WITS UNIVERSITY

Identification of bNAb-Initiating HIV-1 Envelopes to Inform the Design of V3-Glycan-Specific Germline Targeting Immunogens.

Elizabeth Venter

Johannesburg, July 2022

A Dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Medicine.

Declaration

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I Elizabeth Venter (Student number: 1043057) am a student registered for the degree of MSc (MED) in the academic year 2022.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: _____ Date: 05 July 2022

Presentations

Venter, E., Kitchin, D., Moyo-Gwete, T., Pankow, A., Murrell, B., Morris, L., Moore, P. L. (2021) 'Identification of bNAb-initiating HIV-1 envelopes to inform the design of V3-glycan epitope specific immunogens', 10th Infectious Disease in Africa Symposium // Virtual. 2021. Oral presentation.

Abstract

Broadly neutralising antibodies (bNAbs) are able to neutralise diverse HIV-1 strains and are therefore important to elicit by an effective HIV-1 vaccine. bNAbs that target the V3-glycan epitope on the HIV-1 envelope (Env) are among the most common and broad bNAbs elicited during infection, and predominantly use a common IGHV4 germline gene, making this a particularly promising vaccine target. This study aimed to identify Envs which triggered broad V3-glycan-specific responses (bNAb-initiating Envs) in HIV-1 infected donors, which could be used to inform the design of immunogens. Putative bNAb-initiating Envs were identified through PacBio deep sequencing of the *env* genes for two HIV-infected participants (CAP255 and CAP314) who developed V3-glycan bNAbs. The timepoints sequenced were selected based on the emergence of the V3-glycan bNAb precursor, the unmutated common ancestor (UCA). The CAP255 and CAP314 bNAb-initiating Envs were selected based on the absence of two key glycans and the shortening of the V1 loop, respectively, both features are known enhance accessibility for V3-glycan binding. These two Envs were used as the basis for the design of 11 Env trimer immunogens, each incorporating various mutations to increase stability and affinity for V3-glycan bNAbs. We expressed the trimers in mammalian cells and purified them using affinity and size exclusion chromatography. Although CAP314 did not express stable trimers, even with the introduction of additional stabilising mutations, CAP255-derived trimers were more successful, and the CAP255.GT0 and CAP255.GT2 immunogens produced the highest yields. Binding and functional assays showed CAP255.GT2 to be the most stable and well-formed trimer. No UCA interaction were observed with the CAP255.GT2 trimer but early CAP255 intermediate antibodies neutralised the CAP255.GT2 pseudovirus. This study resulted in the design of a successful subtype C trimer that can be used as a basis for future immunisation studies.

Acknowledgements

Firstly I would like to thank my amazing supervisors for all their support. Dr Thandeka Moyo-Gwete for being the most driven scientist and for always pushing me to do my best, Dr Dale Kitchin for teaching me to focus on the little details and for always understanding and reminding me to take breaks and Prof Penny Moore for setting such great examples and for creating the wonderful lab culture in our lab. I'd like to thank them for always reminding me that there is light at the end of the tunnel. I'm so grateful for having had them to guide me through it.

I'd also like to thank the Antibody Immunity Research Unit for all the advice, support, friendly faces, lab bench chats and the best role models.

I'd like to thank Dr Elise Landais at IAVI for providing me with unpublished sequencing data of the PCIN39 lineage for the use in my experiments.

I'd like to thank Dr Alec Pankow at the Icahn School of Medicine at Mount Sinai and Dr Ben Murrell at the Karolinka Institute for assistance in the processing and analysis of our resulting raw PacBio *env* sequencing data.

I'd like to thank Dr Nishal Parbhoo at the School of Agriculture and Life Sciences, UNISA for showing me how to run and analyse DSC experiments and for allowing me the use of their instrument.

I'd also like to thank the University of the Witwatersrand Postgraduate Merit Award, the Poliomyelitis Research Foundation (PRF) and the National Research Foundation (NRF) for their funding support, without whom this project would not have been possible.

Lastly I'd like to thank my family and friends for always being there, even when they were far away. For all the late night snacks and the early morning coffees and for always listening to all my complaints even though most of the time they didn't understand a word I was saying.

Table of Contents

List of Figures

List of Tables

Abbreviations

AA: amino acid
AIDS: acquired immunodeficiency syndrome
ART: antiretroviral therapy
BCR: B-cell receptors
bNAb: broadly neutralising antibody
CCR5: C-C chemokine receptor 5
CCS: circular consensus sequencing
CD4bs: CD4 binding site
cDNA: complementary DNA
CDR: complementarity determining regions
CH: constant heavy
CL: constant light
CRFs: circulating recombinant forms
CT: cytoplasmic tail
CXCR4: C-X-C chemokine receptor 4
D: diversity
DEAE: diethylaminoethyl
DMEM: Dulbecco's modified eagle media
DNA: deoxyribonucleic acid
DSC: differential scanning calorimerty
ELISA: enzyme linked immunosorbent assay
Env: envelope
Fab: fragment antigen binding
FBS: foetal bovine serum
Fc: fragment crystallisable
FP: fusion peptide
GM: growth media
HEK: human embryonic kidney
HIV-1: human immunodeficiency virus 1
HR: heptad repeats
IC50: inhibitor concentration
IgG: Immunoglobulin G
IGK: IG kappa
IGL: IG lambda
J: joining
kb: kilobase
mAb: monoclonal antibody
MPER: membrane proximal region
mRNA: messenger RNA
NGS: next generation sequencing
PBS: phosphate-buffered saline
PCR: polymerase chain reaction
PEImax: polyethylenimine hydrochloride
PEP: post exposure prophylaxis
PMTCT: prevention of mother-to-child transmission
PNG: potential N-linked glycosylation site
PrEP: pre-exposure prophylaxis
RLU: relative luminescence unit
SDS-PAGE: sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC: size exclusion chromatography
SGA: single genome amplification
SHM: somatic hypermutation
SMRT: single molecule, real-time
SPR: surface plasmon resonance
ssRNA: single-stranded ribonucleic acid
sUMI: single unique molecular identifier
TCID50: median tissue culture infectious dose
TD: transmembrane domain
TF: transmitted-founder
Tm: melting temperature
UCA: unmutated common ancestor
V: variable
VH: variable heavy
VL: variable light
VMMC: voluntary medical male circumcision
wpi: weeks post infection
β-ME: beta-mercaptoethanol

1 Introduction

1.1 The HIV-1 pandemic

Human immunodeficiency virus 1 (HIV-1) remains a major global health challenge with approximately 38 million people living with HIV-1 at the end of 2020, 7.8 million of whom were individuals living in South Africa (Statistics South Africa, 2020). Globally 1.5 million new HIV-1 infections were recorded in 2020 (UNAIDS, 2021). HIV-1 attacks CD4 expressing cells of infected persons, leaving them susceptible to infections (Fauci *et al.*, 1984; Fauci, 2003) and when left untreated, leads to acquired immunodeficiency syndrome (AIDS) (Barré-Sinoussi *et al.*, 1983; Gallo *et al.*, 1984).

Great success has been achieved with the development of antiretroviral therapy (ART) (Moosa *et al.*, 2019). ART suppresses viral replication thus preventing AIDS and reducing the probability of onwards transmission of HIV-1 (Gulick *et al.*, 1997; Palella *et al.*, 1998; Moosa *et al.*, 2019). The prevention of mother-to-child transmission (PMTCT) programme has made ART available to further prevent HIV-1 transmission in this group (AVERT, 2018). The use of pre-exposure prophylaxis (PrEP) by individuals who self-identify as high risk, and post-exposure prophylaxis (PEP) within 72 hours of potential HIV-1 exposure have proved to be highly effective prevention strategies (WHO, 2021). However, there are key populations where HIV-1 treatment is often not accessed, such as sex workers in South Africa, who are significantly less likely to be virally suppressed as a result (UNAIDS, 2021). Due to adverse side effects, fear of disclosing HIV-1 status, alcohol and drug use, scarcity of financial resources, unstable living situations and unfavourable work hours which influence the timing of ART, non-adherence to ART regimens is a challenge (Arage *et al.*, 2014; Endalamaw *et al.*, 2018; Engler *et al.*, 2018; Becker *et al.*, 2020). Voluntary medical male circumcision (VMMC) is a key preventative intervention in 15 countries where over 23 million adolescent boys and men have been circumcised since 2007 (WHO, 2021). This strategy is considered a good approach to reach men who do not seek healthcare services often, however education about safe sex and condom use are two other prevention strategies of vital importance to be used in conjunction to prevent transmission (Auvert *et al.*, 2005).

The success of ART regimens and other prevention strategies have resulted in a 43% decrease in new HIV-1 infections in sub-Saharan Africa when comparing the years of 2022 and 2010 (UNAIDS, 2021). However, mathematical modelling has shown that current interventions alone are not sufficient to eradicate the spread of HIV-1 (Freedberg *et al.*, 2015; Martinez *et al.*, 2021), and a vaccine remains an urgent unmet need.

1.1.1 Evolution of HIV-1

There are four distinct HIV-1 groups, each the result of an independent zoonotic spill over event (Keele *et al.*, 2006) **(Figure 1.1)**. Group O represents less than 1% of HIV-1 cases worldwide and is mainly found in Cameroon, Gabon and neighbouring countries (De Leys *et al.*, 1990; Gürtler *et al.*, 1994; Mauclère *et al.*, 1997; Peeters *et al.*, 1997). Groups N and P are less prevalent than Group O and all cases have been documented in individuals from Cameroon (Sharp and Hahn, 2011). Group M is the most widely spread and gave rise to the global HIV-1 pandemic (Korber *et al.*, 2000; Hemelaar, 2012). The genetic divergence of HIV-1 group M led to distinct subtypes predominating in different geographic locations and can provide information about how the virus evolved over time (Taylor *et al.*, 2008) (**Figure 1.2**). There are nine subtypes within this group, each with a designated letter; A-D, F-H, J and K (Myers *et al.*, 1997; Hemelaar, 2012). Over time recombination of various subtypes have led to the emergence of circulating recombinant forms (CRFs). In Southern Africa HIV-1 subtype C is dominant and therefore also responsible for the most infections worldwide (Hemelaar, 2012).

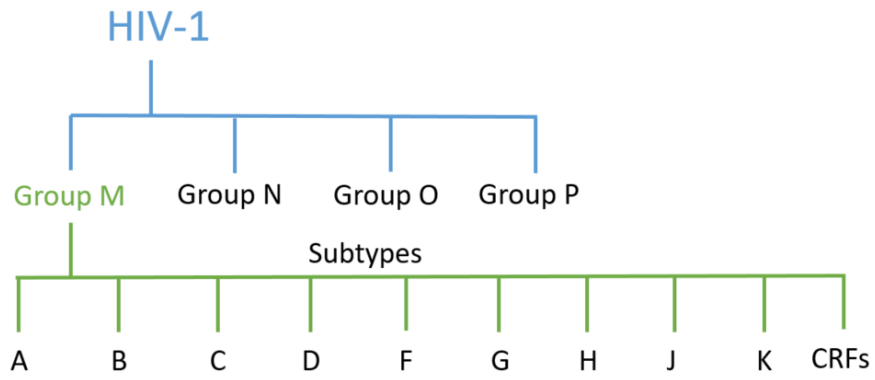

Figure 1.1: Schematic showing the various subtypes of HIV-1
There are four distinct groups of HIV-1 that have been identified to date. Group M represents the pandemic form of HIV-1. Group M is divided into different subtypes based on the percentage of amino acid variation. There are nine different subtypes and multiple CRFs.

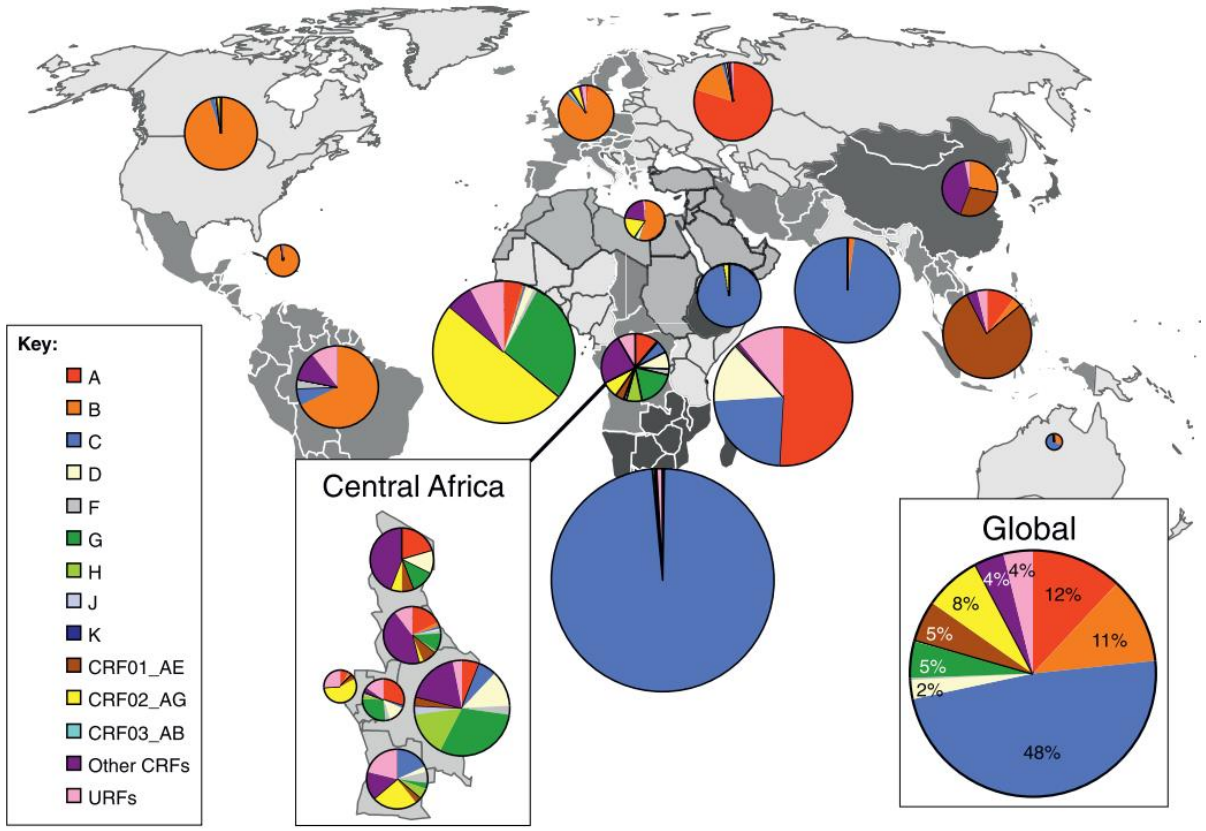

Figure 1.2: Distribution of HIV-1 group M subtypes and recombinants worldwide
The distribution of HIV-1 group M subtypes found in each region represented according to percentage of people living with HIV-1 within that region. Right-hand insert indicates the percentage of people living with each HIV-1 subtype globally. Legend on left-hand side indicate which subtypes and recombinants are represented by each colour. Figure from Hemelaar (2012).

1.2 HIV-1 structure and function

HIV-1 is an enveloped retrovirus containing a 9.8 kilobase (kb) positive sense, single-stranded ribonucleic acid (ssRNA) genome (Ratner *et al.*, 1985; Wain-Hobson *et al.*, 1985) (**Figure 1.3**). The genome contains three major genes, *gag, pol* and *env,* which encode structural proteins (matrix, nucleocapsid, p6 and capsid), viral enzymes (protease, integrase and reverse transcriptase) and envelope (Env) proteins (gp120 and gp41), respectively (Li and De Clercq, 2016).

The first phase of the viral replication cycle begins with entry into the host cell by utilising the matrix, envelope and regulatory protein tat (Costin, 2007; Prabakaran *et al.*, 2007; Wilen *et al.*, 2012). Upon entry, HIV-1 uses reverse transcriptase to convert ssRNA into double stranded deoxyribonucleic acid (DNA), with the assistance of the nucleocapsid, protease, tat and accessory proteins vif and nef assist with this step (Hu and Hughes, 2012). This allows the

viral DNA to enter the nucleus where the matrix, nef, rev, pr and integrase are responsible for the integration of HIV-1 DNA into the host cell (Li and De Clercq, 2016). After successful integration of viral DNA into the host cell DNA, the host cell machinery makes use of transcription to produce HIV-1 viral transcripts which are translated to produce viral proteins with assistance from tat, vif, vpr, nef and nucleocapsid (Sawaya *et al.*, 2000; Joseph *et al.*, 2003; Wang *et al.*, 2008; Hong *et al.*, 2013). The new HIV-1 proteins and viral RNA are assembled into immature HIV-1 virions, which require vif, vpr, vpx, gp41, p6 and reverse transcriptase for viral budding (Briggs and Kräusslich, 2011; Li and De Clercq, 2016). Lastly protease is used to process the immature HIV-1 virion resulting in a mature HIV-1 virus (Li and De Clercq, 2016).

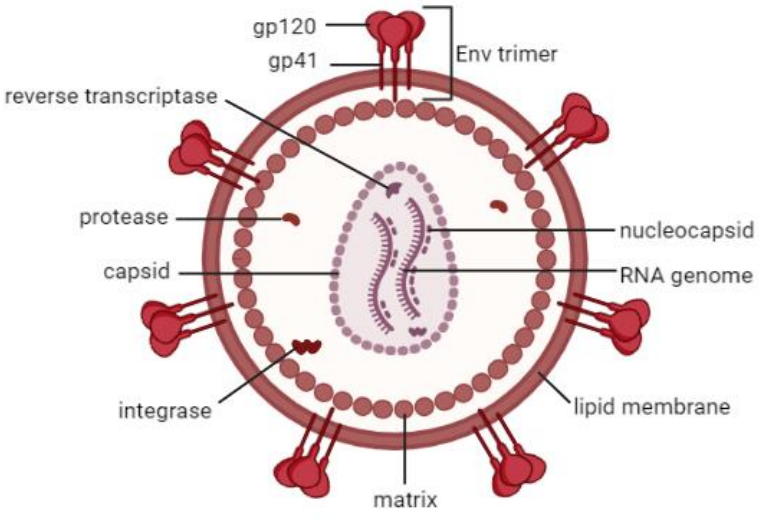

Figure 1.3: Mature HIV-1 virion
A graphical depiction of the HIV-1 virion. The Env trimers consisting of gp120 and gp41 subunits are seen embedded in the lipid membrane. The matrix contains the protease and integrase as well as the capsid protein. The reverse transcriptase, RNA genome and associated nucleocapsid proteins are located within the capsid, consisting of numerous capsid proteins. Figure created using Biorender.

1.2.1 The HIV-1 Env protein

The viral Env glycoprotein is found on the HIV-1 virion surface. The Env is made up of trimers consisting of three surface exposed gp120 glycoproteins and three transmembrane gp41 glycoproteins which act as the stem, anchoring the structure into the viral membrane (Allan *et al.*, 1985; Veronese *et al.*, 1985). Env is expressed as a gp160 pro-protein which is cleaved by the host enzyme furin into its gp120 and gp41 subunits. The gp120 subunit consists of five relatively conserved regions (C1-C5) and five regions with greater sequence variability (V1-V5) (**Figure 1.4**). The transmembrane gp41 is made up of an extracellular domain, consisting of the fusion peptide (FP) of approximately 20 amino acids (AA) at the N-terminus, two heptad

repeats (HR1 & HR2), a transmembrane domain (TD) and a cytoplasmic tail (CT) (Prabakaran *et al.*, 2007). There is also a membrane proximal region (MPER) which connects the gp41 ectodomain to the TD (Salzwedel *et al.*, 1999).

The gp120 protomer interacts with the host cell by binding to the CD4 receptor as well as to the C-C chemokine receptor 5 (CCR5) or C-X-C chemokine receptor 4 (CXCR4) co-receptors (Choe *et al.*, 1996; Feng *et al.*, 1996). This triggers conformational changes which activate gp41, resulting in the FP being driven into the cell membrane. This enables fusion of the viral and host membranes leading to the delivery of the viral genome into the host cell (Prabakaran *et al.*, 2007).

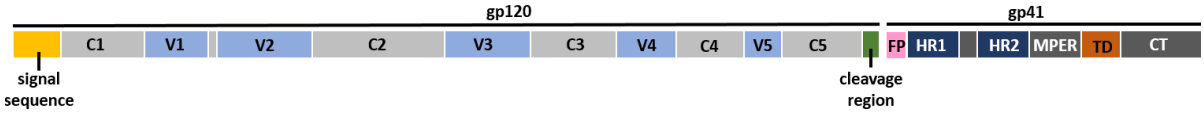

Figure 1.4: Linear representation of the HIV-1 Env
The gp120 and gp41 subunits of gp160 are shown, separated by the cleavage region (green). The gp120 consists of a signal sequence (yellow), five constant regions (C1-C5) represented in light grey and five variable regions (V1-V5) represented in light blue. The gp41 subunit consists of an extracellular domain containing a fusion peptide (FP) (light pink), two heptad repeats (HR1 & HR2) (dark blue), the membrane proximal region (MPER), cytoplasmic tail (CT) (dark grey) and the transmembrane domain (TD) (orange).

1.2.2 HIV-1 Env variation

The HIV-1 Env gp120 is extremely diverse due to reverse transcriptase which results in errors being introduced into the proviral DNA that are subsequently incorporated into the genome of newly infected cells (Zhuang *et al.*, 2002). During the chronic phase of HIV-1 infection, the immune system also contributes to Env diversity through antibody pressure which selects for escape mutations (Wei *et al.*, 2003). Escape mutations occur at sites important for antibody binding and result in a drop in potency and binding or in extreme cases the inability of the antibody to recognise the newly mutated site (Garces *et al.*, 2014; Sok *et al.*, 2016; Bonsignori *et al.*, 2017; Zhou *et al.*, 2018). Escape variants are also driven by extensive T-cell pressure (Price *et al.*, 1997). In contrast to gp120, divergence of the gp41 subunit is limited by functional conservation as it contains complex fusion machinery that undergo conformational changes during the fusion process (Ward and Wilson, 2015).

1.3 Antibody development

Antibodies play an important role in our adaptive immune response where they are responsible for antigen recognition and perform multiple functions such as binding to foreign antigens, neutralisation of viruses and elicitation of fragment crystallisable (Fc) effector functions against infected cells (Murphy *et al.*, 2012; Forthal, 2014). Immunoglobulin G (IgG) are the most common soluble version of these proteins found in the blood (Schur, 1988). IgG heterodimeric proteins consist of two identical heavy chains, made up of a variable region and three constant regions, and two identical light chains, consisting of a variable and constant region. Each antibody can be further divided into a Fc and fragment antigen binding (Fab) region (**Figure 1.5**) (Edelman and Benacerraf, 1962).

