were retrieved from patients who were hospitalised following COVID-19 and had not received immunization against SARS-CoV-2. Each sample was a diagnostic specimen, surplus to diagnostic requirements, taken from an individual patient. Of samples collected those taken forward for this study were selected for the following

criteria: COVID-19 diagnosis with positive PCR test prior to sample collection and sample collection date between 20 and 60 days of diagnosis.

These criteria were chosen so that samples are from patients confirmed to have COVID-19 by diagnostic PCR. At the time of sample collection some patients were diagnosed on symptoms alone, either in the absence of a positive PCR test or in the presence of a negative PCR test. Limiting testing to samples with positive diagnostic PCR test ensured that we did not include samples derived from patients with a different infection with similar symptoms. Samples taken before 20 days and after 60 days were excluded to maximise the probability of a patient seroconverting, allowing investigation into antibodies.

2.1.2 Cell Strains

***E. coli* NEB 5-alpha (New England biolabs, C2987I)**: F´ proA+B+ lacIq Δ(lacZ)M15 zzf::Tn10 (TetR) / fhuA2Δ(argF-lacZ)U169 phoA glnV44 Φ80Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17. Standard cell strain for plasmid amplification. Tetracycline resistance, Reduced endonuclease activity, reduced recombination activity.

VERO E6: *Chlorocebus aethiops* kidney epithelial cell line, genetically modified to express angiotensin-converting enzyme 2 (ACE2) receptor on their surface.

Hek293T: Modified Human embryonic kidney 293 cell line transformed with: SV40T-antigen, Adenovirus E1A and E1B genes.

2.1.3 Plasmids
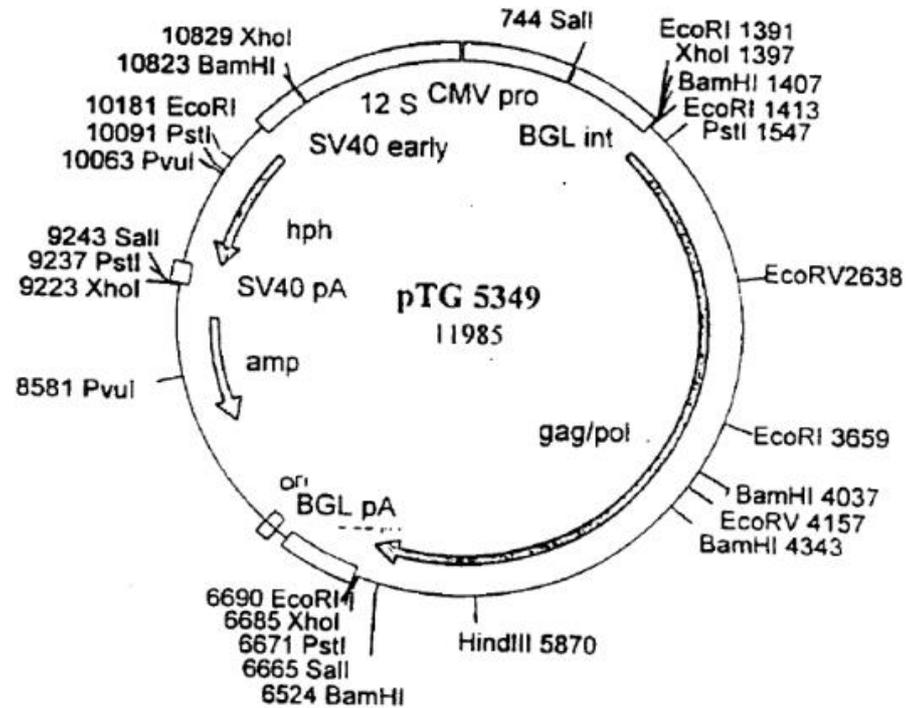
Figure 13. Plasmid map of phCMV5349 (ptg5349) used in pseudotyping

The 'packaging plasmid' which encodes for group specific antigen (gag) and polymerase (pol) genes of murine leukaemia virus (MLV). Gag encodes for a polyprotein which consists of matrix, capsid, nucleocapsid and p12, the essential structural proteins of the MLV pseudotype. Pol consists of enzymatic proteins required for producing pseudotypes. The CMV promoter is used to drive high levels of expression of gag/pol. Ampicillin resistance gene used for selection in plasmid amplification in *E. coli*. The plasmid does not encode for an envelope protein, this is provided in a separate plasmid. (Johnson et al., 2014).
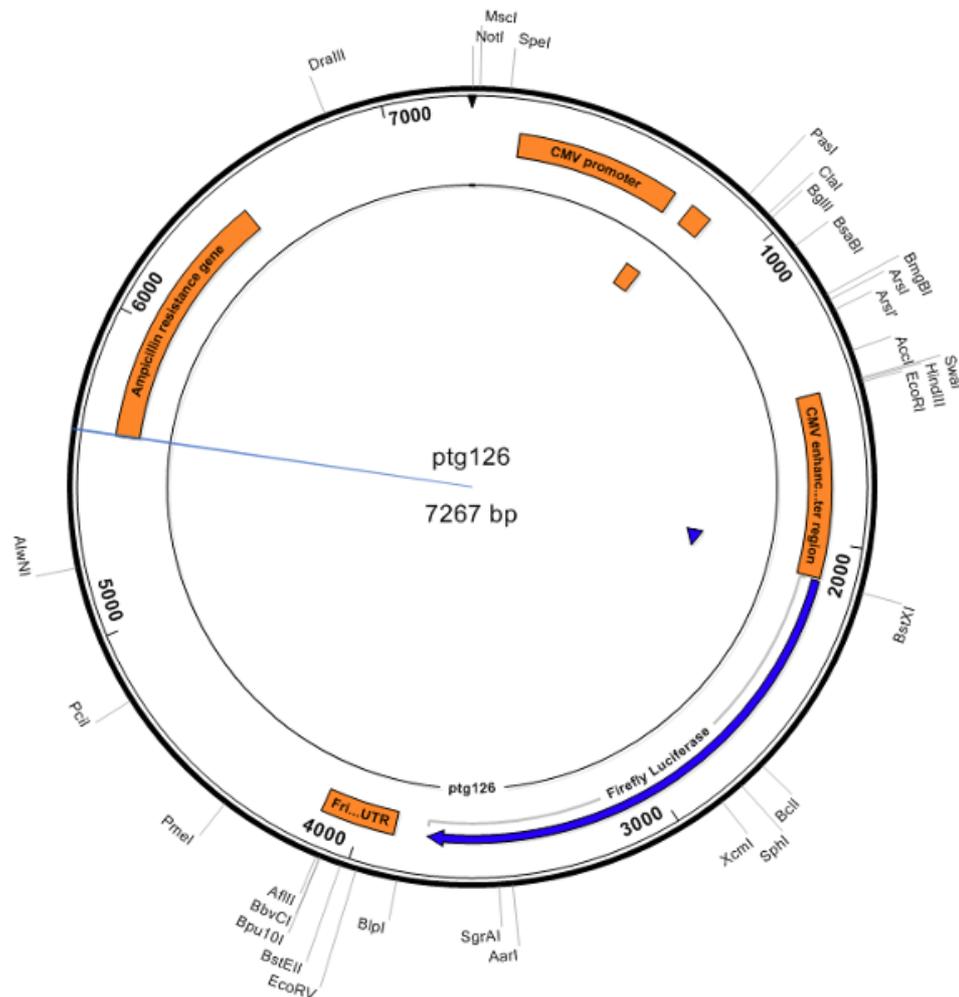
Figure 14. Plasmid map of pTG126 used in pseudotyping
Plasmid map of luciferase encoding plasmid used in pseudotyping for neutralisation of entry assays. RNA transcripts from this plasmid are contained within the released pseudotype, and the encoded luciferase is synthesised once the pseudotype infects a cell. The plasmid is designed to achieve high levels of luciferase expression in the host cell. Features on this plasmid include: ampicillin resistance gene used for strain selection during plasmid amplification in *E. coli,* CMV promoter to drive high levels of RNA expression in mammalian cells, CMV enhancer to increase the activity of the CMV promoter by interacting with host transcription factors.
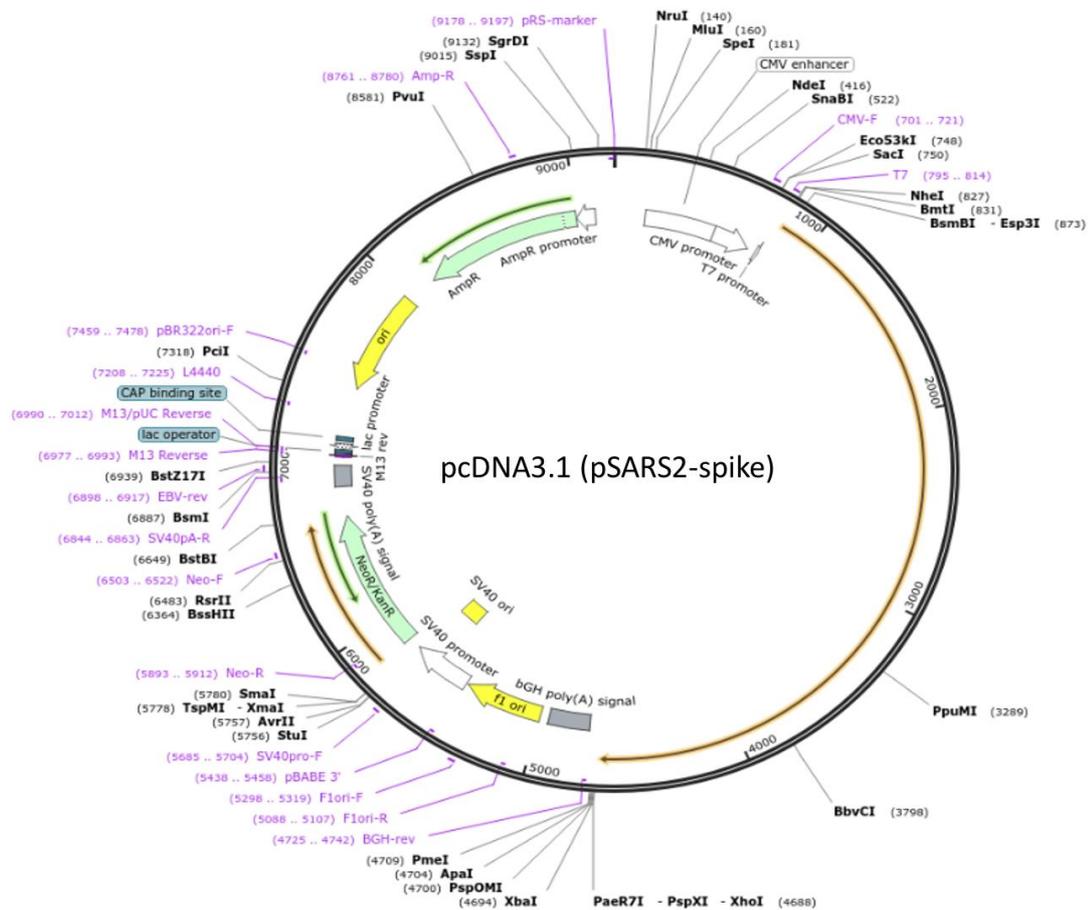
Figure 15. Key features of pSARS2-spike plasmid with spike gene sequence insert shown in orange.

Plasmid map of envelope protein plasmid used in pseudotyping for neutralisation of entry assays. This plasmid is used for synthesis of the SARS CoV-2 spike proteins for pseudotype production in HEK293T cells. The plasmid is designed to be amplified in *E. coli* and to produce high levels of spike protein in mammalian cells. Features on this plasmid include: an origin of replication (ori) for plasmid amplification in *E. coli*, Ampicillin resistance gene (Amp R) for clonal selection in *E. coli*. Also included are the promoters CMV and T7 for high levels of expression. The T7 promoter makes use of the large T antigen present in HEK293T cells during pseudotyping. Finally bovine growth hormone polyadenylation signal (bGH Poly [A]) was included downstream of the spike sequence insert for correct termination of the spike mRNA transcript, mRNA stability and to promote efficient translation in mammalian cells.

2.1.4 Media

MEM non-essential amino acids solution (100X) (NEAA) (Gibco, 1140-035): L-Asparagine (1320.0 mg/L), L-Aspartic acid (1330.0 mg/L), L-Glutamic Acid (1470.0 mg/L), L-Proline (1150.0 mg/L), L-Serine (1050.0 mg/L)

OptiMem (Gibco, 31985-062): L-Arginine HCl (126.22 mg/L), L-Cystine 2HCl (31.29 mg/L), L-Glutamine (292.20 mg/L), Glycine (7.51 mg/L), L-Histidine HCl·H_2O (42.00 mg/L), L-Isoleucine (52.48 mg/L), L-Leucine (52.48 mg/L), L-Lysine HCl (72.57 mg/L), L-Methionine (14.92 mg/L), L-Phenylalanine (32.87 mg/L), L-Threonine (47.62 mg/L), L-Tryptophan (10.21 mg/L), L-Tyrosine 2Na·$2H_2O$ (43.79 mg/L), L-Valine (46.84 mg/L), Choline Chloride (1.00 mg/L), Folic Acid (1.00 mg/L), myo-Inositol (2.00 mg/L), Niacinamide (1.00 mg/L), D-Pantothenic Acid (Ca) (1.00 mg/L), Pyridoxal HCl (1.00 mg/L), Riboflavin (0.10 mg/L), Thiamine HCl (1.00 mg/L), Sodium Chloride (NaCl) (6800.00 mg/L), Potassium Chloride (KCl) (400.00 mg/L), Calcium Chloride ($CaCl_2$) (200.00 mg/L), Magnesium Sulphate ($MgSO_4$) (97.67 mg/L), Sodium Phosphate Monobasic ($NaH_2PO_4 \cdot H_2O$) (125.00 mg/L), Glucose (1000.00 mg/L), Phenol Red (pH Indicator) (10.00 mg/L), Sodium Bicarbonate ($NaHCO_3$) (2200.00 mg/L).

Dulbecco's modified Eagle's media, high glucose (DMEM) (Gibco, 119565-092):

OptiMem (Gibco, 31985-062): Glycine (30 mg/L), L-Arginine hydrochloride (84 mg/L), L-Cystine 2HCl (63 mg/L), L-Glutamine (584 mg/L), L-Histidine hydrochloride-H2O (42 mg/L), L-Isoleucine (105 mg/L), L-Leucine (105 mg/L), L-Lysine hydrochloride (146 mg/L), L-Methionine (30 mg/L), L-Phenylalanine (66 mg/L), L-Serine (42 mg/L), L-Threonine (95 mg/L), L-Tryptophan (16 mg/L), L-Tyrosine disodium salt dihydrate (104 mg/L), L-Valine (94 mg/L), Choline chloride (4 mg/L), D-Calcium pantothenate (4 mg/L), Folic Acid (4 mg/L), Niacinamide (4 mg/L), Pyridoxine hydrochloride (4 mg/L), Riboflavin (4 mg/L), Thiamine hydrochloride (4 mg/L), i-Inositol (4 mg/L), Calcium Chloride (CaCl2) (anhyd.) (200 mg/L), Ferric Nitrate (Fe(NO3)3"9H2O) (0.1 mg/L), Magnesium Sulfate (MgSO4) (anhyd.) (97.67 mg/L), Potassium Chloride (KCl) (400 mg/L), Sodium Bicarbonate (NaHCO3) (3700 mg/L), Sodium Chloride (NaCl) (6400 mg/L), Sodium Phosphate monobasic (NaH2PO4-H2O) (125 mg/L), D-Glucose (Dextrose) (4500 mg/L), Phenol Red (15 mg/L).

2.1.5 Other

Bovine serum albumin (BSA) (Sigma, A3059): Lysophilized BSA powder (≥98%).

Carbonate bicarbonate buffer capsules pH 9.6. 50mMol. - 100 cap. (Sigma, C3041).

Cell Dissociation Buffer, enzyme-free, PBS (PBS) (Gibco, D8537): KCl 0.2 g/L, KH_2PO_4 0.2 g/L, NaCl 0.8 g/L, Na_2HPO_4 1.15 g/L.

Dulbecco's Phosphate Buffered Saline (DPBS) (Gibco, A1285801): NaCl 8.0 g/L, KCl 0.2 g/L, Na_2HPO_4 0.9 g/L, KH_2PO_4 0.2 g/L.

Foetal bovine serum, heat inactivated (FBS) pH 7.2-7.6 (Gibco, 10500064): Albumin (BSA) 20–30 mg/mL, Glucose 5-10 mM.

Luciferase assay substrate (Promega, E151A): D-Luciferin, Adenosine triphosphate, $MgSO_4$, Tris buffer (pH 7.8), Dithiothreitol, Coenzyme A. (Proprietary formula).

Luciferase Cell Culture Lysis 5X Reagent (Promega, E1531): Tris-HCl, Detergents, stabilisers. (Proprietary formula), distilled water.

NeutrAvidin (Thermo scientific, 31000): NeutrAvidin 2 mg/mL.

Phosphate buffered saline (PBS) (Oxoid, BR 0014G): NaCl 8.0 g/L, KCl 0.2 g/L, $Na_2HPO_4.2H2O$ 1.44 g/L, KH_2PO_4 0.24 g/L.

Polyethylenimine (PEI), Linear, MW 25000, Transfection Grade (PEI 25K™) (Polysciences, 23966-1) Milli-Q water.

SARS-CoV-2 (B.1.1.7/Alpha) stabilized spike glycoprotein, biotinylated (Native antigen, REC31924-Biotin-100): PBS+20mM Tris, 1%BSA, trace amounts of ANS and $CaCl_2$ and 0.1% Proclin 950.

SARS-CoV-2 (B.1.351/Beta) stabilized spike glycoprotein, biotinylated (Native antigen, REC31926-Biotin-100): PBS+20mM Tris, 1%BSA, trace amounts of ANS and $CaCl_2$ and 0.1% Proclin 950.

SARS-CoV-2 (B.1.617.2/Delta) stabilized spike glycoprotein, biotinylated (Native antigen, REC31975-Biotin-100): PBS+20mM Tris, 1%BSA, trace amounts of ANS and $CaCl_2$ and 0.1% Proclin 950.

Tetracycline (Gibco, 16376552): Tetracycline Hydrochloride, distilled water.

Trypsin-EDTA (0.05%) (1x) (25300-062): EDTA 0.05%, Na2-EDTA 0.2 g/L, NaCl 8 g/L, KCl 0.5 g/L, Na_2HPO_4 0.2 g/L, KH_2PO_4 0.2 g/L.

Tween20 (Sigma, P1379): Polysorbate 20 (≥40%).

2.2   Methods

2.2.1   General Escherichia coli Cell Microbiology

2.2.1.1   *Escherichia coli* culturing on solid media

E. coli cultures were initially grown on solid LB media and a single colony selected to ensure clonality in downstream applications. The work area was cleaned with 2% Distel solution. Agar plates were prepared by melting a pre-prepared solid LB agar stock, pouring 200mL into a sterile beaker and allowed to cool to below 50 °C. For antibiotic selection LB was supplemented with tetracycline (Gibco, 16376552) (10 µg/mL). The agar must be cooled for safe handling and so the selection antibiotics are not denatured and deactivated once added. Using a Bunsen burner for sterile conditions, 20 mL of agar was pipetted into sterile 90 MM petri dishes. The agar was left to set at room temperature with the lids ajar to prevent condensation. A frozen stock of NEB5 alpha cells was retrieved from –80 °C storage and thawed on ice for 10 minutes. For inoculation, 50 µL of cell suspension was pipetted onto the agar surface. A disposable sterile loop and the quadrant streak method was used to disperse the cells across the agar surface. The quadrant streak method was used to improve the chances of selecting a single isolated clonal colony later. Plates were left to dry for 10 minutes to allow cells to adhere to the media. The plate was incubated inverted overnight at 37 °C for 12 hours in a static incubator. The plates were inverted to prevent condensation on the agar surface which could cause merging of cell colonies. For a negative control a plate was also inoculated with cells without the plasmid insert. Growth of colonies on the negative control plate would likely indicate

ineffective antibiotic selection and as a result the transformed *E. coli* could not be used.

2.2.1.2 Transformation of chemically competent *Escherichia coli*

Plasmids required for pseudotype protocols were made using plasmid amplification in *E. coli*. The plasmids pSARS2-WT, pSARS2-A, pSARS2-B, pSARS2-G and pSARS2-D were used in transformations. Each plasmid individually was transformed into separate *E. coli* cultures. One step to minimise contamination risk was each transformation was carried out on different days. For transformation of chemically competent E. coli, NEB5-alpha cells were used for their high transformation efficiency, high plasmid yield and compatibility with the plasmids used. An aliquot of NEB 5-alpha cells was retrieved from –20 °C short term storage and thawed on ice for 10 minutes while sterile microcentrifuge tubes were chilled on ice. Using a Bunsen burner for sterile technique, the cells were pipetted up and down to resuspend and 50 µL of cells were added to the pre-chilled microcentrifuge tube. Plasmid solutions were diluted to 20 µg/mL and 5 µL of plasmid solution (100 ng of plasmid) was pipetted into the cells and gently mixed. The transformation was incubated on ice for 30 minutes and then heat shocked at 42 °C for exactly 30 seconds, using a water bath. The transformation mix was then held on ice for 5 minutes. After 5 minutes 950 µL of SOC medium was added to the mixture and then incubated for 60 minutes at 37 °C in a shaking incubator at 250 rpm. Antibiotic selection plates were prepared as described in 1.8.1.1, with ampicillin (50 µg/mL) used instead of tetracycline (10 µg/mL). Selection plates were heated to 37 °C prior to inoculation.

2.2.1.3 Plasmid isolation from *Escherichia coli*

For plasmid isolation a midiprep kit (Qiagen, 12181) was used. To achieve sufficient plasmid amplification, a single transformed *E. coli* culture expanded in LB broth. For colony expansion 50 mL of LB broth was added to a 250 mL vented, baffled bottom, Erlenmeyer flask (Corning, 431407). For antibiotic selection the broth was supplemented with ampicillin at 50 µg/mL. Under sterile conditions using a disposable inoculating loop a single culture was picked and used to inoculate the flask. The culture was incubated at 37 °C in a shaking incubator at 250 rpm for 12 hours. Cells were harvested from the overnight culture by centrifugation at 6000 g for 15 minutes at 4 °C. The supernatant was discarded, and the cell pellet retained. The cells were resuspended in 4 mL of resuspension buffer (buffer P1). To extract the plasmids from the cells they were lysed with the addition of 4 mL lysis buffer (buffer P2). The solution was mixed by inverting six times and incubated at 25 °C for 5 minutes to allow for complete cell lysis by the SDS in the buffer. Neutralisation buffer (buffer P3) was retrieved from 4°C storage and 4 mL pipetted into the Lysate which was mixed by inverting six times. Potassium acetate in the neutralisation buffer neutralises the pH of the solution and also causes precipitation of SDS, proteins and chromosomal DNA. Plasmid DNA is not precipitated because it renatures rapidly due to its small size and its supercoiled structure has less surface

area for salt-DNA interactions. The lysate was centrifuged for 4 minutes 15,000 g at 4°C. The spin column was prepared by adding 4 mL of loading buffer (Buffer QBT) which prepared the silica membrane in the column for efficient plasmid binding. The plasmid containing supernatant was then transferred to a spin column.

2.2.2 General Mammalian Cell Microbiology

2.2.2.1 HEK293T passage

Cells lines are passaged to maintain continuous cell growth and maintain cell viability. Passaging is done when cells grow to 70 % to 80 % confluence while in growth phase before cells begin to experience contact inhibition which reduces cell viability. The growth medium is also replaced to remove any waste products of cell metabolism and to maintain optimal growth conditions. A biosafety cabinet (Class II) work area was cleaned first with 2% Distel solution and then with 70% solution of industrially denatured alcohol. A beaker containing 2% Distel was used to dispose of any liquid waste during the protocol. An aliquot of 2% Trypsin and an aliquot of CDMEM (DMEM [Gibco 119565-092], 10% FBS [Gibco 10500064], 1x NEAA [Gibco 1140-035]) were heated to 30 °C. Cells were inspected under an inverted light microscope to check for contamination, to check cell morphology and to estimate confluency before proceeding. The following steps took place inside the biosafety cabinet. The growth media was poured from the 75 cm^2 tissue culture flask carefully so as not to disturb the cell layer. PBS was then used to wash the cells by adding 10 mL to the flask carefully so as not to disturb the cell layer. The flask was gently rocked to distribute the PBS and then it was carefully poured off. To detach the cells 2 mL of trypsin was added to the cells and the flask rocked gently to distribute. After 2 minutes the flask was gently tapped to ensure cell detachment. Trypsinization was halted by adding 10 mL of CDMEM and detached cells were suspended in solution by pipetting up and down 5 to 10 times. To remove the trypsin the cells were then moved to a 10mL falcon tube removed from the biosafety cabinet to be centrifuged at 300 x G (soft acceleration and deceleration) for 7 minutes. Returning to the biosafety cabinet, the supernatant was carefully poured off and the cell pellet retained. The cell pellet was then resuspended in 10 mL CDMEM. A haemocytometer was used to check cell density. In a fresh 75 cm^2 tissue culture flask 10 mL of CDMEM was added. To seed the cells into a new flask 1.5×10^6 cells suspended in solution was added dropwise over the media to aid even dispersal. The flask was gently rocked to aid ensure even dispersal over the culture surface and incubated at 37 °C and 5 % CO$_2$.