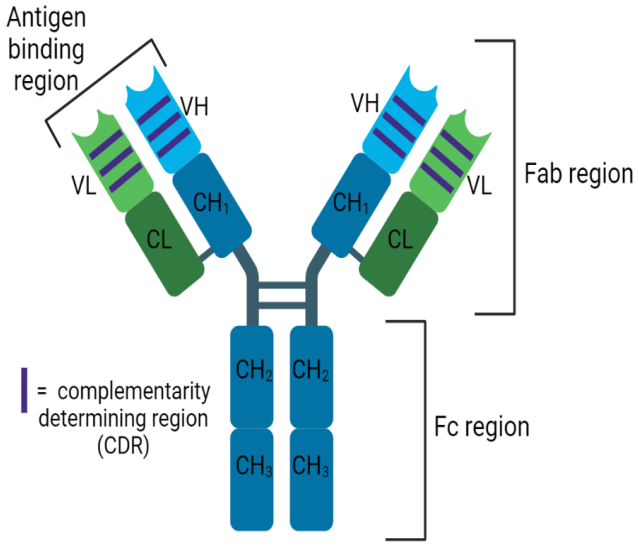

Figure 1.5: Schematic representing the structure of an IgG antibody
The antibody consists of two heavy chains represented in blue and two light chains represented in green. The constant regions for the heavy (CH1, CH2 and CH3) and light chains (CL) are shown in dark blue and dark green respectively, while the variable region of the heavy (VH) and light (VL) chains shown in light blue and light green respectively. The CDR found in the variable regions are indicated using purple bars. The Fab region, consisting of the light chain, heavy chain variable region and one heavy chain constant region (CH1) and Fc region consisting of two heavy chain variable regions (CH2 & CH3) are indicated. Figure created using Biorender.

One complete light chain (VL & CL), the variable region of the heavy chain (VH) and one constant region (CH1) together form the Fab. The variable portion of the Fab region is responsible for the binding of antigens whereas the constant portion is responsible for effector functions (Murphy et al., 2012). The most variable part of the antibody, crucial for diverse antigen specificity, consists of three hypervariable loops, known as complementarity

determining regions (CDR), found in both heavy and light variable regions (Stewart, 2012). The Fc region is formed by two heavy chains consisting of two constant regions (CH2 & CH3) each. This tail region of the antibody is important for interacting with cell surface receptors allowing for the activation of the immune system (Stewart, 2012).

1.3.1 Antibody genetics and somatic recombination

The building blocks necessary for functional antibody heavy and light chains can be found across three primary loci in the human genome in the form of variable (V), diversity (D) and joining (J) IG germline gene segments (Tonegawa, 1983). Genes at the IG heavy chain locus (IGH), located on chromosome 14, are responsible for encoding the heavy chain and genes that encode the light chain can be found on either IG kappa (IGK) locus, located on chromosome 2, or IG lambda (IGL) locus, located on chromosome 22 (Erikson *et al.*,1981; Lefranc, 2001b, 2001a; Watson *et al.*, 2017). Recombination of various V, D and J segments at IGH or V and J segments at IGK and IGL, results in the formation of naïve B-cell receptors (BCR) (Tonegawa, 1983). To further increase diversity of BCRs and secreted antibodies P and N nucleotides can be introduced and deleted at V(D)J junctions, different heavy chains and light chains can be paired together and antibodies that have come into contact with antigens can undergo affinity maturation mediated by somatic hypermutation (SHM) (Tonegawa, 1983; Flaherty, 2012). Antibodies are secreted soluble protein versions of membrane bound BCRs. Naïve BCRs are extraordinarily diverse, due to V(D)J recombination, which ensures the likelihood of a naïve antibody, that has never encountered an antigen, being present that will be able to bind, potentially very weakly, to a new antigen (Watson *et al.*, 2017). SHM introduces somatic mutations in the variable regions of antibodies, predominantly at targeted hotspots within the hyper variable CDRs, and thereafter a process of selection occurs in germinal centres where BCRs with high affinity and specificity against particular antigens are selected (Watson *et al.*, 2017). HIV-1 directed antibodies have been shown to have high levels of SHM compared to other immunoglobulins due to chronic infection and a continuously evolving target (Wei *et al.*, 2003; Scheid *et al.*, 2009, 2011; Klein *et al.*, 2013).

1.4 Broadly neutralising antibodies in HIV-1 infection

Upon HIV-1 infection the immune system elicits a robust antibody response towards the Env protein, the sole target of neutralising antibodies, due to it being the only virally encoded antigen located on the HIV-1 surface (Allan *et al.*, 1985; Wyatt and Sodroski, 1998). Most of the antibodies generated are directed towards regions of gp120 and gp41 that are not exposed on trimeric viral Envs and are thus unable to neutralise the virus (Burton *et al.*, 2004). The antibodies likely arise as a result of binding to highly immunogenic regions of dissociated gp120 and gp41, and are thus unable to bind fully formed trimers (Parren *et al.*, 1997; Wyatt and Sodroski, 1998).

However, all HIV-1 infected individuals develop a neutralising antibody response against the virus during infection (Mouquet, 2014). Neutralising antibodies prevent infection of the host cell by blocking the interaction between Env and cell receptors or by preventing conformational changes needed for fusion and thus inhibiting entry of the infectious agent (Walker and Burton, 2008).

The rapid evolution of the HIV-1 Env spike over the course of the infection results in the continual evasion of antibody-mediated neutralisation (Wibmer *et al.*, 2013; Doria-Rose *et al.*, 2014). Over time, the evolutionary pressure of antibodies throughout the pandemic has driven the Env to develop additional immune evasion mechanisms (Ward and Wilson, 2015). This includes changes in variable loops used to mask the functionally important regions of gp120, such as self-derived glycan shielding which acts through occluding underlying protein epitopes from antibody binding by covering the trimer in a dense array of glycans (Wyatt *et al.*, 1998; Wei *et al.*, 2003). The variability of Env glycosylation occurs due to variation in the types of glycans and the positions at which they are expressed on the Env surface (Yang *et al.*, 2017). A conformationally dynamic trimer is another change undergone by Env to make it difficult for antibodies to recognise epitopes (Kwong *et al.*, 2002). Multiple trimers displayed on the membrane surface vary depending on the above mentioned immune evasion mechanisms, this heterogeneity is important for infection of the host cell and is responsible for HIV-1's successful immune evasion (Merk and Subramaniam, 2013).

Neutralising antibodies that develop in an HIV-1 infected individual are mostly strain-specific meaning they are only able to recognise the viral Envs circulating in that particular individual

(Gray *et al.*, 2011a). However, there are approximately 20% of HIV-1 infected individuals that develop broadly neutralising antibodies (bNAbs), which develop the ability to neutralise diverse, global HIV-1 strains despite Env diversity (Landais and Moore, 2018). A vaccine that is able to induce bNAbs would likely provide protection against diverse viral isolates, effectively disrupting HIV-1 transmission chains.

1.4.1 Development of broadly neutralising antibodies

The development of bNAbs in HIV-1 infected individuals takes years, and is predominantly driven by the interplay between the Env and the antibody response over time (Bhiman *et al.*, 2015) (**Figure 1.6**). In the case of broad lineages the engagement of antigen by a naïve B-cell receptor leads to the emergence of a bNAb precursor, also known as an unmutated common ancestor (UCA). Upon UCA binding to the bNAb-initiating virus the process of bNAb development is triggered (Bhiman *et al.*, 2015) (**Figure 1.6**). The bNAb-initiating virus is thought to be a distinct viral variant from the virus which establishes infection in an individual, known as the transmitted-founder (TF) virus due to UCAs and early lineage members only seen in later timepoints of infection (Bhiman *et al.*, 2015). These ideas have been confirmed through longitudinal antibody and plasma breadth studies which have shown that UCAs emerge 3-16 months post-infection whereas mature bNAbs take 1-2 years to develop (Sather *et al.*, 2009; Gray *et al.*, 2011a; Landais and Moore, 2018).

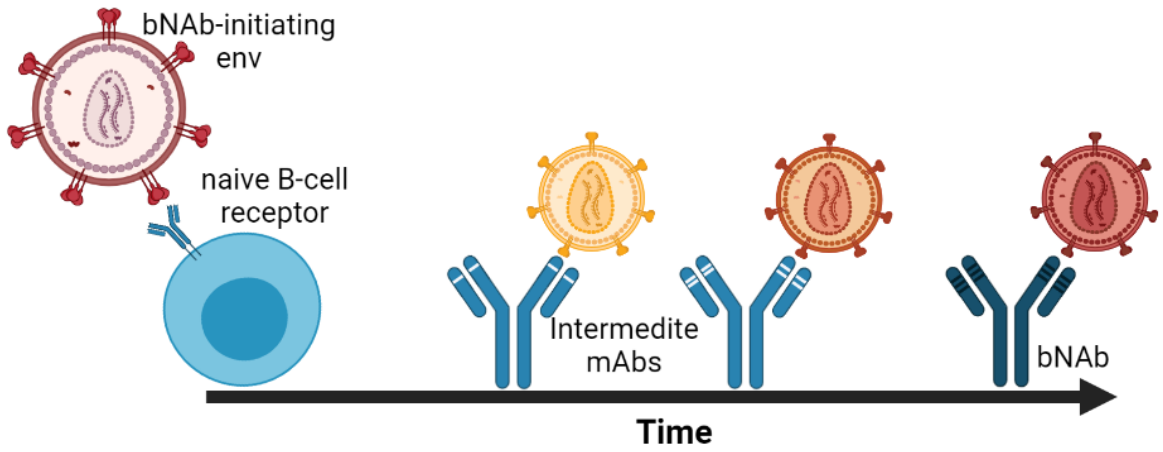

Figure 1.6: Development of bNAbs during HIV-1 infection
Schematic showing the engagement of the bNAb initiating virus by a naïve B-cell receptor. Over time the virus acquires mutations (yellow, orange and red virions) but intermediate antibodies continue to mature to recognise the virus through SHM, some resulting in bNAbs. Figure created using Biorender.

The engagement of the UCA by the bNAb-initiating virus induces a rapid diversification and expansion of the antibody lineage, driven by viral escape mutations and selection of intermediate antibodies that recognise the changing virus (Liao *et al.*, 2013; Doria-Rose *et al.*, 2014; MacLeod *et al.*, 2016). Analyses of memory B-cell repertoires have shown that early intermediates that acquire different mutations result in multi-limb maturation, giving rise to sub-lineages (MacLeod *et al.*, 2016). There are also "off-track" antibodies that develop within bNAb lineages, these antibodies have substantial SHM but fail to develop breadth (Bhiman *et al.*, 2015; Sacks *et al.*, 2019). Antibody branches that are successful in recognising newly mutated viruses will continue to mature through increased SHM as they continue to adapt to new viral variants (MacLeod *et al.*, 2016). These antibody lineages mature due to extensive affinity maturation and the introduction of functionally relevant mutations in the UCA sequence that results in bNAbs over time (Wiehe *et al.*, 2018).

Factors that contribute to the development of bNAbs in HIV-1 infected individuals include increased germinal centre activity, a more diverse B cell repertoire (Cohen *et al.*, 2014), a high viral load upon infection and viral Env diversity (Landais *et al.*, 2016). High levels of antigenic stimulation are required for the extensive SHM often seen in bNAbs. Following initiation on ART regimens HIV-1 infected individuals rarely develop broad antibody responses. Viral properties also influence bNAb development as seen with subtype C infection which is associated with a higher probability of developing a V3-glycan-directed response and enhanced breadth whereas subtype B infection is predisposed to elicitation of a CD4 binding site (CD4bs) directed response (Landais *et al.*, 2016; Rusert *et al.*, 2016).

A decline in T regulatory cells have also been associated with the development of bNAbs, potentially by enabling B-cell intermediates that have the potential for autoreactivity to persist within the immune system (Moody *et al.*, 2016). An increase in germinal centre activity has also been implicated in the development in bNAbs, with individuals that developed bNAb responses having a high frequency of peripheral T follicular helper cells in early infection (Cohen *et al.*, 2014). In addition, the role that superinfection plays in the development of breadth is still unclear. It has been suggested that HIV-1 superinfection enhances breadth through additive responses to each individual virus and by increasing intra-host viral diversity (Cortez *et al.*, 2012; Williams *et al.*, 2018) but in a study observing superinfected individuals

over time, no correlation between development of bNAbs and superinfection was found (Sheward *et al.*, 2018).

1.4.2 Broadly neutralising antibody epitopes on the HIV-1 Env

The Env is the most diverse HIV-1 protein due to amino acid substitution, insertions/deletions and shifting of glycan sites, however there are conserved Env regions across most HIV-1 strains and these regions are targeted by bNAbs (Wibmer *et al.*, 2015). The isolation of potent bNAbs and advances in epitope mapping technologies has made it possible to identify six conserved regions on the HIV-1 Env protein: the V2/apex, the V3-glycan site, the CD4bs, the gp120-gp41 interface which includes the fusion peptide, the silent face and the MPER (**Figure 1.7**). The V2/apex site is found at the trimer apex and is formed by the converging V1/V2 domain and associated glycans (Julien *et al.*, 2013). The densely packed glycans in this region and hypervariable V1 and V2 loops are employed for viral protection from the humoral response (McLellan *et al.*, 2011). However, the bNAbs that are able to penetrate this site include some of the most potent antibodies described to date such as PG9 (Walker *et al.*, 2009), PG16 (Walker *et al.*, 2009), CAP256-VRC26 (Doria-Rose *et al.*, 2014) and PGT145 (Walker *et al.*, 2011). The CD4bs is also protected by glycans however, bNAbs targeting this site generally do not interact with them, with the exception of the glycan at N276, but rather mimic CD4 binding by interacting with residues on the highly conserved gp120 CD4 binding loop (Balla-Jhagjhoorsingh *et al.*, 2013; Zhou *et al.*, 2013). bNAbs directed to the CD4bs, such as VRC01, take longer to develop due to their high levels of SHM but provide the best combination of breadth and potency (Wu *et al.*, 2010). The gp120-gp41 interface is found towards the bottom of the trimer and most antibodies binding to this site, such as 35O22 and PGT151, require the quaternary trimeric structure to bind as they cannot bind to individual gp120 or gp41 protomers (Falkowska *et al.*, 2014; Huang *et al.*, 2014). One of the most densely glycosylated regions found on the Env trimer is the silent face (Zhou *et al.*, 2018). Despite this heavy glycosylation, bNAbs, SF12 and VRC-PG05, that bind the silent face have been identified (Kwong and Mascola, 2018; Zhou *et al.*, 2018; Schoofs *et al.*, 2019). The MPER epitope is a helical linear epitope preceding the transmembrane region (Cardoso *et al.*, 2005; Pinto *et al.*, 2019). Two of the broadest bNAbs identified, 10E8 and 4E10, target the MPER (Stiegler *et al.*, 2001; Huang *et al.*, 2012). Very broad bNAbs have been found to develop towards the MPER epitope due to its high level of conservation however, these bNAbs tend to lack potency (Sok

and Burton, 2018). Lastly, the V3-glycan site is composed of numerous, overlapping glycan-dependant epitopes which lie proximal to the V2/apex (Kong *et al.*, 2013; Wibmer *et al.*, 2015). It is also known as the N332 supersite, as it is targeted by bNAbs whose epitopes are centred on the glycan at position N332 (Sok *et al.*, 2014). PGT121 and PGT128 are among the bNAbs that are highly dependent on the N332 glycan (Walker *et al.*, 2011). V3-glycan site directed bNAbs are the most common antibodies to be elicited during HIV-1 infection and this site is therefore an attractive vaccine target (Jacob *et al.*, 2015).

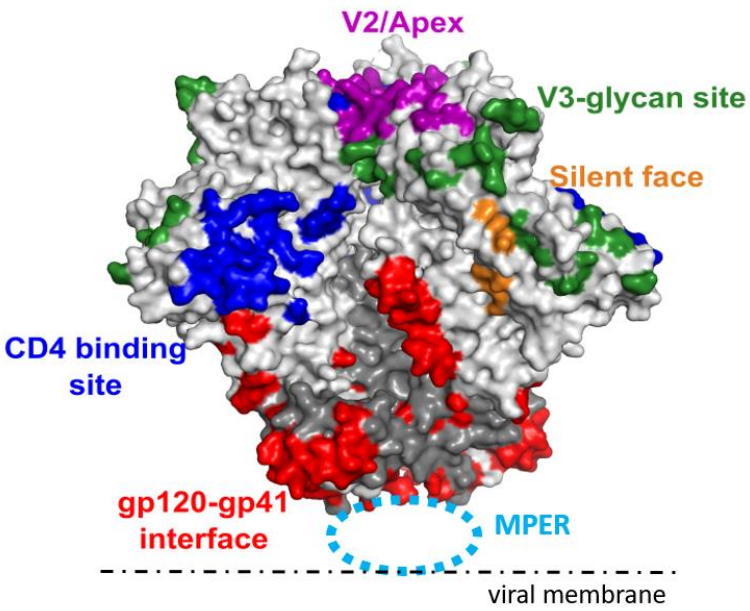

Figure 1.7: Conserved viral epitopes targeted by HIV-1 bNAbs This model is based on the BG505 SOSIP.664 trimer (PDB ID:4ZMJ). The trimer consists of three gp120 subunits (light grey) and three gp41 subunits (dark grey). The six bNAb epitopes are coloured as follows: V3-glycan (dark green), the V2/apex (purple), CD4bs (blue), the gp120-gp41 interface (red) and the silent face (orange). The MPER (light blue), which is the sixth site, is located between gp41 and the viral membrane but is not present on the soluble trimer. The viral membrane is indicated as a black dotted line. This figure was generated using Pymol (version 2.3.4).

1.4.3 Broadly neutralising antibodies directed to the V3-glycan epitope

Most bNAbs that bind the V3-glycan site have long CDRH3s that are able to penetrate the glycan shield and access an underlying protein epitope known as the ^{325}GDIR^{328} motif (Pejchal *et al.*, 2011; Walker *et al.*, 2011) (**Figure 1.8**). Epitope masking by glycans or the V1 loop results in steric hindrance of bNAbs directed to the V3-glycan epitope (Kwong *et al.*, 2002; Wei *et al.*, 2003; Krachmarov *et al.*, 2006; Pantophlet *et al.*, 2008). bNAbs that target this site generally make use of the same residues and glycans, however their angles of approach and binding mechanisms differ (Wibmer *et al.*, 2015). In addition, some level of glycan "promiscuity" have been observed in bNAbs that target the V3-glycan site, making them able to target glycans at

N301 and N295 in the absence of a N332/334 glycan or able to recognise both high-mannose and complex-type glycans (Mouquet *et al.*, 2012; Sok *et al.*, 2014).

V3-glycan-targeting bNAbs also require lower levels of SHM compared to bNAbs which target other sites such as the CD4bs (Corti and Lanzavecchia, 2013). The most potent V3-glycan directed antibody isolated to date, BG18, has been shown to have moderate SHM and no indels (Freund *et al.*, 2017).

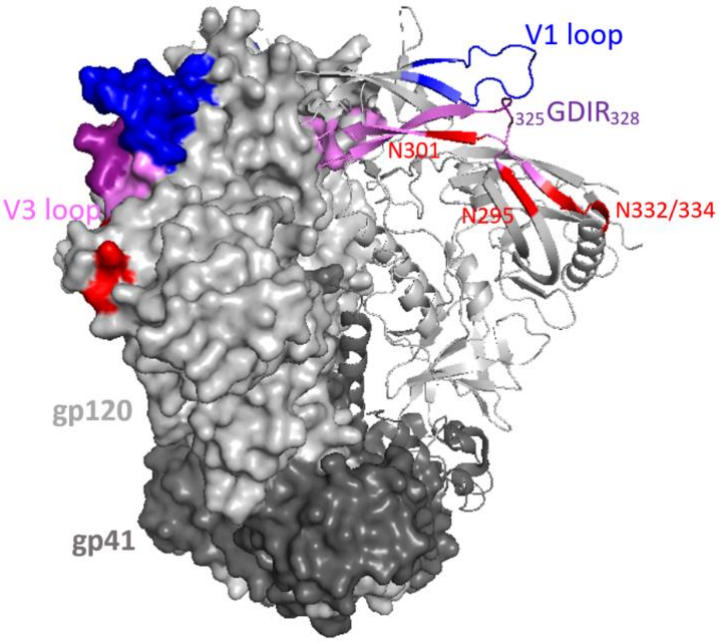

Figure 1.8: Features influencing V3-glycan bNAb binding
This model is based on the BG505 SOSIP.664 trimer (PDB ID:4ZMJ). The trimer consists of three gp120 subunits (light grey) and three gp41 subunits (dark grey). Two gp120 and gp41 subunits are represented as surface exposed structures and one gp120 and gp41 are represented as a ribbon structure. The V1 loop is highlighted in dark blue and the V3-glycan loop is shown in light pink. Important features within the V3 loop such as the $_{325}$GDIR$_{328}$ motif (dark purple), N295, N301 and N332/334 glycan sites (red) are highlighted. This figure was generated using Pymol (version 2.3.4).

A large number of V3-glycan bNAbs have been shown to use IGHV4 germline genes, specifically IGHV4-34 which is known to give rise to autoreactive antibodies (Schickel *et al.*, 2017; Roskin *et al.*, 2020). SHM of codon 26 in IGHV4-34 has been shown to decrease autoreactivity and is seen in a large number of HIV-1 infected individuals (Roskin *et al.*, 2020). When comparing HIV-1 uninfected individuals with HIV-1 bNAb producers, significantly lower frequencies of SHM of codon 26 was observed in IGHV4-34 gene segments which suggests that potentially autoreactive antibodies may be implicated in the development of bNAbs in HIV-1 infected individuals (Roskin *et al.*, 2020).

1.5 HIV-1 Immunogens

1.5.1 Vaccine development in HIV-1

As the elicitation of HIV-1 bNAbs is deemed necessary for the development of a successful preventative vaccine candidate, an understanding of bNAb development during HIV-1 infection has been a major focus of HIV-1 vaccine research. There are multiple vaccines currently in clinical trials which aim to elicit bNAb responses. The most successful vaccine candidate to date is from the RV144 study done in Thailand which made use of a recombinant canarypox vector, modified to carry HIV genes, and two HIV-1 protein boosts (Rerks-Ngarm *et al.*, 2009). This vaccine elicited a wide-ranging response against HIV-1, including Env binding antibodies, antibody-directed cell-mediated cytotoxicity and CD4 lymphoproliferation. However, it's important to note that the vaccine did not elicit any neutralising antibody responses. These responses resulted in up to 31% vaccine efficacy in low risk participants (Rerks-Ngarm *et al.*, 2009). Another study, HVTN 702, was modelled after RV144 and adapted to a subtype C virus, as this is the most prevalent subtype in Southern Africa. This vaccine trial was discontinued in early 2020 due to a lack of efficacy (NIH, 2020). Another vaccine trial, the HVTN 705 study, started in 2017 in sub-Saharan Africa making use of a subtype C gp140 and adenovirus vector showed no safety concerns but did not result in sufficient protection provided against HIV-1 and was halted (NIH, 2017; NIAID, 2021). Despite the lack of protection seen in HVTN 705, the HVTN 706 study, started in 2019 in Europe and the Americas using a subtype C and mosaic gp140 and adenovirus vector is currently in late-stage trials (NIH, 2019). Results from the AMP trials, which investigated the use of passive immunisation through VRC01 bNAb infusions, showed effectiveness in preventing the acquisition of HIV-1 strains sensitive to the bNAb (Corey *et al.*, 2021; SAMRC, 2021). This trial provided evidence that bNAbs present at high enough levels are able to prevent HIV-1 infection. A challenge that will need to be overcome in vaccine development is the sustained elicitation of the correct antibodies at high enough concentrations.