2.2.2.2 VERO E6 passage

Passaging of VeroE6 cells was similar to the passage of HEK293T cells except VeroE6 cells were more adherent and required more trypsin, incubation at 37 °C to aid trypsinization and firmer tapping to detach the cell layer. For VeroE6 passage, first a biosafety cabinet (Class II) work area was cleaned with 2% Distel solution and then with 70% solution of industrially denatured alcohol. A beaker containing 2% Distel

was used to dispose of any liquid waste during the protocol. An aliquot of 2% Trypsin and an aliquot of CDMEM were heated to 30 °C. Cells were inspected under an inverted light microscope to check for contamination, check cell morphology and to estimate confluency before proceeding. The following steps took place inside the biosafety cabinet. The growth media was poured from the 75 cm^2 tissue culture flask carefully so as not to disturb the cell layer. PBS was then used to wash the cells by adding 10 mL to the flask carefully so as not to disturb the cell layer. The flask was gently rocked to distribute the PBS and then it was carefully poured off. To detach the cells 2-10 mL of trypsin was added to the cells and the flask rocked gently to distribute. The cells were incubated at 37 °C to aid trypsinization. After 10 minutes the flask was firmly tapped to ensure cell detachment. Trypsinization was halted by adding 10 mL of CDMEM and detached cells were suspended in solution by pipetting up and down 5 to 10 times. To remove the trypsin the cells were then moved to a 10mL falcon tube removed from the biosafety cabinet to be centrifuged at 300 x G (soft acceleration and deceleration) for 7 minutes. Returning to the biosafety cabinet, the supernatant was carefully poured off and the cell pellet retained. The cell pellet was then resuspended in 10 mL CDMEM. A haemocytometer was used to check cell density. In a fresh 75 cm^2 tissue culture flask 10 mL of CDMEM was added. To seed the cells into a new flask 1.0×10^6 cells suspended in solution was added dropwise over the media to aid even dispersal. The flask was gently rocked to aid ensure even dispersal over the culture surface and incubated at 37 °C and 5 % CO_2.

2.2.3   Assays

2.2.3.1   Viral Pseudotyping

To tests the efficacy of polyclonal antibodies present in serum collected from patients with severe COVID-19, pseudotypes were used as a safer alternative to live SARS CoV-2 virus. To produce the peusdotypes the cell line HEK293T was used as they have a high transfection efficiency, are permissive cell line, and are routinely used in pseudotype production. A total of 1.5×10^6 HEK293T cells were seeded into Primaria-treated 100 mm x 20 mm tissue culture dishes (Corning, 35380) with 10 mL CDMEM (DMEM [Gibco 119565-092], 10% FBS [Gibco 10500064], 1x NEAA [Gibco 1140-035]). The cells were incubated at 37 °C and 5 % CO_2 for 24 hours. After incubation cells were examined under microscope to check cell confluence was between 50-60 % for optimal transfection and to ensure there was no contamination. For transfection, the growth media was removed and replaced with to 7 mL OptiMem (Gibco 31985-062). Plasmids were prepared by adding 2 µg of each plasmid phCMV5349 (MLV-Gag/Pol), pTG126 (MLV-Luciferase) and the required pcDNA3.1+ plasmid (Thermo fisher, V79020) encoding the relevant SARS2-spike variant sequence into 300 µL of OptiMem. Separately a 6% solution of PEI 25kD (Polysciences 23966-1) in 300 µL Optimem was prepared. The PEI solution was then added to the plasmid-OptiMem to give a final volume of 600 µL containing 2 µg of each plasmid and 3% PEI. This method of preparing the plasmids is used to ensure even distribution of plasmids and PEI which allows for more uniform PEI-plasmid complex formation.

The plasmids solution was incubated for 1 hour at room temperature, then added dropwise evenly over the prepared cells, and mixed by gentle agitation. Cells were incubated at 37 °C and 5 % CO_2 for 48 hours. Supernatants containing the pseudotypes were collected and filtered through a 0.45 μM filter and stored at 4 °C overnight. Two negative controls were made by following the protocol but with elimination of one of the plasmids in each. In the first negative control only the MLV-Gag/Pol and MLV-Luciferase plasmids were used in the plasmid preparation step and the plasmid encoding the glycoprotein was omitted. This negative control confirms that pseudotype entry into the target cells is dependent on expression of the spike glycoprotein. In the second negative control MLV-Gag/Pol and the required pcDNA3.1-SARS2-spike plasmid were used, and MLV-Luciferase was omitted. This negative control was used to establish background luminescence not related to the expression of luciferase from the MLV-Luciferase plasmid.

2.2.3.2   Indirect ELISAs

To measure the binding affinity of antibodies to the SARS CoV-2 spike glycoprotein present in patient sera samples an indirect ELISA (Enzyme-Linked Immunosorbent Assay) was used (Figure 16.). First the ELISA plate was prepared with spike glycoprotein then patient sera samples were incubated with the immobilised spike. Binding of antibodies was detected using HRP conjugated secondary antibodies and TMB (3,3',5,5'-Tetramethylbenzidine) and the colorimetric reaction measured with a microplate reader. All pipetting and liquid handling steps were carried out by an Opentrons OT-2 liquid handler pipetting robot to increase precision and accuracy. First the clear Nunc Maxisorp 112uL flat bottom 384 well plate (Thermo fisher, 464718) was coated by adding 20 μL of 5.0 μL/mL neutravidin, dissolved in 0.05 M carbonate bicarbonate buffer, to each well. The plates were sealed and incubated at 4 °C for 12 hours for optimal adsorption of neutravidin molecules to the well surface. Excess neutravidin buffered solution was removed and washed three times with 80 μL of PBST (0.05% Tween20 in 1x PBS solution) in each well. After washing plates were blocked by adding 50 μL 3% BSA in PBST solution to each well, sealed and incubated at 4 °C overnight. Plates were washed three times with 80 μL PBST and 20 μL SARS CoV-2 spike antigen diluted to 1 μg/mL in 0.05 M carbonate bicarbonate buffer. Full length spike glycoproteins used for separate assays were SARS-CoV-2 (B.1.1.7/Alpha) stabilized spike glycoprotein, biotinylated (Native antigen, REC31924-Biotin-100), SARS-CoV-2 (B.1.351/Beta) stabilized spike glycoprotein, biotinylated (Native antigen, REC31926-Biotin-100), and SARS-CoV-2 (B.1.617.2/Delta) stabilized spike glycoprotein, biotinylated (Native antigen, REC31975-Biotin-100). The plates were incubated at 25 °C for 30 minutes before the spike solution was removed and the plates were washed three times with 80 μL of PBST. Patient serum samples were defrosted on ice for 10 minutes and diluted 1:200 in 3% BSA in PBST solution. To duplicate wells, 20 μL of diluted serum was added and incubated for 30 minutes at 25 °C. Diluted serum was removed and the wells were washed three times with 80 μL PBST. To detect patient antibodies that had bound to the spike glycoprotein 20 μL of goat anti-Human IgG (Fc specific) HRP (horseradish

peroxidase) conjugated antibody (Sigma, A0170-1ML) diluted to 10 µg/mL in PBST was added to each well and incubated at 25 °C for 30 minutes. After incubation the plates were washed three times with 80 µL PBST and 100 µL of 0.4 mg/mL TMB ELISA substrate (Thermo fisher, 34021) was added to each well. After 20 minutes incubation at 25 °C, 25 µL of 2N H_2SO_4 was added to halt the HRP enzymatic reaction by denaturing the HRP. The plates were incubated for 10 minutes at 25 °C before absorbance was measured at 450 nm in a microplate reader. Three positive controls of serum with known binding characteristics were used alongside a negative control of no serum. To compare the binding characteristics quantitively 12 dilutions of a standard control serum (Working standard NIBSC Anti-SARS-CoV-2 antibody diagnostic calibrant. NIBSC, 20/162) were used. Starting at 1:200 a 12-point 1:2 serial dilution was performed with a final concentration of 1:409,600.
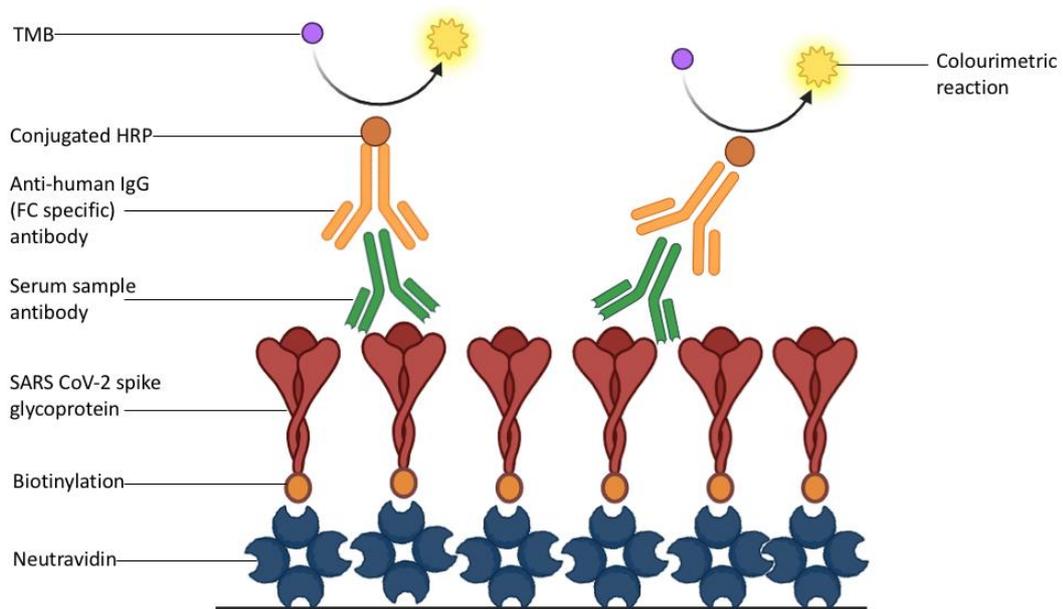
Figure 16. Diagram of antibody-spike binding ELISA
Neutravidin is immobilised on the plate surface where it captures the biotin tag of the spike protein. The captured spike proteins are orientated by the biotin tag to allow binding of serum samples antibodies. Anti-human IgG is shown binding to the FC region of the human antibodies. The conjugated HRP enzyme catalyses the colourimetric reaction of TMB substrate which can be measured to determine the total antibody binding of the serum sample.

2.2.3.3 Neutralisation of Cell Entry Assays

The neutralisation of entry (NOE) assay was developed to detect spike specific antibodies impacting on entry of SARS CoV-2 pseudotype into the cell. Vero E6 cells were prepared by dilution to a concentration of 1×10^5 VeroE6 cells mL^{-1} in CDMEM using the same method as described in VERO E6 passage method. In a sterile biosafety cabinet 100 µL of the VERO E6 cell suspension (1×10^4 cells) were added to

each well of a white 96-well tissue culture treated plate (Flat Bottom Polystyrene TC-treated, Corning, 3879). The cells were incubated at 37 °C and 5 % CO2 for 24 hours to grow to an appropriate confluency. Serum samples were diluted 1:100 in PBS into and 10 µL of this dilution was added to 90 µL aliquots of pseudotype suspension and incubated for 40-60 minutes. This incubation step was necessary to allow time for the antibodies present in the serum to bind to the spike glycoprotein. Serum and pseudotypes aliquots were manually agitated prior to pipetting to ensure even dispersal of solution contents. Media was carefully removed from the VeroE6 cells without disturbing the cell monolayer and, in triplicate, 90 µL of the serum/pseudotype mixture was pipetted to individual wells of the white 96-well plate and incubated at 37 °C and 5% CO2 for 72 hours. The inoculum was removed from each well and replaced with 50 µL of 1 x cell lysis solution (Promega, E1531), the plate was rocked for 30 minutes and shaken vigorously for 15 seconds. Luciferase activity was measured by recording luminescence at 560 nm and 3600 gain after adding 50 µL luciferase substrate (Promega, E151A). During measurement the plate was shaken at 300rpm in an orbital pattern and emissions were recorded over a 1 second period. Luciferase was added automatically by the plate reader injecting 50 µL Luciferase assay substrate at 300 µL/second. Positive controls with no serum and negative controls of pseudotype particles without a glycoprotein and pseudotype particles were used on each plate. a valid assay required two criteria. First the absence of neutralisation in the positive control and, second, low luminescence in the negative controls. This assay was repeated with four serial dilutions of sera with a two-fold dilution series, from 1:100 to 1:1200. Triplicate repeats were used for every condition to increase precision and accuracy and to ensure reproducibility. Triplicates of the controls were used as these measurements were used for normalisation in data analysis. The decision to not use titration of pseudotypes was based on consistency of pseudotype input for each plate and the use of internal plate controls.

2.2.4 Serum sample preparation

Surplus diagnostic specimens from inpatients admitted to Nottingham University Hospitals National Health Service (NHS) Trust Queen's Medical Centre. Approval for use of anonymized clinical data was provided by the NHS Health Research Authority (HRA) and Health and Care Research Wales (reference number 20/HRA/4843). Samples were determined not to be relevant materials in line with the Human Tissue Authority. Risk assessments were approved by the United Kingdom Health and Safety Executive (reference number CBA1.470.20.1). Samples were obtained via an agreement with the Nottingham University Hospitals National Health Service (NHS) Trust Queen's Medical Centre clinical pathology department.

Serum samples were initially stored at 4° C for 24 hours and then inactivated with the WHO–approved protocol: In a class I biosafety cabinet samples were treated with

1% triton x100 (SLS, X100) diluted in PBS. Samples were then incubated for 4 hours at 25 °C. treated samples were stored at -20 °C for short term storage.

Chapter 3. Assays and Results

3.1 Antibody binding ELISA

3.1.1 Background

3.1.1.1 ELISA overview

The Enzyme linked immunosorbent assay (ELISA) is a commonly used immunological assay to detect and quantify proteins, antibodies, and hormones in biology. It is a highly sensitive assay and highly versatile with a high number of enzyme conjugated antibodies, detection substrates, and conjugated proteins widely available. The ELISA technique employed in this study was a modified sandwich ELISA, first described in 1971 by Peter Perlmann and Eva Engvall (Engvall & Perlmann, 1971). In this assay the surface was prepared with NeutrAvidin to capture biotinylated spike glycoproteins. a sample of serum from patients with severe COVID-19 was added to the well and antibodies bound to the captured antigen. Bound antibodies were detected using a HRP conjugated detection antibody, and the substrate for the conjugated detection enzyme, TMB, was added which produced a colour change. This colour change could be measured and the relative binding of antibodies in the samples could be calculated. Accurate quantification of binding antibodies in samples using binding antibody units (BAU) was possible by using a control serum with known binding characteristics and generating a 12-point curve to interpolate results.

3.1.1.2 Spike protein

SARS CoV-2 entry into host cells is mediated by the spike glycoprotein. Following binding of Spike to the cell surface receptor ACE2 the virus is able to undergo a complex multistep process to achieve membrane fusion and infect the cell. Antibodies that bind to the S glycoprotein can influence effectiveness of viral entry to host cells. Antibodies binding to the S RBD in particular can prevent interaction with ACE2 and inhibit viral entry, effectively neutralising the virus. Antibodies that bind to other regions of S may also inhibit entry by affecting receptor binding or influencing the conformation change of S necessary for membrane fusion and inhibiting infection. Antibodies that bind to S and affect receptor interaction and conformational change may however also act to enhance elements of the multistep process involved in membrane fusion and enhance infection, known as antibody dependent enhancement.

In SARS CoV-2, the virus spike is present on the envelope as a homotrimer. In this ELISA the captured antigen used was an S monomer. This is a limitation of the assay as sample antibodies were raised against spike homotrimers and may bind to epitopes that only exist in the homotrimeric form. These assays were conducted with

S monomers due to availability at the time of experiments. The virus emerged and spread rapidly, evolving into new variants at a rapid rate and so commercial availability of appropriate S variant proteins was limited at the time of these experiments.

3.1.2 ELISA setup

NeutrAvidin was used to capture biotinylated spike glycoproteins. Biotinylated S proteins were used because the orientation of the protein can be controlled with the placement of the biotinylation and biotinylation does not affect the properties of spike protein, which is important in this assay.

To detect and quantify binding, goat anti-human IgG antibody conjugated with HRP was used with TMB. Goat derived antibodies were chosen to reduce the possibility of cross-reactivity. HRP and TMB were used for detection due to the high sensitivity of HRP, which amplifies the signal making small amount of bound antigen detectable. Once added to HRP the TMB undergoes a colour change from colourless, to blue once the reaction begins and yellow on the addition of H_2SO_4, which denature the HRP and stops the reaction. The reaction can be quantified by measuring absorbance at 450 nm using a plate reader. The absorbance is proportional to the yellow colour, which is proportional to the bound HRP-conjugated antibodies, which in turn in proportional to the bound antibodies from the serum sample. This allows for precise quantification of bound anti-spike antibodies in the patient serum sample.

To accurately quantify antibody binding a standard reference serum was used - working standard NIBSC Anti-SARS-CoV-2 antibody diagnostic calibrant, NIBSC, 20/162. The standard reference serum was calibrated to an initial concentration of 1000 BAU/mL. This was serial diluted in a 12 step, two-fold dilution series, resulting in a series of concentrations: 1000, 500, 250, 125, 62.5, 21.25, 15.63, 7.81, 3.91, 1.95 and 0.49 BAU/mL. These concentrations were used in the ELISA and the results used to generate a standard curve for quantification of antibody binding in samples. The absorbance values from the samples were compared to the standard curve to interpolate antibody binding concentrations in BAU/mL.

3.1.3 Aims

The aims of these experiments were to quantify the level of antibody binding to SARS CoV-2 spike glycoprotein variants in patients with severe COVID-19. Severe COVID-19 cases are often associated in a robust secondary immune response and characterised by high levels of antibodies circulating. The binding characteristics and quantify of binding antibodies can vary between individuals. A secondary aim was to investigate how antibodies raised in response to a severe infection bound to different variants of the SARS CoV-2 S glycoprotein. The rapid emergence and spread of highly infective variants presented a significant possibility of re-infection with a new variant. Determining if antibodies raised in response to one variant were cross

reactive with different circulating variants and their binding affinity is crucial in understanding if the virus has developed immune escape mechanisms.

3.1.4 Results
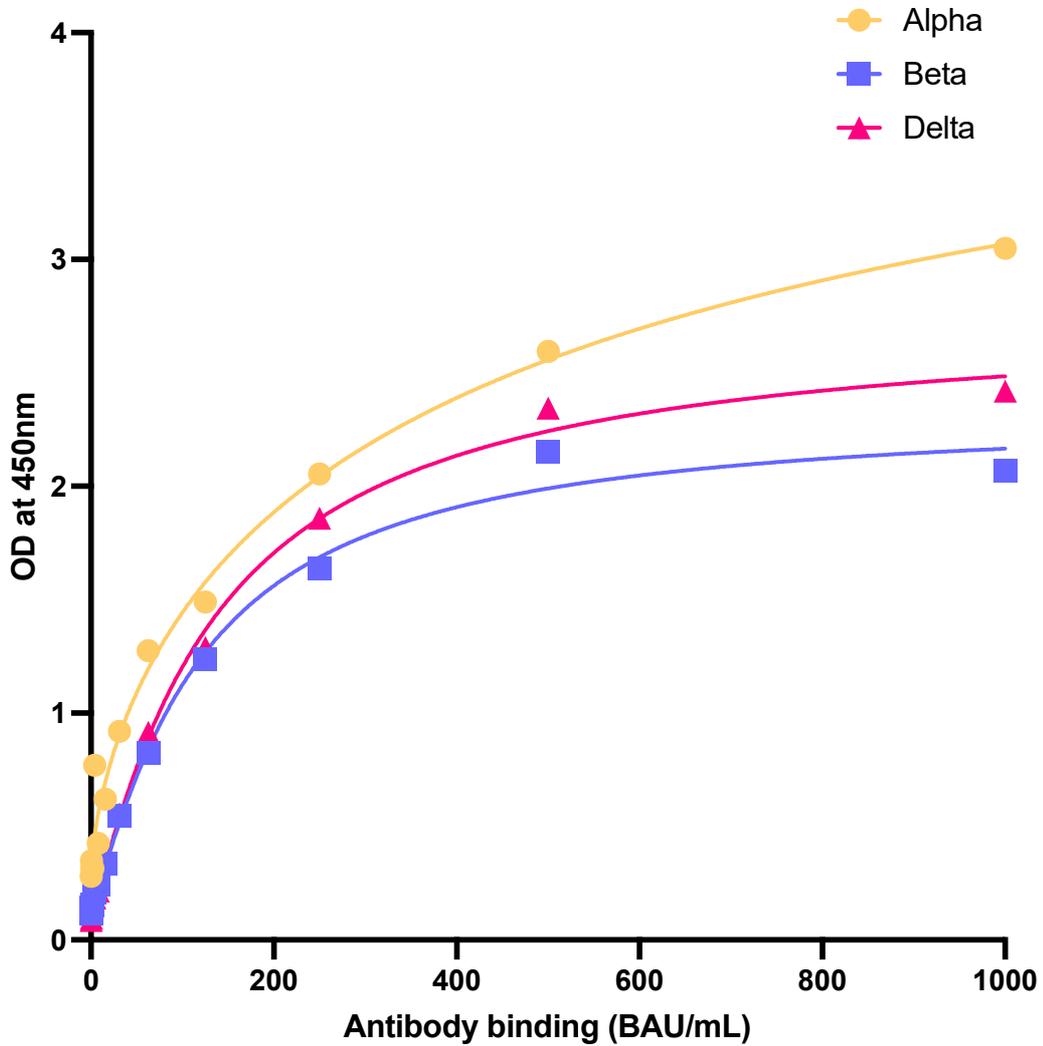
Figure 17. Sigmoidal 4-parameter logistic model binding curves for NIBSC Anti-SARS-CoV-2 antibody diagnostic calibrant 20/162, against SARS CoV-2 variants. An ELISA was performed, and absorbance (OD450) was measured, the results were fitted using a Sigmoidal 4-parameter logistic regression model. The generated curves were used to quantify antibody binding of samples by converting OD measured to BAU/mL in patient samples.

Figure 18. Antibody binding activity (BAU/mL) for patient sera samples as quantified by ELISA against spike variants.

Serum samples from 6 individuals admitted to hospital with COVID-19 were tested for antibody binding affinity against three SARS CoV-2 spike protein variants: Alpha, Beta and Delta. An ELISA was performed, and absorbance (OD450) was measured to quantify binding affinity, compared to a 12-point curve standard control to determine BAU/mL. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. ns is non-significant, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Two-way ANOVA with post hoc Tukey's test was used to analyse the anybody binding activity between spike variants within each sample (Figure 18.).

Sample 974 showed antibody binding across all three variants below 150 BAU/mL indicating low levels of antibody binding and similar cross reactivity with different spike variants. There was no significant difference in antibody binding between variants.

Sample 977 showed a difference in antibody binding between the three variants. A value of 1016.56 BAU/mL suggests a high level of antibody binding with the spike Alpha variant. 977 showed significantly lower binding for Beta (P < 0.0001) and Delta (P < 0.0001) compared to Alpha with values of 211.81 and 556.75 BAU/mL, respectively. For 977 there was significant difference in antibody binding between Beta and Delta (P < 0.0001).

Sample 980 had similarly low total antibody binding across all three variants at 111.98, 44.53 and 35.22 BAU/mL for Alpha, Beta and Delta, respectively. Sample 980 had no significant difference in binding across the variants.

Sample 981 showed a difference in cross reactivity across the three variants with moderate antibody binding to Alpha at 557.71 BAU/mL. 981 showed low antibody binding for Beta at 60.82 BAU/mL and Delta at 291.35 BAU/mL. Sample 981 had significant difference in antibody binding between Alpha and Beta (P <0.0001), Alpha and Delta (P = 0.0003) and between Beta and Delta (P = 0.0008).

Sample 1024 had high binding of antibodies to all three variants at 980.69, 1084.72 and 1019.15 BAU/mL for Alpha, Beta and Delta, respectively. This would indicate strong cross reactivity of binding antibodies across the three variants. Sample 1024 showed a significant difference in antibody binding only between Alpha and Delta (P = 0.0003). For sample 1024 the Beta and Delta variants only contain 1 replicate so while there does appear to be significant difference between the variants the results for this ELISA are unreliable for this sample.

Sample 1100 had a high level of antibody binding with 759.65 BAU/mL, low levels of antibody binding for the Beta variant at 184.95 BAU/mL and moderate to low levels of antibody binding for Delta at 364.17 BAU/mL. sample 1100 showed significant differences in antibody binding between Alpha and Beta (P <0.0001), Alpha and Delta (P = 0.0073) and, between Beta and Delta (P = 0.0073).

3.2   Neutralisation of binding ELISA

3.2.1   Overview

In order to assess the neutralisation capabilities of patient sera on the interactions between SARS CoV-2 spike and ACE2, an additional ELISA was developed. While the binding ELISA was effective at distinguishing binding antibodies, it does not evaluate the functional capabilities of the antibodies to block spike binding to ACE2. Antibodies that are capable of neutralising binding between spike and ACE2 would likely have a neutralising effect *in vivo*, since interaction between ACE2 and spike is one of the entry pathways of SARS CoV-2. A neutralisation of binding (NOB) ELISA was designed to model the conditions of the early stages of the complex multistage process of SARS CoV-2 cell entry and the effectiveness of patient antibodies in neutralising this process.

3.2.2   Assay overview

The NOB ELISA is a competitive ELISA where patient antibodies, present in serum samples, would bind to spike glycoprotein and interfere with spike RBD interaction with ACE2 peptidase domain (Figure 19.). In a high-binding, flat-bottom, clear

polystyrene 96-well plate (Corning, 9018) each well was first coated with Neutravidin (Thermo Scientific C31000). The neutravidin was used to immobilise recombinant hexa-his tagged human ACE2 (Sino Biological, 10108-H08H). Patient sera was then incubated with recombinant SARS-CoV-2 spike S1 mouse FC tagged protein (Sino Biological 40591-V05H1) where binding antibodies present in the serum samples would bind to spike. The spike-antibody mix was then added to the wells containing immobilised ACE2. In the absence of antibodies that neutralise spike-ACE2 interactions, spike should interact with ACE2 and remain in the wells for detection. In the presence of antibodies that neutralise spike-ACE2 interactions, fewer spike proteins should interact with ACE2 and remain in the wells for detection. Quantification of the spike protein remaining in the wells was achieved by detection of the mouse FC tag of the spike protein using goat anti-mouse IgG HRP-conjugated antibody (Chemicon AP 308P). The HRP conjugate was used to catalyse colourimetric reaction of TMB and the colour change quantified by measuring absorbance at 450 nm in a plate reader.
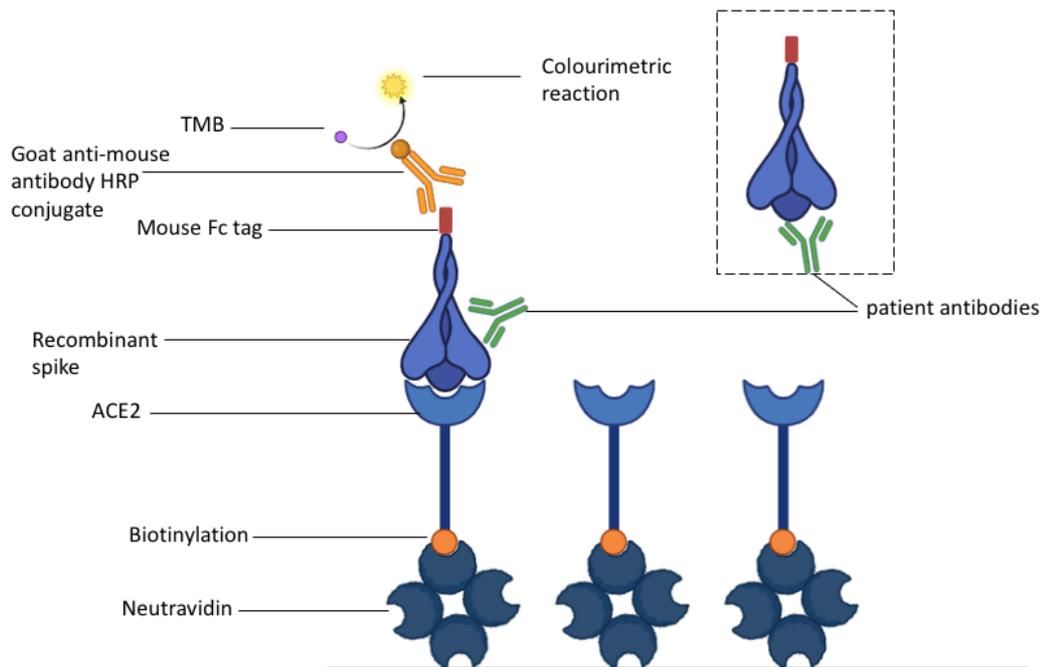
Figure 19. Diagram of neutralisation of binding ELISA
Neutravidin is immobilised on the plate surface where it captures the biotin tag of the ACE2 protein. The captured ACE2 proteins are orientated by the biotin tag to allow interaction of the peptidase domain with spike RBD. Spike is incubated with patient samples and patient antibodies bind. The antibody-spike mixture is added to the immobilised ACE2. Spike proteins that can bind with ACE2 will remain after washing steps and patient antibody-spike complexes that cannot bind to ACE2 (inside dashed box) will be removed by washing steps. Goat anti-mouse antibodies are shown binding to the mouse Fc tag of spike. The conjugated HRP enzyme catalyses the colourimetric reaction of TMB substrate which can be measured to determine the relative spike interacting with the immobilised ACE2.

3.2.3 Optimisation

To develop the method for the assay, it was necessary to determine appropriate amounts of each protein to be used. NeutrAvidin was used at 100 ng per well as this was commonly used in ELISA in the lab and performed well in similar assays.

ACE2

To determine the appropriate amount of ACE2 to use in the assay, varying levels of ACE2 would be added to NeutrAvidin coated plates and binding detected with a TMB reaction (Figure 20.). To detect the relative levels of ACE2 immobilised by the neutravidin, mouse anti-his tag HRP conjugated antibody (Sino biological, 105327-MM02T-H) were used to detect the presence of the hexa-his tag of the ACE2 protein. The HRP conjugate would catalyse the reaction of TMB and the absorbance would be relative to ACE2 immobilised by the NeutrAvidin and retained in the well. First, a five step two-fold dilution series of was performed starting with 80 ng of ACE2, resulting in final protein levels of 80 ng, 40 ng, 20, ng, 10 ng and 5 ng total protein added to individual wells in quadruplicate.
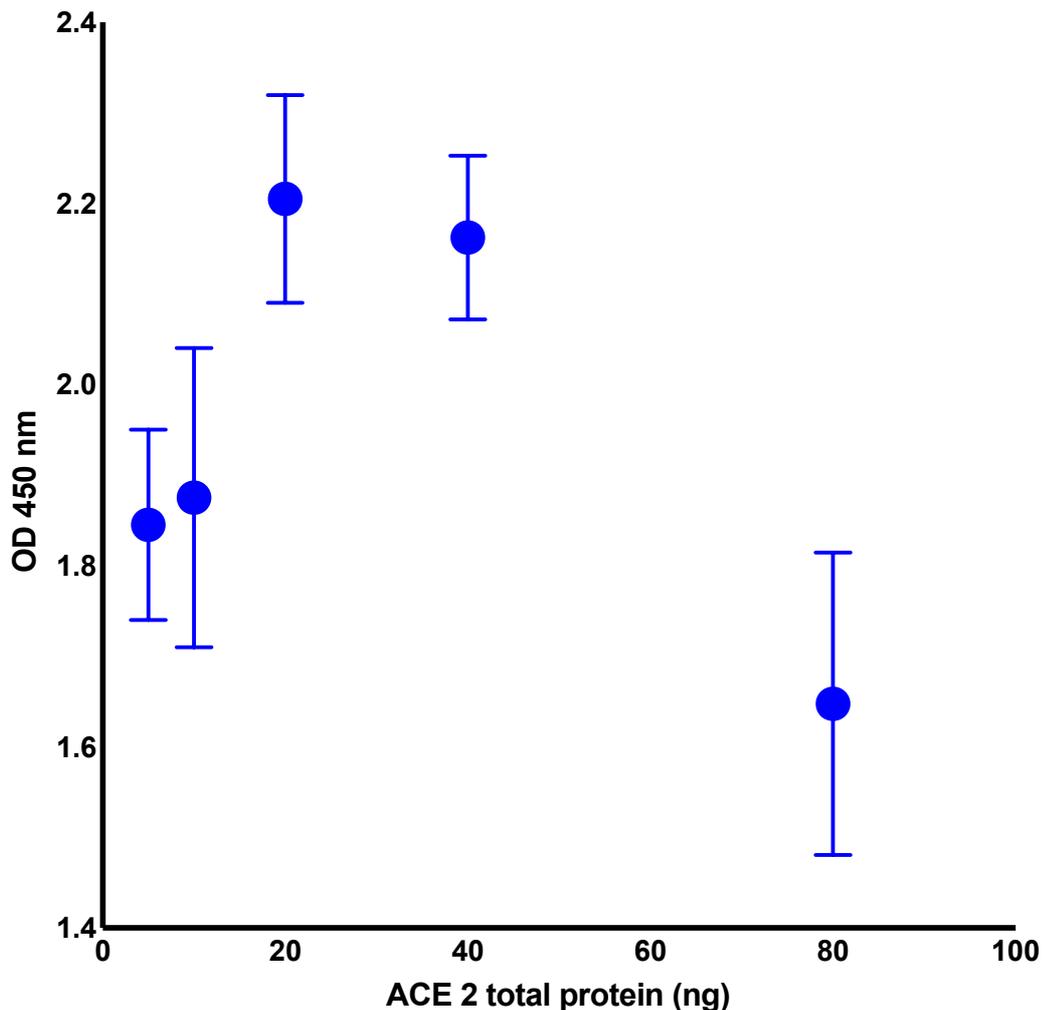
Figure 20. Biotinylated ACE2 immobilised by NeutrAvidin coated wells in a 96 well plate.

To calibrate the NOB ELISA, high binding 96 well plate wells were coated with NeutrAvidin, followed by incubation with biotinylated ACE2 to immobilise the ACE2 on the well surface. ACE2 was immobilised at total protein levels of 80, 40, 20 and 10 ng. to test the levels of immobilised his tagged ACE2 protein, a HRP conjugated anti-his antibody and TMB were used.

These results indicated that ACE2 is successful being immobilised by NeutrAvidin at expected levels and the assay signal was within expected ranges. This enabled progress to the next step, identifying amounts of spike and ACE2 protein to be used, based on their interactions in the assay.

Spike

Once it was established that ACE2 was successfully immobilised by NeutrAvidin, it was necessary to determine the appropriate amount of SARS CoV-2 spike needed. In this assay format, the degree of ACE2-spike binding was unknown so a range of ACE2 and spike total protein were tested against each other (Figure 21.).

Spike RBD should bind to the peptidase domain of ACE2 and remain in the wells after washing. The bound spike was detected using HRP conjugated anti-spike antibodies with TMB to develop the assay signal. The colour change was quantified by measuring OD at 450nm after stopping the reaction with sulphuric acid.

Wells of a high binding 96 well plate were coated with neutravidin to immobilise biotinylated ACE2 on the well surface. ACE2 was serial diluted in a four step, two-fold dilution series, resulting in final total protein of 80 ng, 40 ng, 20 ng and 10 ng per well. Each concentration was tested in combination with SARS CoV-2 spike protein, which was serial diluted in in a four step, two-fold dilution series, resulting in final total protein of 80 ng, 40 ng, 20 ng and 10 ng per well.

Figure 21. Neutralisation of binding (NOB) ELISA calibration for ACE2 and spike proteins.

To calibrate the NOB ELISA, high binding 96 well plate wells were coated with NeutrAvidin, followed by incubation with biotinylated ACE2 to immobilise the ACE2 on the well surface. ACE2 was immobilised at total protein levels of 80, 40, 20 and 10 ng. To test SARS CoV-2 spike binding to immobilised ACE2 spike protein, mouse Fc tagged spike protein was added to wells at a total protein level of 80, 40, 20 and 10 ng. Interaction between ACE2 and spike was achieved by detection of remaining spike using HRP conjugated anti-mouse FC antibody and TMB. Values shown in cells are means of signal strength of assay, measured by OD at 450 nm.

These results showed no significant difference between results and the signal strength was far below what was expected (1.0 – 3.0 OD at 450 nm). This indicated that the assay was not proceeding as intended. Possible explanations for this were: a problem with the TMB reaction, absence of antibody binding to the mouse FC tag of the recombinant spike, no quantifiable interaction between ACE2 and spike in the assay. The TMB reaction relied on two components, active HRP conjugated to the antibody which catalysed the TMB colourimetric reaction and properly functioning TMB. It was possible that the TMB had degraded or the HRP had lost catalytic activity due to degradation or denaturation. To test this a similar amount of HRP conjugated anti-mouse FC antibody that was used in the assay was added to an aliquot of TMB, which underwent the expected colour change, indicating these components were functioning normally. The ELISA relied on ACE2 interacting with spike to retain spike protein in the well

Spike validation

In the preliminary calibration experiments there was no apparent correlation between protein levels of ACE2 and spike and assay signal. Previous calibration experiments indicated that ACE2 was successfully immobilised on the plate surface by detection of the his tag on the ACE2 protein. This experiment validated the his tag and biotinylation of the ACE2 protein, but not the subsequent binding activity of ACE2 with spike. Since expected ACE2-spike interactions were not observed in the assay, it was possible that the lot of ACE2 protein used in the assay was not interacting with spike as expected. To investigate the interaction of the spike protein with ACE2, the assay was repeated with a separate second lot of functionally identical tagged spike protein in a similar ELISA (Figure 22.).
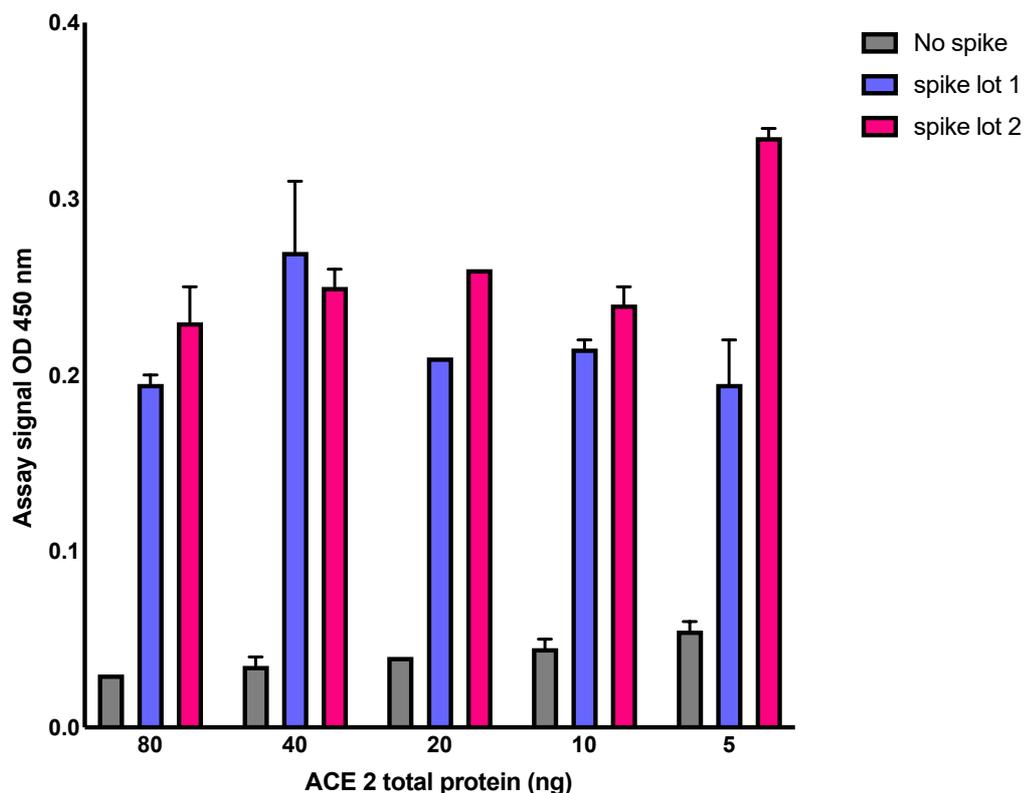
Figure 22. Spike recombinant protein lot validation.
The NOB ELISA was not functionating. To determine if the lot of spike being used in the assay was the cause of the error, the assay was repeated with a range of ACE2 protein levels with the original lot of spike, a second functionally identical lot of spike and no spike.

In this validation assay lower background levels were recorded (<0.1) than the previous assay (~0.2). while there was a slight difference in assay signal for each lot of spike, assay signal was still below expected levels of OD (1-3). The results were broadly similar between each lot, indicating that the lot of spike used was not the causative factor in lower than expected interactions between ACE2 and spike. There was no correlation between ACE2 protein concentration and assay signal, this could indicate that the lot of ACE2 used could be the source of lower than expected levels of protein interaction.

Previous assays indicated ACE2 was not interacting in the assay as expected. To determine if the particular lot of ACE2 was the causative factor, the assay was repeated to compare the primary lot of ACE2 with a second functionally identical lot of ACE2 (Figure 23.). A difference in assay signal between the two lots would indicate that the primary ACE2 lot used in previous steps was not interacting with spike as expected. This could be due to denaturation of ACE2 that occurred either in transport or storage.
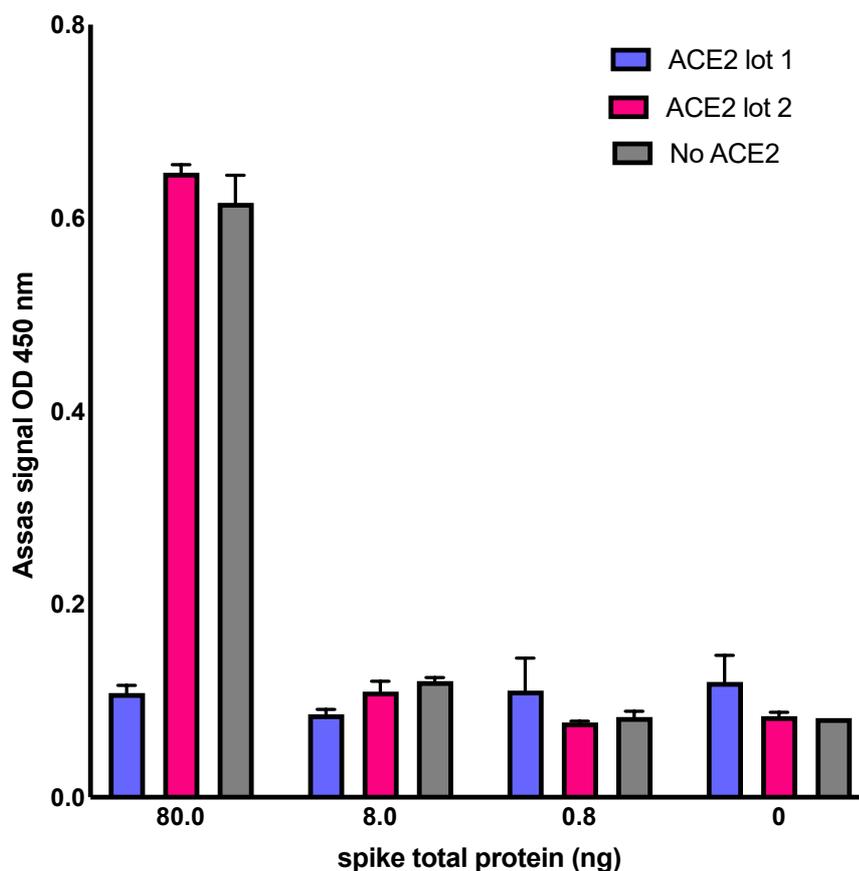
Figure 23. ACE2 recombinant protein lot validation.
The NOB ELISA was not functionating. To determine if the lot of ACE2 being used in the assay was the cause of the error, the assay was repeated with a range of spike protein levels with the original lot of ACE2, a second functionally identical lot of ACE2 and no ACE2. Binding of spike to immobilised ACE2 was measured using HRP conjugated goat anti-mouse antibody and TMB colourimetric reaction to detect the mouse Fc tag of the spike protein.

Summary

It was difficult to achieve consistent results and therefore the Assay optimisation was discontinued in favour of NOE assay and modified sandwich ELISA.

61

3.3 Neutralisation of Cell Entry Assay

3.3.1 Background

3.3.1.1 Overview

The mechanism of SARS CoV-2 entry into host cells is a critical step in infection. This step is mediated by the S glycoprotein binding to ACE2 receptors of host cells, which begins the process of membrane fusion. Spike mediated entry into host cells is a complex multistep process and interruption of any of these steps may inhibit viral entry to varying degrees. All three variants tested used the same entry receptors and same entry mechanism. The NOE assay is a critical tool in investigating viruses, particularly in assessing the capacity of antibodies to neutralise virus entry into cells. This assay allows for the use of a reductionist approach in a controlled setup with the safety of using pseudotypes instead of live virus. At the time of beginning these experiments little was known about SARS CoV-2 and best practices on safe handling were still being developed so the use of viral pseudotypes allowed investigation into the virus safely. These NOE assays provide insight into how the polyclonal antibodies of patients with severe COVID-19 affect entry of the pseudotype into cells, particularly across emerging variants that possess mutations conferring resistance to antibody neutralisation and enhanced cell entry.

3.3.1.2 Assay overview

The NOE assay was performed using a pseudotyping system. A non-replicating murine leukaemia virus (MLV) vector was pseudotyped with the spike glycoprotein of SARS CoV-2 variants. There are a number of reasons this approach was selected. First, pseudotypes are safer to handle than live replicating SARS CoV-2 virus, particularly given the recent emergence of this virus and the global pandemic it caused. This allowed for the experiments to be conducted at biosafety level II, making it more practical to carry out. Second the pseudotype approach enabled the glycoprotein variants to be changed relatively simply. Secondly, this method effectively simulates the entry mechanism of the SARS CoV-2 virus, using the same host receptors. Therefore, it can be used to investigate the effect that antibodies would have on live virus in a reliable and reproducible way.

To measure and quantify viral entry and infer effectiveness of patient antibodies in preventing viral entry luciferase was used as a reporter. When the pseudotype successfully enters the cell the luciferase enzyme is expressed inside the cells. Lysing the cells exposes this enzyme and exposing the enzymes to a controlled dose of luciferase assay substrate results in emission of light. The luminescence recorded is proportional to the expression of luciferase enzyme which, in turn, is proportional to pseudotype entry. Measurement of the intensity of luminescence enables the measurement of deviation from controls which was used to determine the effect of antibodies on pseudotype entry. A lower measured luminescence would indicate lower levels of viral entry and therefore neutralisation of entry by the antibodies and

a higher measured luminescence would potentially indicate antibody dependent enhancement of entry. Understanding the effect of antibodies raised in response to an infection is valuable in assessing immune protection and can give insights in development of vaccination strategies.