1.5.2 HIV-1 Env trimer design

Multiple studies have investigated using Env trimers as potential immunogens. The expression of stable Env trimers has long been a challenge, with the first purified Env SOS gp140 (Binley *et al.*, 2002) and SOSIP gp140 (Sanders *et al.*, 2002) forming the basis for the

SOSIP trimers widely used today. The BG505 SOSIP trimer is considered to be the gold standard of SOSIP trimers (Sanders *et al.*, 2013), derived from a subtype A TF virus isolated from a 6-week old infant (Wu *et al.*, 2006). A SOSIP.664 trimer bears sequence modifications A501C and T605C resulting in the formation of an extra disulphide bond and I559P mutation to increase trimer stability, deletion of the MPER to improve solubility and homogeneity and the replacement of the natural furin cleavage site sequence with the optimal furin cleavage site, RRRRRR (Sanders *et al.*, 2013). Subsequently, other studies have based the design of their Env trimers on the BG505 SOSIP framework. A successful subtype B Env trimer, B41, has previously been derived from a TF virus isolated from a HIV-1 infected plasma donor (Pugach *et al.*, 2015). Subtype C trimers have been suggested to differ from other subtypes with regards to their structural and glycosylation properties due to the neutralisation profiles observed by subtype C viruses (Moore *et al.*, 2009). Previously expressed subtype C trimers have been shown to be less stable and more difficult to express when compared to subtype A trimers (Julien *et al.*, 2015). Julien *et al.* (2015) introduced the previously described SOSIP mutations (Sanders *et al.*, 2013) into 15 subtype C candidates and found that only two were able to express well-formed stable trimers, suggesting that *env* sequences play a role in the expression of SOSIP trimers.

There are several clinical trials currently evaluating SOSIP trimers as vaccine candidates. The BG505 SOSIP gp140 is currently being tested in a Phase 1 IAVI trial at various concentrations with the AS01B adjuvant (W001). The phase 1 ACTHIVE-001 trial is investigating the use of a ConM SOSIP.v7 gp140, a trimer based on the consensus of subtype A, B, C and G consensus sequences, with a monophosphoryl lipid A adjuvant. Another phase 1 trial, EAVI2020_01, is evaluating prime-boost combinations using various forms of SOSIP gp140s namely, ConM SOSIP, ConS SOSIP, EDC ConS uncleaved prefusion-optimised (UFO), ConS UFO and mosaic SOSIPs. No preliminary results for any of these trials have been released to date. Due to success seen with messenger RNA (mRNA) vaccines and COVID-19 an HIV-1 vaccine trial making use of the mRNA platform was started, HVTN302 (NIH, 2022). The three trimers used for the HVTN302 trials are based on the BG505 MD39 native like trimer (Steichen *et al.*, 2016), preliminary results are expected at the end of 2023.

1.5.3 Germline-targeting immunogens

One potential vaccine strategy consists of using a priming immunogen, also known as a germline-targeting immunogen, to engage bNAb precursors, followed by several successive boosting immunogens to drive the development of bNAbs (Stamatatos *et al.*, 2017). Germline-targeting immunogens are designed to elicit precursors that make use of a specific set of germline genes depending on the epitope that is being targeted. For example, previous work has shown the restricted use of IGVH1-2 germline genes by VRC01-class bNAbs that target the CD4bs (Wu *et al.*, 2010; Zhou *et al.*, 2015). The most successful germline-targeting immunogen developed to date is the CD4bs specific engineered outer domain germline-targeting 8 (eOD-GT8) immunogen (Jardine *et al.*, 2016). This immunogen is currently in use in a phase 1 trial, IAVI G001, where participants receive two doses of the eOD-GT8 immunogens two months apart. The preliminary results from this trial showed that 97% of participants who received an eOD-GT8 immunogen developed VRC01-class precursor BCRs (IAVI, 2021). This study demonstrated the successful use of a priming immunogen, however subsequent boosters will be needed to generate bNAbs and that will require substantial maturation as they are VRC01-class antibodies. Another germline-targeting immunogen currently in clinical trials, IAVI C101 phase 1, is BG505 SOSIP.GT1.1 gp140, modelled on BG505 but containing 18 AA substitutions and a 7-residue deletion in the V2 loop making it able to engage precursors of bNAbs that target the CD4bs and V2/apex (Medina-Ramírez *et al.*, 2017). Results of the IAVI C101 phase 1 trial are expected towards the end of 2022.

1.5.4 V3-glycan germline-targeting immunogens

Germline-targeting immunogens specific for V3-glycan antibodies have been developed based on the previously designed BG505 SOSIP with additional germline-targeting mutations introduced into the V1 and V3 loops. These include the $11MUT_B$ trimer which had 11 new mutations introduced to increase the binding affinity of the trimer to PGT121-class precursor V3-glycan specific antibodies (Steichen *et al.*, 2016). A collection of six PGT121-class germline-reverted antibodies were used as selection agents across multiple gp120 mutagenesis libraries which led to the identification of the $11MUT_B$ trimer as having the highest affinity for PGT121-class precursors. The $11MUT_B$ mutations were introduced into the BG505 SOSIP trimer and showed to have excellent yield, thermal stability and antigenic profile. This

resulted in N332-directed germline responses observed when tested in human IGH and IGK knock-in mouse models (Steichen *et al.*, 2016). Steichen *et al.* (2019) used a bNAb directed to the V3-glycan site called BG18, to identify multiple inferred germline precursors through analysis of interactions observed between BG18 bound to BG505. The resulting precursors were used to guide the design of a germline-targeting immunogen. Three of the BG18 precursors, exactly matching their initial search criteria when scanning NGS data, were isolated from naïve human B cells using N332 germline-targeting trimers. These trimers were modelled on the $11MUT_B$ trimer (Steichen *et al.*, 2016), showing that the immunogen is capable of binding bNAb precursor-like B cells (Steichen *et al.*, 2019).

Another immunogen named RC1, designed by Escolano *et al.* (2019) and based on the $11MUT_B$ trimer, elicited V3-glycan specific precursors but differed through the removal of a glycan at position N156. This immunogen was successful in eliciting V3-glycan specific serological responses in wild-type mice. Additional animal studies took place in wild-type rabbits and macaques. The study showed that RC1 is able to elicit V3-glycan targeting antibodies in wild-type animal models capable of producing polyclonal responses and that a RC1 variant is a suitable candidate priming immunogen for sequential immunisation strategies (Escolano *et al.*, 2019).

A third approach focussed on an HIV-1 infected individual who had developed a DH270 bNAb lineage from which Saunders *et al.* (2019) designed an Env trimer, 10.17DT. This group removed the N133 and N138 glycans from the viral trimer which resulted in a SOSIP trimer that showed high binding affinity for DH270-lineage antibodies. Env trimer nanoparticles were used to immunise human IGH and IGK knock-in mice and were able to elicit V3-glycan specific serological responses (Saunders *et al.*, 2019).

These studies all show proof-of-principle for a germline-targeting V3-glycan immunogen, however only one germline-targeting V3-glycan immunogen published to date has been derived from a subtype C virus (Saunders *et al.*, 2019). In addition, these studies used germline-targeting immunogens that were not based on bNAb-initiating Envs, which may have intrinsic features resulting in better engagement of naïve bNAb precursor B-cells.

1.6 Study aim and objectives

The potential of a subtype C trimer based on naturally circulating viral sequences was investigated. We identified three HIV-1 infected participants, CAP255, CAP314 and CAP177, who developed broad responses to the V3-glycan epitope (Gray *et al.*, 2011b; Moore *et al.*, 2012). CAP255 and CAP177 were participants in the Centre for the AIDS Programme of Research in South Africa (CAPRISA) 002 acute infection cohort study in KwaZulu-Natal from 2005-2009, while CAP314 was a participant enrolled in the CAPRISA 004 Tenofovir trial in KwaZulu-Natal from 2007-2010. Throughout these studies, longitudinal blood samples were collected at multiple timepoints. CAP255 was previously shown to have developed a broad response towards the V3-glycan site at 109 weeks post infection (wpi) (Gray *et al.*, 2011b). Our group has previously isolated two bNAbs, CAP255.G3 and CAP255.C5, that belonged to divergent branches of the broad CAP255 V3-glycan directed lineage at 149 wpi (Kitchin *et al.*, unpublished data). Two potential UCAs of the CAP255 V3-glycan lineage were inferred from 13 wpi sequencing data, which only differed from each other in the junctional region of their light chain sequences (Kitchin *et al.*, unpublished data). The light chain of CAP255 UCA01 was more similar to the mature monoclonal antibody (mAb) CAP255.G3 while the light chain of CAP255 UCA02 was more similar to the mature mAb CAP255.C5. Deep unpaired sequencing of the CAP255 heavy chains previously done on the Illumina MiSeq platform also led to the identification of two CAP255 intermediate antibody heavy chains, CAP255 Int A and CAP255 Int B, at 39 wpi (Kitchin *et al.*, unpublished data). These heavy chains were paired with the mature CAP255.C5 light chains as they were members of the CAP255.C5 branch of the lineage. The putative bNAb-initiating *env* was identified at 13 wpi through single genome amplification (SGA) sequencing previously done by our lab (Kitchin *et al.*, unpublished data). Similarly, we have identified and isolated antibodies from a V3-glycan-specific bNAb lineage from CAP314 and identified the UCA at 22 wpi using deep sequencing of the lineage. Two early CAP314 intermediates, CAP314 54 and CAP314 89, were isolated at 54 wpi and two late intermediates, CAP314 56 and CAP314 75, were isolated at 90 and 115 wpi respectively. No bNAbs have yet been isolated from CAP177, but the laboratory previously showed that CAP177 plasma develops a V3-glycan response at approximately 45 wpi (Gray *et al.*, 2011b). Due to the bNAb response specificity to the V3-glycan site, these three CAPRISA participants were selected for this study.

1.6.1 Project aim

This project aimed to identify and characterise bNAb-initiating Envs from CAP255, CAP314 and CAP177, which triggered V3-glycan-specific bNAb responses in HIV-1 infected donors, and use them as a basis for the design of germline-targeting immunogens.

1.6.2 Specific objectives

- **To identify potential bNAb-initiating Envs in CAPRISA donors: CAP177 and CAP314.** Plasma samples from CAP314 and CAP177 will be selected based on development of breadth and timepoints at which bNAbs and UCAs had previously been identified, and will be sequenced using longitudinal deep sequencing on the Pacific Biosciences (PacBio) platform. A putative bNAb-initiating virus from CAP255 had been identified prior to this study using a combination of SGA and PacBio derived *env* sequences.

- **To design and characterise V3-glycan-specific germline-targeting immunogens based on bNAb-initiating Envs.** The putative bNAb-initiating Envs of CAP255 and CAP314 will be used as the basis for the design of stable SOSIP Env trimers. Mutations that had previously been introduced into V3-glycan specific germline-targeting immunogens will be inserted into the CAP255 and CAP314 bNAb-initiating Envs. Binding and functional assays will be used to determine the stability and conformation of the resulting Env trimers.

- **To assess the ability of UCAs from V3-glycan targeting broad lineages to engage the bNAb-initiating Envs.** Binding and neutralisation assays will be used to determine the ability of the putative bNAb-initiating virus derived Envs to engage V3-glycan-directed UCAs.

2 Materials and Methods

2.1 Study participants and sample collection

Three HIV-1 positive CAPRISA participants were selected, based on their known bNAb responses toward the V3-glycan epitope. CAP314 was a participant enrolled in the CAPRISA 004 Tenofovir trial, whereas both CAP177 and CAP255 were enrolled in the CAPRISA 002 acute infection cohort study. The participants enrolled in the above mentioned cohorts were ART-naïve upon enrolment at 6-13 wpi and longitudinal blood samples were collected from participants over time at longitudinal follow-ups. CAP255 had previously undergone viral and antibody sequencing and the putative bNAb-initiating virus had been determined prior to this study.

2.2 *Env* deep sequencing

2.2.1 Sample processing, RNA extraction and cDNA synthesis

QIAamp Viral RNA Mini Kits (Qiagen) were used to extract approximately 40 000 HIV-1 RNA copies as per the manufacturer's instructions from selected CAP314 and CAP177 samples. A targeted amplicon sequencing approach on the PacBio platform was utilised to sequence the full gp160 region of HIV-1 (**Figure 2.1**). *Env* amplicon libraries were prepared for sequencing on the PacBio Sequel system making use of its single molecule, real-time (SMRT) sequencing technology to examine the long coding regions of individual *env* genes. The technology provides exceptional read lengths for long gene sequences, such as *env*, without compromising on throughput or accuracy. For this approach, we designed 16 complementary DNA (cDNA) primer sequences (synthesised by Integrated DNA Technologies, IDT) (**Table 2.1 and Table 2.2**), comprised of two primer binding regions for the two different rounds of polymerase chain reaction (PCR), a sample ID sequence, a single unique molecular identifier (sUMI) sequence and a gene specific sequence (**Figure 2.2**). The sUMI region of the primer consisted of an 8 random nucleotide tag/molecular barcode, to uniquely tag each cDNA molecule. Addition of the tag occurred during the cDNA synthesis step and was used during data analysis to distinguish between different *env* variants present at specific timepoints. A

sample ID, consisting of 6 nucleotides, also forms part of the cDNA primer so that different samples can be pooled on the same PacBio SMRT cell, but allowing sample specific sequences to be disaggregated during the data analysis steps.

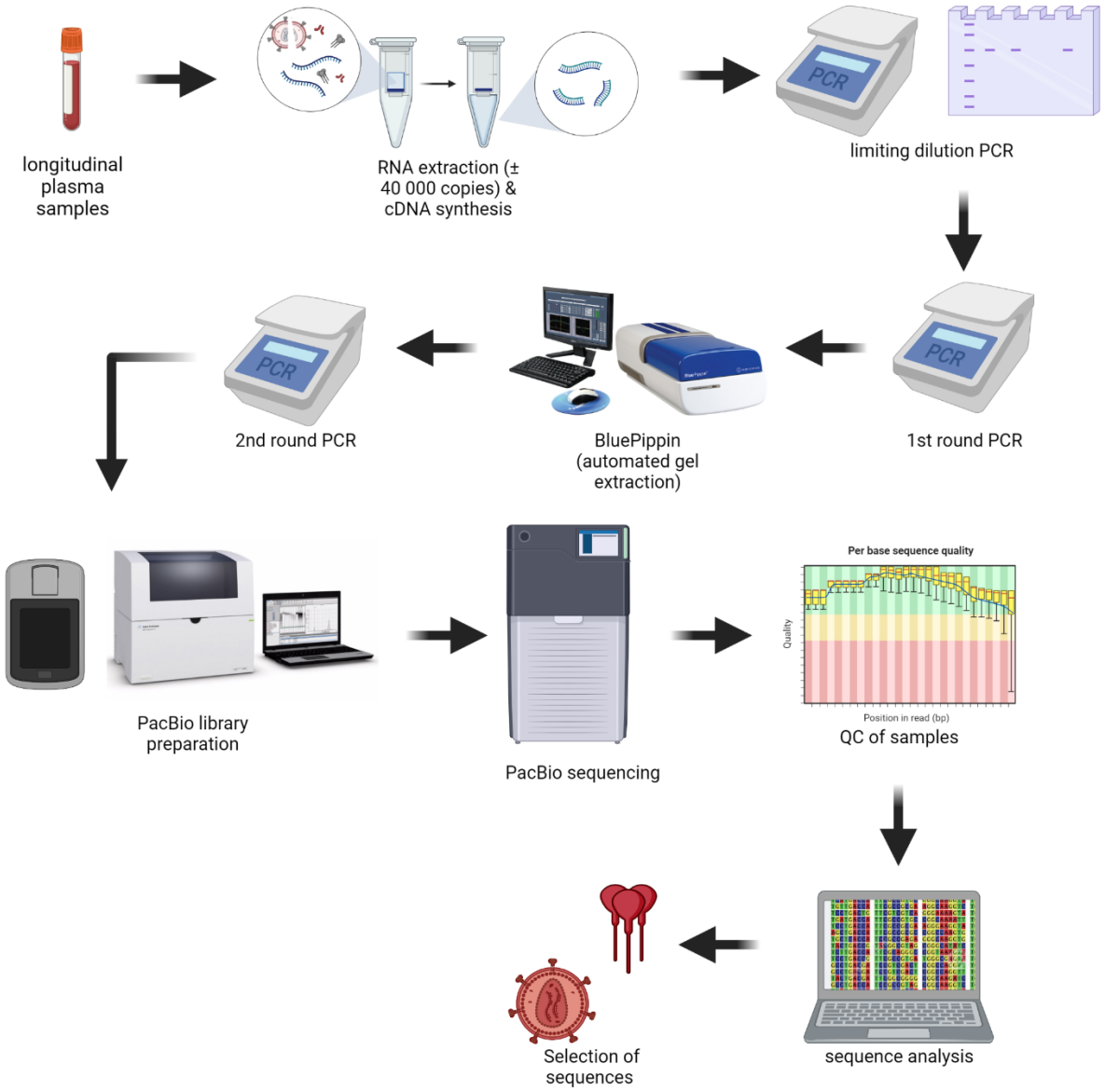

Figure 2.1: *Env* deep sequencing pipeline
A schematic representing the various steps followed to generate PacBio sequencing data to inform the selection of sequences for trimer synthesis. The initial steps include generation of cDNA from viral RNA, followed by limiting dilution PCR to estimate the viral copy number per sample. First and second round PCRs are shown, each with their respective BluePippin, Qubit and Tapestation QC steps, to ensure the correct region has been selected. The resulting amplicon sequencing library was sequenced on a PacBio Sequel I instrument. Finally the sequencing data was used to select potential sequences of interest to be made into trimers and pseudoviruses. Figure created using BioRender.

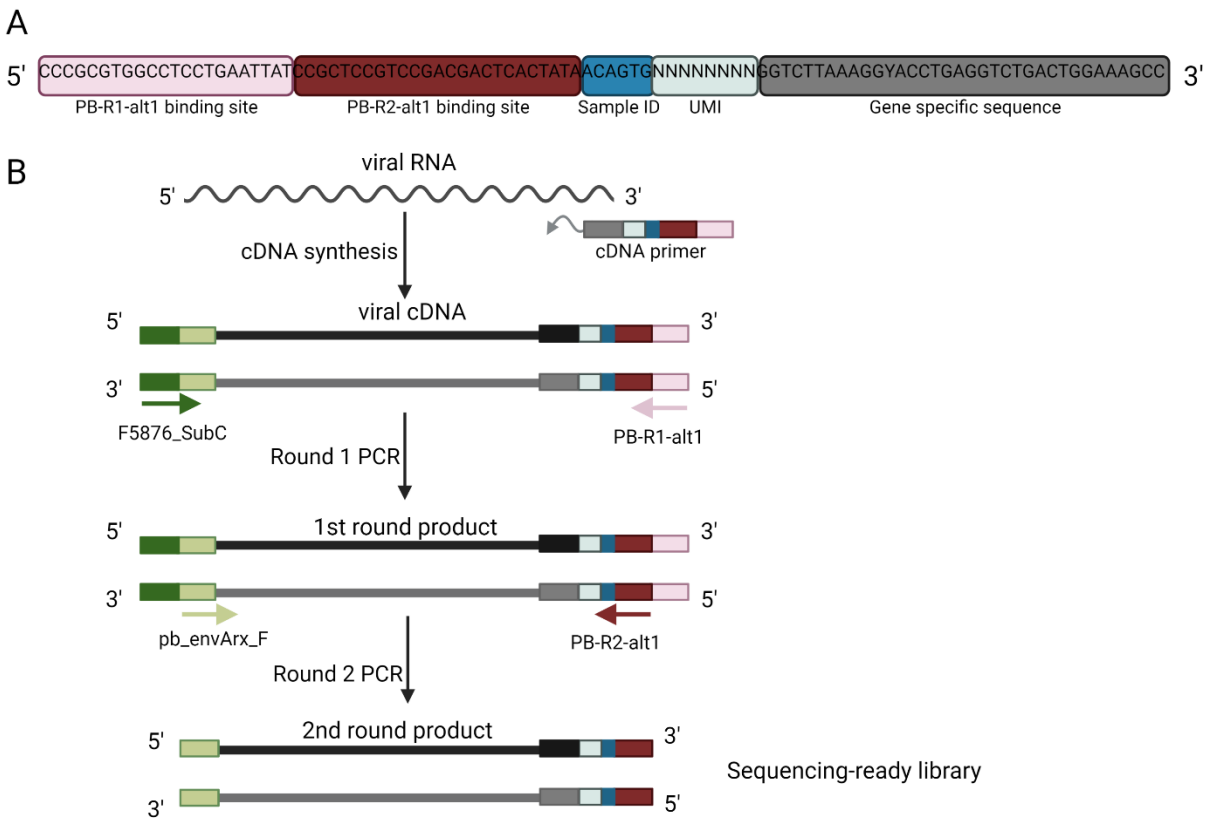

Figure 2.2: Regions of the cDNA *env* primer
(A) Regions of the cDNA *env* primer are shown with the binding sites of the first round reverse primer (PB-R1-alt1), and the second round reverse primer (PB-R2-alt1) highlighted. The position of the sample ID and UMI are nested between the PB-R2-alt1 and gene specific sequences. **(B)** Schematic showing cDNA synthesis, followed by first and second round PCR reactions and their respective primer binding sites. Figure created using BioRender.

After the RNA extraction 25 µl of RNA, 2.5 µl dNTPs (10 mM each) and 2 µl cDNA primer (10 µM) were heated for 5 minutes at 65°C and placed on ice for 2 minutes. A second mastermix containing 2 µl Thermastop (5 U/µl) (Sigma-Aldrich), 10 µl Superscript IV buffer (5×) (Invitrogen), 2.5 µl Dithiothreitol (DTT) (100 mM), 2.5 µl RNaseOUT (40 U/ µl) (Invitrogen) and 2.5 µl Superscript IV (10 U/ µl) (Invitrogen) was added to the first reaction. Another incubation at 50°C for 90 minutes followed. Thereafter, 1 µl Superscript IV (10 U/ µl) was added to the reaction, followed by another incubation at 50°C for 90 minutes. The reaction was heated to 85°C for 15 minutes and 1 µl RNaseH was added. The reaction was incubated at 37°C for 20 minutes, followed by a Mag-Bind total pure NGS (Omega Bio-tek) bead clean-up using a 1:1 beads:sample ratio, as per the manufacturer's instructions, with cDNA eluted in 25 µl nuclease free water and stored in aliquots at −80°C.