3.3.1.3   NOE assay setup

3.3.1.3.1   Viral pseudotyping

HEK293T cells were used for generating retroviral pseudotypes. This cell was commonly used for pseudotyping for a number of reasons. First, HEK293T cells have a high transfection efficiency due to the SV40 large T antigen genetic modification. The cells express the SV40 large T antigen which enhances replication of plasmids with the SV40 origin of replication. This origin of replication was present on both the cDNA3.1+ plasmid and phCMV5349 (MLV-Gag/Pol) plasmids used in pseudotyping. This ensured that the high levels of pseudotypes were produced in the cells. As HEK293T cells are derived from human cells they perform human post-translational modification. This is essential to mimic the spike glycoprotein in SARS CoV-2 infection of humans in the production of viral pseudotypes. An additional benefit to HEK293T cells is they are relatively easy to maintain in standard tissue culture conditions. HEK293T cells also express furin, an enzyme essential for priming the spike protein during virion maturation.

3.3.1.3.2   Plasmids

phCMV5349 (MLV-Gag/Pol) is the plasmid that encodesfor proteins and enzymes to create a functional viral core of the pseudotype. Gag is a polyprotein that is cleaved into the structural proteins that form the capsid, matrix and nucleocapsid. Pol encodes for the enzymes essential for viral replication including reverse transcriptase, integrase and protease. Viral envelope proteins are not encoded on this plasmid for safety and incorporation of selected envelope proteins in the pseudotype. The envelope protein is encoded on the separate plasmid pcDNA3.1+ in this experiment. This ensures the MLV virus is non-replicative. This means the envelope protein in the pseudotypes can be changed by including a different pcDNA3.1+ plasmid encoding for a different envelope protein in the transfection.

The pcDNA3.1+ plasmids used included a human adapted sequence encoding for variants of the SARS CoV-2 S protein. The plasmid encodes a cytomegalovirus immediate-early promoter for high levels of gene expression in HEK293T cells. The SV40 origin of replication is also included in the cDNA3.1+ plasmid to enhance plasmid replication and thereby enhance production of the selected envelope protein. The plasmid was also selected for its suitability for plasmid amplification in *E. coli* cells. It replicates efficiently in *E. coli* cells and the incorporation of the ampicillin resistance gene allows for selection during cloning and plasmid amplification.

3.3.1.3.3 Target cells

Although Vero E6 cells are derived from a *Chlorocebus aethiops* kidney epithelial cell line they are a suitable target cell for the NOE assay. It is a stable and well characterised cell line. The VERO E6 cells used express high levels of ACE2, the cell surface receptor used by SARS CoV-2 spike for cell entry. The cell line is commonly used for pseudotype assays as they are very permissive to a range of viruses. finally, the cell line closely mimics the entry dynamics of SARS CoV-2 in human cells, making it a good model for neutralisation of entry by antibodies.

3.3.2   Aims

The primary objective of this study was to evaluate the neutralisation efficiency of antibodies from patients with severe COVID-19 against three significant SARS CoV-2 spike protein variants: Alpha, Beta and Delta. These experiments were conducted using murine leukaemia virus (MLV) based SARS CoV-2 Spike pseudotype based neutralisation of entry assays. The specific aims of these experiments were to assess the neutralisation of efficiency of polyclonal antibodies in the sera of patients with severe COVID-19 against different SARS CoV-2 spike variants. This required first determining the capacity of patient derived polyclonal antibodies in serum to neutralise cell entry of MLV pseudotypes expressing either Alpha, Beta or Delta spike glycoproteins. The second aim was to compare the binding affinity of antibodies, determined by ELISA, with their neutralisation efficiency across three SARS CoV-2 variants. This aim focused on comparing antibody binding, determined by ELISA, with the neutralisation efficiency recorded in the NOE assays for each spike variant used. These experiments were designed to provide insight into the effectiveness of polyclonal antibody responses in patients with severe COVID-19, particularly in the context of the rapidly evolving viral spike protein in variants and vaccination strategies centred around the spike glycoprotein.

3.3.3   Results

Figure 24. Neutralisation of cell entry assay for serum 974

Serum sample from an individual admitted to hospital with COVID-19 was tested for neutralisation of cell entry (NOE) for three MLV based pseudotypes, each expressing one of three variants of the SARs CoV-2 spike protein. NOE assay was performed, and luminescence was measured to quantify cell entry. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. non-significant results are not shown for clarity, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Sample 974 showed an increase of signal strength across all dilutions as the dilution factor increased for Alpha at 0.45, 0.24, 1.65 and 2.91 fold for 1:200, 1:400, 1:800 and 1:1600, respectively (Figure 24.). There was a significant difference between 1:200 and 1:1600 (P = 0.0008) and between 1:400 and 1:1600 dilutions (P = 0.0002). At 1:1600 dilution there was a significant difference in assay signal between Alpha and Beta variants (P < 0.0001) and Alpha and Delta variants (P <0.0001). At 1:1600 dilution there was significant difference in assay signal between Alpha and Beta (P < 0.0001) and between Alpha and Delta (P <0.0001).

Figure 25. Neutralisation of cell entry assay for serum 977

Serum sample from an individual admitted to hospital with COVID-19 was tested for neutralisation of cell entry (NOE) for three MLV based pseudotypes, each expressing one of three variants of the SARs CoV-2 spike protein. NOE assay was performed, and luminescence was measured to quantify cell entry. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. non-significant results are not shown for clarity, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Sample 977 showed a significant increase in assay signal as the dilution increased for the Alpha variant, for 1:200 and 1:800 (P < 0.0001) and 1:800 and 1:1600 (P = 0.0224) (Figure 25.). For the Delta variant the signal decreased as dilution increased, from -0.10 at 1:200, 0.32 at 1:400, 0.88 at 1:800 and 1.43 fold change in signal at 1:1600 dilution. These changes were significant for 1:200 to 1:400 (P = 0.0012), 1:400 to 1:800 (P = 0.00131) and 1:400 and 1:1600 (P = 0.0047). For this sample there was a significant difference in assay signal between Alpha and Delta at all concentrations (P < 0.0001).
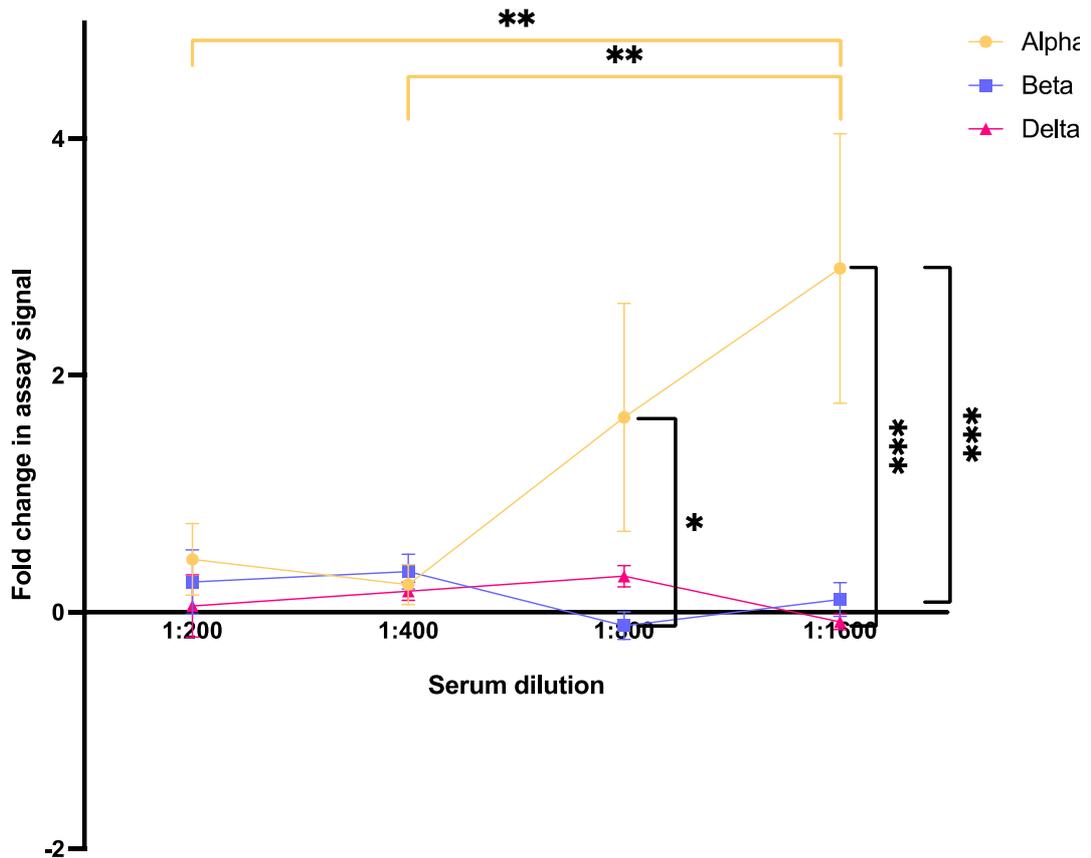
66

Figure 26. Neutralisation of cell entry assay for serum 980

Serum sample from an individual admitted to hospital with COVID-19 was tested for neutralisation of cell entry (NOE) for three MLV based pseudotypes, each expressing one of three variants of the SARs CoV-2 spike protein. NOE assay was performed, and luminescence was measured to quantify cell entry. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. non-significant results are not shown for clarity, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Sample 980 showed decreased assay signal with the Delta variant across all dilutions with 0.46, 0.33, 0.09 and 0.33 fold decreases at 1:200, 1:400, 1:800 and 1:1600 respectively (Figure 26.).

Figure 27. Neutralisation of cell entry assay for serum 981

Serum sample from an individual admitted to hospital with COVID-19 was tested for neutralisation of cell entry (NOE) for three MLV based pseudotypes, each expressing one of three variants of the SARs CoV-2 spike protein. NOE assay was performed, and luminescence was measured to quantify cell entry. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. non-significant results are not shown for clarity, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Sample 981 showed little significant difference between dilutions for all three variants (Figure 27.). At 1:1600 dilution there was a significant difference in the signals between Alpha (1.76 fold increase) and Beta (0.004 fold decrease) (P = 0.0017). similarly, at 1:1600 dilution there was a significant difference in the signal between Alpha and Delta (0.12 fold decrease) (P = 0.0029).
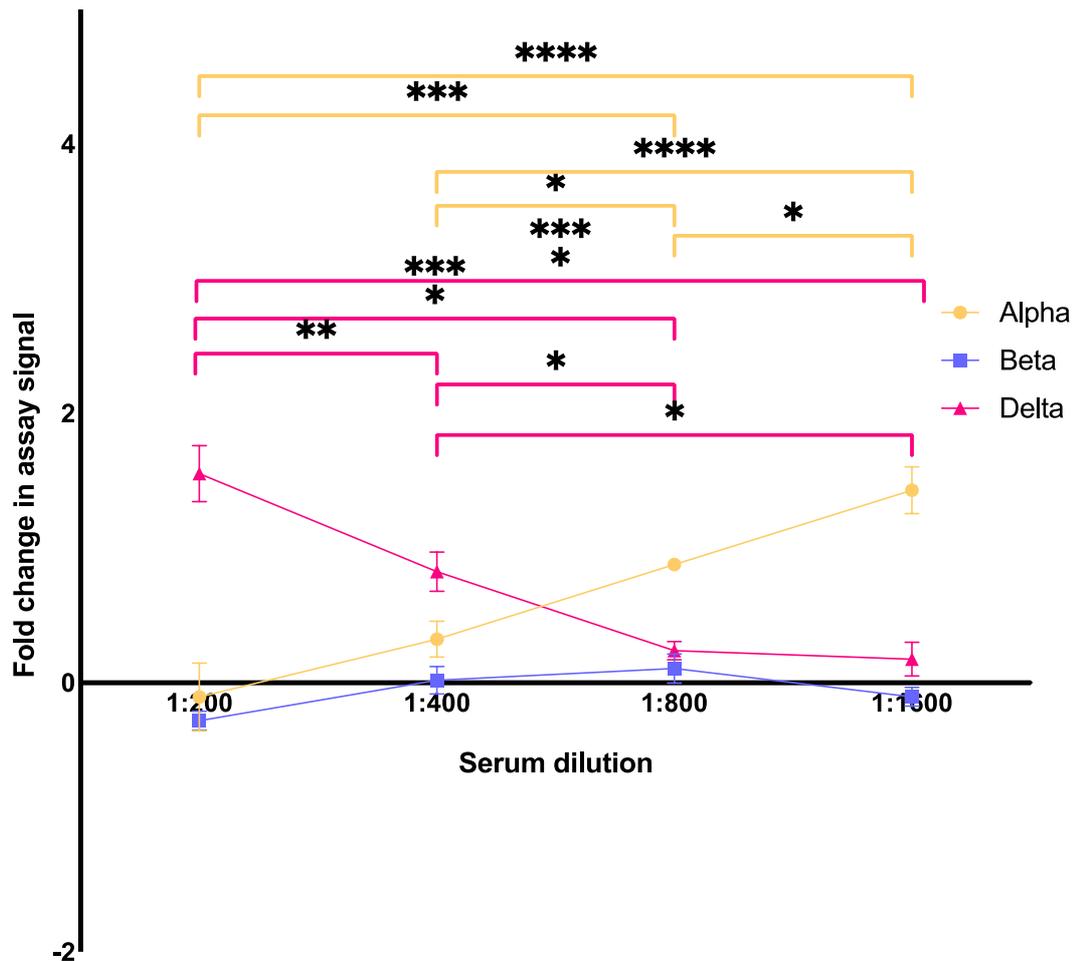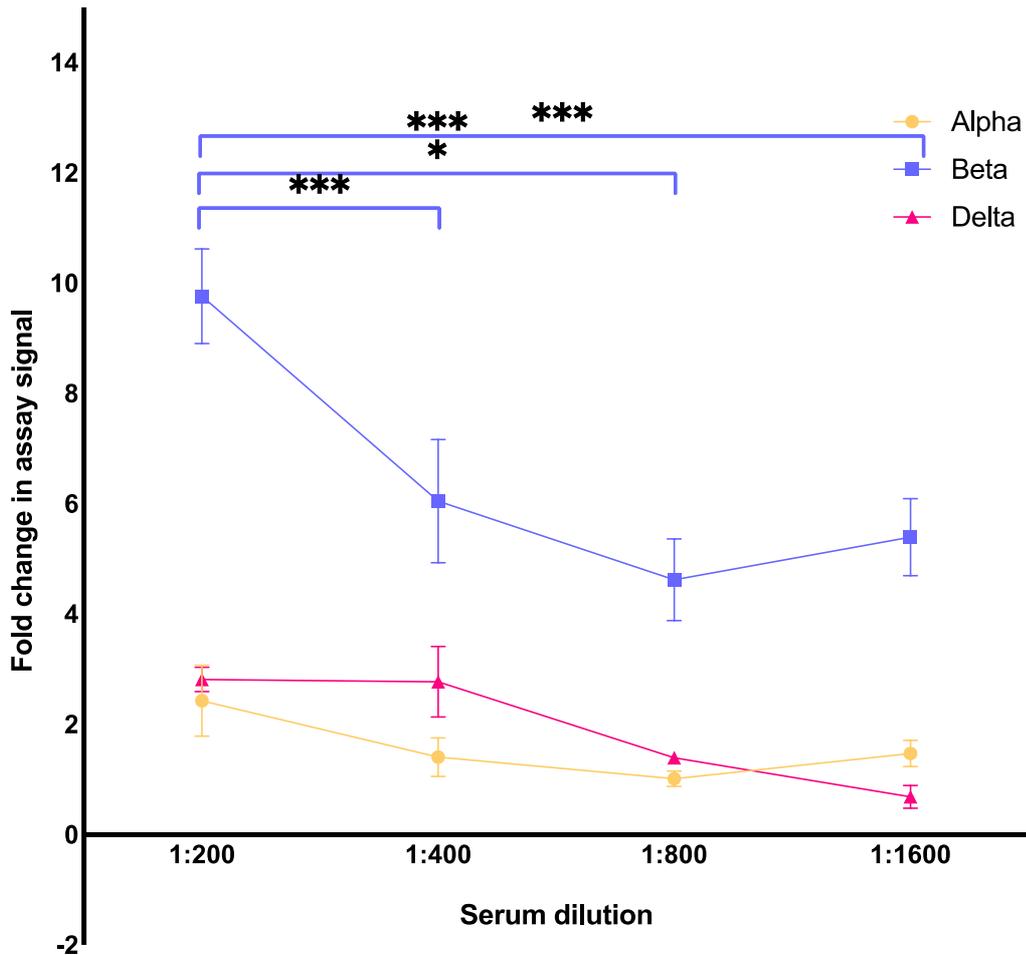
Figure 28. Neutralisation of cell entry assay for serum 1024

Serum sample from an individual admitted to hospital with COVID-19 was tested for neutralisation of cell entry (NOE) for three MLV based pseudotypes, each expressing one of three variants of the SARs CoV-2 spike protein. NOE assay was performed, and luminescence was measured to quantify cell entry. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. non-significant results are not shown for clarity, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Sample 1024 had no significant differences between dilutions for both Alpha and Delta variants, across all dilutions (Figure 28.). For Beta, there was a significant difference in change in assay signal between 1:200 (9.76 fold increase) and all other dilutions (P < 0.0001). At all concentrations there was a significant difference in assay signal between Beta and both Alpha (P < 0.0001) and Delta (P ≤ 0.0002).
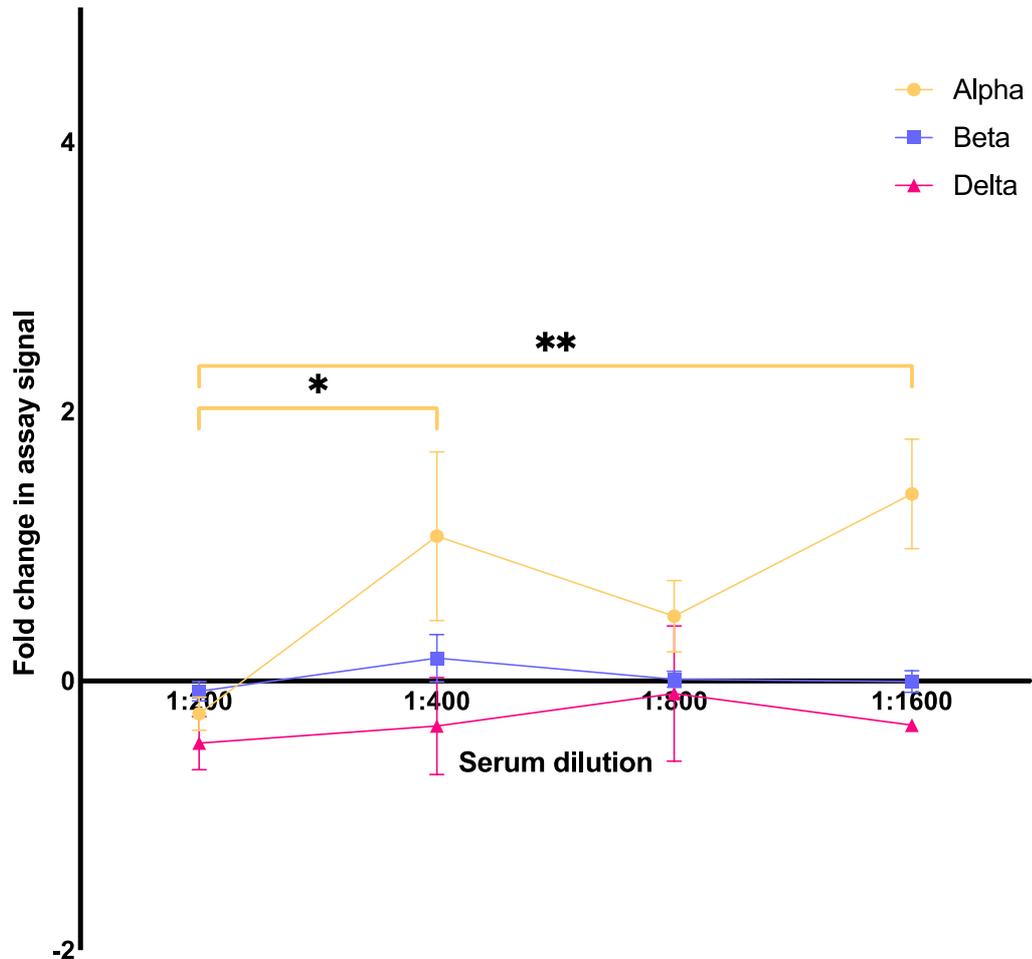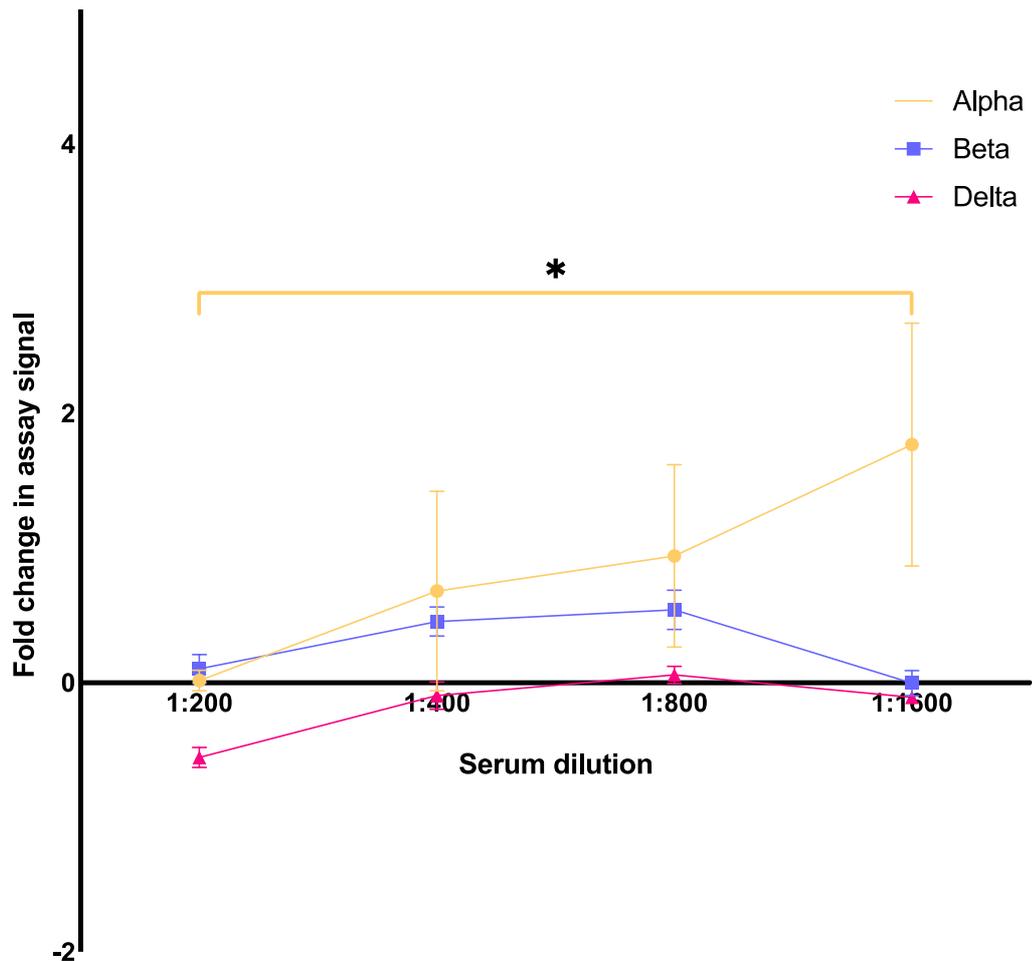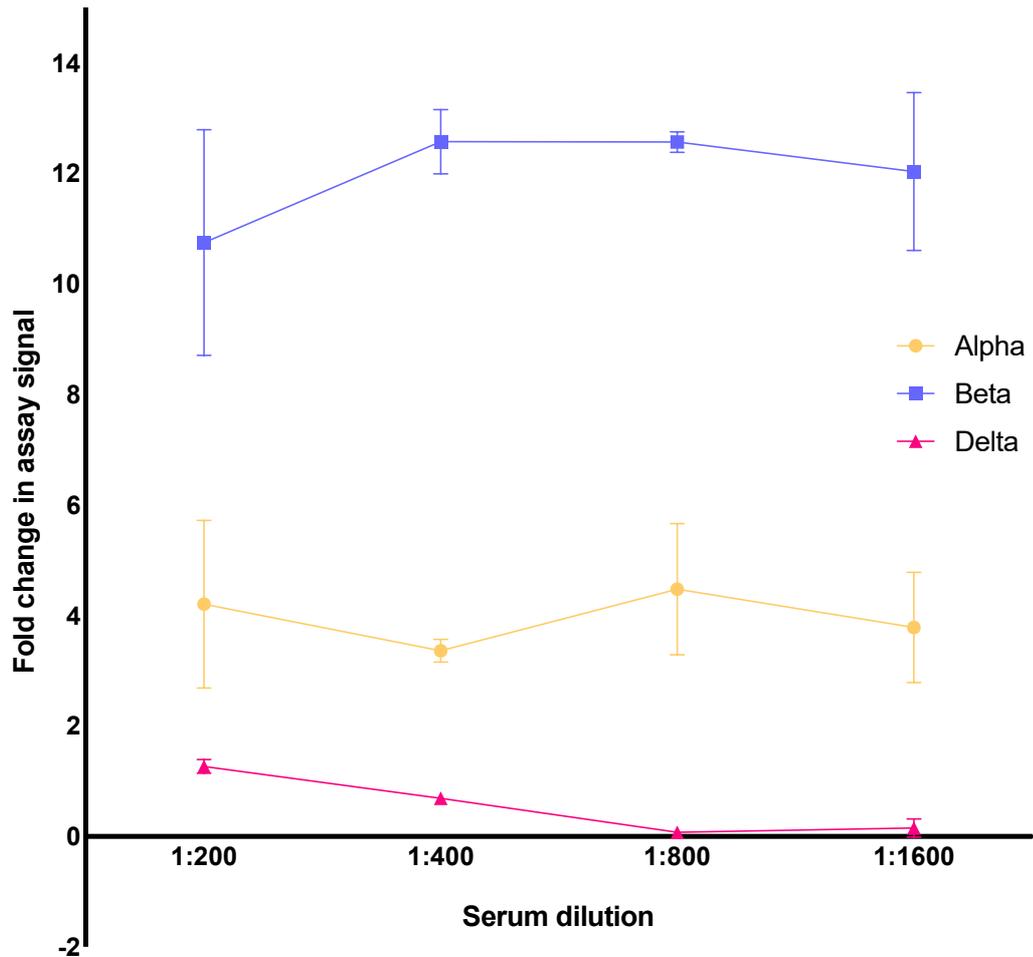
Figure 29. Neutralisation of cell entry assay for serum 1100

Serum sample from an individual admitted to hospital with COVID-19 was tested for neutralisation of cell entry (NOE) for three MLV based pseudotypes, each expressing one of three variants of the SARs CoV-2 spike protein. NOE assay was performed, and luminescence was measured to quantify cell entry. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. non-significant results are not shown for clarity, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Sample 1100 showed no significant difference in change in signal strength between dilutions for each of the three variants (Figure 29.). Signal strength change with Beta was between 10.75 and 12.56 fold increase for all dilutions. Signal strength increase with Alpha decreased as dilution increased from 1.27 fold increase at 1:200, to 0.015 fold increase at 1:1600. There was a significant difference in increase in signal strength across all dilutions between Alpha and Beta (P < 0.0001) and between Beta and Delta across all dilutions (P < 0.0001).

3.4   LEXSY expression system and plan

The use of commercially sourced proteins in the indirect ELISA, was convenient, however, it presented some limitations such as cost and restricted availability of variants and mutants for further experiments. In-house production of SARS CoV-2 spike proteins would allow for greater flexibility in experimental setup and reduced cost would enable for more experiments within budget. A key factor that was considered in this endeavour was primarily human-like post translational modification (PTM) of the protein as this would be essential in antibody interactions in planned experiments. Other key considerations included time investment required, cost of setting up the expression system, ease of use and equipment availability was considered as the expression system should be relatively easy to integrate into the lab making use of pre-existing equipment and colleagues knowledge. Finally yield of protein was an important factor as this could be balanced against the other factors.

Several expression systems were considered, and each had advantages and disadvantages. Bacterial cells such as *E. coli* were inexpensive and easy to use however, they could not perform human like PTMs. Insect cells could perform some PTM although, importantly, the glycosylation patterns did not match human cells. Human cells could perform the suitable PTMs making them a good choice however they are costly to use and the protein yield would be low and require a time investment to calibrate the process to achieve suitable yields. An uncommon protein expression system LEXSY uses *leishmania tarentolae* cells for protein synthesis. *L. tarentolae* are a eukaryotic unicellular kinetoplast pathogenic to the Moorish gecko. These cells are not pathogenic to humans and so can be handled at biosafety level I. Ease of use was comparable to *E. coli*, media and supplements were relatively inexpensive, capable of producing high yields of protein with very close PTM to human cells.

To determine the suitability of LEXSY for protein synthesis a review of literature and bioinformatic analysis of human ACE2 and SARS CoV-2 spike proteins was carried out. The aim of this analysis was to identify the PTMs of the proteins, their relevance in planned experiments and suitability of the expression system in performing these PTMs.

In the experiments the use for which they were considered, the only PTMs the cells could not perform was the addition of N-terminal sialic acid to a protein. In the process of cell entry the spike protein uses the N-terminal sialic acid of ACE2 to form weak interactions, augmenting the binding between the peptidase domain on ACE2 and the RBD of spike. This meant that LEXSY cells could not be used for ACE-2 synthesis as the N-terminal sialic acid of ACE-2 was used by SARS-CoV-2 spike as part of the multistep cell entry process.

LEXSY was determined to be suitable for producing spike protein and using this system mutants and variants could be produced in-house.

To further investigate antibody binding to spike mutants it would be advantageous to produce a set of spike mutant proteins that had either single substitutions or combinations of substitutions present in the individual variants Alpha, Beta and Delta. this could elucidate the specific mutations or combinations of mutations that are responsible for the differences in antibody binding in ELISA. For example, in the Alpha mutant contained 9 mutations in the spike protein including the substitutions D614G and N501Y. Using LEXSY, it would be possible to produce mutants with D614G alone, N501Y alone and D614G with N501Y and other combinations. Making these mutants would be a relatively cost effective and simple process using site-directed mutagenesis of the spike gene in PCR.

3.4.1 Optimisation

Problems with the freeze thaw process made use of the cells difficult at first. Compared to other cells used previously the major difference was that LEXSY cells were monoflagellated eukaryotes. Disruption of the cell membrane is a common cause of problems, particularly in this type of cell. It was suspected that the thawing process could have been detrimental to the stability of the flagella and reduced the number of viable cells. Private communications with the QMC reproductive health department on cryopreservation protocol helped optimization of the process with LEXSY cells. The standard LEXSY protocols for thawing were modified - instead of thawing the cells on ice, a Mr frosty (Nalgene, CRY8400) controlled rate freezer was used. The isopropanol in the container controlled the rate of freezing and thawing which reduced damage to cell membranes and the flagellum. An empty Mr. frosty was placed in a - 80 °C freezer for 3-4 hours to reach - 80 °C. LEXSY cryostock stored at - 80 °C were retrieved from storage with the empty Mr. Frosty and the vial of cells was placed in the empty Mr. Frosty, sealed and left at 25 °C to slowly thaw the cells in a more controlled environment. This modification to protocol resulted in a much higher success rate with these cells.

Issues encountered with these cells primarily were caused by the requirement of specific supplements to the media and their availability during and soon after the COVID pandemic. A predictable supply of these supplements could not be secured This made it impossible to continue cultures already in progress and introduced a complication in planning future work with this cell line. Because of these issues use of LEXSY for protein synthesis was discontinued in favour of pursuing work that could be planned and executed with a predictable procurement pipeline.
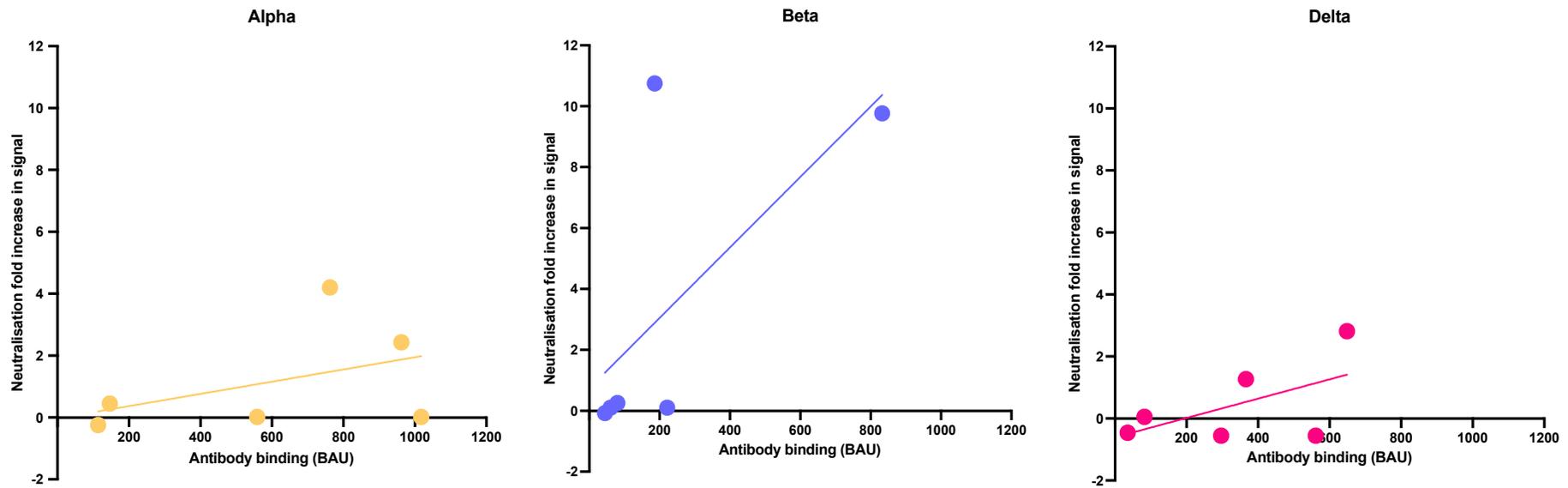
Figure 30. Scatter plot illustrating the relationship between total antibody binding and neutralisation of cell entry of patient sera across 3 variants

Linear regression model was fitted to the data for each SARS CoV-2 variant. Equations for the line for Alpha Y = 0.001972 × X - 0.02573 (R^2 = 0.1891), Beta Y =0.01157 × X + 0.7342 (R^2 = 0.0436) and Delta Y = 0.003112 × X − 0.6050 (R^2 = 0.3192). For all three variants the relationship between total antibody binding and neutralisation of cell entry was non-significant.

3.5   Novel computational approach to mutagenesis

To further develop this study, it was planned to further investigate existing mutations in different variants, using the LEXSY expression system. The aim of this approach was to synthesise spike proteins of specific variants and to also produce spike proteins that contained combinations of mutations within each variant. For example, for the Alpha variant which, compared to the Wuhan strain, contains the mutations N501Y, A570D, T716I, S982A, D1118H and P681H, among others, we would synthesised proteins with one mutation alone and proteins with combinations of two or mutations, e.g. N501Y & A570D, or S982A & D1118H & P681H. This would have allowed us to identify which mutations or combinations of mutations were the biggest contributing factors to immune evasion. While the experiment would have provided valuable understanding to how the existing variants interact with antibodies, being able to predict potential mutations that would enhance the ability of SARSCoV2 to evade the immune system would be invaluable in vaccine design. To do this we would need to explore novel mutations in the spike protein.

To predict where novel mutations might occur in the gene encoding the spike protein, we initially proposed adopting a bioinformatic approach to examine synonymous mutations. The approach would have involved a creating a program to mimic the changes that would be observed during viral replication in human cells. The approach would have required accounting for a large number of factors that influence mutation at a given site including: GC content and localisation, codon usage bias, sequence context in the form of flanking sequences of a given nucleotide, RNA editing by host enzymes, oxidative damage by the host immune response, RNA-protein interactions, mutation prone motifs and polyuridine tracts among many others. The complexity and interactions of all these mutation factors would have required the investment of considerable time and effort to make the results of such a program reliable.

Random mutagenesis was considered as an alternative approach for its greater simplicity. Here random variation in the sequences is achieved through the use of error prone (EP) PCR in which mutations are created by performing PCR using suboptimal conditions such as imbalance in dNTP concentrations, use of low fidelity polymerase and addition of manganese ions. These suboptimal conditions cause errors in the replication of the initial sequence resulting in mutations. Its simplicity is one of the reasons it has been widely used by researchers to investigate spike proteins. However, a drawback to the approach is the reliance on a single consensus sequence for the initial sequence template.

We designed an approach that would augment the standard EP PCR approach by using multiple synonymous sequences as the initial templates for EP PCR rather than a single consensus sequence. Synonymous sequences differ from the consensus

sequence at the nucleotide level but encode the same amino acids and we choose this approach to broaden the diversity of possible mutations for study. It was additionally selected because for a virus to develop an amino acid substitution mutation it will usually first undergo a series of discreet synonymous mutations before the resulting substitution mutation. For example, the codon CGC (Arg) changing to CAG (Gln) requires the mutation of two bases simultaneously. This mutation is more likely to occur in a stepwise manner through two mutations in at least two rounds of replication: CGC (Arg) to CGG (Arg) and then CGG (Arg) to CAG (Gln). The use of synonymous sequences as initial templates for EP PCR makes the occurrence of a novel substitution mutation at a given codon more likely.

A literature search was conducted to identify suitable programs that could be used to generate synonymous sequences from a consensus sequence. After no suitable program was identified we consulted bioinformaticians in the School Life Sciences and it was determined the best approach would be to develop a custom program.

Methods: I wrote a python program called 'synonymaker'. This program uses an amino acid sequence, codon bias table and a codon usage scoring table to generate sequences likely to occur in nature (Figures 31-34.). The program takes the 3 inputs and generates non-redundant synonymous sequences each with a codon usage score and ranks them by codon usage score. Since the program was designed to augment EP PCR, which is a commonly used method for investigating proteins, the program was designed to be modular and simple to use. To adapt the program for any virus and any species, the inputs for codon bias and codon usage scoring can be changed by the editing of simple excel tables and text documents. The consensus sequence of interest is pasted into the program when prompted and then selection of the desired number of consensus sequences to be returned. The program produces a simple text file output of the generated sequences and associated codon usage score. These generated synonymous sequences can then be converted from RNA sequences to DNA sequences, ordered from a oligonucleotide synthesis supplier and used in EP PCR.

Figure 31. Sequence input file
With accessibility in mind, the sequence input file is a simple text or fasta file format. Only the amino acid sequence need be input and the program does not need to be edited.
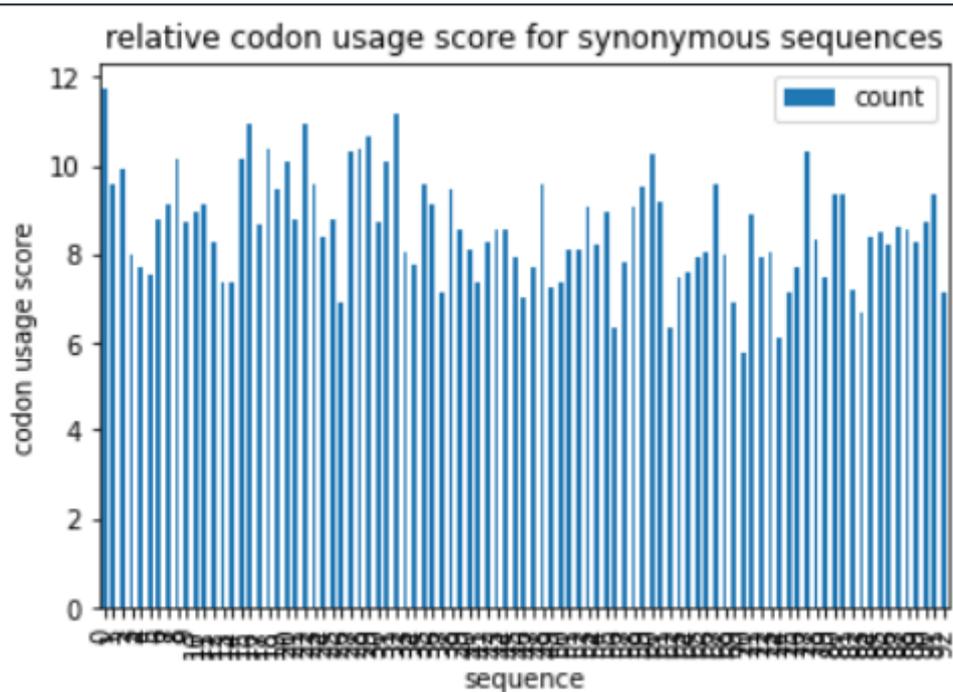
Figure 32. Codon usage score output graph

Once a series of synonymous sequences have been generated each is scored according to codon usage bias table.
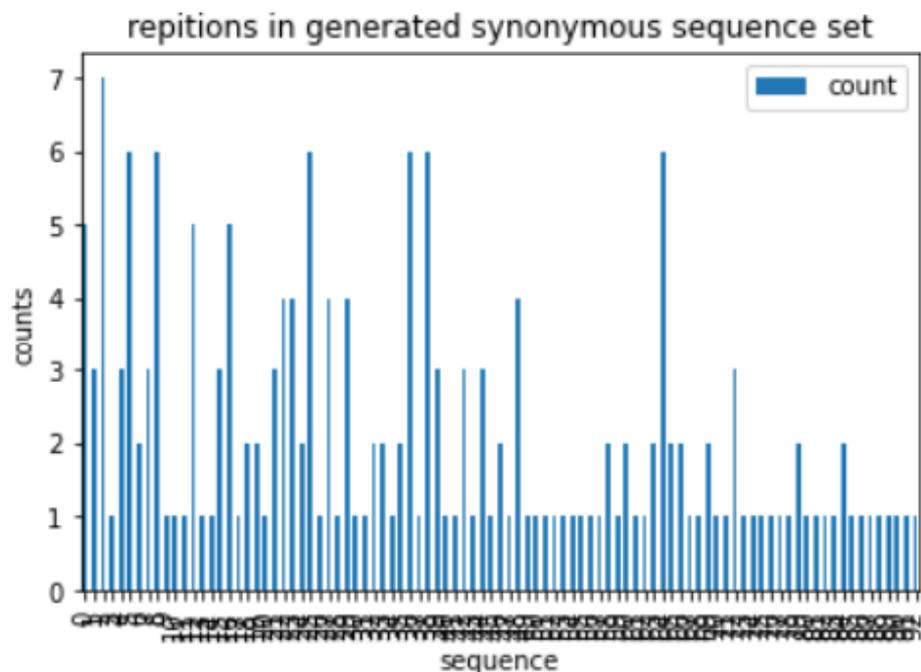
Figure 33. Redundant sequences in file

As the program generates random sequences, taking into account codon usage bias with first order approximation, the output set of sequences will contain redundant sequences. The preliminary output file is checked for redundancy and a non-redundant output is generated.

	A	B
1	seq	count
2	UUUGUUUUAUUACAACAAGGU	11.73
3	UUUGUUUUAUUACAACAGGGU	10.95
4	UUUGUUUUAUUACAGCAAGGU	10.95
5	UUUGUAUUAUUACAACAAGGU	10.68
6	UUUGUUUUGUUACAGCAAGGU	10.38
7	UUUGUUUUGUUACAACAGGGU	10.38
8	UUUGUCUUAUUACAACAAGGU	10.33
9	UUUGUUUUAUUACAACAAGGA	10.23
10	UUUGUUUUAUUACAGCAGGGU	10.17
11	UUCGUUUUAUUACAGCAAGGU	10.15
12	UUUGUUUUAUUACAACAAGGC	10.12
13	UUUGUAUUGUUACAACAAGGU	10.11
14	UUUGUAUUAUUACAACAGGGU	9.9
15	UUUGUAUUAUUACAGCAAGGU	9.9
16	UUCGUAUUAUUACAACAAGGU	9.88
17	UUUGUUUUGUUGCAACAGGGU	9.81
18	UUUGUUUUAUUGCAGCAGGGU	9.6
19	UUUGUGUUAUUACAACAGGGU	9.56
20	UUUGUGUUAUUACAGCAAGGU	9.56
21	UUUGUCUUAUUACAGCAAGGU	9.55
22	UUUGUUUUAUUGCAACAAGGC	9.55
23	UUUGUUUUAUUACAGCAAGGA	9.45

Figure 34. Non-redundant sequences, with codon usage scores ranked

To achieve a high level of diversity in the initial templates with the smallest number of sequences for the starting point of EP PCR, the selection of specific synonymous sequences would be necessary. Selecting highly dissimilar sequences, both to the consensus sequence and to each other would be desirable to maximise diversity. To do this the sequences could be compared using a phylogenetic tree, which uses similarities and differences in the genetic code to illustrate relatedness. Selecting a small range of less related sequences from the phylogenetic tree by using branch lengths would enable the selection of a diverse range of synonymous sequences.

While the development of this software was successful, it was not possible to carry out the remainder of this experiment as planned as it required using the LEXSY system, which was not successfully implemented during my research. However, this program can augment EP PCR to generate mutations in a method that more closely models nature. The modularity of the program allows it to be used for any gene of interest with any expression system including those commonly such as *E. coli* and yeast.

Chapter 4. Discussion and future work

4.1 Samples

Five patients in this study were male. Very limited information was available for patient 981 and gender was not recorded. Patients 980 and 1100 were both over 75 years of age and the three other patients were over 55 years old, with the age of 981 unknown. Patient 1100 had a pre-existing severe health condition which made them clinically vulnerable. For all patients there was no recorded history of vaccination against SARS CoV-2. This was not unexpected as the patients in this study were hospitalised with severe COVID-19, this would bias the sample set towards people in vulnerable groups such as the elderly, clinically vulnerable and it would also select against those who has received a vaccination. Vaccination was either not available or limited at the time the patients were initially infected.

Samples taken from patients infected early in the pandemic. During this time the circulating variants of SARS CoV-2 in the UK were Wild-Type, Alpha, Beta and Gamma. Wild-Type was the dominant variant initially however by the time of sample collection Alpha was the dominant variant with Beta responsible for a small number of infections and Gamma a rare variant in the UK. It is likely that the patients in this study were infected with either the Wild-type or Alpha variant, however it was not possible to identify the infecting variant in each case.

A limitation of using patient sera samples in assays is the presence of factors such as albumin, rheumatoid factors, lipids and other proteins. These factors may have non-specific interactions with elements of the ELISA such as proteins or antibodies which would affect the background signal and measured signal.

Sample collection

Organising identification of patients and obtaining samples from hospital departments was difficult during the pandemic. Hospitals were experiencing a number of difficulties during the pandemic. Hospitals experienced significant staffing issues due to isolation requirements. Due to COVID-19 hospitals had a surge in inpatients with highly a communicable disease that required robust containment measures and often required extensive treatment such as intubation. At the start of this study there was no way to identify patients with a specific condition easily. The method available was manually searching through all tests performed by one department, opening the file and interrogating the test type and result. This manual method had to be repeated for each individual department that performed diagnostic testing and would hold excess serum samples. To circumnavigate this we organised with the clinical virologist in our lab, Queen's Medical Centre's clinical chemistry department, outside IT contractors and other staff to develop a method to electronically tag patients in the hospital databases with a COVID-19 tag. This was a complex and time-consuming process but enabled for more efficient identification and collection of samples. However there were some issues with this method and its

implementation. Some patients tagged were not tested for COVID-19 and were diagnosed on symptoms and some patients tested negative and were diagnosed on symptoms. This could have been due to lack of resources for testing early in the pandemic, difficulties in record keeping during the pandemic response or improper implementation of the tagging system. As the tagging system was a new development and introduced solely for identification of samples for this study it is completely understandable that its correct implementation was not a high priority for frontline hospital staff at the height of the pandemic. Another issue encountered was that the developed tag was non-expiring, so it persisted after the patient was discharged and remined if the patient was readmitted, even if readmission was for a different condition. Linking the samples with patients and patient detail was also a time-consuming process which required the assistance of hospital staff and manual access of patient records. This necessitated preparation of sample numbers to be cross referenced by hospital staff and liaising with staff to collect data, again during the pandemic response.

Samples collected with this method included some samples from patients with: no evidence of COVID-19, no positive test for COVID-19 and no COVID-19 diagnosis at time of sample collection but previous COVID-19 diagnosis. Multiple samples from the same patient on the same day were also collected. This meant that of the samples collected, only a small number could be carried forward for further testing.