Table 2.1: List of sample ID sequences

Primer name	Sequence
PB-A1	ACAGTG
PB-A2	CACTCA
PB-A3	GGTAGC
PB-A4	TAGCTT
PB-A5	CTATAC
PB-A6	ATCACG
PB-A7	ACTGAT
PB-A8	TGACCA
PB-A9	GCTCAT
PB-A11	CGATGT
PB-A12	ATGCTG
PB-A13	ACGATC
PB-A14	GTCATC
PB-A15	CGAGTA
PB-A16	GACAGA
PB-A17	TAGAGC

Table 2.2: List of primer sequences

Primer name	Description	Sequence
PB_AX_nef67_degen_PID_cDNA	cDNA primer	CCCGCGTGGCCTCCTGAATTATCCGCTCCGTCCGACGACTCACTATAACAGTGNNNNNNNNGGTCTTAAAGGYACCTGAGGTCTGACTGGAAAGCC
F5876_SubC	First round PCR forward primer	AGAGCCCTGGAACCATCCAGGAAG
pb_envArx_F	Second round PCR forward primer	GGCTTAGGCATCTCCTATAGCAGGAAGAA
PB-R1-alt1	Universal first round reverse primer	CCCGCGTGGCCTCCTGAATTAT
PB-R2-alt1	Universal second round reverse primer	CCGCTCCGTCCGACGACTCACTATA
EnvN	First round reverse primer, positive control	TTGCCAATCAAGGAAGTAGCCTTGTGT
E01	Second round reverse primer, positive control	TCCAGTCCCCCCTTTTCTTTTAAAAA

2.2.2 HIV-1 *env* amplification

In order to quantify the approximate number of cDNA templates that resulted from the cDNA synthesis, a limiting dilution PCR was used. This was to estimate the amount of cDNA that would be needed in the first round PCR to ensure that no more than 200 copies of cDNA were added per sample, since this can lead to nonspecific amplification and recombination

between variant sequences. Serial dilutions of 1:4, 1:16, 1:32, 1:64 and 1:256 were prepared using cDNA from the previous step. A nested PCR was set up with 3 replicate reactions per dilution. Each reaction, consisting of 5 µl PrimeSTAR GXL buffer (5×) (Takara Bio), 2 µl dNTPs (2.5 mM each), 0.625 µl F5876_SubC (10 µM), 0.625 µl PB-R1-alt1 (10 µM), 0.5 µl PrimeSTAR GXL (1.25 U/ µl) (Takara Bio) and 1 µl of diluted cDNA template, were performed according to the PCR parameters shown in **Table 2.3**. To ensure that the reagents and primers were working 1 µl CAP45 TF Env, a full length gp160 cloned into a pJET1.2 vector (Thermo Fisher Scientific), (1:100 000 plasmid dilution at starting concentration 4300 ng/µl) was used as a control template and 1 µl dH_2O was used as the negative control. The control template reaction contained 0.625 µl of the EnvN primer (10 µM). See **Table 2.2** for a list of all primer sequences.

Table 2.3: Limiting dilution round 1 PCR cycling parameters

	Cycles	Temperature	Time
Initial denaturation	**1×**	**98°C**	**2 min**
Denaturation	**35×**	**98°C**	**10 s**
Annealing		**60°C**	**15 s**
Extension		**68°C**	**4 min**
Final extension	**1×**	**68°C**	**7 min**
Hold	**1×**	**4°C**	**∞**

Each second round reaction consisted of 1 µl of the first round PCR, 5 µl PrimeSTAR GXL buffer (5×), 2 µl dNTPs (2.5 mM each), 0.625 µl pb_envArx_F (10 µM), 0.625 µl PB-R2-alt1 (10 µM), and 0.5 µl PrimeSTAR GXL (1.25 U/ µl). The PCR was performed according to the parameters stipulated in **Table 2.4**. The control template reaction contained 0.625 µl of the E01 primer (10 µM).

Table 2.4: Limiting dilution round 2 PCR cycling parameters

	Cycles	Temperature	Time
Initial denaturation	**1×**	**98°C**	**2 min**
Denaturation	**35×**	**98°C**	**10 s**
Annealing		**60°C**	**15 s**
Extension		**68°C**	**3 min**
Final extension	**1×**	**68°C**	**7 min**
Hold	**1×**	**4°C**	**∞**

The products were visualised on a 1% agarose gel and the cDNA copy number for each sample was calculated using the QUALITY programme (https://indra.mullins.microbiol.washington.

edu/quality/). The method allows the user to input total number of reactions and the number of PCR positives at each dilution to estimate the cDNA copy number for each sample. The programme is a variant of the minimum χ^2 (MC) method that has previously been described by Taswell (1981), the method on which the QUALITY programme is based is described by Rodrigo *et al.* (1997).

The first round of PCR was set up using the limiting dilution results to normalise the cDNA template per reaction to 200 copies. Each reaction consisted of 5 µl PrimeSTAR GXL buffer (5×) (Takara Bio), 2 µl dNTPs (2.5 mM each), 0.625 µl F5876_SubC (10 µM), 0.625 µl PB-R1-alt1 (10 µM), 0.5 µl PrimeSTAR GXL (1.25 U/ µl) (Takara Bio) and 2 µl of bead purified cDNA. The PCR parameters in **Table 2.5** were followed. As a control template, 1 µl CAP45 TF Env (1:100 000 plasmid dilution at starting concentration 4300 ng/µl) was used and 1 µl dH_2O was used as the negative control. The control template reaction received a 0.625 µl addition of EnvN (10 µM). PCR replicates were pooled and 30 µl was run on the BluePippin (Sage Science) automated gel extraction instrument, set to a capture size of 3100 bp, to purify first round products.

Table 2.5: Round 1 PCR cycling parameters

	Cycles	Temperature	Time
Initial denaturation	**1×**	**98°C**	**2 min**
Denaturation	**22×**	**98°C**	**10 s**
Annealing		**60°C**	**15 s**
Extension		**68°C**	**4 min**
Final extension	**1×**	**68°C**	**7 min**
Hold	**1×**	**4°C**	**∞**

Four second round reactions were set up using 2 µl of the first round reactions as template. For each reaction 5 µl Superfi buffer (5×) (Thermo Fisher Scientific), 0.5 µl dNTPs (10 mM each), 1.25 µl pb_envArx_F (10 µM), 1.25 µl PB-R2-alt1 (10 µM), and 0.25 µl Platinum Superfi polymerase (1.25 U/ µl) (Thermo Fisher Scientific) were added to 2 µl of template and the parameters shown in **Table 2.6** were used to perform the PCR. A total of 5 µl reagent was added to the reaction consisting of 1 µl Superfi buffer (5×) (Thermo Fisher Scientific), 0.5 µl dNTPs (10 mM each), 1.25 µl pb_envArx_F (10 µM), 1.25 µl PB-R2-alt1 (10 µM), and 0.25 µl Platinum Superfi polymerase (1.25 U/ µl) (Thermo Fisher Scientific).

A final PCR reaction was carried out as shown in **Table 2.7**. The template control and negative control of the first round PCR were used as templates for the second round controls and 1.25 µl E01 (10 µM) was added into the control template reaction. The controls did not receive additional reagents.

Table 2.6: Round 2 PCR cycling parameters

	Cycles	Temperature	Time
Initial denaturation	**1×**	**98°C**	**30 s**
Denaturation	**22×**	**98°C**	**10 s**
Annealing		**65°C**	**10 s**
Extension		**72°C**	**3 min**
Final extension	**1×**	**72°C**	**7 min**
Hold	**1×**	**4°C**	**∞**

Table 2.7: Spike in PCR cycling parameters

	Cycles	Temperature	Time
Denaturation	**1×**	**98°C**	**1 min**
Annealing	**1×**	**65°C**	**10 s**
Extension	**1×**	**72°C**	**10 min**
Hold	**1×**	**4°C**	**∞**

The products were visualised on a 1% agarose gel and PCR replicates were pooled per sample. The samples, from multiple timepoints for each individual, were pooled in equimolar amounts with a target of 1.5-2 µg total DNA per pooled amplicon library. The pooled second round reactions were bead purified, as per the manufacturer's instructions using a 0.7:1 bead:sample ratio of Mag-Bind total pure NGS beads (Omega Bio-tek). The amplicons were eluted with 40 µl Tris-HCl (10 mM), concentrations were measured on a Qubit 3 (Invitrogen) and run on a TapeStation 4200 (Agilent) using a DNA High Sensitivity D5000 kit. Pooled samples were submitted to the NICD sequencing core for sequencing on the PacBio Sequel I instrument.

2.2.3 Env analysis

Circular consensus sequencing (CCS) reads derived from individual cDNA copies were generated from raw PacBio sequencing data, using the onboard PacBio software. Sequences from each timepoint were binned according to their sUMIs, generating variant specific consensus sequences, using custom Julia scripts. The resulting fasta sequences were trimmed

manually and aligned using MAFFT (version 7) (Katoh and Standley, 2013). Sequences containing large deletions (≥10 AA) and stop codons were deleted and the remaining sequences were dereplicated (to collapse identical variant sequences) using VSearch (version 2.17.0) (Rognes *et al.*, 2016). To determine if the remaining sequences showed any interesting mutations in regions of interest the lanl highlighter tool (https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html) was used. Sequences were grouped according to their timepoints and broken down into variable and constant regions using Aliview (version 1.17.1) (Larsson, 2014). Changes observed in areas known to influence V3-glycan dependant bNAb binding were assessed. Trends identified in regions of interest were recorded and compared between timepoints. Sequences that showed changes, such as loss of glycans at the CD4bs and V3-glycan site, insertions, deletions and shortening of V1 loops, at sites that are known to influence bNAb binding were selected to be made into trimers and pseudoviruses.

2.3 HIV-1 Env construct design and transfection

2.3.1 Design and expression of SOSIP.664 gp140 trimers

Selected CAP255 and CAP314 *env* sequences were made into SOSIP.664 gp140 constructs by introducing the following changes described by Sanders *et al.* (2013): A501C, T605C, I559P, $_{508}$REKR$_{511}$ to RRRRRR and a stop codon after gp41$_{ECTO}$ residue D664. A linker and His-tag were added to the SOSIP trimers by adding GSGSGGSGHHHHHHHH to the C-terminus, preceding the stop codon (Sanders *et al.*, 2013). Additional changes that had previously been shown to stabilise trimers and enhance binding affinity of V3-glycan antibodies were added to the CAP314 and CAP255 trimers (Steichen *et al.*, 2016, 2019; Escolano *et al.*, 2019; Saunders *et al.*, 2019). The mammalian codon-optimised constructs were obtained from GenScript and cloned into the CMVR expression vector (NIH HIV Reagent Program, Division of AIDS, NIAID) using restriction sites *Xba*I and *Bam*HI. XL10-Gold ultracompetent cells were transformed using the above mentioned gene constructs, cultured overnight and single colonies were purified using a ZymoPURE II plasmid maxiprep kit (Zymo Research). Sanger sequencing was used to ensure the cloning was successful by confirming the sequence and insert orientation.

Human embryonic kidney (HEK) 293F suspension cells were seeded at 1×10^6 cells/ml in a final volume of 1000 ml Freestyle media (Thermo Fisher Scientific) and incubated overnight at 37°C, 10% CO_2, 70% humidity in a shaking incubator. The following day, 500 µg of gp140 plasmid, 300 µg of Furin and 40 ml of OptiPRO SFM (Thermo Fisher Scientific) were combined. Separately, freshly prepared linear polyethylenimine hydrochloride (PEImax) (Polysciences) transfection reagent was added to 40 ml OptiPRO. The DNA-OptiPro mixture was filter sterilised using a 0.2 µm filter and added to the PEImax-OptiPRO mixture. This was left to incubate at room temperature for 20 minutes. After 20 minutes the mixture was added to the HEK 293F cells and incubated at room temperature for a further 20-30 minutes. The flask was then incubated in the shaking incubator for 6-7 days at 37°C, 10% CO_2, 70% humidity.

2.3.2 Expression of monoclonal antibodies

HEK 293F cells were seeded at 1×10^6 cells/ml in a final volume of 400 ml Freestyle media and incubated overnight at 37°C, 10% CO_2, 70% humidity in a shaking incubator. A volume of 1.2 ml freshly prepared PEImax was added to 10 ml OptiMEM (Thermo Fisher Scientific) and incubated for 5 minutes. In a separate flask 10 ml OptiMEM was thoroughly mixed with 200 µg of purified light chain and 200 µg of purified heavy chain plasmid DNA. The OptiMEM-DNA mixture was directly filtered into the OptiMEM-PEImax mixture using a 0.2 µm filter and incubated for 20 minutes. The mixture was then added to the HEK 293F cells and incubated for 10-20 minutes at room temperature. The cells were incubated at 37°C, 10% CO_2, 70% humidity in a shaking incubator for 6 days.

2.4 Protein purification

2.4.1 Env protein purification

A Nickel-affinity column was prepared for each His-tagged SOSIP.664 trimer using a Bio-rad Econo-pack column (Lasec). The column was washed with wash buffer (20 mM Imidazole, 20 mM Tris-HCl, 150 mM NaCl) prior to running the filter sterilised (0.2 µm filter) supernatant over the column at a flow rate of 0.5 – 1 ml/min. Thereafter, the column was washed with wash buffer, followed by the elution of the trimer using elution buffer (400 mM Imidazole, 150 mM NaCl, 20 mM Tris-HCl). The eluent was concentrated using a 50 kDa Vivaspin cloumn

and the affinity purified Env proteins were further purified using size exclusion chromatography (SEC) on a Superdex 200 column (HiLoad 16/600, Cytiva). The trimer fractions were collected, pooled and run through a 447-52D negative selection column to remove ill-formed trimers with exposed V3 loop regions. The column was washed with 2× coupling buffer (0.1 M Sodium bicarbonate, 0.5 M NaCl, pH 9) and 2× low pH buffer (0.1 M acetic acid, 0.5 M NaCl, pH 4.5), alternating between washes, prior to the addition of the pooled fractions. The flow through was collected and a 50 kDa Vivaspin column was used to buffer exchange the Env trimers into 1× phosphate-buffered saline (1× PBS). Protein concentrations were determined using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific). Aliquoted trimers were stored at -80°C until needed.

2.4.2 Antibody purification

A protein A column which has affinity for the Fc region of IgG1 antibodies was prepared to purify each antibody through gravity-flow using a Bio-rad Econo-pack column (Lasec). The column was washed with 1× PBS and the filter sterilised (0.2 µm filter) supernatant was run over the column at a flow rate of 1 – 1.5 ml/min. Once all the supernatant had run through, the column was washed with 1× PBS. The antibody was eluted from the column using elution buffer (0.1 M glycine, 0.15 M NaCl, pH 2.5) directly into neutralisation buffer (1 M Tris, pH 8) to prevent the acidity of the elution buffer from altering the antibody's structure. The resulting flow through was dialyzed overnight in 1× PBS. The antibodies were concentrated using a 50 kDa Vivaspin column to 1 mg/ml. The concentrations of eluted antibodies were determined using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific), aliquoted and stored at -80°C.

2.5 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

A sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gel was used to assess the purity of SOSIP trimers using non-reducing and reducing conditions. The reducing conditions resulted from the addition of beta-mercaptoethanol (β-ME) to the loading buffer. The samples and 6× loading buffer were mixed together and placed on a heating block at 90°C for 5 minutes. The samples were loaded onto a pre-cast polyacrylamide gel (Novex Invitrogen). A Sharp Pre-stained Protein Standard ladder (Novex Invitrogen) was used as a

reference for the size of protein bands seen. The gels were run at 160 V for approximately 2 hours, stained overnight on a shaking incubator using CBB safe stain (Takara Bio) and destained using water before being visualised on the GelDoc system (BioRad).

2.6 Enzyme linked immunosorbent assay

To assess binding of the selected monoclonal antibodies to the purified SOSIP trimers an enzyme linked immunosorbent assay (ELISA) was used. The purified SOSIP trimers were diluted with 1× PBS to a final concentration of 2 µg/ml and 50 µl of diluted trimer was added to each well of nickel coated plates (Thermo Fisher Scientific). The plates were incubated at room temperature for 2 hours, then 200 µl of blocking buffer (5% non-fat milk in 1× PBS) was added per well and incubated at room temperature for 2 hours. Plates were washed three times using 200 µl wash buffer (1× PBS) for each wash. Thereafter, selected antibody dilutions were prepared (100 ug/ml starting concentration for UCAs and 10 µg/ml for all other antibodies). The prepared antibody dilutions were added and titrated down the plate in 3-fold dilutions. The plates were incubated at room temperature for 1 hour. A wash step followed, after which 50 µl of a 1:3000 dilution of a secondary horseradish peroxidase anti-IgG antibody was added to each well and incubated for 1 hour at room temperature. The wash step was repeated and 100 µl of 1-Step Ultra TMB substrate (Thermo Fisher Scientific) was added to each well and allowed to develop for 5 minutes. A blue colour change could be observed in wells where the trimer had been bound by the antibodies. The reaction was stopped using 50 µl 1 M sulphuric acid resulting in a yellow colour for positive samples. The absorbance of samples were measured at 450 nm using a SpectraMax ABS Plus Microplate Reader (Molecular Devices) to give the optical density (OD_{450}nm). All ELISAs were repeated twice. The optical densities for the antibody titrations against each of the trimers was plotted using GraphPad Prism 9. To assess whether there were significant differences in the binding of the mature (CAP255.C5 and CAP255.G3) or intermediate (CAP314 54, CAP314 56 and CAP314 75), conformational (PGT151 and 35O22), non-neutralising (17b, F105 and 447-52D) and known non-binding (10E8 and PG9) antibodies to each trimer, a one-way ANOVA was performed to determine whether there were any statistically significant differences in the mean area under the ELISA binding curves between the different antibodies. A post-hoc

Dunnett test was used to assess whether the binding of each of the mature, conformational and non-neutralising antibodies to the trimer was statistically significant, relative to the known non-binding antibody controls. In the above statistical analyses the binding of all antibodies were compared to the mean area under the curve for the combined curves of the two known non-binding antibodies. All statistical analyses were performed using GraphPad Prism 9.

2.7 Surface plasmon resonance

Surface plasmon resonance (SPR) was utilised to determine whether the UCAs bound to the selected CAP255 SOSIP trimers. The instrument was primed with running buffer (1× PBS, 0.05% Tween20 and 0.5% bovine serum albumin) and the Biotin sensor was loaded into the OpenSPR (Nicoya) instrument. Glycine-HCl was injected at 200 µl/min in both channel 1 and 2. After establishing the baseline, a flow rate of 40 µl/min was selected and 0.5 µM Streptavidin was injected. The valve position was changed to channel 2 after the new baseline had been established and 4 µg/ml Protein A was injected. An injection of 20 µg/ml CAP255.C5 followed where after the valve was changed to channel 1 and 2 and the flow rate was changed to 20 µl/min before injecting 1 µM CAP255.GT2. The sensor was washed with Glycine-HCl after each coating of trimer. The experiment was repeated with a range of CAP255.GT2 and CAP255.GT0 concentrations (1 µM – 10 nM) against both 20 µg/ml CAP255.C5 and 20 µg/ml CAP255 UCA01. The resulting data were analysed using the OpenSPR software, TraceDrawer.

2.8 Differential scanning calorimetry

In order to determine the stability of selected trimers, differential scanning calorimerty (DSC) was performed. The thermal denaturation of the selected trimers was determined using a Nano DSC (TA instruments) instrument. The protein concentration of the above trimers were adjusted to 0.6 – 1.7 mg/ml and 650 µl of each sample was loaded into the cell. The samples were tested from 5-95°C at a scan rate of 1°C/min. Buffer correction, normalisation and baseline subtraction were applied to each sample and the resulting curves were analysed using NanoAnalyze (TA instruments). The data were fitted using a Gaussian model.

2.9 Pseudovirus design and production

Pseudoviruses expressing the selected trimer matched Envs were made to be tested in neutralisation assays. Non-codon optimised *rev-vpu-env* gene cassettes in pcDNA3.1 were obtained from GenScript, transformed using XL10-Gold ultracompetent cells and purified using a QIAprep spin miniprep kit (Qiagen). Sanger sequencing was used to confirm Env sequence identity.

HIV-1 pseudoviruses were prepared as previously described by Montefiori (2009). HEK 293T cells were seeded at 2 × 10^6 cells/10 ml in growth media (GM) (Dulbecco's modified eagle media (DMEM), 10% the foetal bovine serum (FBS), 25 mM HEPES buffer, 50 µg/ml gentamycin) and incubated at 37°C overnight. The following day, 4 µg of HIV-1 Env plasmid and 4 µg of pSG3Δenv backbone plasmid were mixed in DMEM. The DNA was then mixed with PEImax and incubated at room temperature for 30-45 minutes. The mixture was added to the HEK 293T cells and incubated at 37°C at 5% CO_2 for 48-72 hours. The pseudovirus-containing supernatant was harvested and filtered using a 0.45 µM nitrocellulose filter, FBS concentration adjusted to 10% of the total volume and 1 ml aliquots were made. The virus was stored at -80°C and thawed as needed. A median tissue culture infectious dose (TCID50) assay, as described by Montefiori (2009), was performed to determine what dilution of the virus would result in a relative luminescence unit (RLU) reading that falls in an acceptable range (20 000-60 000 RLU) for the viral controls of a neutralisation assay.

2.10 Neutralisation assays

The neutralisation potency of various antibodies were determined using a standardised TZM-bl neutralisation assay as previously described by Montefiori (2009). GM was added to all the wells of a flat bottom 96-well plate. Antibodies were added to the appropriate wells and titrated in a series of 4-fold dilutions. The UCAs were diluted to a starting concentration of 100 µg/ml, whereas the starting concentration of intermediate and mature antibodies ranged between 3-50 µg/ml depending on their potency. Previously produced Env-pseudovirus were thawed and diluted with GM to give RLUs between 20 000-60 000. The diluted pseudoviruses

were added to all wells excluding the cell controls. The plates were incubated at 37°C for 45-90 minutes. TZM-bl cells (0.5×10^6 cells/ml) were suspended in media and mixed with diethylaminoethyl (DEAE) dextran to make the cells susceptible to infection, and added to all wells. The plates were incubated for 24 hours at 37°C at 5% CO_2. All wells received 130 µl of GM. After a further 24 hours 100 µl of GM was removed and 100 µl of Bright Glo Reagent (AnaTech) was added. The plates were incubated for 2 minutes at room temperature. A total of 150 µl GM and Bright Glo Reagent (Promega) were mixed together 5-10 times and transferred to a 96 well black plate. The plates were read using a VICTOR XLight Luminometer (PerkinElmer) and the neutralisation potency of the antibodies were determined by comparing the light signal received from the wells containing antibody with that received from the cell and virus controls. The neutralisation potency of antibodies was expressed as the half maximal inhibitory concentration (IC_{50}), at which a 50% reduction of RLU with respect to the virus control wells was seen. The IC_{50} values of all antibodies were plotted using GraphPad Prism 9.

3 Results

3.1 Germline gene usage of V3-glycan targeting antibodies

A large number of V3-glycan bNAbs that have been previously isolated make use of the IGHV4 germline genes (Moyo *et al.*, 2020). The mature V3-glycan-directed bNAbs from CAP314 and CAP255 used the same germline genes, which we investigated to identify any common conserved germline motifs seen across V3-glycan bNAbs.

Sequences of mature V3-glycan targeting bNAbs from other broad donors that have been added to the Genbank (https://www.ncbi.nlm.nih.gov/genbank/) database and that use IGHV4 genes (AIIMS-P01, DH563, PCDN38B, PCDN27D, PCDN33A, PGDM21, PGT121-128, PGT130, PGT131, PGT135-136, VRC22.01, VRC28.01, VRC29.01-.04, VRC41.01, VRC41.02) were selected, grouped with available CAP255 antibody sequences, and compared to sequences from bNAbs that target the V3-glycan epitope but use different germline genes (PGDM12 using IGHV3 and BF520 using IGHV1 germline genes). These sequences were also compared to bNAbs with IGHV4 germline gene usage but that target different epitopes (CH103 which targets the CD4bs).

The sequences for all mature V3-glycan targeting bNAbs that are known to utilise IGHV4 genes were obtained from Genbank, grouped according to the three allele groups represented (IGVH4-34, IGVH4-39 and IGVH4-54) and separated into their framework and CDRH regions using IMGT/V-quest (Brochet *et al.*, 2008) (**Figure 3.1)**. The sequences showed that some motifs in the framework regions of IGHV4 antibodies were conserved from the germline sequences (**Figure 3.1A-C**). These motifs were not seen in the framework regions of other mature V3-glycan-directed bNAbs which had different germline usage. These motifs, including LSLTC in framework 1 and R-XX-PG in framework 2, also appeared in affinity matured IGHV4 antibody sequences that were not directed to the V3-glycan site (**Figure 3.1A-B)**. There were no conserved motifs in the CDRH1 and CDRH2 regions (**Figure 3.1D-E**). By investigating these five regions of the heavy chain sequences of V3-glycan targeting bNAbs, no trends were observed exclusively in V3-glycan antibodies that make use of IGHV4 genes.

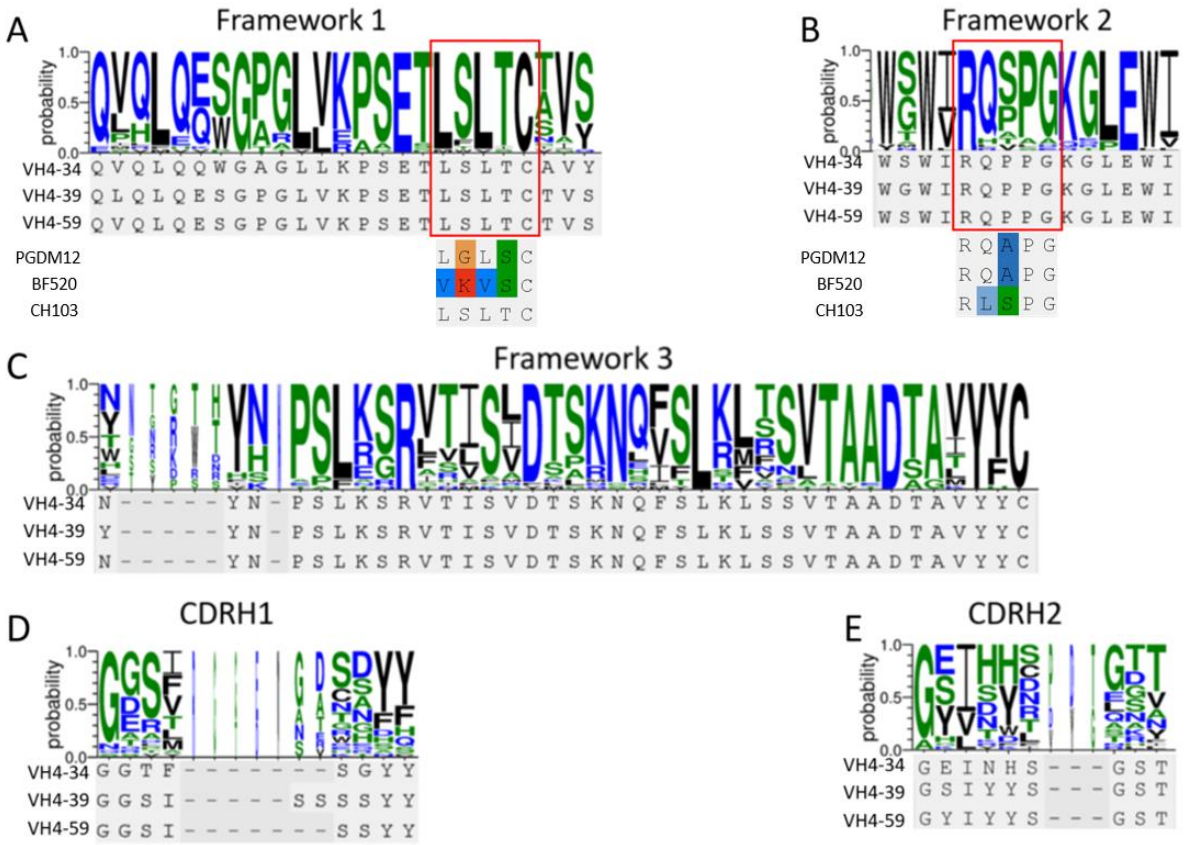

Figure 3.1: Logograms of aligned regions of V3-glycan bNAb heavy chain sequences
The aligned **(A)** framework 1, **(B)** framework 2 and **(C)** framework 3 regions of the heavy chain sequences of all available V3-directed bNAbs (AIIMS-P01, CAP255.C5, CAP255.G3, DH563, PCDN38B, PCDN27D, PCDN33A, PGDM21, PGT121-128, PGT130, PGT131, PGT135-136, VRC22.01, VRC28.01, VRC29.01-04, VRC41.01, VRC41.02) which utilise the IGVH4 germline genes (VH4-34, VH4-39 and VH4-59). Changes seen in the mature **(D)** CDRH1 and **(E)** CDRH2 regions of V3-directed bNAbs mentioned previously compared to the IGVH4 germline sequences. Motifs observed are highlighted using red blocks and compared to control bNAb sequences which are V3-directed but use IGVH3-11 (PGDM12) and IGVH1-2 (BF520) germline genes and which are CD4bs directed but make use of IGVH4-61 germline genes (CH103).

3.2 Identification of putative bNAb-initiating sequences using *env* deep sequencing

To identify which sequences resulted in the development of the V3-glycan targeting lineages for participants CAP314 and CAP177, deep *env* sequencing was performed using the PacBio platform. We planned to identify the putative bNAb-initiating Envs through the investigation of motifs implicated in V3-glycan binding, at timepoints where each lineage is thought to have emerged. To narrow down the window in which these Envs were thought to occur, the timeframe we focused on was just prior to UCA emergence (17-24 wpi) and before

development of breadth (Gray *et al.*, 2011b; Moore *et al.*, 2012; Kitchin *et al.*, unpublished data; Scheepers *et al.*, unpublished data) (**Figure 3.2**). This information was used to extrapolate ten timepoints for which previously collected blood plasma samples were selected for sequencing (**Table 3.1**). The five selected CAP314 timepoints, that cover the 11 weeks before the UCA was identified, were selected to increase the likelihood of finding the bNAb-initiating virus in at least one of the timepoints (**Figure 3.2**).

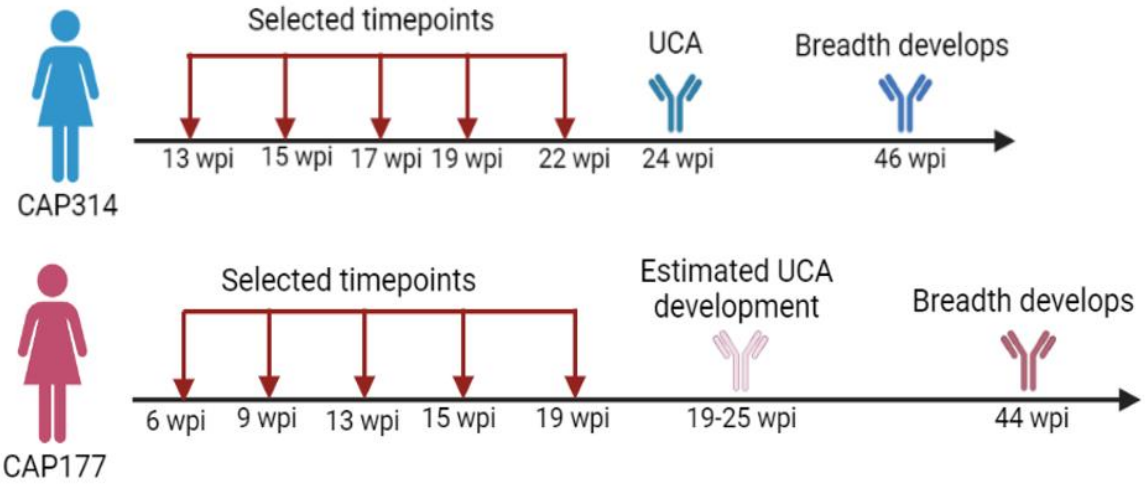

Figure 3.2: Selection of CAP314 and CAP177 timepoints for deep *env* sequencing
Schematic showing the CAP314 and CAP177 timepoints selected for next generation sequencing (NGS) in relation to the timepoint at which the UCA was first identified, or estimated to have emerged for CAP177, and the development of breath for the CAP314 and CAP177 lineages. Figure created using BioRender.

Table 3.1: CAP314 and CAP177 timepoints selected for deep *env* sequencing

Participant	Weeks post infection (wpi)
CAP314 (UCA emerges at 24 wpi)	13
	15
	17
	19
	22
CAP177 (Time of UCA emergence unknown)	6
	9
	13
	15
	19

Since CAP177 has no identified UCA, the timepoints for sequencing were selected based on an estimated timeline of CAP177 UCA emergence, informed by previous CAP314 and CAP255

studies (Gray *et al.*, 2011b; Moore *et al.*, 2012; Kitchin *et al.*, unpublished data; Scheepers *et al.*, unpublished data). The estimated emergence of the CAP177 UCA was calculated to be between 19-25 wpi, before plasma breadth developed at approximately 44 wpi (**Figure 3.2**). Therefore, the timeframe for identification of a putative CAP177 bNAb-initiating Env was estimated to be between 6-19 wpi and five timepoints within this timeframe were selected (**Table 3.1**).

The PacBio sequencing method used to generate these sequences was very high throughput, which resulted in a large number of generated *env* sequences. An average of 180 sequences per CAP314 timepoint and 105 sequences per CAP177 timepoint were generated for subsequent analysis after quality control and de-replication of identical variants present at the same timepoint (**Table 3.2**).

Table 3.2: Number of *env* sequences for both participants at each timepoint of the data processing pipeline

CAPRISA participant	Weeks post-infection	PacBio sequences	Post QC	Post de-replication
CAP314	13	271	188	164
	15	112	91	82
	17	217	164	117
	19	637	404	309
	22	394	325	282
CAP177	6	228	172	78
	9	224	146	86
	13	198	150	102
	15	384	339	197
	19	99	80	61

3.2.1 Evolution of the CAP314 V3-glycan lineage over time

The CAP314 NGS data and previously sequenced SGA sequences were combined and together showed how the virus evolved over the timepoints assessed. The sequences confirmed a very clear shift from a potential N-linked glycosylation site (PNG) (resulting from a N-X-T/S glycan motif) at position N334 to N332 through a N334S substitution (**Figure 3.3A**). The shift from a N334 to a N332 glycan has been implicated as an escape mutation to block binding of early

strain-specific antibodies (Moore *et al.*, 2012). This PNG shift was seen in a single sequence at 13 wpi, however, at 19 wpi the N332 PNG was present in 90% of all sequences. Finally, at 22 wpi there were no sequences with N334 present as all sequences had evolved to possess the PNG at position N332. A population of sequences were identified which had a loss of the PNG at position N460 and/or N463. These populations were seen throughout all timepoints. A crystallised structure of a gp120 in complex with V3-directed bNAbs (PGT121, PGT122, PGT124, PGT128, PGT135 and 10-1074) was used to determine if the loss of the glycan at position N460/N463 would affect the interaction between V3-glycan-directed antibodies and the viral Env (**Figure 3.3B**). From the modelling, it is clear that the glycans at N460 & N463 are positioned distal from the V3-glycan site and therefore make no contact with V3-glycan-directed bNAbs.

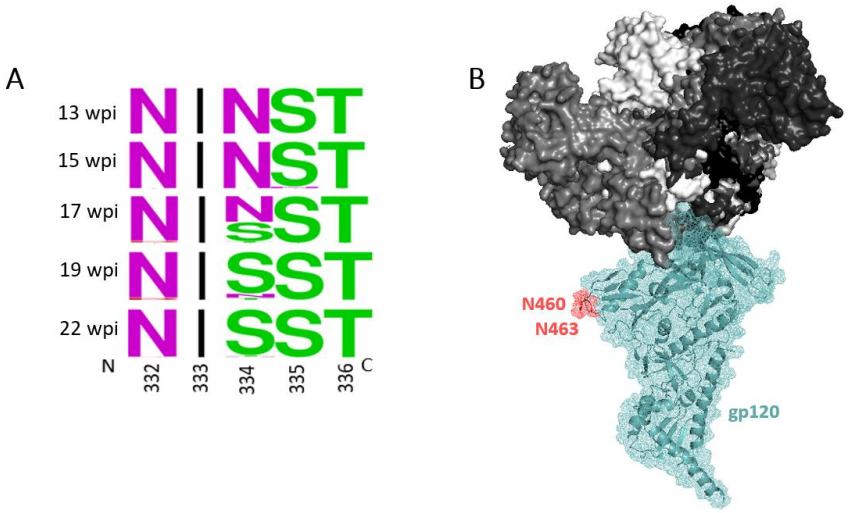

Figure 3.3: Distinct populations observed through deep sequencing of CAP314
(A) The evolution of CAP314 sequences from a PNG (N-X-T/S) at N334 to a PNG at N332 over time due to a N334S substitution. **(B)** A crystallised structure of gp120 (blue) (PBD ID: 5T3Z) shown in complex with known V3-glycan targeting bNAbs (shown in varying shades of grey/black) namely; PGT121 (PBD ID: 4FQ1), PGT122 (PBD ID: 4JY5), PGT124 (PBD ID: 4R2G), PGT128 (PBD ID: 5C7K), PGT135 (PBD ID: 4JM2) and 10-1074 (PBD ID: 5T3Z). The N460 and N463 glycans are highlighted in red. Structures were modelled using Pymol (version 2.3.4).

Previous studies have shown that longer V1 loop lengths occlude the V3-glycan site which made this an important feature to take note of when identifying the putative bNAb-initiating Env (van den Kerkhof *et al.*, 2016; Anthony *et al.*, 2017; Bonsignori *et al.*, 2017). Over time, the CAP314 V1 loop length decreased from 13 wpi, where all V1 loop sequences had a uniform length of 26 AA, to 15 wpi, where there were 7 populations with varying V1 loop lengths of 17-26 AA (**Figure 3.4A-C)**. The 26 AA population was shown to be the dominant population

until 19 wpi when four populations were seen to represent a similar percentage of sequences with V1 lengths of 17 AA, 21 AA, 23 AA and 26 AA (**Figure 3.4D**). At 22 wpi the dominant population changed to that of 23 AA length, observed in 45% of sequences (**Figure 3.4E**).

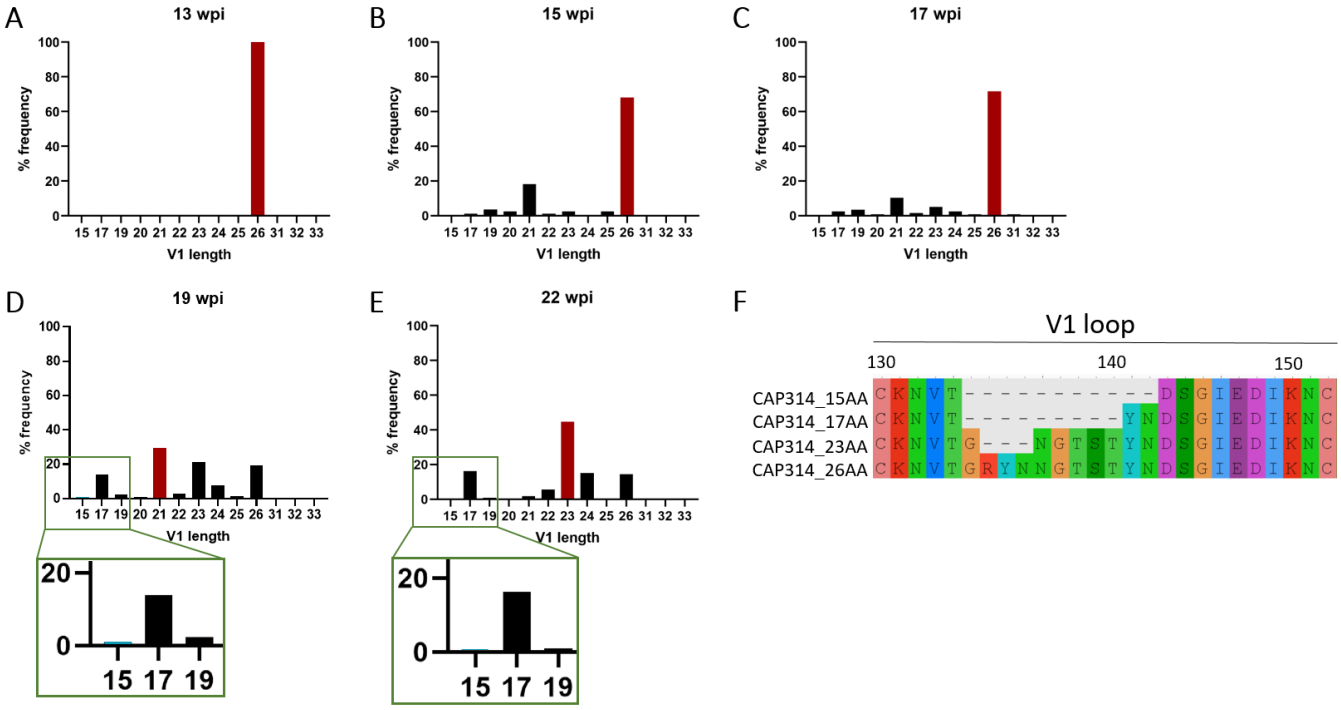

Figure 3.4: V1 loop length of CAP314 V3-glycan-directed lineage Envs decrease over time
The frequency of V1 loop lengths observed at each timepoint. Dark red is used to indicate the population with the highest frequency at each specific timepoint. **(A)** At 13 wpi all *env* sequences had a V1 loop length of 26 AA. **(B)** The most frequent V1 loop length remained 26AA at 15 wpi, however populations with shorter V1 loop lengths were also detected. **(C)** At 17 wpi there were a wide variety of populations with varying V1 loop length present, however 26 AA was still the dominant population. **(D)** The population with the highest frequency shifted to 21AA at 19 wpi and the V1 length of sequences are distributed more evenly amongst populations of 17AA, 23AA and 26AA. A small percentage of sequences are seen with a V1 loop of 15AA (highlighted in blue). A zoomed in portion of the graph is highlighted in green to shown the frequency of the V1 length of 15AA compared to other V1 lengths. **(E)** The population with the highest percentage is shown at 23 AA. An even distribution of sequences are seen between V1 loops of 17AA, 24AA and 26AA. A zoomed in portion of the graph is highlighted in green to shown the frequency of the V1 length of 15AA compared to other V1 lengths. **(F)** The V1 loop sequences of Envs from 19 wpi with lengths of 15AA, 17AA and 23AA that were selected as potential bNAb-initiating viruses aligned to the original 26 AA sequence at 13 wpi.

There was a small population of sequences observed at 19 and 22 wpi which had a V1 loop length of 15 AA (**Figure 3.4D & E**). An average V1 loop length of 27 AA was observed in *env* sequences from the Los Alamos database highlighting the unusual nature of the short V1 loops of the CAP314 Envs identified with a V1 loop length of 15 AA. A shorter V1 loop length was used as a selection criteria for selecting the CAP314 bNAb-initiating virus, and three

representative sequences with shorter V1 loop lengths of 15 AA, 17 AA and 23 AA were selected as putative bNAb-initiating Env sequences (**Figure 3.4F**).

3.2.2 Identification of a putative bNAb-initiating Env of the CAP314 CD4bs directed lineage.

Although the focus of this study was V3-glycan bNAbs, there were two other lineages that developed and co-circulated in CAP314 at the same time as the V3-glycan lineage (Scheepers *et al.*, unpublished data). Both of these were CD4bs directed, and we sought to identify the Env that initiated these responses (Scheepers *et al.*, unpublished data). The PNGs at N460 and N463, which were not present in a population of sequences across all sequenced timepoints, are located within the CD4bs and therefore sequences possessing these glycans were investigated further. Across the five timepoints sequenced, there were six populations identified with PNGs present at N276, N460 and N463 (**Figure 3.5A-E**).

The removal of the N276 glycan has previously been shown to improve CD4bs bNAb binding (LaBranche *et al.*, 2019) and a loss of this PNG was seen from 17 wpi in a small number of sequences (**Figure 3.5C-E**). At 19 and 21 wpi there was a group of sequences that showed a loss of all three PNGs at positions N276, N460 and N463 (**Figure 3.5D-E**). The loss of these glycans has been implicated in enhanced binding by CD4bs antibodies (McGuire *et al.*, 2013). Only one sequence out of 303 sequences showed a lack of all three above mentioned PNGs at 19 wpi and therefore, was selected as the putative CD4bs bNAb-initiating Env. As the focus of this study was on V3-glycan germline-targeting immunogens, we did not pursue this further, but future studies testing the characteristics of this Env will need to be conducted.