The refinement of the tagging method its implementation would be an excellent additional to NHS standard operating procedures. It would enable for a faster more efficient way to track, report and access samples of patients, not only for research but also for reporting and tracing trends. As demonstrated in the rapidly spreading and evolving COVID-19 pandemic, the ability to respond to a new pathogen by is essential. Widespread use of this method would enable researchers' better access to samples from patients infected with a newly emerged pathogen. This would allow for a greater ability to research a newly emerged pathogen, test and develop diagnosis techniques and understand the pathogenicity of a newly emerged pathogen.

4.2   Antibody binding ELISA

Antibodies produced as a result of an infection will be targeted at specific epitopes on the antigen surface. In SARS CoV-2 the primary target for neutralising antibodies is the spike protein. While the targeting of antibodies is random, through affinity maturation and clonal selection of B cells, antibodies with the highest binding affinity are selected for. This selection results in targeting of specific highly immunogenic sites on the protein surface (Janeway, Travers, Walport, & Bhlomchik, 2001). Within the RBD of spike there are 23 epitopic sites (ES) identified Antibodies from a variety of sources have been shown to bind to these specific ES, show in figure 31 (Figure 35.)(Jiang et al., 2023). Antibodies that bind to the RBD will bind to one of these discreet ESs. A change to the electrostatic charge, hydrophobicity, polarity, or size of

the side chain of an AA in one of these ES will likely disrupt antibody binding to that site. Therefore, antibodies that neutralise one variant, targeting a specific ES, will likely have reduced interactions with a variant that has a substitution mutation at the ES which alters the characteristics of the ES. This likely explain why patients in this study show differential binding and neutralisation of SARS CoV-2 spike variants (Figure 36.).
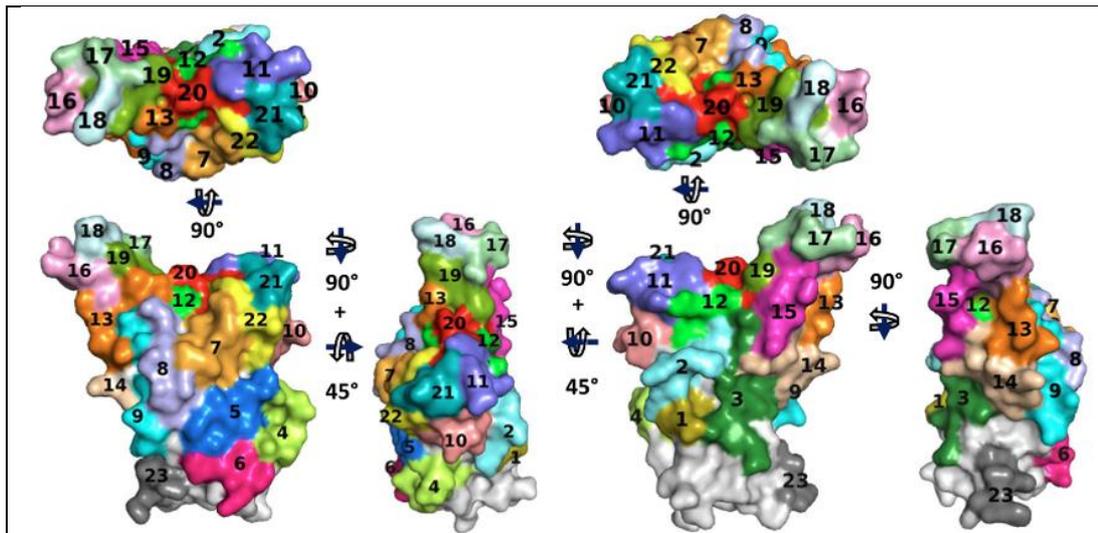
Figure 35. Representation of SARS CoV-2 RBD of spike protein with identified epitopic sites, recognised by binding antibodies.

SARS CoV-2 spike RBD rotated to show all surfaces and mapped epitopic sites (ES). The ES include both linear and conformational sites, each ES is colour coded and numbered (Jiang et al., 2023).
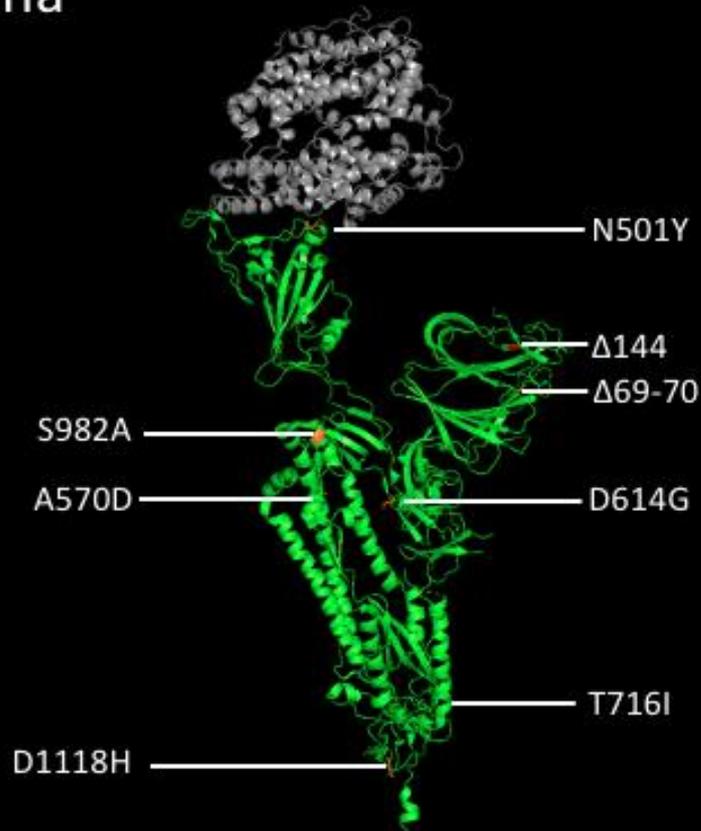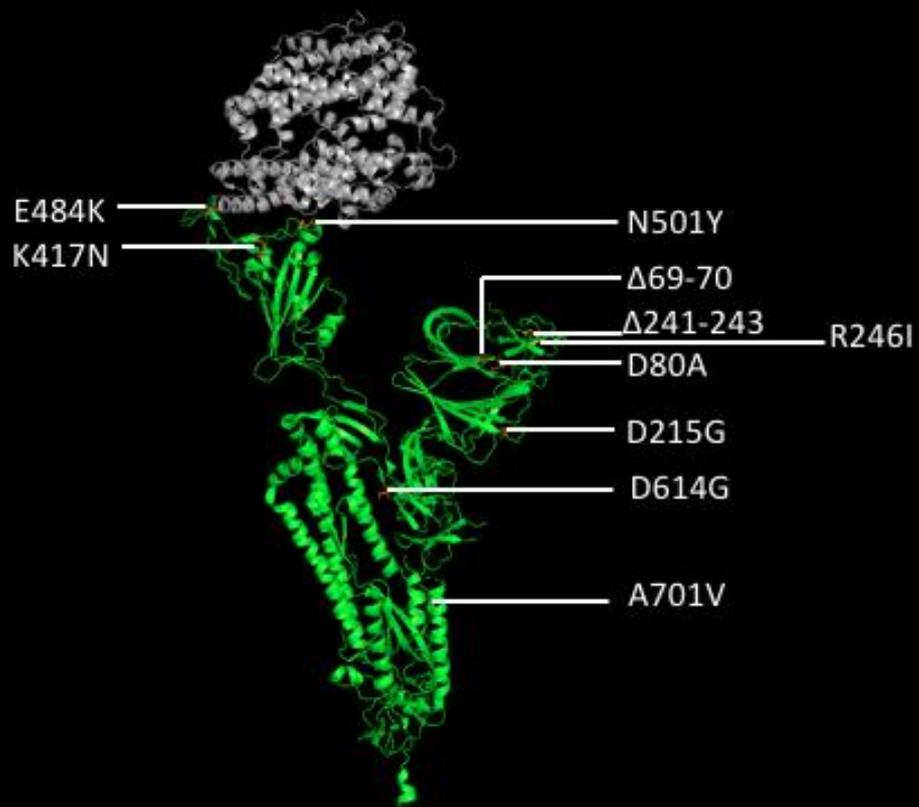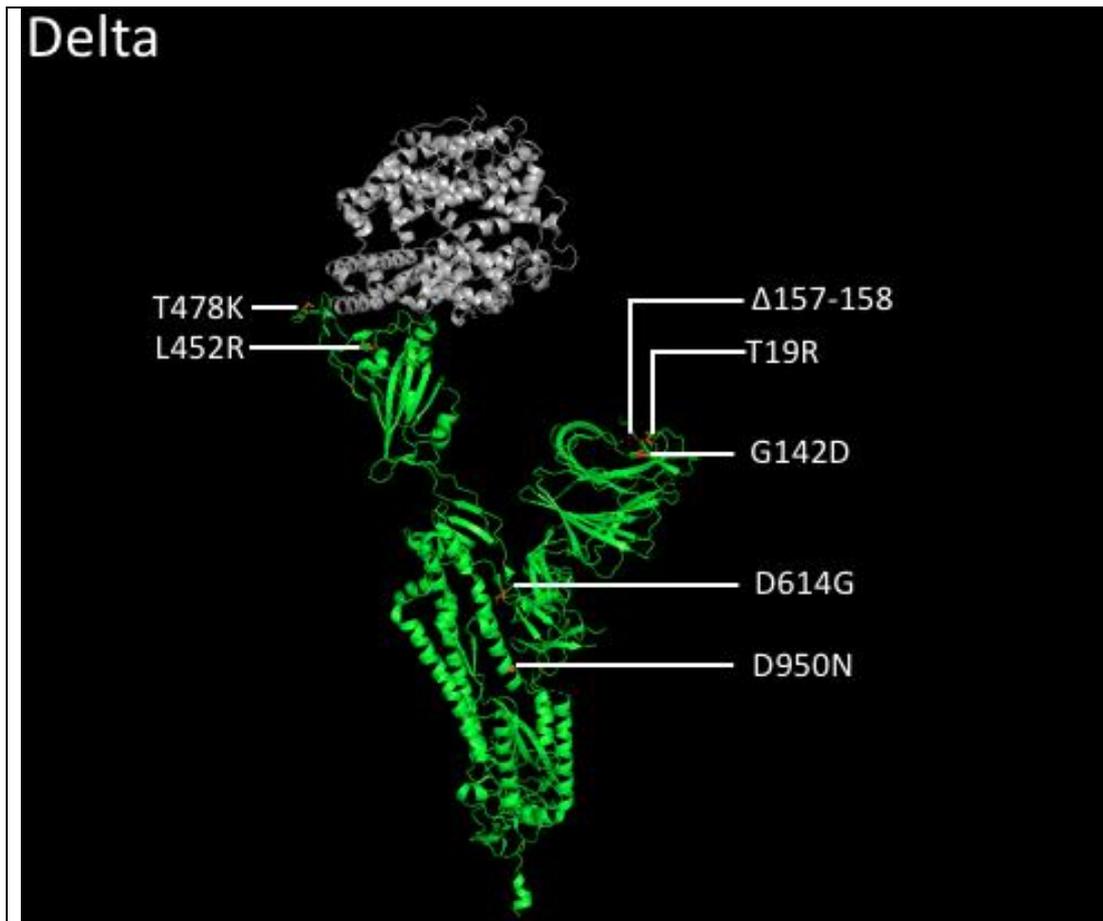
Figure 36. Three-dimensional structure of SARS CoV-2 spike protein (green), bound to ACE2 (grey), with spike mutations and deletions labelled.
The positions of the mutations (orange) and deletions (red) are shown to highlight their position in the 3D structure of the folded protein and their proximity to ACE2.

Sample 974 had particularly low antibody binding levels across all three variants in the ELISA and these did not show any significant degree of neutralisation of the pseudovirus in the NOE assay. Against the Alpha variant this sample showed increased cell entry with decreasing concentration of antibodies. A similar pattern of enhancement was not observed with the Beta or Delta variants.

Delta and Beta share no mutation that could account for this. It is possible that this result is due to a mutation present in the Alpha variant. The mutation Δ69-70 is present in the Alpha variant but not the Beta or Delta variants. This deletion has been reported to enhance immune evasion. Another mutation not shared with the other two variants but is associated with enhanced ability of the virus to enter cells is the P681H mutation in the S1/S2 cleavage site of Alpha variant spike. While this mutation is not present in the other variants, a similar mutation is present in Delta, P681R, at the same position. Both mutations at the 681 positions replace proline with an AA with a positively charged side chain.

Sample 977 had high binding antibody levels with the Alpha variant, moderate binding with Delta and low binding affinity with Beta. In the NOE assay this sample was mildly neutralising against the Beta variant and showed different patterns of apparent ADE with both Alpha and Delta variant. Against the Delta variant the serum apparently enhanced infection more at lower concentrations and decreased with dilution. This would suggest ADE as a mechanism for increased viral entry at lower concentrations as described by Coutant *et al.*, 2024, where they described 38 monoclonal antibodies that enhanced infection. The opposite effect was observed against the Alpha variant where at the lowest dilution the sera was mildly neutralising but as the dilution factor was increased the sera began to enhance infectivity of the pseudotypes. This suggests dilution dependent antibody enhancement of pseudotype entry acts differently for different variants of the virus for the antibodies in this sample. The low levels of antibody binding along with apparent enhancement of Alpha variant cell entry would indicate that these antibodies were enhancing the infection of this individual. Serum sample 977 exhibited low cross reactivity with Beta. This could be due to the presence of the mutations K417N and E484K in the RBD of the S1 subunit which has been reported to reduce neutralization by antibodies, aiding in immune evasion. This sample also had low cross reactivity with Delta, mutations to the RBD, L452R and T478K both enhanced the binding affinity with ACE2 which could account for the greater levels of cell entry measured.

Sample 980 had low levels of antibody binding across all three variants. The neutralising effects of the antibodies were reasonably effective against the Delta variant but either non-neutralising or enhancing against both Alpha and Beta variants in the NOE assays. The neutralising effect against Delta pseudotype would suggest that if the patient from which the antibodies derived was exposed to SARS CoV-2 Delta variant while these antibodies were circulating, they would aid in providing a robust protective immune response.

Sample 981 had variable antibody binding across all three variants, with moderate binding to Alpha, low levels of binding to beta and moderate to low levels of binding to Delta. Cross referencing the NOE assay it appears that the moderate binding of antibodies were non-neutralising at low dilutions but at higher dilutions acted to enhance pseudotype entry for both Alpha and Beta. Neutralisation of Delta by this sample was moderate at low concentrations and decreased as dilution increased. These results suggest that while the antibodies enhanced cell entry of the Alpha variant pseudotypes and were non-neutralising for beta variant, if the patient from which the antibodies derived was exposed to SARS CoV-2 Delta variant while these antibodies would aid in providing a robust protective immune response. Sample 981 exhibited the lowest cross reactivity with Beta. The beta variant included tshe K417N and E484K mutation in the S1 subunit. These mutations were associated with reduced neutralization by antibodies, aiding in immune evasion. This could indicate that there mutations played a role in the reduced cross reactivity observed.

Sample 1024 had the highest total antibody binding across all three variants and enhanced cell entry in all conditions across all three variants, with significantly higher levels of enhancement at high concentrations against the Beta variant. These results indicate that this individual would be particularly susceptible to reinfection with all circulating variants at the time due to the enhancing nature of the non-neutralising antibodies. Symptoms would likely be severe upon reinfection with Alpha, Beta or Delta variants of SARS CoV-2.

Sample 1100 showed a similar antibody binding pattern across the variants to sample 981. While both enhanced Alpha across all dilutions, sample 1100 dramatically enhanced infectivity of Beta variant pseudotypes with over 12-fold signal increase at three of four dilutions. A less severe but still significant increase in infectivity was seen against the Alpha variant where increases in infectivity ~ fourfold were observed. These results indicate that this individual would be particularly vulnerable to infection and would likely develop severe symptoms if infected with the Beta variant of SARS CoV-2. Similarly, the antibodies are likely to enhance a reinfection with the Alpha SARS CoV-2 variant and the patient would likely experience severe symptoms of COVID-19.

Guo et al.performed a similar binding ELISA where they investigated patient antibodies binding to full length spike, and both S1 and S2 subunits in separate assays. While this allowed them to distinguish antibody binding localisation to a small degree, the value of this additional information is not clear without further investigation into the neutralisation effects of these antibodies.

Luo et al. used a similar method to the binding ELISA in this study with only some deviations from the method described for the binding ELISA. Instead of the spike monomers used in this study, Ji Luo *et al*. immobilised trimeric spike protein to the cell surface. Use of monomeric proteins may not capture the effect of antibodies that bind to epitopic sites that that are present only in the trimeric form of the protein. To determine if the use of trimeric versus monomeric spike proteins in an antibody binding ELISA could reveal more information about antibody binding it would be necessary to perform both types of ELISA with the same samples and interrogate the results. Luo et al. found positive correlation between binding levels of anti-spike antibodies and severity of COVID-19 symptoms. While this study did not quantify and investigate correlation between antibody binding and symptom severity, a retrospective analysis of patient records for those involved in this study could determine if there was a correlation between findings in the NOE assays, binding ELISA and severity of symptoms.

The ELISA used in this study investigated total antibody binding. This limitation meant it was not possible to discriminate between signal from monoclonal antibodies that bound with high affinity to the spike protein and a population of polyclonal

antibodies that each bound with low affinity, as both would produce a high signal. Analysing the combined results of the ELISA and NOE assays it is possible to further discriminate the effect that the binding antibodies had on virus entry.

NeutrAvidin is a modified version of streptavidin, a molecule originally derived from the bacteria *Streptomyces avidinii*. Streptavidin has a high binding affinity with biotin which can be conjugated to antigens to capture them on the well surface in ELISAs. Streptavidin can bind up to four molecules of biotin. NeutrAvidin has similar properties to streptavidin with some advantages. NeutrAvidin is deglycosylated and has a more neutral isoelectric point than streptavidin, these properties reduced non-specific binding and result in an improved signal to noise ratio in the ELISA.

4.3   Neutralisation of binding ELISA

The neutralisation of binding ELISA was designed to assess the capacity for patient derived antibodies, present in serum samples, to prevent interaction between ACE2 and SARS CoV-2 spike protein. The assay could not be successfully completed during the course of this study. Despite extensive troubleshooting and validation efforts we could not identify the reason behind this.

Due to the success of the functionally similar binding ELISA, it was expected that the NOB ELISA would work with the designed protocol. As the results in the previous chapter show, we successfully validated immobilisation of ACE2 using NeutrAvidin and eliminated that an issue. It became apparent that the expected interaction between spike and ACE2 did not occur in the assay. To ensure the reliability of the assay components and to investigate the reason behind the unexpected results, multiple critical components were evaluated. Multiple lots of ACE2 and spike were tested to rule out potential variability between lots such as denaturation, specifically of the spike RBD or ACE2 PD, both of which would affect binding. Despite our efforts, the reasoning behind the lack of anticipated interaction in the assay could not be determined during the time of this study.

Shalash *et al*. performed a similar NOB ELISA with some differences in assay design and sensitivity. The ELISA in this study entailed immobilising ACE2 in wells of a high-binding ELISA plate via NeutrAvidin coating, incubation of patient sera with spike protein and addition of the spike-antibody complexes to the immobilised ACE2. Shalash *et al*. immobilised spike RBD on the ELISA plate wells, rather than ACE2. Rather than using full length spike proteins, the other ELISA used only RBD. This change would eliminate some interactions of antibodies with other regions of the spike protein which could affect ACE2 binding. To determine if this difference is significant, a comparison of assay results from both methods on the same sets or sera would be required.

A notable limitation of the assay used in this study is the reliance on absorbance measurements without employing a calibration curve. This would limit quantitative

interpretation of assay results across samples and across spike variants. However, further development of the assay would have entailed using the same 12 dilutions of the standard control serum (Working standard NIBSC Anti-SARS-CoV-2 antibody diagnostic calibrant. NIBSC, 20/162) used in the binding ELISA, against the spike variants tested. This would have allowed comparison of the NOB characteristics of patient sera samples quantitively.

Several limitations were identified in the NOB ELISA, developed to evaluate the efficacy of antibodies in inhibition of spike-ACE2 binding. The assay was designed to detect only those antibodies that obstructed ACE2 binding by the spike protein. Other antibodies that may neutralise the virus by impacting different stages of the cell entry process,, mediated by the spike protein, were not detected by this assay. The assay did not account for antibodies that have a neutralising effect via Fc receptor mediated functions. Fc receptor mediated neutralisation could play a significant role in neutralisation *in vivo* in the patients studied. The assay lacked the ability to discern between a low number of high affinity antibodies and a high number of low affinity antibodies or any intermediates. The ELISA was conducted in a non-physiological environment, which does not mimic conditions in the human body where the virus generally enters cells – the respiratory tract. As a result, the in vivio performance of antibodies may differ from the results indicated by the assay and the results may not directly translate to the immune protected offered by the antibodies in the patient in question.

4.4   Neutralisation of cell entry assay

A limitation of the NOE assay is the target cells used. VERO E6 cells, the target cells, did not express the host expressed TMPRSS2 surface enzyme used by SARS CoV-2 in cell entry via ACE2. Therefore, the results of these NOE assays may not be fully representative of the viral entry process in TMRPSS2 expressing cells. Given this, determining the threshold for meaningful antibody neutralisation or enhancement is challenging without further experiments.

Another limitation of *in vitro* assays is the controlled environment the infection takes place in. In this assay one cell type, the VERO E6 target cells, are present. In this setup the only effect of binding antibodies that can be studied is the effect that binding has on cell entry via disabling effective functioning of the spike protein. While this is certainly a mechanism thought which antibodies may neutralise SARS CoV-2 live virus in an infected human, it is not the only method by which antibodies effect virus clearance and neutralisation of cell entry. *In vivo* antibodies work alongside other immune cells such as B cells, phagocytotic cells, natural killer cells and the complement pathway to neutralise the virus. Since this assay cannot account for the interaction of antibody-virus complexes with these cells, the conclusions drawn from these results are limited and do not fully describe the impact of the antibodies studied on neutralisation of the virus.

Another limitation of using VERO E6 cells as the target cells is that they are of non-human origin and from a non-respiratory tissue and the primary target cells of SARS CoV-2 is human respiratory cells. This can make extrapolating results of these assay to human biology difficult, beyond the direct effect measured.

VERO E6 cells are also used in a monolayer so the cultures lack three-dimensional structure. While this may pose a problem for some applications, in investigating the effects of antibody biding on cell entry this is advantageous as the cell culture produces a reliable environment with predictable and reasonably consistent monolayer. This helps with reproducibility across assays.

The use of immortalised cell line presents a unique problem. As these cells experience genetic and polymorphic drift this can lead to variations in the cell behaviour, from the origin cells.

The NOE assay in this study suffered from some problems, particularly lower than expected signal for positive controls in some assays with the Alpha variant pseudotype. These low signals in positive controls could have been due to suboptimal calibration of pseudotype production, however the high relative signals in wells treated with patient sera would seem to indicate that a sufficient amount of correctly functioning pseudotypes were produced for the assays. The difference in positive controls that particularly affected Alpha, but not Beta and Delta could suggest that the Alpha variant was simply less infective of VERO E6 cells due to the glycoprotein variant used. It was not possible to investigate this issue during this study, mainly due to incredibly limited availability of consumables – particularly tissue culture treated vessels such as plates and flasks. The COVID-19 pandemic severely affected supply of these goods and with waiting times of up to one year for some essential equipment. Equipment in the lab was strictly rationed and repeated assays to investigate the cause of these low positive controls in the Alpha pseudotypes was not possible. Alternative equipment for some steps of this assay was investigated and the methods re-calibrated to continue the experiments with alternatives. Sourcing these alternatives also took a long time and testing and recalibration of steps of this assay, especially pseudovirus production, were time consuming. Despite replacing some of the most severely limited equipment, suitable and available alternatives for consumables for all aspects of the assay could not be sourced. Efforts to continue the study only shifted the bottleneck and as the other necessary consumables were depleting, rationing took an effect again and slowed progress.

Comparing NOE assay results with binding ELISA results showed no correlation between total antibody binding activity to spike and effectiveness at neutralising pseudovirus cell entry.

Shimizu *et al*. performed a neutralisation of cell entry assay using a human dendritic cell line as target cells for infection with SARS CoV-2 spike displaying pseudotypes.