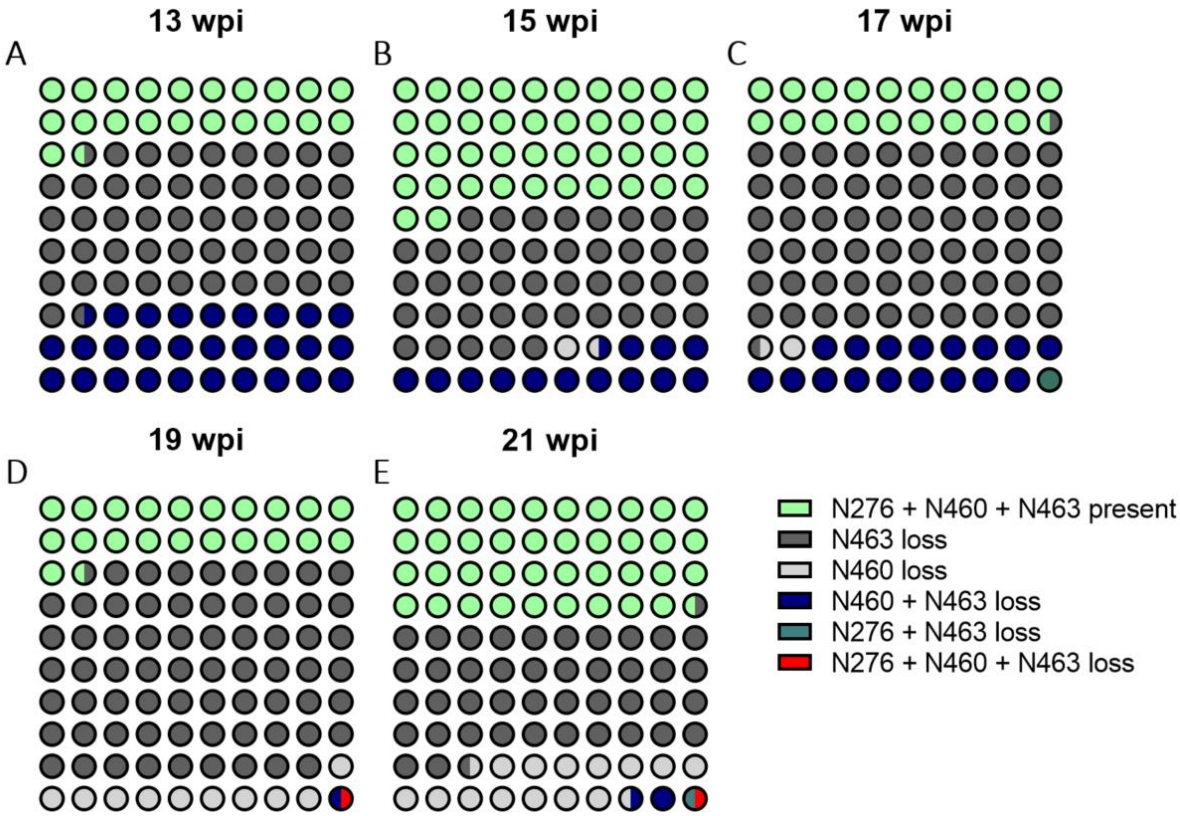

Figure 3.5: Glycan holes develop in the CD4bs throughout infection in CAP314
The percentage of populations present at each timepoint is represented depending on the presence or absence of glycans at positions N276, N460 and N463. **(A)** At 13 wpi three populations are present, no glycans lost (light green), loss of N463 (dark grey) which has the highest frequency and loss of both N460 and N463 (dark blue). **(B)** The emergence of an additional group is seen at 15 wpi, with the loss of the N460 glycan (light grey), at a very low frequency. **(C)** At 17 wpi no glycan at position N463 (dark grey) is still seen in the majority of sequences. There is another population of sequences which showed loss of glycans at N276 as well as N463 (light blue). **(D)** The emergence of a new minority population at 19 wpi shows the formation of glycan holes due to the lack of glycans at N276, N460 and N463 in the same sequences (red). The group which showed a loss of glycans at N460 as well as N463 (dark blue) had decreased compared to the previous timepoints. **(E)** There are a total of six populations present at 22 wpi. The sequences with all glycans present (light green) and with no glycan at N463 (dark grey) represent the majority. There is a small population of sequences that show a loss of glycan N276, some with a loss of N463 as well (light blue) and others with a loss of all three glycans (red).

3.2.3 Identification of a putative CAP177 bNAb-initiating Env

The *envs* from CAP177, the second CAPRISA participant who showed a broad plasma response toward the V3-glycan epitope, showed one group of HIV-1 sequences present at 6 wpi but two distinct groups of viral sequences from 9 wpi (**Figure 3.6**). The distinct viral population which emerged at 9 wpi represented approximately 10% of sequences at this timepoint. The two viral populations each represented 50% of sequences at 13 wpi.

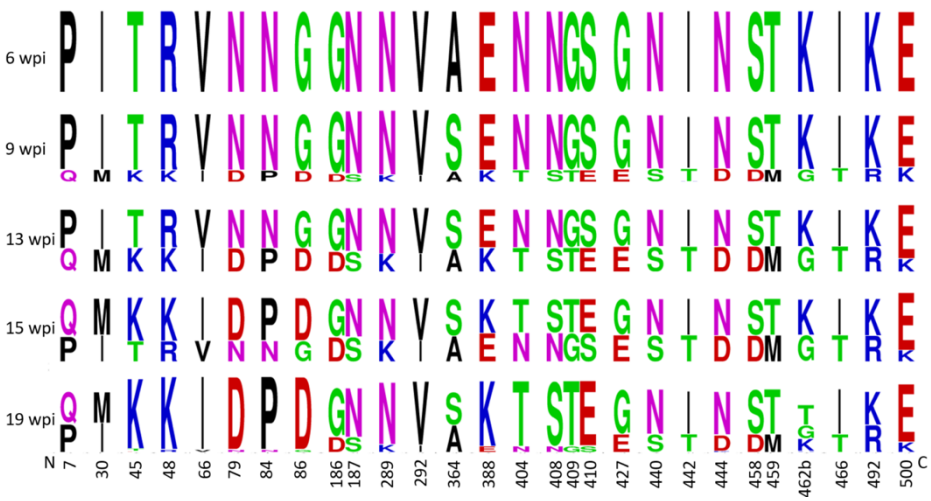

Figure 3.6: Two distinct viral populations observed through CAP177 deep *env* sequencing
Single mutations representing the two genetically distinct HIV-1 variants. Residues shown represent single positions at which the new HIV-1 variant, responsible for superinfection, collectively differed from the initial HIV-1 variant. Residues are numbered according to the HXB2 reference sequence.

At 19 wpi a new dominant recombinant group emerged that had amino acids previously observed in both viral populations (**Figure 3.6**). This new viral population shared AA in the C1, V4 and C5 regions with the second viral variant and amino acids seen in the first viral variant within the V2 and C2 regions. There were no mutations observed in areas known to impact V3-glycan bNAb binding, and so we did not pursue CAP177 further. The UCA for the CAP177 V3-glycan-specific, broad lineage in CAP177 has not been isolated and as a result the identification of a bNAb-initiating virus was expected to be challenging. Further studies will have to be conducted with respect to isolating bNAbs and the UCA of the CAP177 lineage to determine which timepoints to focus on for identification of the CAP177 bNAb-initiating virus.

3.3 Design and characterisation of CAP255 bNAb-initiating soluble trimers for use as V3-glycan germline-targeting candidates

Previous sequencing data from CAP255 was used to select the potential bNAb-initiating Env prior to the start of this study (Kitchin *et al.*, unpublished data). The bNAb-initiating Env was selected due to a glycan hole (Figure 3.7), a result of a lack of PNGs at positions N442 and N448, which have been shown to open up the V3-glycan site allowing antibodies to bind more readily (Li *et al.*, 2008; Canducci *et al.*, 2009). This motif was present across limited early timepoints before the PNGs appeared in later sequences from 17 wpi.

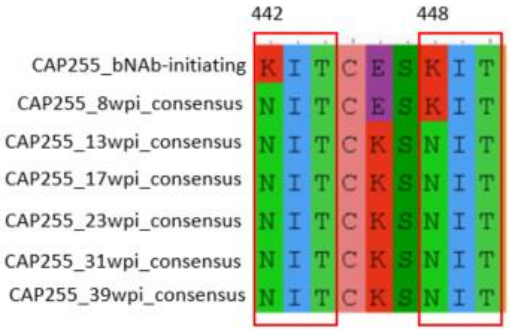

Figure 3.7: CAP255 bNAb-initiating Env

CAP255 consensus sequences at each timepoint aligned to the CAP255 bNAb-initiating Env, highlighting the missing glycans at position 442 and 448, indicated with the red blocks.

SOSIP modifications previously described by Sanders *et al.* (2013) were introduced to the CAP255.WT sequence, an early CAP255 *env* sequence which was identified at 8 wpi. A total of five additional CAP255 SOSIP.664 trimers were designed to include varying mutations, with the overarching aim of enhancing access by V3-glycan bNAb precursors (**Table 3.3**). CAP255.GT0 is the putative bNAb-initiating virus sequence, which contained N442A and N448A substitutions, removing PNGs from these sites. CAP255.GT0 was used as a basis for the design of the other four germline-targeting modified trimers (**Figure 3.8A & B**). The CAP255.GT1 trimer had additional mutations added to improve trimer stability and binding affinity of V3-glycan-specific antibodies by removing glycosylation sites and reducing epitope occlusion as shown by Steichen *et al.* (2016). The mutations introduced into CAP255.GT2 were informed by Steichen *et al.* (2019) and focused on mutations in the V1 and V4 regions, removing glycosylation sites and improving binding affinity and trimer stability. CAP255.GT3 had mutations introduced to reduce steric hindrance and improve binding affinity (Saunders *et al.*, 2019) along with mutations to reduce off-target responses (Escolano *et al.*, 2019). The V1 loop of CAP255.GT4 was completely removed in order to remove steric hindrance and a mutation was introduced in the C3 region to reduce off target responses (Escolano *et al.*, 2019). All mutations introduced into the four modified trimers were to enhance binding of V3-glycan UCAs and intermediates and increase the stability of the trimer (**Table 3.3**).

The expressed CAP255 trimers were purified using SEC which showed elution of trimer peaks between 58 and 64 minutes (**Figure 3.8C**). There was little dimer/monomer formation although aggregation was seen at approximately 55 minutes. CAP255.GT2 had the highest yield, followed by the CAP255.GT0 and CAP255.WT trimers. The CAP255.WT trimer expressed well with only the original stabilising SOSIP.664 mutations and no further stabilisation mutations. CAP255.GT1 did not express well and had to be transfected multiple times for use

in downstream analysis and characterisation. The CAP255.GT4 trimer did not form a well-defined peak when purified but had a high enough yield to still be collected for downstream experiments.

Table 3.3: CAP255 trimer modifications

Trimer name	Mutations introduced	Rationale	Reference
CAP255 GT0	N442A, N448A	Identified in sequencing data as potential bNAb initiating virus motif.	This study
CAP255 GT1	N442A, N448A	Identified in sequencing data as potential bNAb initiating virus motif.	This study
	N135A, N139I	Removal of glycosylation sites.	(Steichen et al., 2016)
	A134Y, S136P, N137F, T138L, V140N, T320F, Q328M	Improve binding affinity.	
	E290A, S363Q	Improve yield and stability.	
CAP255 GT2	N442A, N448A	Identified in sequencing data as potential bNAb initiating virus motif.	This study
	N135A, N139R	Removal of glycosylation sites.	(Steichen et al., 2019)
	A134Y, S136P, N137K, T138L, V140S, T141M, T320F, I323V, I324L, I326V, Q328M	Improve binding affinity.	
	I414L, T415I	Improve yield and stability.	
CAP255 GT3	N442A, N448A	Identified in sequencing data as potential bNAb initiating virus motif.	This study
	N133D	Remove steric hindrance.	(Saunders et al., 2019)
	N156Q	Improve binding affinity.	
	R344N	Reduce off target responses.	(Escolano et al., 2019)
CAP255 GT4	N442A, N448A	Identified in sequencing data as potential bNAb initiating virus motif.	This study
	R344N	Reduce off target responses.	(Escolano et al., 2019)
	131-157 deletion	Remove steric hindrance.	

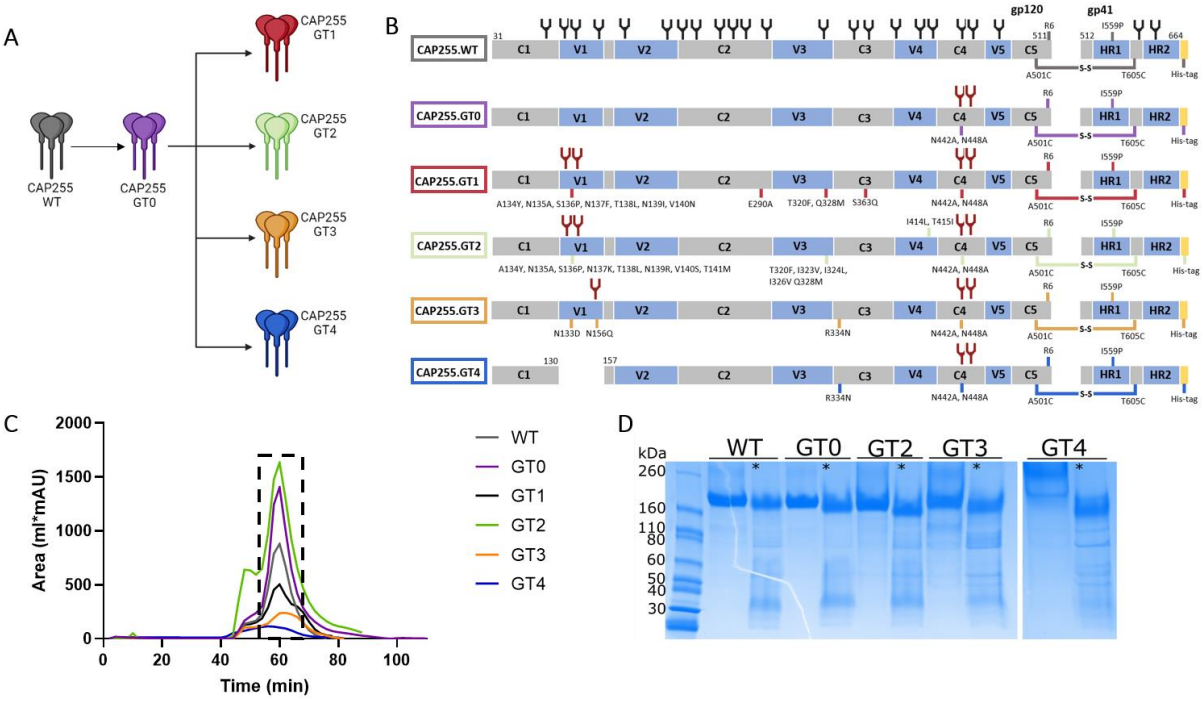

Figure 3.8: Expression of CAP255.GT0, WT and modified trimers
(A) Schematic diagram of the six selected trimers that were expressed. **(B)** Linear representation of the six CAP255 SOSIP.664 gp140 Env proteins. Modifications compared to a WT gp160 sequence are highlighted. The transmembrane and cytoplasmic tail domains were deleted to result in a soluble gp140 and a His-tag was added at position 664. A disulphide bond was formed between the gp120 and $gp41_{ECTO}$ by introducing the A501C and T605C substitutions. An I559P mutation was introduced to promote trimerisation and an optimal cleavage site, RRRRRR, was added. These mutations were introduced based on previous research by Sanders *et al.* (2013). Env sub-domains are represented as follows: conserved domains in light grey (C1-C5), variable domains (V1-V5) and heptad repeats (HR1-HR2) in light blue. PNGs present on the WT Env are indicted in black, loss of PNGs, relative to those present on the WT Env, due to mutations added are shown in red. The N442A and N448A mutations were added to all GT trimers as these are the signatures of the putative bNAb-initiating Env. Other mutations introduced were based on previous studies summarised in **Table 3.3**. PNGs were predicted using NetNGlyc version 1.0 (Gupta and Brunak, 2001). **(C)** The SEC profiles of all nickel-affinity purified CAP255 trimers. The size at which trimer peaks were formed are indicated by the dotted black box. A Superdex 200 HiLoad 16/600 column was used. **(D)** The CAP255 trimers were run on a SDS-PAGE using a 14% pre-cast gel under reducing (*) and non-reducing conditions. β-ME was used as the reducing agent and the gel was run at 160 V for approximately 1 hour. The Novex Sharp Pre-stained Protein Standard ladder was used.

The collected trimer fractions were run through a negative selection column and SDS-PAGE was used to confirm that the collected trimer fractions from the resulting flow through ran to the correct size under reducing and non-reducing conditions (**Figure 3.8D**). Under non-reducing conditions, a single band was present at approximately 140 kDa for CAP255.WT, CAP255.GT0 and CAP255.GT2 which corresponded with the correct size of a modified gp140 protomer (**Figure 3.8D)**. There were some additional bands visible for CAP255.GT3 and

CAP255.GT4 which indicated degradation of the trimer. Under reducing conditions (indicated with an *), all trimers except CAP255.GT4 showed two bands at approximately 120 kDa and just above 30 kDa which corresponded with the gp120 and truncated gp41 subunits, respectively. This indicated successful cleaving of the trimers.

To determine if the trimers were in the optimal, closed conformation, trimer binding ELISAs were performed using conformation-dependent antibodies, non-neutralising antibodies and gp120-specific bNAbs. CAP255.C5 and CAP255.G3, bNAbs isolated from the CAP255 participant at 149 wpi, were used as the positive controls as they are known to bind to CAP255 trimers at the V3-glycan site. Two bNAbs were included as negative controls, 10E8, which binds to the MPER region that is absent in SOSIP trimers, and PG9, which requires a glycan at position N160 that is absent in CAP255 trimers. Conformational antibodies, 35O22 and PGT151, were used to determine if the trimers were folded correctly since they only bind to intact, fully formed trimers. The conformational antibodies were able to bind to CAP255.WT and significantly bind CAP255.GT0 (p value < 0.05) compared to the negative controls but these trimers were not bound by the non-neutralising antibodies, F105, 17b and 447-52D, which bind to regions that are occluded in the correctly formed, closed trimer conformation (**Figure 3.9A & B**). There was a significant difference between the binding of mature bNAbs and negative control bNAbs (p value < 0.001 and p value < 0.0001). The conformational antibodies were able to bind to CAP255.GT2, however the binding was not significant compared to the binding of negative control bNAbs, and most non-neutralising antibodies did not bind, with the exception of 17b which showed slight binding to the trimer (**Figure 3.9D**). Lower binding of conformational antibodies with no significant difference compared to the negative controls and no binding of non-neutralising antibodies was observed for CAP255.GT1 (**Figure 3.9C**). The CAP255 bNAbs bound all trimers except CAP255.GT3 and CAP255.GT4 (**Figure 3.9E & F**). There was also no binding of conformational antibodies to CAP255.GT3 and CAP255.GT4 **(Figure 3.9E & F**).

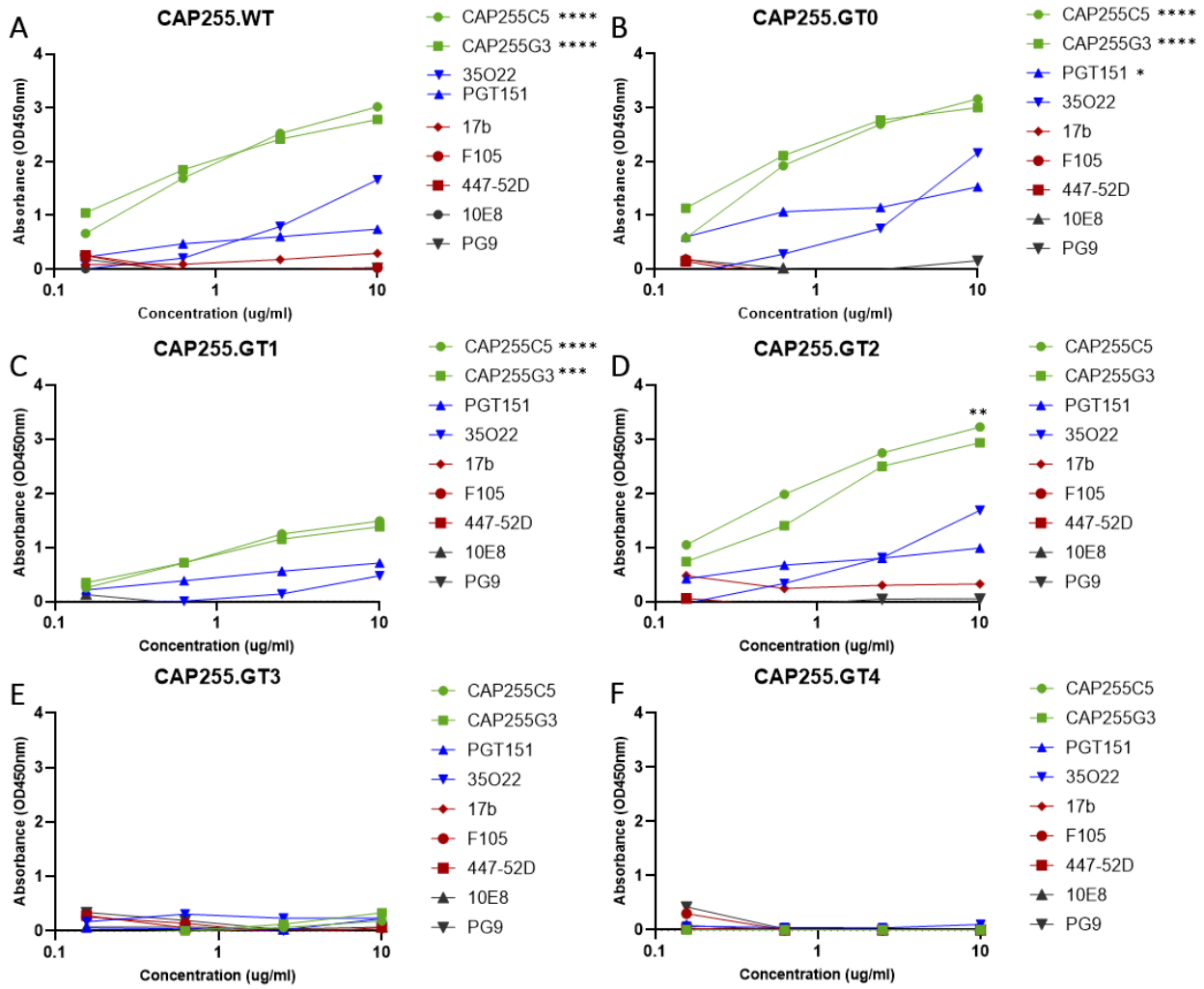

Figure 3.9: Characterisation of modified CAP255 trimers by ELISA
CAP255 trimers were coated on nickel plates. All CAP255 trimers were run against both CAP255.C5 and CAP255.G3 bNAbs (green), conformational antibodies, PGT151 & 35O22 (blue), non-neutralising antibodies, 17b, F105 & 447-52D (red) and negative controls, 10E8 & PG9 (black). Starting concentration for all antibodies was 10 µg/ml. Representative binding curves of above mentioned antibodies to purified **(A)** CAP255.WT, **(B)** CAP255.GT0, **(C)** CAP255.GT1, **(D)** CAP255.GT2, **(E)** CAP255.GT3 and **(F)** CAP255.GT4 SOSIP.664 His-tagged trimers. Significant differences between binding of antibodies compared to negative control antibodies are indicated with * (p <0.05), *** (p <0.001) and **** (p <0.0001). Binding data was analysed using GraphPad Prism (version 9.3.1.471).

From the binding data as well as the SEC profiles we determined that CAP255.GT2 is the most stable and well-formed modified trimer out of the six CAP255 trimers that had been designed. As a result we continued with this trimer as a possible germline-targeting immunogen candidate.

Next, ELISAs were performed with multiple V3-glycan UCAs from various V3-glycan bNAb lineages, including the two potential CAP255 UCAs, CAP255 UCA01 and CAP255 UCA02 (Kitchin *et al.*, unpublished data). The CAP314 UCA (Scheepers *et al.*, unpublished data), PCDN

UCA (MacLeod *et al.*, 2016) and PC39 UCA (Landais *et al.*, unpublished data) were also included as these UCAs all matured into bNAbs directed to the V3-glycan site. All five UCAs displayed only negligible binding to five of the six trimers (**Figure 3.10A-E**). The CAP255.GT2 trimer was the exception, where very slight binding was observed by the CAP255 UCA01 (**Figure 3.10D**). However, none of the differences in binding between UCAs were statistically significant for any of the trimers ($p > 0.05$ for all trimers).

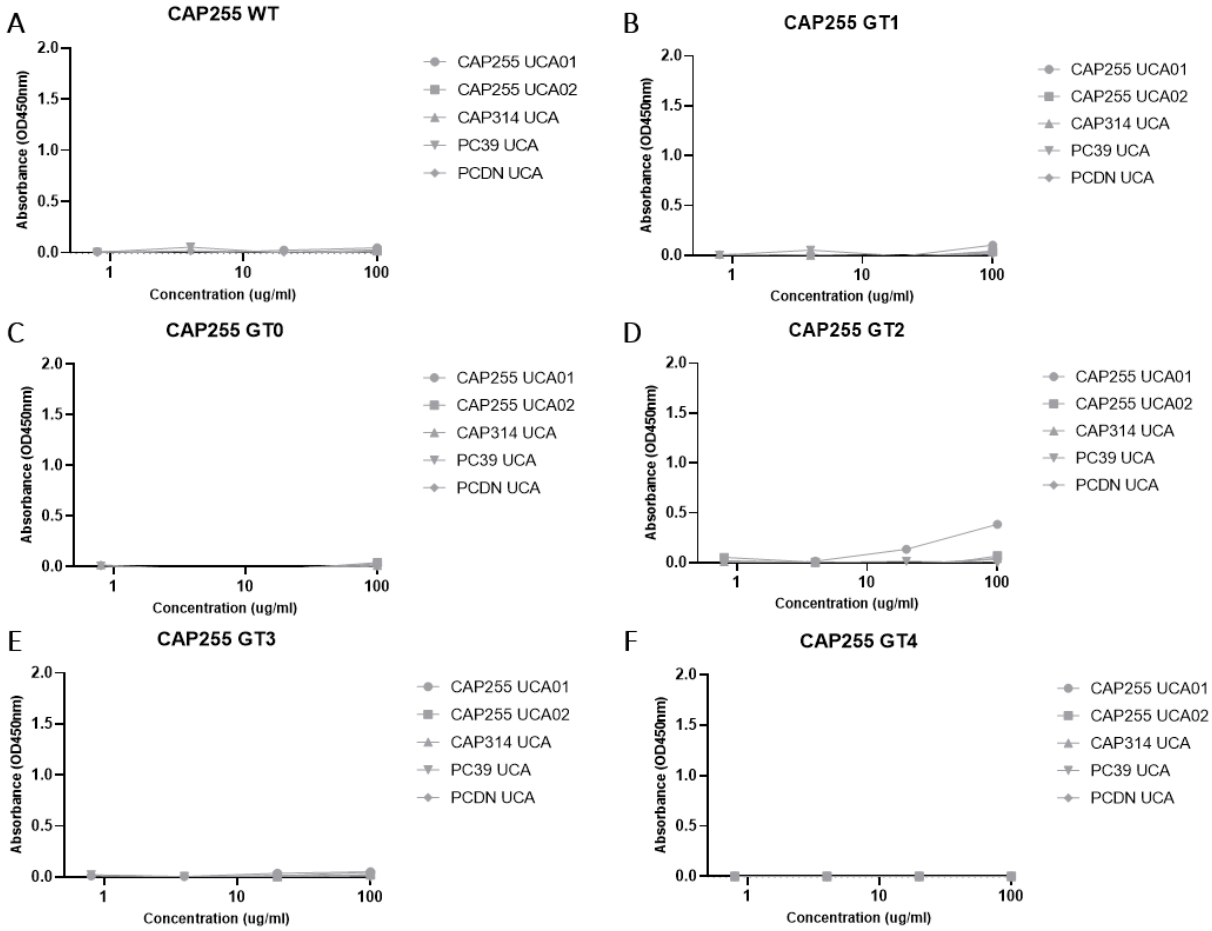

Figure 3.10: CAP255 SOSIP trimers have weak affinity for V3-glycan targeting UCAs
CAP255 trimers were coated on nickel plates. All CAP255 trimers were ran against CAP255 UCA01, CAP255 UCA02, CAP314 UCA, PC39 UCA and PCDN UCA at a starting concentration of 100 µg/ml. Representative binding curves of above mentioned antibodies to purified **(A)** CAP255.WT, **(B)** CAP255.GT0, **(C)** CAP255.GT1, **(D)** CAP255.GT2, **(E)** CAP255.GT3 and **(F)** CAP255.GT4 SOSIP.664 His-tagged trimers. No significant differences were observed between UCA antibodies ($p > 0.05$). Binding data was analysed using GraphPad Prism (version 9.3.1.471).

In order to confirm the presence of any potential binding interaction between the CAP255 trimers and UCAs, SPR was performed on selected trimers and the CAP255 UCAs. UCAs are known to bind weakly to their targets (Wu *et al.*, 2011) and SPR is a highly sensitive technique and thus able to detect very weak binding interactions (Jönsson *et al.*, 1991). As a positive

control, CAP255.GT2 (modified germline-targeting immunogen) (**Figure 3.11A**) and CAP255.GT0 (putative bNAb-initiating virus sequence) (**Figure 3.11B**) were tested for binding against the mature CAP255.C5 bNAb. CAP255.C5 showed strong affinity for both trimers with KD values of 4.51×10^{-9} and 2.4×10^{-8}, respectively. Both trimers were also tested against CAP255 UCA01 which showed no binding activity towards either one of the CAP255 trimers (**Figure 3.11C & D**).

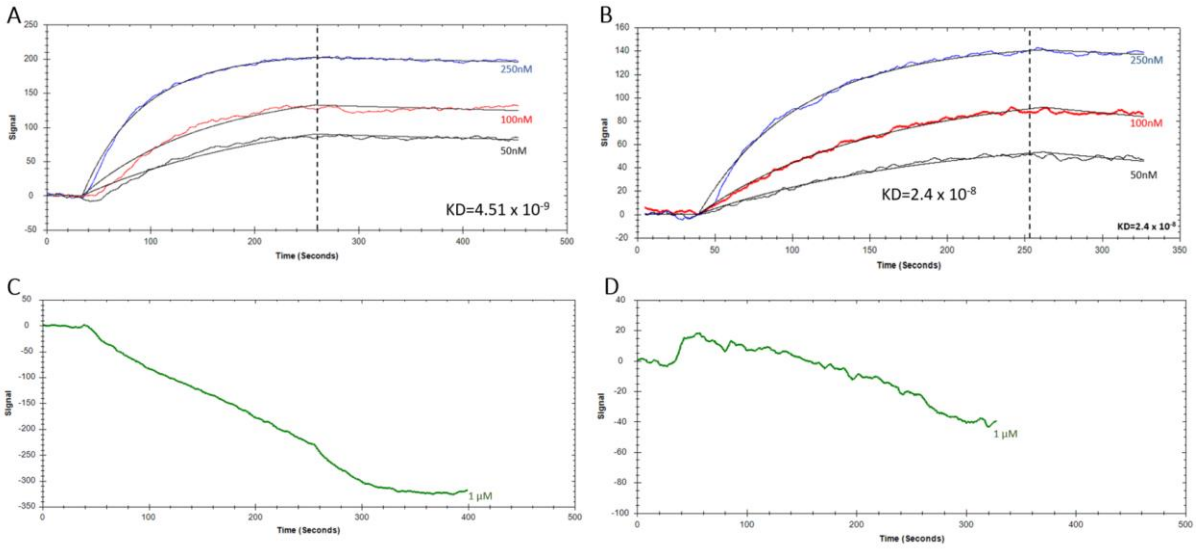

Figure 3.11: Surface plasmon resonance of CAP255.GT2 and CAP255.GT0 against CAP255.C5 and CAP255 UCA01
(A & B) Representative SPR curve showing the association from 0 seconds to 250 seconds followed by dissociation at 250 seconds (indicated by dotted line) of the (**A**) CAP255.GT2 trimer and (**B**) CAP255.GT0 trimer to the CAP255.C5 bNAb. Golden chips were loaded with trimer protein of varying concentrations (50 nM, 100nM and 250 nM). The same experiment was conducted with (**C**) CAP255.GT2 and (**D**) CAP255.GT0 and no binding to CAP255 UCA01 at the highest trimer protein concentration (1 µM) was observed. Each SPR experiment was repeated in duplicate.

As the CAP255.GT2 trimer was the best candidate to move forward with, we used DSC to determine the thermal stability of the purified trimer compared to that of the purified CAP255.WT and CAP255.GT0 trimers. The DSC profile for CAP255.WT had a melting temperature (Tm) of 58.6°C (**Figure 3.12A**). This was a similar denaturation profile to the CAP255.GT0 trimer which had a Tm of 58°C (**Figure 3.12B**). CAP255.GT2 had a higher Tm of 61°C compared to the CAP255.WT and CAP255.GT0 trimers (**Figure 3.12C**). These data suggest that CAP255.GT2 is a more stable trimer compared to the other two CAP255 trimers tested. All CAP255 SOSIP trimers tested had unfolding events occurring in two phases; the separation of the trimer protomers represented by the first peak followed by the degradation of monomers, represented by the second peak (**Figure 3.12**).

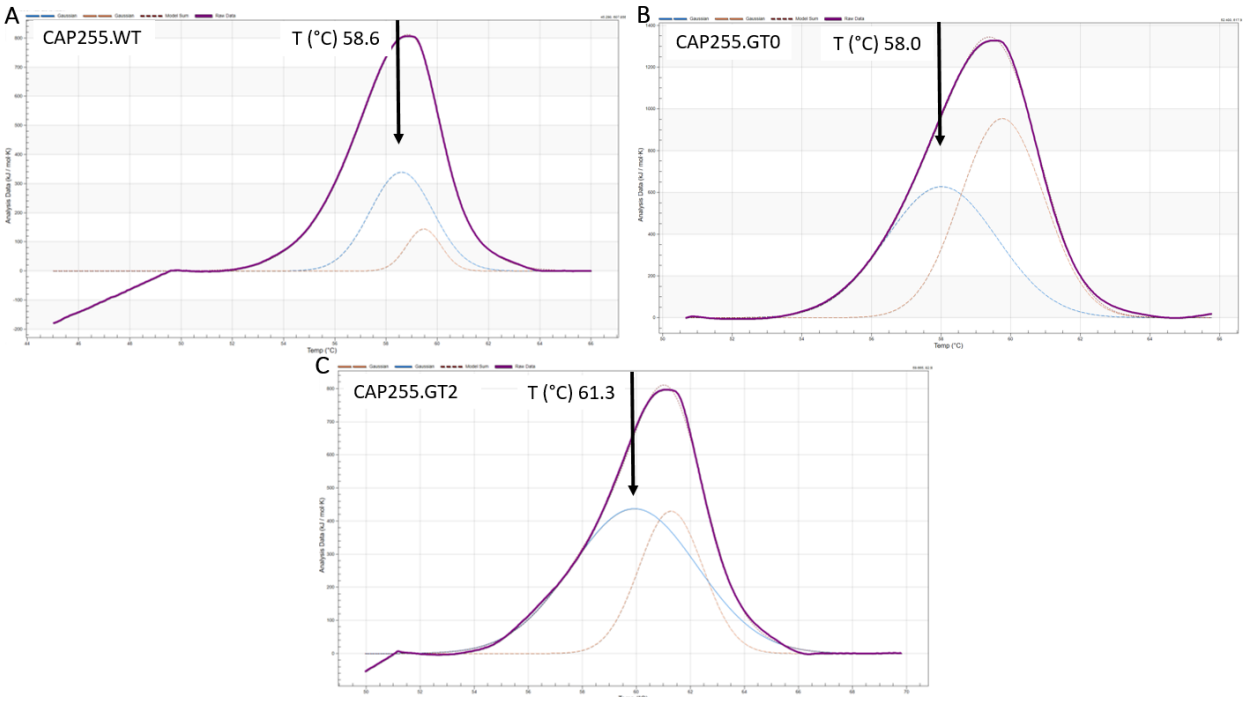

Figure 3.12: CAP255.GT2 is the most thermal stable CAP255 trimer
DSC analysis of (**A**) purified CAP255.WT, (**B**) CAP255.GT0 and (**C**) CAP255.GT2 trimers. The raw data (purple) was fitted using a Gaussian model and showed two events happening. The first event as the trimer broke into monomers (blue) and the degradation of the monomers (orange) as the second event. The raw data of the trimers were fitted and analysed using NanoAnalyze.

3.4 Neutralisation of putative CAP255 bNAb-initiating pseudoviruses by CAP255 early intermediates

To test whether various bNAbs and intermediate antibodies, which were isolated at various timepoints between UCA emergence and development of breadth, would neutralise the selected CAP255 Envs, pseudovirus-based neutralisation assays were performed. The UCAs of various broad V3-glycan lineages; CAP255 UCA01, CAP255 UCA02, CAP314 UCA, PCDN UCA and PC39 UCA, were tested for neutralisation against CAP255.WT (early virus sequence), CAP255.GT0 (putative bNAb-initiating virus sequence) and CAP255.GT2 (modified germline-targeting immunogen sequence) pseudoviruses (**Figure 3.13A**). There was no neutralisation observed by UCAs to any of the three CAP255 trimers tested.

While early CAP255 intermediates isolated at 39 wpi (CAP255 Int A and Int B) were not able to neutralise CAP255.GT0, CAP255 Int B was able to neutralise CAP255.WT and CAP255.GT2 **(Figure 3.13B)**.

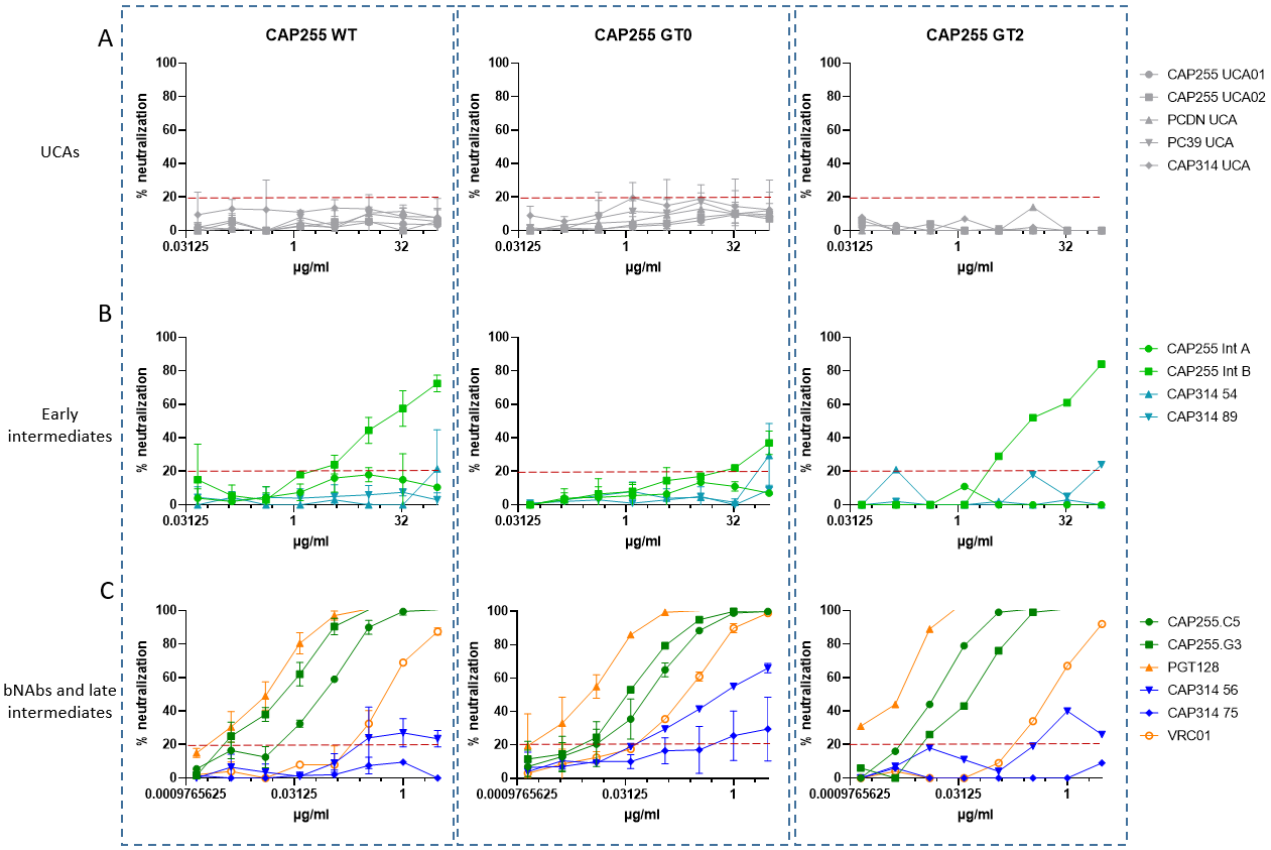

Figure 3.13: CAP255 early intermediates are able to neutralise CAP255 pseudoviruses
Representative neutralisation curves of CAP255.WT, CAP255.GT0 and CAP255.GT2 pseudoviruses ran against: **(A)** CAP255 UCA01, CAP255 UCA02, PCDN UCA, PC39 UCA and CAP314 UCA at a starting concentration of 100 µg/ml. **(B)** Early CAP255 intermediates, CAP255 Int A and CAP255 Int B represented in light green, and early CAP314 intermediates, CAP314 54 and CAP314 89 represented in light blue. The starting concentration of all intermediate antibodies was 100 µg/ml. **(C)** Mature CAP255 bNAbs, CAP255.C5 and CAP255.G3 (dark green), and late CAP314 intermediates, CAP314 56 and CAP314 75 (dark blue). VRC01 and PGT128 were included as positive controls (yellow). The starting concentration of all bNAbs and late intermediates was 3 µg/ml. All assays were done in duplicate except for the CAP255.GT2 pseudovirus assay since the virus had trouble growing. Error bars on the figure represents standard deviation between repeats. Results were analysed using GraphPad Prism (version 9.3.1.471).

The CAP314 intermediates, CAP314 54 and CAP314 89, isolated at 54 wpi did not show any neutralisation activity against any CAP255 pseudoviruses. The CAP255 mature bNAbs were able to neutralise all CAP255 pseudoviruses with high potency (**Figure 3.13C**). Both VRC01 and PGT128, included as positive controls, showed high levels of neutralisation towards CAP255.WT, CAP255.GT0 and CAP255.GT2. The CAP314 late intermediates, CAP314 56 and CAP314 75, isolated at 90 and 115 wpi respectively, were able to neutralise CAP255.GT0. CAP314 56 showed a 3-fold higher potency when neutralising CAP255.GT0 compared to CAP255.WT and CAP255.GT2.

3.5 Design and characterisation of CAP314 bNAb-initiating soluble trimers

The PacBio sequencing data generated in this study was used to identify potential CAP314 bNAb-initiating *envs* (**Figure 3.3**). These *envs* consisted of three CAP314 representative sequences with decreasing V1 loop lengths of 23AA (CAP314.GTL), 17AA (CAP314.GTM) and 15AA (CAP314.GTS). Informed by these data and previous CAP255 trimer experiments which showed that the modified CAP255.GT2 trimer was the most stable and well-formed trimer, CAP314 trimer modifications were selected (**Table 3.4**). Some CAP255.GT2 modifications could not be introduced to the CAP314.GTS2 and CAP314.GTM2 sequences due to their shortened V1 loops and thus lack of amino acids at positions 136-143 and 136-141, respectively, in the V1 loop.

Table 3.4: CAP314 trimer modifications

Trimer name	Mutations introduced	Rationale	Reference
CAP314.GTS, CAP314.GTM & CAP314.GTL	A501C, T605C	Disulphide bond formation.	(Sanders *et al.*, 2013)
	I559P	Promote trimerisation.	
	REKR508-511RRRRRR	For cleavage enhancement.	
	Stop codon at 664	Improve solubility	
CAP314.GTS2	V134Y, T135A, T320F, I323V, I324L, I326V, Q328M	Improve binding affinity.	(Steichen et al., 2019)
	T415I	Improve yield and stability.	
CAP314.GTM2	V134Y, T135A, T320F, I323V, I324L, I326V, Q328M	Improve binding affinity.	
	T415I	Improve yield and stability.	
CAP314.GTL2	V134Y, T135A, G136P, N137K, G138L, T139R, T141M, T320F, I323V, I324L, I326V, Q328M	Improve binding affinity.	
	T415I	Improve yield and stability.	

Altogether, six CAP314 trimers were designed which represented three unmodified putative bNAb-initiating *envs* and three modified *envs* based on the selected putative bNAb-initiating *envs* as well as CAP255.GT2 (**Figure 3.14A & B**). The six trimers were expressed and their SEC profiles showed high levels of aggregation as seen at 45 minutes (**Figure 3.14C**). The trimer

peaks were poorly defined for most of the proteins tested, seen at 53 minutes. The peaks seen at 62 minutes representing monomer/dimer formation were also high. The only trimer with a well-defined trimer peak was CAP314.GTS2 (**Figure 3.14C, red**).