The cell line was modified to express ACE2, TMPRSS and Fc receptors on the cell surface. This makes the cell line particularly suited to investigating ADE as these cell surface proteins play a role in ADE of SARS CoV-2. Rather than investigating serum antibodies of patients with severe COVID-19, Jun Shimizu *et al*. focused on therapeutic monoclonal antibodies and post-vaccination sera. Jun Shimizu *et al*. found that the monoclonal antibodies they tested demonstrated ADE at low concentrations in ACE2 and FcR positive cells. The ADE measured was only present within a narrow concentration range which indicated that ADE was a particular risk in sub-neutralising doses. The study by Jun Shimizu *et al*. also investigated the antibody response against three SARS CoV-2 variants the Wuhan (WT) strain, Delta and Omicron.

The study by Jun Shimizu *et al*. more closely mimics *in vivo* conditions by replicating the pseudovirus assay with live virus neutralisation of entry assays. The combination of pseudotype and live virus approach allows for comparison between the two techniques. a notable difference was a higher incidence of ADE in live virus assay, which was not observed in the pseudotype assays using the same antibodies. The live virus assays also showed a greater difference in neutralisation across variants when compared to the pseudotype assay. Jun Shimizu *et al*. quantification of cell entry was achieved by qPCR of virus RNA in cells. This approach is a more direct measure of cell entry than the luciferase reporter system used in this study, however it is more time consuming and increases the complexity of the assay, introducing more sources of error.

4.5   LEXSY expression system

While we were unable to use the LEXSY system to express SARS CoV-2 spike proteins during the course of this study, Son et al. successfully expressed spike RBD using the LEXSY system in February 2024.

In the use of the LEXSY system, cells were defrosted and appeared to be healthy and motile under a microscope but did not propagate. This was likely due to the freeze thaw cycle resulting in damage to the to the kinetoplast. The cells have a single large mitochondrion, associated with the single flagellum which makes this structure particularly vulnerable to damage from ice crystals. There were a number of issues related to the delivery of the cell and it was suspected that the first batch of cells delivered to the lab were had defrosted enroute while held in customs over the weekend and refrozen in unknown conditions before delivery to the lab. To mitigate a repeat of this, I went as far as personally contacting the customs warehouse foreman to ensure the levels of dry ice were closely monitored and refreshed frequently. With this second batch of cells successful propagation was achieved, however due to significant issues with supply of essential media supplements, the experiments could not progress.

4.6   Novel computational approach to mutagenesis

During this study it was not possible to validate the approach of synonymous sequences as initial templates for EP PCR due to time restrictions. Future work for validation would involve comparisons between the mutation diversity of the spike protein generated from EP PCR using the consensus sequence versus the synonymous sequences. These experiments could be done either in a functioning LEXSY expression system or in another expression system. We would expect that the synonymous sequences undergoing EP PCR would result in a more diverse range of mutations in end products than those with a single consensus sequence as an initial template.

The use of random mutagenesis to discover impactful mutations in SARS CoV-2 spike proteins has been performed by Zahradník et al. where the team found a novel combination of mutations that greatly increased the binding affinity of spike RBD to ACE2. The use of synonymous sequences as initial templates would yield a greater diversity of mutations faster while using the same resources.

References

Andre, M., Lau, L.-S., Pokharel, M. D., Ramelow, J., Owens, F., Souchak, J., Akkaoui, J., Ales, E., Brown, H., Shil, R., Nazaire, V., Manevski, M., Paul, N. P., Esteban-Lopez, M., Ceyhan, Y., & El-Hage, N. (2023). From Alpha to Omicron: How Different Variants of Concern of the SARS-Coronavirus-2 Impacted the World. *Biology*, *12*(9), 1267. https://doi.org/10.3390/biology12091267

Azhar, E. I., Hui, D. S. C., Memish, Z. A., Drosten, C., & Zumla, A. (2019). The Middle East Respiratory Syndrome (MERS). *Infectious Disease Clinics of North America*, *33*(4), 891–905. https://doi.org/10.1016/j.idc.2019.08.001

Berman, H., Henrick, K., & Nakamura, H. (2003). Announcing the worldwide Protein Data Bank. *Nature Structural & Molecular Biology*, *10*(12), 980–980. https://doi.org/10.1038/nsb1203-980

Bloom, J. D., Beichman, A. C., Neher, R. A., & Harris, K. (2023). Evolution of the SARS-CoV-2 Mutational Spectrum. *Molecular Biology and Evolution*, *40*(4). https://doi.org/10.1093/molbev/msad085

Breitinger, U., Farag, N. S., Sticht, H., & Breitinger, H.-G. (2022). Viroporins: Structure, function, and their role in the life cycle of SARS-CoV-2. *The International Journal of Biochemistry & Cell Biology*, *145*, 106185. https://doi.org/10.1016/j.biocel.2022.106185

Cao, Y., Hiyoshi, A., & Montgomery, S. (2020). COVID-19 case-fatality rate and demographic and socioeconomic influencers: worldwide spatial regression analysis based on country-level data. *BMJ Open*, *10*(11), e043560. https://doi.org/10.1136/bmjopen-2020-043560

Cao, Y., Yang, R., Lee, I., Zhang, W., Sun, J., Wang, W., & Meng, X. (2021). Characterization of the SARS-CoV-2 E Protein: Sequence, Structure, Viroporin, and Inhibitors. *Protein Science*, *30*(6), 1114–1130. https://doi.org/10.1002/pro.4075

Carabelli, A. M., Peacock, T. P., Thorne, L. G., Harvey, W. T., Hughes, J., de Silva, T. I., Peacock, S. J., Barclay, W. S., de Silva, T. I., Towers, G. J., & Robertson, D. L. (2023). SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nature Reviews Microbiology*. https://doi.org/10.1038/s41579-022-00841-7

Cevik, M., Bamford, C. G. G., & Ho, A. (2020). COVID-19 pandemic—a focused review for clinicians. *Clinical Microbiology and Infection*, *26*(7), 842–847. https://doi.org/10.1016/j.cmi.2020.04.023

Chakraborty, C., Bhattacharya, M., & Dhama, K. (2023). SARS-CoV-2 Vaccines, Vaccine Development Technologies, and Significant Efforts in Vaccine Development during the Pandemic: The Lessons Learned Might Help to Fight against the Next Pandemic. *Vaccines*, *11*(3), 682. https://doi.org/10.3390/vaccines11030682

Chang, C., Sue, S.-C., Yu, T., Hsieh, C.-M., Tsai, C.-K., Chiang, Y.-C., Lee, S., Hsiao, H., Wu, W.-J., Chang, W.-L., Lin, C.-H., & Huang, T. (2006). Modular organization of SARS coronavirus nucleocapsid protein. *Journal of Biomedical Science*, *13*(1), 59–72. https://doi.org/10.1007/s11373-005-9035-9

Cui, J., Li, F., & Shi, Z.-L. (2019). Origin and evolution of pathogenic coronaviruses. *Nature Reviews. Microbiology*, *17*(3), 181–192. https://doi.org/10.1038/s41579-018-0118-9

Diamond, M. S., & Kanneganti, T.-D. (2022). Innate immunity: the first line of defense against SARS-CoV-2. *Nature Immunology*, *23*(2), 165–176. https://doi.org/10.1038/s41590-021-01091-0

Donnelly, C. A., Ghani, A. C., Leung, G. M., Hedley, A. J., Fraser, C., Riley, S., Abu-Raddad, L. J., Ho, L.-M., Thach, T.-Q., Chau, P., Chan, K.-P., Lam, T.-H., Tse, L.-Y., Tsang, T., Liu, S.-H., Kong, J. H., Lau, E. M., Ferguson, N. M., & Anderson, R. M. (2003). Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *The Lancet*, *361*(9371), 1761–1766. https://doi.org/10.1016/S0140-6736(03)13410-1

Ejemel, M., Li, Q., Hou, S., Schiller, Z. A., Tree, J. A., Wallace, A., Amcheslavsky, A., Kurt Yilmaz, N., Buttigieg, K. R., Elmore, M. J., Godwin, K., Coombes, N., Toomey, J. R., Schneider, R., Ramchetty, A. S., Close, B. J., Chen, D.-Y., Conway, H. L., Saeed, M., … Wang, Y. (2020). A cross-reactive human IgA monoclonal antibody blocks SARS-CoV-2 spike-ACE2 interaction. *Nature Communications*, *11*(1), 4198. https://doi.org/10.1038/s41467-020-18058-8

Engvall, E., & Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry*, *8*(9), 871–874. https://doi.org/10.1016/0019-2791(71)90454-X

Finkel, Y., Mizrahi, O., Nachshon, A., Weingarten-Gabbay, S., Morgenstern, D., Yahalom-Ronen, Y., Tamir, H., Achdout, H., Stein, D., Israeli, O., Beth-Din, A., Melamed, S., Weiss, S., Israely, T., Paran, N., Schwartz, M., & Stern-Ginossar, N. (2021). The coding capacity of SARS-CoV-2. *Nature*, *589*(7840), 125–130. https://doi.org/10.1038/s41586-020-2739-1

Fisman, D., Rivers, C., Lofgren, E., & Majumder, M. S. (2014). Estimation of MERS-Coronavirus Reproductive Number and Case Fatality Rate for the Spring 2014 Saudi Arabia Outbreak: Insights from Publicly Available Data. *PLoS Currents*. https://doi.org/10.1371/currents.outbreaks.98d2f8f3382d84f390736cd5f5fe133c

Forni, D., Cagliani, R., Clerici, M., & Sironi, M. (2017). Molecular Evolution of Human Coronavirus Genomes. *Trends in Microbiology*, *25*(1), 35–48. https://doi.org/10.1016/j.tim.2016.09.001

Gaunt, E. R., Hardie, A., Claas, E. C. J., Simmonds, P., & Templeton, K. E. (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. *Journal of Clinical Microbiology*, *48*(8), 2940–2947. https://doi.org/10.1128/JCM.00636-10

Guo, Y., Huang, L., Zhang, G., Yao, Y., Zhou, H., Shen, S., Shen, B., Li, B., Li, X., Zhang, Q., Chen, M., Chen, D., Wu, J., Fu, D., Zeng, X., Feng, M., Pi, C., Wang, Y., Zhou, X., … Rao, Z. (2021). A SARS-CoV-2 neutralizing antibody with extensive Spike binding coverage and modified for optimal therapeutic outcomes. *Nature Communications*, *12*(1), 2623. https://doi.org/10.1038/s41467-021-22926-2

Harcourt, J., Tamin, A., Lu, X., Kamili, S., Sakthivel, S. K., Murray, J., Queen, K., Tao, Y., Paden, C. R., Zhang, J., Li, Y., Uehara, A., Wang, H., Goldsmith, C., Bullock, H. A., Wang, L., Whitaker, B., Lynch, B., Gautam, R., … Thornburg, N. J. (2020). Severe Acute Respiratory Syndrome Coronavirus 2 from Patient with Coronavirus Disease, United States. *Emerging Infectious Diseases*, *26*(6), 1266–1273. https://doi.org/10.3201/eid2606.200516

Horita, N., & Fukumoto, T. (2023). Global case fatality rate from COVID-19 has decreased by 96.8% during 2.5 years of the pandemic. *Journal of Medical Virology*, *95*(1). https://doi.org/10.1002/jmv.28231

Jackson, C. B., Farzan, M., Chen, B., & Choe, H. (2022). Mechanisms of SARS-CoV-2 entry into cells. In *Nature Reviews Molecular Cell Biology* (Vol. 23, Issue 1, pp. 3–20). Nature Research. https://doi.org/10.1038/s41580-021-00418-x

Janeway, C., Travers, P., Walport, M., & Bhlomchik, M. (2001). *Immunobiology 5 : the immune system in health and disease* (5th ed.). Garland Science. https://doi.org/https://www.ncbi.nlm.nih.gov/books/NBK27160/

Janeway, C., Travers, P., Walport, M., & Shlomchik, M. (2001). *Immunobiology: the immune system in health and disease* (Vol. 2). Garland Pub. New York.

Jiang, J., Boughter, C. T., Ahmad, J., Natarajan, K., Boyd, L. F., Meier-Schellersheim, M., & Margulies, D. H. (2023). SARS-CoV-2 antibodies recognize 23 distinct epitopic sites on the receptor binding domain. *Communications Biology*, *6*(1), 953. https://doi.org/10.1038/s42003-023-05332-w

Jiao, T., Huang, Y., Sun, H., & Yang, L. (2024). Research progress of post-acute sequelae after SARS-CoV-2 infection. *Cell Death & Disease*, *15*(4), 257. https://doi.org/10.1038/s41419-024-06642-5

Johnson, S. F., Collins, J. T., D'Souza, V. M., & Telesnitsky, A. (2014). Determinants of Moloney Murine Leukemia Virus Gag-Pol and Genomic RNA Proportions. *Journal of Virology*, *88*(13), 7267–7275. https://doi.org/10.1128/JVI.03513-13

Kaduskar, O., Gurav, Y. K., Deshpande, K., Desphande, G. R., Yadav, P., Rakhe, A., Tilekar, B. N., Gomade, P., Salunke, A., Patil, C., Dange, V., Salve, P., Patsute, S., Abraham, P., & Sapkal, G. N. (2022). Understanding the dynamics of IgM & IgG antibodies in COVID-19-positive patients. *The Indian Journal of Medical Research*, *155*(5 & 6), 565–569. https://doi.org/10.4103/ijmr.IJMR_675_21

Khetran, S. R., & Mustafa, R. (2024). Correction: Mutations of SARS-CoV-2 Structural Proteins in the Alpha, Beta, Gamma, and Delta Variants: Bioinformatics Analysis. *JMIR Bioinformatics and Biotechnology*, *5*, e64915. https://doi.org/10.2196/64915

Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J. W., Kim, V. N., & Chang, H. (2020). The Architecture of SARS-CoV-2 Transcriptome. *Cell*, *181*(4), 914-921.e10. https://doi.org/10.1016/j.cell.2020.04.011

Lapuente, D., Winkler, T. H., & Tenbusch, M. (2023). B-cell and antibody responses to SARS-CoV-2: infection, vaccination, and hybrid immunity. *Cellular & Molecular Immunology*, *21*(2), 144–158. https://doi.org/10.1038/s41423-023-01095-w

Leder, P. (1982). The Genetics of Antibody Diversity. *Scientific American*, *246*(5), 102–115. http://www.jstor.org/stable/24966594

Lee, N., Chan, P. K. S., Ip, M., Wong, E., Ho, J., Ho, C., Cockram, C. S., & Hui, D. S. (2006). Anti-SARS-CoV IgG response in relation to disease severity of severe acute respiratory syndrome. *Journal of Clinical Virology*, *35*(2), 179–184. https://doi.org/10.1016/j.jcv.2005.07.005

Li, X., Yuan, H., Li, X., & Wang, H. (2023). Spike protein mediated membrane fusion during SARS-CoV-2 infection. *Journal of Medical Virology*, *95*(1). https://doi.org/10.1002/jmv.28212

Lim, Y., Ng, Y., Tam, J., & Liu, D. (2016). Human Coronaviruses: A Review of Virus–Host Interactions. *Diseases*, *4*(3), 26. https://doi.org/10.3390/diseases4030026

Liu, Q., Chi, S., Dmytruk, K., Dmytruk, O., & Tan, S. (2022). Coronaviral Infection and Interferon Response: The Virus-Host Arms Race and COVID-19. *Viruses*, *14*(7), 1349. https://doi.org/10.3390/v14071349

Luo, J., Klett, J., Gabert, J., Lipp, T., Karbach, J., Jäger, E., Borte, S., Hoffmann, R., & Milkovska-Stamenova, S. (2022). A Quantitative ELISA to Detect Anti-SARS-CoV-2 Spike IgG Antibodies in Infected Patients and Vaccinated Individuals. *Microorganisms*, *10*(9), 1812. https://doi.org/10.3390/microorganisms10091812

Macpherson, A. J., Geuking, M. B., & McCoy, K. D. (2011). Immunoglobulin A: a bridge between innate and adaptive immunity. *Current Opinion in Gastroenterology*, *27*(6), 529–533. https://doi.org/10.1097/MOG.0b013e32834bb805

Marra, M. A., Jones, S. J. M., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S. N., Khattra, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girn, N., Griffith, O. L., Leach, S. R., Mayo, M., McDonald, H., Montgomery, S. B., … Roper, R. L. (2003). The Genome Sequence of the SARS-Associated Coronavirus. *Science*, *300*(5624), 1399–1404. https://doi.org/10.1126/science.1085953

Mathieu, E., Ritchie, H., Rodés-Guirao, L., Appel, C., Giattino, C., Hasell, J., Macdonald, B., Dattani, S., Beltekian, D., Ortiz-Ospina, E., & Roser, M. (2020). Coronavirus Pandemic (COVID-19). *Our World in Data*.

McBride, R., Van Zyl, M., & Fielding, B. (2014). The Coronavirus Nucleocapsid Is a Multifunctional Protein. *Viruses*, *6*(8), 2991–3018. https://doi.org/10.3390/v6082991

Moss, P. (2022). The T cell immune response against SARS-CoV-2. In *Nature Immunology* (Vol. 23, Issue 2, pp. 186–193). Nature Research. https://doi.org/10.1038/s41590-021-01122-w

Mueller, S. N., & Rouse, B. T. (2008). Immune responses to viruses. In *Clinical Immunology* (pp. 421–431). Elsevier. https://doi.org/10.1016/B978-0-323-04404-2.10027-2

Parija, S. C. (2023). *Textbook of Microbiology and Immunology*. Springer Nature Singapore. https://doi.org/10.1007/978-981-19-3315-8

Payne, S. (2017a). Chapter 17 - Family Coronaviridae. In S. Payne (Ed.), *Viruses* (pp. 149–158). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-803109-4.00017-9

Payne, S. (2017b). Family Coronaviridae. In *Viruses* (pp. 149–158). Elsevier. https://doi.org/10.1016/B978-0-12-803109-4.00017-9

Plümper, T., & Neumayer, E. (2022). Lockdown policies and the dynamics of the first wave of the Sars-CoV-2 pandemic in Europe. *Journal of European Public Policy*, *29*(3), 321–341. https://doi.org/10.1080/13501763.2020.1847170

Qing, E., Kicmal, T., Kumar, B., Hawkins, G. M., Timm, E., Perlman, S., & Gallagher, T. (2021). Dynamics of SARS-CoV-2 Spike Proteins in Cell Entry: Control Elements in the Amino-Terminal Domains. *MBio*, *12*(4). https://doi.org/10.1128/mBio.01590-21

Redondo, N., Zaldívar-López, S., Garrido, J. J., & Montoya, M. (2021). SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns. *Frontiers in Immunology*, *12*. https://doi.org/10.3389/fimmu.2021.708264

Santos-Mendoza, T. (2023). The Envelope (E) Protein of SARS-CoV-2 as a Pharmacological Target. *Viruses*, *15*(4), 1000. https://doi.org/10.3390/v15041000

Sette, A., & Crotty, S. (2021). Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell*, *184*(4), 861–880. https://doi.org/10.1016/j.cell.2021.01.007

Shalash, A. O., Azuar, A., Madge, H. Y. R., Modhiran, N., Amarilla, A. A., Liang, B., Khromykh, A. A., Watterson, D., Young, P. R., Toth, I., & Skwarczynski, M. (2021). Detection and Quantification of SARS-CoV-2 Receptor Binding Domain Neutralization by a Sensitive Competitive ELISA Assay. *Vaccines*, *9*(12), 1493. https://doi.org/10.3390/vaccines9121493

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., & Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature*, *581*(7807), 221–224. https://doi.org/10.1038/s41586-020-2179-y

Shimizu, J., Sasaki, T., Koketsu, R., Morita, R., Yoshimura, Y., Murakami, A., Saito, Y., Kusunoki, T., Samune, Y., Nakayama, E. E., Miyazaki, K., & Shioda, T. (2022). Reevaluation of antibody-dependent enhancement of infection in anti-SARS-CoV-2 therapeutic antibodies and mRNA-vaccine antisera using FcR- and ACE2-positive cells. *Scientific Reports*, *12*(1), 15612. https://doi.org/10.1038/s41598-022-19993-w

Siu, Y. L., Teoh, K. T., Lo, J., Chan, C. M., Kien, F., Escriou, N., Tsao, S. W., Nicholls, J. M., Altmeyer, R., Peiris, J. S. M., Bruzzone, R., & Nal, B. (2008). The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. *Journal of Virology*, *82*(22), 11318–11330. https://doi.org/10.1128/JVI.01052-08

Son, L., Kost, V., Maiorov, V., Sukhov, D., Arkhangelskaya, P., Ivanov, I., Kudryavtsev, D., Siniavin, A., Utkin, Y., & Kasheverov, I. (2024). Efficient Expression in Leishmania tarentolae (LEXSY) of the Receptor-Binding Domain of the SARS-

CoV-2 S-Protein and the Acetylcholine-Binding Protein from Lymnaea stagnalis. *Molecules*, *29*(5), 943. https://doi.org/10.3390/molecules29050943

Tangye, , Stuart G., Phillips, J. H., & Lanier, L. L. (2000). The CD2-subset of the Ig superfamily of cell surface molecules: receptor–ligand pairs expressed by NK cells and other immune cells. *Seminars in Immunology*, *12*(2), 149–157. https://doi.org/10.1006/smim.2000.0217

Tiwari, R., Dhama, K., Sharun, K., Iqbal Yatoo, Mohd., Malik, Y. S., Singh, R., Michalak, I., Sah, R., Bonilla-Aldana, D. K., & Rodriguez-Morales, A. J. (2020). COVID-19: animals, veterinary and zoonotic links. *Veterinary Quarterly*, *40*(1), 169–182. https://doi.org/10.1080/01652176.2020.1766725

Tortorici, M. A., & Veesler, D. (2019). *Structural insights into coronavirus entry* (pp. 93–116). https://doi.org/10.1016/bs.aivir.2019.08.002

van der Hoek, L., Pyrc, K., Jebbink, M. F., Vermeulen-Oost, W., Berkhout, R. J. M., Wolthers, K. C., Wertheim-van Dillen, P. M. E., Kaandorp, J., Spaargaren, J., & Berkhout, B. (2004). Identification of a new human coronavirus. *Nature Medicine*, *10*(4), 368–373. https://doi.org/10.1038/nm1024

Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., & Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*, *181*(2), 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058

Wang, Y., Grunewald, M., & Perlman, S. (2020). Coronaviruses: an updated overview of their replication and pathogenesis. *Coronaviruses: Methods and Protocols*, 1–29.

Wen, J., Cheng, Y., Ling, R., Dai, Y., Huang, B., Huang, W., Zhang, S., & Jiang, Y. (2020). Antibody-dependent enhancement of coronavirus. *International Journal of Infectious Diseases*, *100*, 483–489. https://doi.org/10.1016/j.ijid.2020.09.015

Wiegand, T., Nemudryi, A., Nemudraia, A., McVey, A., Little, A., Taylor, D. N., Walk, S. T., & Wiedenheft, B. (2022). The Rise and Fall of SARS-CoV-2 Variants and Ongoing Diversification of Omicron. *Viruses*, *14*(9), 2009. https://doi.org/10.3390/v14092009

Woo, P. C. Y., Lau, S. K. P., Chu, C., Chan, K., Tsoi, H., Huang, Y., Wong, B. H. L., Poon, R. W. S., Cai, J. J., Luk, W., Poon, L. L. M., Wong, S. S. Y., Guan, Y., Peiris, J. S. M., & Yuen, K. (2005). Characterization and Complete Genome Sequence of a Novel Coronavirus, Coronavirus HKU1, from Patients with Pneumonia. *Journal of Virology*, *79*(2), 884–895. https://doi.org/10.1128/JVI.79.2.884-895.2005

Woo, P. C. Y., Lau, S. K. P., Lam, C. S. F., Lau, C. C. Y., Tsang, A. K. L., Lau, J. H. N., Bai, R., Teng, J. L. L., Tsang, C. C. C., Wang, M., Zheng, B.-J., Chan, K.-H., & Yuen, K.-Y. (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. *Journal of Virology*, *86*(7), 3995–4008. https://doi.org/10.1128/JVI.06540-11

World Health Organization 2023 data.who.int. (n.d.-a). *WHO Coronavirus (COVID-19) dashboard > Cases [Dashboard]*. Https://Data.Who.Int/Dashboards/Covid19/Cases.

World Health Organization 2023 data.who.int. (n.d.-b). *WHO Coronavirus (COVID-19) dashboard > Variants [Dashboard].* Https://Data.Who.Int/Dashboards/Covid19/Variants.

Wu, A., Peng, Y., Huang, B., Ding, X., Wang, X., Niu, P., Meng, J., Zhu, Z., Zhang, Z., Wang, J., Sheng, J., Quan, L., Xia, Z., Tan, W., Cheng, G., & Jiang, T. (2020). Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China. *Cell Host & Microbe*, *27*(3), 325–328. https://doi.org/10.1016/j.chom.2020.02.001

Xu, Y., Zhu, J., Liu, Y., Lou, Z., Yuan, F., Liu, Y., Cole, D. K., Ni, L., Su, N., Qin, L., Li, X., Bai, Z., Bell, J. I., Pang, H., Tien, P., Gao, G. F., & Rao, Z. (2004). Characterization of the Heptad Repeat Regions, HR1 and HR2, and Design of a Fusion Core Structure Model of the Spike Protein from Severe Acute Respiratory Syndrome (SARS) Coronavirus. *Biochemistry*, *43*(44), 14064–14071. https://doi.org/10.1021/bi049101q

You, Y., Yang, X., Hung, D., Yang, Q., Wu, T., & Deng, M. (2024). Asymptomatic COVID-19 infection: diagnosis, transmission, population characteristics. *BMJ Supportive & Palliative Care*, *14*(e1), e220–e227. https://doi.org/10.1136/bmjspcare-2020-002813

Zahradník, J., Marciano, S., Shemesh, M., Zoler, E., Harari, D., Chiaravalli, J., Meyer, B., Rudich, Y., Li, C., Marton, I., Dym, O., Elad, N., Lewis, M. G., Andersen, H., Gagne, M., Seder, R. A., Douek, D. C., & Schreiber, G. (2021). SARS-CoV-2 variant prediction and antiviral drug design are enabled by RBD in vitro evolution. *Nature Microbiology*, *6*(9), 1188–1198. https://doi.org/10.1038/s41564-021-00954-4

Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E., & Fouchier, R. A. M. (2012). Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *New England Journal of Medicine*, *367*(19), 1814–1820. https://doi.org/10.1056/NEJMoa1211721

Zhang, Q., Xiang, R., Huo, S., Zhou, Y., Jiang, S., Wang, Q., & Yu, F. (2021). Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. *Signal Transduction and Targeted Therapy*, *6*(1), 233. https://doi.org/10.1038/s41392-021-00653-w

Zhou, P., Yang, X. Lou, Wang, X. G., Hu, B., Zhang, L., Zhang, W., Si, H. R., Zhu, Y., Li, B., Huang, C. L., Chen, H. D., Chen, J., Luo, Y., Guo, H., Jiang, R. Di, Liu, M. Q., Chen, Y., Shen, X. R., Wang, X., … Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7

Zhou, Z., Ali, A., Walelign, E., Demissie, G. F., El Masry, I., Abayneh, T., Getachew, B., Krishnan, P., Ng, D. Y. M., Gardner, E., Makonnen, Y., Miguel, E., Chevalier, V., Chu, D. K., So, R. T. Y., Von Dobschuetz, S., Mamo, G., Poon, L. L. M., & Peiris, M. (2023). Genetic diversity and molecular epidemiology of Middle East Respiratory Syndrome Coronavirus in dromedaries in Ethiopia, 2017–2020. *Emerging Microbes & Infections*, *12*(1). https://doi.org/10.1080/22221751.2022.2164218

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733. https://doi.org/10.1056/NEJMoa2001017

&#xfffd;

Appendices

Figure 17

Sigmoidal 4-parameter logistic model binding curves for NIBSC Anti-SARS-CoV-2 antibody diagnostic calibrant 20/162, against SARS CoV-2 variants.

Sample ID (BAU)	alpha A:1	alpha A:2	beta B:1	beta B:2	delta C:1	delta C:2
20/162 1000	3.03	3.07	1.93	2.21	2.38	2.46
20/162 500	2.59	2.60	2.04	2.26	2.35	2.34
20/162 250	2.07	2.04	1.63	1.65	1.83	1.81
20/162 125	1.53	1.45	1.22	1.26	1.24	1.19
20/162 62.5	1.28	1.27	0.80	0.85	0.91	0.92
20/162 31.24	1.09	0.75	0.49	0.60	0.53	0.58
20/162 15.6	0.64	0.63	0.31	0.36	0.33	0.36
20/162 7.8	0.47	0.38	0.22	0.27	0.20	0.23
20/162 3.9	0.54	1.00	0.21	0.21	0.18	0.19
20/162 1.95	0.33	0.30	0.16	0.15	0.12	0.13
20/162 0.97	0.33	0.30	0.15	0.14	0.13	0.12
20/162 0.49	0.37	0.33	0.11	0.12	0.09	0.09
No Serum	0.28	0.28	0.09	0.14	0.09	0.08

Figure 18

Figure 18.   Antibody binding activity (BAU/mL) for patient sera samples as quantified by ELISA against spike variants.

Serum samples from 6 individuals admitted to hospital with COVID-19 were tested for antibody binding affinity against three SARS CoV-2 spike protein variants: Alpha, Beta and Delta. An ELISA was performed, and absorbance (OD450) was measured to quantify binding affinity, compared to a 12-point curve standard control to determine BAU/mL. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. ns is non-significant, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Sample ID (BAU)	BAU (X)	Alpha A:1	Alpha A:2	Beta B:1	Beta B:2	Delta C:1	Delta C:2
20/162 (1000)	1000	3.03	3.07	1.93	2.21	2.38	2.46
20/162 (500)	500	2.59	2.6	2.04	2.26	2.35	2.34

20/162 (250)	250	2.07	2.04	1.63	1.65	1.83	1.89
20/162 (125)	125	1.53	1.45	1.22	1.26	1.24	1.34
20/162 (62.5)	62.5	1.28	1.27	0.8	0.85	0.91	0.93
20/162 (31.25)	31.25	1.09	0.75	0.49	0.6	0.53	0.58
20/162 (15.6)	15.625	0.62	0.62	0.31	0.36	0.33	0.36
20/162 (7.8)	7.8125	0.47	0.38	0.22	0.27	0.2	0.23
20/162 (3.9)	3.90625	0.54	1	0.21	0.21	0.18	0.19
20/162 (1.95)	1.953125	0.33	0.3	0.16	0.15	0.12	0.13
20/162 (0.97)	0.976563	0.33	0.3	0.15	0.14	0.13	0.13
20/162 (0.49)	0.488281	0.37	0.33	0.11	0.12	0.09	0.09
No Serum (0)	0	0.28	0.28	0.09	0.14	0.09	0.08
974	-	2.79	2.95	1.65	1.68	2.53	2.51
977	-	2.95	3.03	1.8	1.99	2.61	2.81
980	-	2.87	2.9	2.01	2	2.47	2.53
981	-	2.58	2.79	0.74	0.73	1.13	1.57
1024	-	2.56	2.79	1.48	1.64	1.89	2.16
1100	-	2.72	2.88	1.64	1.66	2.54	2.3

Figure 20.

Biotinylated ACE2 immobilised by NeutrAvidin coated wells in a 96 well plate.

Biotin and ace 2 binding

80 40 20 10

Ace2 (µg/mL)	Spike (µg/mL)	A:Y1	A:Y2	A:Y3	A:Y4
80	2.0000	1.18	1.65	1.94	1.82
40	1.0000	2.32	2.30	2.09	1.94
20	0.5000	2.25	1.89	2.44	2.24

10	0.2500	1.85	2.35	1.63	1.67
5	0.1250	1.57	1.84	1.89	2.08
PBS	0.0000	0.15	0.14	0.13	0.12
PBS	0.0000	0.16	0.14	0.23	0.16

Figure 21

Neutralisation of binding (NOB) ELISA calibration for ACE2 and spike proteins.

To calibrate the NOB ELISA, high binding 96 well plate wells were coated with NeutrAvidin, followed by incubation with biotinylated ACE2 to immobilise the ACE2 on the well surface. ACE2 was immobilised at total protein levels of 80, 40, 20 and 10 ng. To test SARS CoV-2 spike binding to immobilised ACE2 spike protein, mouse Fc tagged spike protein was added to wells at a total protein level of 80, 40, 20 and 10 ng. Interaction between ACE2 and spike was achieved by detection of remaining spike using HRP conjugated anti-mouse FC antibody and TMB. Values shown in cells are means of signal strength of assay, measured by OD at 450 nm.

Spike total protein (ng)	ACE2 80 ng			ACE2 40 ng			ACE2 20 ng			ACE2 10 ng		
	1	2	3	1	2	3	1	2	3	1	2	3
80 ng	0.254	0.261	0.265	0.218	0.223	0.227	0.240	0.243	0.244	0.223	0.226	0.226
40 ng	0.242	0.248	0.245	0.224	0.230	0.229	0.208	0.212	0.210	0.219	0.224	0.224
20 ng	0.262	0.266	0.267	0.239	0.243	0.246	0.220	0.224	0.223	0.212	0.217	0.216
10 ng	0.280	0.287	0.288	0.237	0.242	0.241	0.265	0.273	0.272	0.239	0.244	0.244

Figure 22

Figure 22.   Spike recombinant protein lot validation.

The NOB ELISA was not functionating. To determine if the lot of spike being used in the assay was the cause of the error, the assay was repeated with a range of ACE2 protein levels with the original lot of spike, a second functionally identical lot of spike and no spike.

ACE2 (ng)	No spike			Spike lot 1			Spike lot 2		
	Rep 1	Rep 2	Rep 3	Rep 1	Rep 2	Rep 3	Rep 1	Rep 2	Rep 3
80	0.018	0.021	0.022	0.225	0.230	0.245	0.255	0.258	0.285
40	0.027	0.029	0.033	0.290	0.305	0.305	0.270	0.285	0.285
20	0.037	0.041	0.044	0.210	0.220	0.235	0.265	0.275	0.280
10	0.048	0.052	0.055	0.235	0.245	0.260	0.270	0.280	0.285
5	0.055	0.065	0.075	0.195	0.205	0.200	0.320	0.335	0.345

Figure 23

Figure 23.   ACE2 recombinant protein lot validation.

The NOB ELISA was not functionating. To determine if the lot of ACE2 being used in the assay was the cause of the error, the assay was repeated with a range of spike protein levels with the original lot of ACE2, a second functionally identical lot of ACE2 and no ACE2. Binding of spike to immobilised ACE2 was measured using HRP conjugated goat anti-mouse antibody and TMB colourimetric reaction to detect the mouse Fc tag of the spike protein.

Spike (ng)	Lot 1			Lot 2			No ACE2		
	Rep 1	Rep 2	Rep 3	Rep 1	Rep 2	Rep 3	Rep 1	Rep 2	Rep 3
80.0	0.091	0.106	0.103	0.611	0.648	0.641	0.589	0.618	0.623
8.0	0.083	0.078	0.079	0.092	0.104	0.104	0.118	0.122	0.120
0.8	0.111	0.127	0.122	0.078	0.087	0.075	0.093	0.086	0.091
0	0.136	0.122	0.132	0.067	0.072	0.071	0.078	0.081	0.081

Figure 24

Figure 24.   Neutralisation of cell entry assay for serum 974

Serum sample from an individual admitted to hospital with COVID-19 was tested for neutralisation of cell entry (NOE) for three MLV based pseudotypes, each expressing one of three variants of the SARs CoV-2 spike protein. NOE assay was performed, and luminescence was measured to quantify cell entry. Two-way ANOVA (a = 0.05) with post-hoc Tukey's test was used to measure the statical significance of antibody binding for each sample between variants. non-significant results are not shown for clarity, P * < 0.05, P ** < 0.01, P *** < 0.001, P **** < 0.0001.

Sample 974	Alpha A:1	Alpha A:2	Alpha A:3	Beta B:1	Beta B:2	Beta B:3	Delta C:1	Delta C:2	Delta C:3
Positive control	2345	1238	1310	6867	5336	4640	5825	4831	6454
1:200	3043	1401	2611	5088	10003	6033	8731	5683	3591
1:400	2462	2055	1518	9141	6850	6583	7350	6901	5876
1:800	2597	7408	2877	6220	4666	4057	8439	6827	7011
1:1600	2779	8949	7255	7791	5244	5600	5895	4603	5222
Negative control	1	23	20	27	49	111	25	50	35

Figure 25

Sample 977	Alpha A:1	Alpha A:2	Alpha A:3	Beta B:1	Beta B:2	Beta B:3	Delta C:1	Delta C:2	Delta C:3
Positive control	1405	1812	1453	5825	4831	6454	3618	3090	3371
1:200	1070	2161	947	4711	4259	3344	9849	7938	7667
1:400	1754	2446	1958	6228	6528	4648	7041	5472	5737
1:800	3025	2863	2818	7231	5123	6530	3835	4585	4014
1:1600	3247	3843	4161	4330	5624	5374	4146	3142	4529
Negative	9	22	35	25	50	35	62	49	61

contro l								

Figure 26

Sampl e 980	Alpha A:1	Alpha A:2	Alpha A:3	Beta B:1	Beta B:2	Beta B:3	Delta C:1	Delta C:2	Delta C:3
Positiv e contro l	1405	1812	1453	11481	8101	9591	1514	1493	3031
1:200	913	1550	1087	10279	7851	8843	426	1123	1215
1:400	2492	2010	5112	13061	9646	9537	516	745	1983
1:800	1481	2686	2711	9014	10952	9537	2541	1260	353
1:1600	2628	3645	4790	8892	10456	—	1132	1045	—
Negati ve contro l	9	22	35	273	21	25	24	210	490

Figure 27

Sampl e 981	Alpha A:1	Alpha A:2	Alpha A:3	Beta B:1	Beta B:2	Beta B:3	Delta C:1	Delta C:2	Delta C:3
Positiv e contro l	1405	1812	1453	5825	4831	6454	5825	4831	6454
1:200	1403	1802	1540	5166	6484	7208	1877	3333	2452
1:400	4847	1777	1180	8765	9004	7057	4036	5457	5989
1:800	1928	1994	5081	7885	7984	10424	6530	5322	6243
1:1600	2183	3739	6891	5144	5170	6736	5329	4853	—
Negati ve contro l	9	22	35	25	50	35	25	50	35

Figure 28

Sampl e 1024	Alpha A:1	Alpha A:2	Alpha A:3	Beta B:1	Beta B:2	Beta B:3	Delta C:1	Delta C:2	Delta C:3

Positive control	481	1656	1119	693	1076	1927	696	2803	1629
1:200	3389	2314	4552	11228	14576	12087	6009	5927	7072
1:400	3201	1977	2415	7708	6384	10785	5458	4995	8361
1:800	1978	1991	2414	5399	8306	6175	3946	4370	3709
1:1600	2775	2911	2116	6400	9118	7071	2927	2214	3369
Negative control	70	49	52	91	69	34	81	67	46

Figure 29

Sample 1100	Alpha A:1	Alpha A:2	Alpha A:3	Beta B:1	Beta B:2	Beta B:3	Delta C:1	Delta C:2	Delta C:3
Positive control	1436	1785	446	384	1321	1510	3618	3090	3371
1:200	9492	5823	3299	16096	8825	11670	6874	8291	7488
1:400	5418	4712	5492	15047	12979	14242	5131	5909	5869
1:800	4206	6285	9084	14266	13693	14292	3060	4313	3453
1:1600	3444	6221	7455	10916	16054	13607	3222	3400	4947
Negative control	31	37	42	38	35	36	62	49	61

Synonymaker

```python
#!/usr/bin/env python3

# -*- coding: utf-8 -*-


"""

Purpose:

- Take a protein sequence (one-letter amino-acid codes) and generate many DNA
sequences that all translate to the same protein (synonymous codons chosen at
random)

- Remove duplicate DNA sequences and count how often each appears (redundancy)
```

```python
    - Score each DNA sequence using a codon-usage table

    - Save results and simple bar-chart PDFs
    """


import argparse

import random

from pathlib import Path

from collections import Counter


import pandas as pd

import matplotlib.pyplot as plt




# ====== FILE I/O (READ/WRITE) ======


def read_codon_table_xls(path: str) -> dict[str, list[str]]:
    """

    Read the Excel codon table you provide.
    Expected format:
      - Each column header is an amino-acid one-letter code (A, R, N, …, *)
      - Under each column: the DNA codons for that amino acid (use T not U)
        - Returns a dictionary {AA: [list of codons]} for later use.
    """

    df = pd.read_excel(path, dtype=str)

    codon_dict = {}

    for aa in df.columns:

        # Clean empty cells, standardize to uppercase DNA (T not U), keep only
triplets
```

```python
        vals = [v.strip().upper() for v in df[aa].dropna().tolist() if v and
v.strip()]

        vals = [v.replace("U", "T") for v in vals if len(v) == 3]

        if vals:

            codon_dict[str(aa).strip().upper()] = vals

    return codon_dict


def read_protein_fasta(path: str) -> str:
    """

    Read your protein sequence from FASTA or plain text.

    - Lines starting with '>' are ignored (FASTA headers)

    - Spaces and newlines are removed

    - One-letter amino-acid codes are assumed

    """

    lines = Path(path).read_text(encoding="utf-8").splitlines()

    seq = "".join([ln.strip() for ln in lines if not
ln.strip().startswith(">")])

    seq = seq.replace(" ", "").upper()

    if not seq:

        raise ValueError("Protein sequence is empty.")

    return seq


def write_lines(path: str, lines: list[str]) -> None:
    """Write a list of strings to a text file, one per line."""

    Path(path).write_text("\n".join(lines) + "\n", encoding="utf-8")
```

```python
# ====== SYNONYMOUS DNA GENERATION ======


def aa_to_random_codon(aa: str, table: dict[str, list[str]], rng: random.Random)
-> str:

    """

    For a given amino acid, pick one codon at random from your codon table.

    This makes sequences "synonymous" but different at the DNA level.

    """

    if aa not in table or not table[aa]:

        raise ValueError(f"No codons available for amino acid '{aa}'.")

    return rng.choice(table[aa])




def generate_synonymous_set(protein: str, table: dict[str, list[str]],

                            n: int, seed: int | None = None) -> list[str]:

    """

    Create N different DNA sequences that all translate to the same protein.

    Each position uses a random codon that matches the amino acid at that
position.

    """

    rng = random.Random(seed)

    # Safety check: ensure every amino acid in your protein exists in the codon
table

    missing = sorted({aa for aa in set(protein) if aa not in table})

    if missing:

        raise ValueError(f"AAs missing from codon table: {missing}")

    return ["".join(aa_to_random_codon(aa, table, rng) for aa in protein) for _
in range(n)]
```

```python
# ====== DUPLICATES + REDUNDANCY ======


def deduplicate(seq_list: list[str]) -> list[str]:

    """Remove exact duplicates to get a "non-redundant" set of DNA sequences."""

    seen, out = set(), []

    for s in seq_list:

        if s not in seen:

            seen.add(s)

            out.append(s)

    return out




def redundancy_counts(seq_list: list[str]) -> pd.DataFrame:

    """

    Count how many times each DNA sequence appears (higher = more redundancy).

    Returns a table with columns: sequence, count.

    """

    cnt = Counter(seq_list)

    df = pd.DataFrame(cnt.items(), columns=["sequence",
"count"]).sort_values("count", ascending=False)

    return df.reset_index(drop=True)




# ====== CODON-USAGE SCORING (OPTIONAL) ======


def read_codon_scores_tsv(path: str) -> dict[str, float]:

    """

    Read a simple table of codon weights you provide (TSV: CODON<TAB>SCORE).

    Examples of SCORE: usage %, usage per 1000, or RSCU.
```

```python
    """

    df = pd.read_csv(path, sep="\t", header=None, names=["codon", "score"],
dtype={"codon": str})

    d = {}

    for _, row in df.iterrows():

        c = str(row["codon"]).strip().upper().replace("U", "T")

        d[c] = float(row["score"])

    return d




def score_sequence(seq: str, score_map: dict[str, float]) -> float:

    """

    Split a DNA sequence into codons and sum their weights from your score
table.

    """

    if len(seq) % 3 != 0:

        raise ValueError("Sequence length is not a multiple of 3.")

    s = 0.0

    for i in range(0, len(seq), 3):

        codon = seq[i:i+3]

        if codon not in score_map:

            raise ValueError(f"Codon not found in score table: {codon}")

        s += score_map[codon]

    return s




def score_sequences(seq_list: list[str], score_map: dict[str, float]) ->
pd.DataFrame:

    """Score every sequence and return a table with columns: sequence, score."""
```

```python
    rows = [{"sequence": s, "score": score_sequence(s, score_map)} for s in
seq_list]

    return pd.DataFrame(rows).sort_values("score",
ascending=False).reset_index(drop=True)


# ====== SIMPLE BAR PLOTS ======


def plot_bar(df: pd.DataFrame, x: str, y: str, title: str, out_pdf: str) ->
None:

    """

    Make a basic bar chart and save as a PDF.

    Useful for a quick visual check (e.g., which sequences dominate, which score
highest).

    """

    ax = df.plot(kind="bar", x=x, y=y, legend=False)

    ax.set_xlabel(x)

    ax.set_ylabel(y)

    ax.set_title(title)

    plt.tight_layout()

    plt.savefig(out_pdf)

    plt.close()


# ====== COMMAND-LINE WRAPPER ======


def main():
    # Define the inputs/outputs you can set when running the script

    p = argparse.ArgumentParser(description="Synonymous DNA generator +
redundancy + scoring.")
```

```python
    p.add_argument("--codon_table_xls", default="ctab.xls",
                   help="Excel codon table (columns=AA; rows=codons).")

    p.add_argument("--protein_fasta", default="seq.fas",
                   help="Protein FASTA/plain-text file (one-letter codes).")

    p.add_argument("--n", type=int, default=200,
                   help="How many DNA sequences to generate.")

    p.add_argument("--seed", type=int, default=None,
                   help="Random seed for reproducibility (optional).")

    p.add_argument("--out_all", default="syn_out.txt",
                   help="All generated sequences (duplicates kept).")

    p.add_argument("--out_nr", default="nr_output.txt",
                   help="Non-redundant sequences (duplicates removed).")

    p.add_argument("--redundancy_csv", default="redundancy_count.csv",
                   help="CSV of sequence counts (redundancy).")

    p.add_argument("--redundancy_pdf", default="redundancy_plot.pdf",
                   help="Bar chart PDF of the most common sequences.")

    p.add_argument("--score_tsv", default=None,
                   help="Optional: TSV with codon<tab>score (usage %, per-1000,
or RSCU).")

    p.add_argument("--scores_csv", default="sequence_scores.csv",
                   help="CSV of per-sequence scores (if scoring is used).")

    p.add_argument("--scores_pdf", default="codon_score_plot.pdf",
                   help="Bar chart PDF of top sequence scores (if scoring is
used).")

    args = p.parse_args()


    # 1) Load inputs

    table = read_codon_table_xls(args.codon_table_xls)

    protein = read_protein_fasta(args.protein_fasta)
```

```python
    # 2) Generate synonymous DNA sequences and save them

    seqs = generate_synonymous_set(protein, table, args.n, seed=args.seed)

    write_lines(args.out_all, seqs)



    # 3) Remove duplicates; save counts and generate a bar chart (top 50 for
legibility)

    nr = deduplicate(seqs)

    write_lines(args.out_nr, nr)

    red_df = redundancy_counts(seqs)

    red_df.to_csv(args.redundancy_csv, index=False)

    plot_bar(

        red_df.head(50),

        x="sequence",

        y="count",

        title="Repetitions in generated synonymous sequence set",

        out_pdf=args.redundancy_pdf,

    )



    # 4)  Score sequences using your codon-usage weights; save results and plot

    if args.score_tsv:

        score_map = read_codon_scores_tsv(args.score_tsv)

        sc_df = score_sequences(nr, score_map)

        sc_df.to_csv(args.scores_csv, index=False)

        plot_bar(

            sc_df.head(50),

            x="sequence",

            y="score",

            title="Codon-usage score (higher = closer to reference)",

            out_pdf=args.scores_pdf,
```

```
    )


    # 5) Console summary of where results were written

    print(f"Generated: {args.out_all}")

    print(f"Non-redundant: {args.out_nr}")

    print(f"Redundancy table: {args.redundancy_csv} | Plot:
{args.redundancy_pdf}")

    if args.score_tsv:

        print(f"Scores: {args.scores_csv} | Plot: {args.scores_pdf}")




if __name__ == "__main__":

    main()
```

Synonymaker script.
Comments explaining the function of each section of code are highlighted in grey