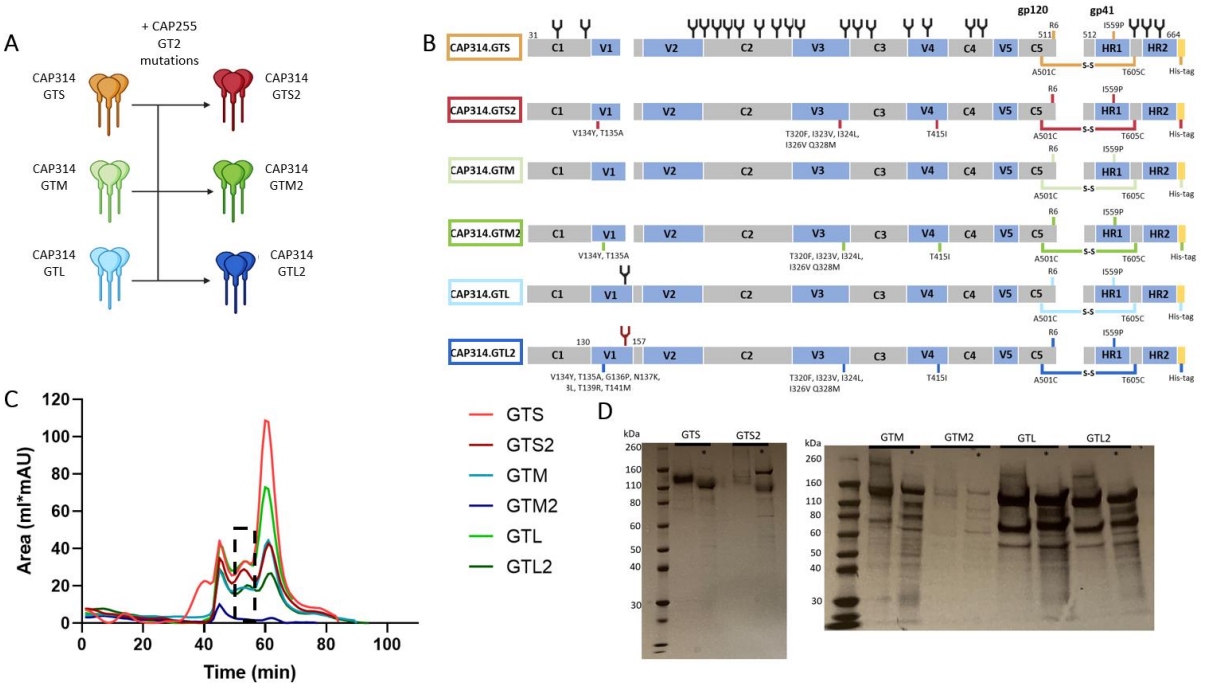

Figure 3.74: Expression of CAP314 putative bNAb-initiating and modified trimers
(A) A schematic representing the three WT putative bNAb-initiating CAP314 trimers and the three GT2 modified CAP314 trimers based on the WT *envs*. **(B)** Linear representation of the six CAP314 SOSIP.664 gp140 Env proteins. Modifications compared to a WT gp160 sequence are highlighted. The transmembrane and cytoplasmic tail domains were deleted to result in a soluble gp140 and a His-tag was added at position 664. A disulphide bond was formed between the gp120 and gp41$_{ECTO}$ by introducing the A501C and T605C substitutions. An I559P mutation was introduced to promote trimerisation and an optimal cleavage site, RRRRRR, was added. These mutations were introduced based on previous research by Sanders *et al.* (2013). Env sub-domains are represented as follows: conserved domains in light grey (C1-C5), variable domains (V1-V5) and heptad repeats (HR1-HR2) in light blue. PNGs present on WT sequences are shown in black, PNGs lost due to added mutations, relative to the WT sequence, are indicated in red. Other mutations introduced were based on previous studies summarized in **Table 3.4**. PNGs were predicted using NetNGlyc (version 1.0) (Gupta and Brunak, 2001). **(C)** The SEC profiles of all nickel-affinity purified CAP314 trimers. The size at which trimer peaks were formed are indicated by the dotted black box. A Superdex 200 HiLoad 16/600 column was used. **(D)** The CAP314 trimers were run on a SDS-PAGE using a 14% pre-cast gel under reducing (*) and non-reducing conditions. β-ME was used as the reducing agent and the gel was run at 160 V for approximately 1 hour. A Novex Sharp Pre-stained Protein Standard ladder was used.

All CAP314 trimers had very low yields. For example, compared to CAP255.GT2, CAP314.GTS2 had an 85-fold lower yield. As the yields of CAP314 trimers were so low, they were not run through a negative selection column after SEC. All of the trimers ran to the correct size (140 kDa) in non-reducing conditions on SDS-PAGE (**Figure 3.14D**). CAP314.GTS, CAP314.GTS2 and

CAP314.GTM2 were the only trimers that showed a singular band of this size suggesting that the other trimers were not stable and correctly formed. In reducing conditions CAP314.GTM, CAP314.GTM2, CAP314.GTL and CAP314.GTL2 all had multiple bands of varying sizes. CAP314.GTS did not have a gp41 band visible and CAP314.GTS2 had very faint bands visible of the correct size and less degradation (**Figure 3.14D).** The SEC profile and SDS-PAGE analysis of the modified CAP314 trimers suggest the *env* sequences from participant CAP314 do not readily form stable trimers and would potentially need to have more stabilising mutations introduced (**Figure 3.14C & D**).

The conformation of trimers were validated by performing trimer binding ELISAs. Due to CAP314 trimers expressing at such low levels, ELISAs were only performed with CAP314.GTS2, CAP314.GTL and CAP314.GTL2 trimers. The positive controls, VRC01 and PGT128, bound to all three CAP314 trimers. There was a significant difference observed in binding between CAP314 late intermediates compared to non-neutralising antibodies in all three trimers (p <0.05 – p <0.0001) (**Figure 3.15**). PGT151, a conformational antibody, showed very little binding to all three trimers. Another conformational antibody, 35O22, showed high binding to all three. CAP314.GTS2 was not bound by any non-neutralising antibodies but was also not bound by the early CAP314 intermediate, CAP314 54 from 54 wpi (**Figure 3.15A**). CAP314.GTL was bound by two non-neutralising antibodies and all CAP314 intermediates (**Figure 3.15B**). CAP314.GTL2 was slightly bound by one non-neutralising antibody but was bound by all three CAP314 intermediates (**Figure 3.15C**). No binding of UCAs to any CAP314 trimers were observed (**Figure 3.15D & F**).

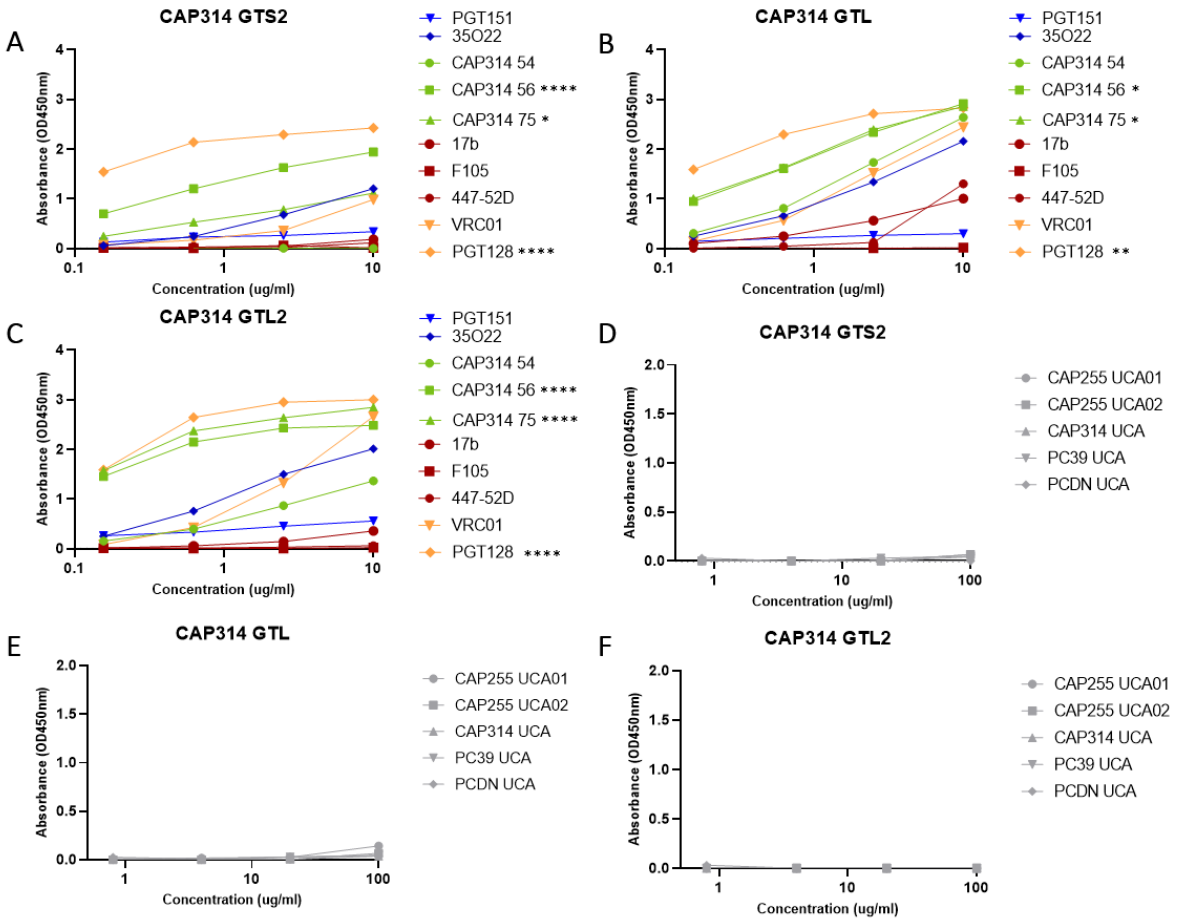

Figure 3.15: Characterisation of modified CAP314 trimers by ELISA
CAP314 trimers were coated on nickel plates. Three CAP314 trimers were run against CAP314 intermediates, CAP314 54, CAP314 56 and CAP314 75 (green), conformational antibodies, PGT151 & 35O22 (blue), non-neutralising antibodies, 17b, F105 & 447-52D (red) and positive controls, VRC01 and PGT128, (yellow). Starting concentration for all antibodies was 10 µg/ml. Representative binding curves of above mentioned antibodies to purified **(A)** CAP314.GTS2, **(B)** CAP314.GTL and **(C)** CAP314.GTL2 SOSIP.644 His-tagged trimers. These three CAP314 trimers were also run against CAP255 UCA01, CAP255 UCA02, CAP314 UCA, PC39 UCA and PCDN UCA at a starting concentration of 100 µg/ml. Representative binding curves of the above mentioned antibodies to purified **(D)** CAP314.GTS2, **(E)** CAP314.GTL and **(F)** CAP314.GTL2 SOSIP.664 His-tagged trimers. Statistical significance in binding between antibodies compared to non-neutralising antibody binding for panels (A), (B) and (C) are indicated with * (p <0.05), ** (p <0.01) and **** (p <0.0001). No statistically significant differences between the binding of UCAs to any of the trimers were observed. Binding data was analysed using GraphPad Prism (version 9.3.1.471).

3.6 Neutralisation of putative CAP314 bNAb-initiating pseudoviruses by CAP314 intermediates.

The same five UCAs (CAP255 UCA01, CAP255 UCA02, CAP314 UCA, PCDN UCA and PC39 UCA) that were tested against CAP255 pseudoviruses were tested against CAP314.GTS2 (putative

bNAb-initiating Env with CAP255.GT2 mutations introduced) **(Figure 3.16A)**. The Envs with the shortest V1 loops were focused on in neutralisation assays since they were thought to have the highest likelihood of being the putative bNAb-initiating viruses. The CAP314.GTS pseudovirus failed to express and thus only CAP314.GTS2 was tested in neutralisation. There was no neutralisation of CAP314.GTS2 observed when tested against the UCAs The only early intermediate that was able to neutralise CAP314.GTS2 was CAP314 54 **(Figure 3.16B)**. Both CAP255 early intermediates showed no neutralisation. The positive controls, VRC01 and PGT128, were able to neutralise CAP314.GTS2 **(Figure 3.16C)**. Both CAP314 late intermediates, CAP314 56 (isolated at 90 wpi) and CAP314 75 (isolated at 115 wpi), showed potent neutralisation activity towards CAP314.GTS2. CAP314 56 had a 20-fold increase in neutralisation potency towards CAP314.GTS2 compared to CAP255.GT0. There was a 5-10-fold drop in potency of the CAP255 bNAbs, CAP255.C5 and CAP255.G3, towards CAP314.GTS2 compared to the response observed against all three CAP255 pseudoviruses.

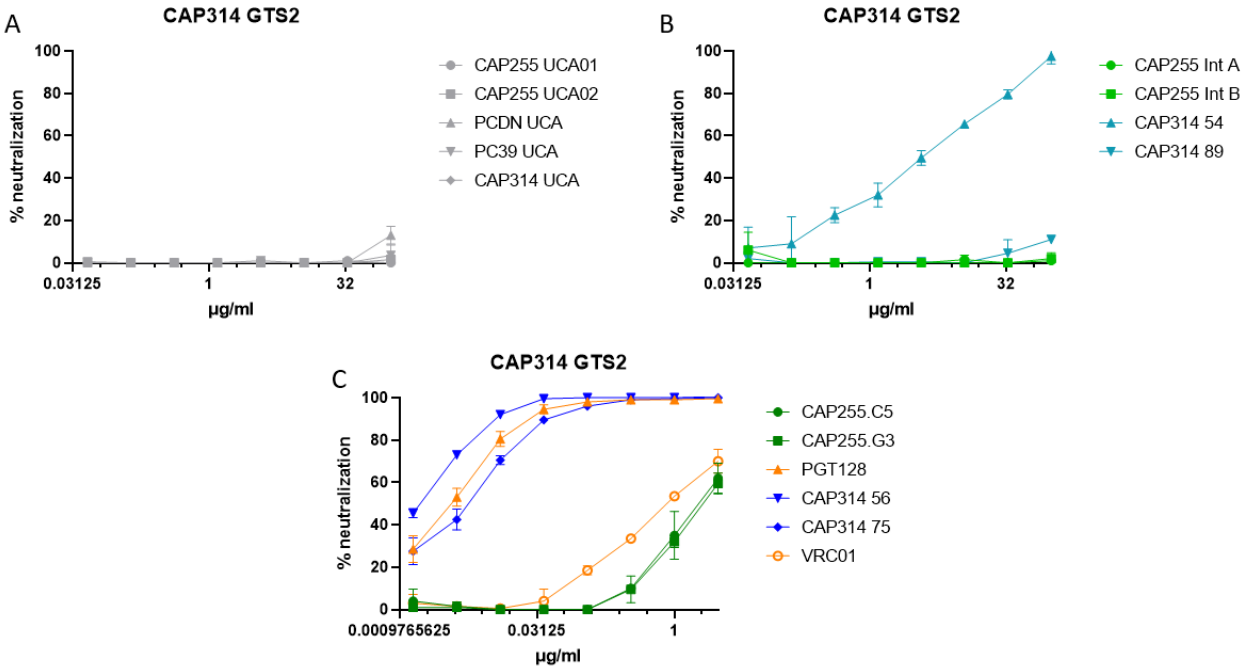

Figure 3.16: CAP314 early and late intermediates are able to neutralise autologous pseudoviruses
The neutralisation of the CAP314.GTS2 pseudoviruses represented as neutralisation curves. **(A)** CAP314.GTS2 was run against CAP255 UCA01, CAP255.UCA02, CAP314 UCA, PCDN UCA and PC39 UCA. All UCAs had a starting concentration of 100 µg/ml. **(B)** The early intermediates CAP255 Int A, CAP255 Int B (light green), CAP314 54 and CAP314 89 (light blue) had a starting concentration of 100 µg/ml. **(C)** The late CAP314 intermediates, CAP314 56 and CAP314 75 (dark blue) and CAP255 bNAbs, CAP255.C5 and CAP255.G3 (dark blue) had a starting concentration of 3 µg/ml together with the positive controls, PGT128 and VRC01 (yellow). Error bars on the figure represents standard deviation between repeats. Results were analysed using GraphPad Prism (version 9.3.1.471).

4 Discussion

Limited success in controlling the spread of HIV-1 has been achieved through current HIV-1 prevention methods, however an effective HIV-1 vaccine, capable of eliciting bNAbs will be needed to end the pandemic. Due to the dominance of subtype C and the tendency of V3-glycan antibody responses preferentially developing in subtype C HIV-1 infected individuals (Landais *et al.*, 2016), we aimed to develop a V3-glycan targeting immunogen based on a naturally occurring subtype C Env. HIV-1 infected donors, CAP255, CAP314 and CAP177, were selected for this study because they developed broad responses to the V3-glycan epitope. Both CAP255 and CAP314 have been shown to utilise IGHV4-34 genes for V3-glycan directed bNAbs while the gene usage of CAP177 is unknown, since no bNAbs have yet been isolated. The V3-glycan site is thought to be the most commonly targeted bNAb epitope due to the multiple angles of approach utilised by antibodies which target this region and the ability of antibodies to use a variety of glycans at this site. The levels of SHM required for the acquisition of breadth in V3-glycan bNAbs are relatively low when compared to CD4bs bNAbs (Simonich *et al.*, 2016). This indicates that elicitation of V3-glycan precursors through germline-targeting immunogens represents a more achievable target. Although IGHV4 is a commonly used germline gene among V3-glycan directed bNAbs (Moyo *et al.*, 2020), no conserved germline encoded motifs were present across all mature V3-glycan bNAbs that had been identified in this study after undergoing affinity maturation. This is different to the IGHV1-2 enrichment that is required to result in VRC01-like bNAbs (Jardine et al., 2013) and shows conservation of $W50_{HC}$, $N58_{HC}$ and $R71_{HC}$ germline encoded contact residues (Scharf *et al.*, 2013; Zhou *et al.*, 2015). Structural studies of IGVH4 V3-glycan bNAbs bound to their epitopes would be needed to identify any commonalities in contact residues that may be germline encoded. No germline encoded commonalities were readily apparent from the sequence analysis of the 31 bNAbs that were examined in this study.

The first aim of this project was to identify bNAb-initiating Envs from the above mentioned participants using NGS techniques. The CAP255 putative bNAb-initiating Env was identified as an Env clone which lacked two key PNGs, N442 and N448, in the C4 region of the Env. These two glycans have been shown to prevent bNAb binding through glycan shielding (Li *et al.*, 2008; Canducci *et al.*, 2009; Kumar *et al.*, 2011; Cenci *et al.*, 2014; Schoofs *et al.*, 2019). The

removal of N448 has been shown to enhance antigenicity of neutralising epitopes in the V3 loop (Li *et al.*, 2008). The absence of these glycans likely created a glycan hole which may have facilitated UCA binding to the V3-glycan site since glycan holes have been shown to play an important role in autologous neutralising responses (Wei *et al.*, 2003; McCoy *et al.*, 2016; Ringe *et al.*, 2019; Schorcht *et al.*, 2022).

The CAP314 sequences possessed decreasing V1 loop lengths over time, with the V1 loop being noticeably shorter than that of other HIV-1 Envs such as CAP255. Shortened V1 loop lengths have been implicated in the exposure of the V3-glycan site (van den Kerkhof *et al.*, 2016; Anthony *et al.*, 2017; Bonsignori *et al.*, 2017), potentially allowing for easier binding of BCRs. These studies supported the selection of the CAP314 putative bNAb-initiating viruses at 19 wpi.

A co-circulating, broad CD4bs lineage has previously been identified for CAP314 (Scheepers *et al.*, unpublished data). Multiple timepoints showed a loss of PNGs N460 and N463 in the V5 loop of the CAP314 *envs* and a small group of sequences showed a loss of the PNG at N276. The removal of both N460 and N463 have been confirmed to increase binding affinity of bNAbs as well as UCAs when removed along with N276 (McGuire *et al.*, 2013; Wang *et al.*, 2015). Removal of the N276 glycan alone has also been shown to improve the binding of certain CD4bs directed antibodies (LaBranche *et al.*, 2019). These three glycans have also been removed from a previously designed germline-targeting immunogen, eODGT8, which has been proven to successfully target naïve B-cells to elicit CD4bs directed antibodies (Jardine *et al.*, 2016). These sequences could potentially be recognised by a CD4bs-specific UCA more easily due to the presence of these glycan holes. This supports the selection of a 19 wpi sequence that lacked all three PNGs as the putative bNAb-initiating virus for the CAP314 CD4bs lineage.

Sequences from CAP177 at 9 wpi showed two distinct viral populations, indicating a dual infection. This is due to a superinfection, confirming previous work done based on SGA derived sequences (Moore *et al.*, 2012). We did not identify any sequences with motifs that would enhance accessibility to the V3-glycan epitope to be readily bound by naïve BCRs. This could be a result of the inaccurate selection of timepoints due to the estimation of CAP177 UCA emergence informed by the emergence of CAP255 and CAP314 lineages. Future studies will need to focus on the identification and isolation of the CAP177 V3-glycan lineage UCA and

monoclonal antibodies as this could lead to a more accurate identification of the CAP177 bNAb-initiating Env sequence. An inferred CAP177 UCA may also be needed to screen potential bNAb-initiating Envs for binding.

Deep *env* sequencing from selected CAP314 timepoints resulted in 672 sequences from 13 – 19 wpi and showed that there was a glycan shift from position N334 to N332 in the V3 region. Although this has been previously reported in a study from our laboratory (Moore *et al.*, 2012), sequences acquired in this study were at greater depth than the previous study which only had 44 sequences from 12 – 19 wpi . This dramatic shift to a N332 glycan seen from 15 to 19 wpi has been implicated as an escape mechanism that evolved to block binding of early strain-specific antibodies (Moore *et al.*, 2012). The in-depth sequencing at each timepoint detected low frequency variants that were previously not detected which used SGAs and sampled the viral sequence diversity at a lower depth. These data show how the use of in-depth NGS will be instrumental in studying changes observed in HIV-1 *envs*, such as those resulting from immune pressure.

After identification of putative bNAb-initiating Envs, they were expressed as SOSIP trimers for testing. The CAP255.WT trimer, based on an early 8 week virus, formed an inherently stable trimer in the native conformation without the introduction of any additional mutations other than the SOSIP.664 stabilising mutations (Sanders *et al.*, 2013). This was contrasting to previously designed subtype C trimers that have been shown to form large amounts of aggregates and dimer/monomers (Guenaga *et al.*, 2015; Julien *et al.*, 2015). This is likely due to the fact that trimer stabilising mutations introduced in the C1 and C2 regions, T106E, M271I and F288L, to improve expression and yield of a previously designed V3-glycan germline-targeting immunogens (Steichen *et al.*, 2016) were naturally present in the CAP255.WT trimer. Mutations introduced to reduce antibody responses to off-target epitopes; D230N, S241N and N289, (Escolano *et al.*, 2019) and remove steric hindrance; N138T, (Saunders *et al.*, 2019) were also found to be naturally occurring in the CAP255.WT Env sequence. This would suggest that CAP255.WT is more likely to express as a stable, well-formed trimer due to these inherent properties. The CAP255 putative bNAb-initiating Env, which lacked PNGs N442 and N448, resulted in a very stable trimer that expressed better than the CAP255.WT trimer, indicating that the removal of these two PNGs potentially had a positive effect on trimer expression. The CAP255.GT2 trimer was based on the bNAb-initiating Env with

additional mutations introduced (Steichen *et al.*, 2019) and showed to have the highest expression of all trimers. Although all three CAP255 trimers were well-formed stable trimers, they showed no binding by the CAP255 UCAs. It is known that some UCAs have very low binding affinity for their bNAb-initiating trimer (Doria-Rose et al., 2014) and weekly neutralise the corresponding Env-pseudotyped virus (Bhiman et al., 2015). However, there are other V3-glycan and MPER lineages where UCA binding and neutralisation of autologous viral isolates were not observed (Bonsignori et al., 2017; Krebs et al., 2019). The CAP255 UCA may belong to this latter group where the interaction of the UCA and the cognate bNAb-intiating Env was below the level of detection of the assays we have used in this study. There is also some thought in the field that the activation of bNAb precursor naïve B-cells may require the multivalent interaction of multiple B cell receptors on the cell surface with the viral Env antigens, which cannot be assessed using the soluble monoclonal antibody version of the UCA, as was done in this study. It is important to note that like many UCAs used in immunogen design studies both CAP255 UCA01 and UCA02 were inferred UCAs. This creates a level of uncertainty when selecting UCA sequences due to the difficulties surrounding the accurate inference of germline sequences (Sliepen *et al.*, 2015; LaBranche *et al.*, 2019). Another factor to take into consideration is the selection of the putative bNAb-initiating Env. There is evidence that minority viral variants often initiate bNAb lineages (Bhiman *et al.*, 2015; Wibmer *et al.*, 2016) and although deep sequencing of CAP255 resulted in a bigger sample size of viral variants represented, we cannot be confident that the resulting sequences were representative of all sequences circulating at that specific timepoint due to significant viral diversity observed at these timepoints.

CAP255.GT2 expressed as the most stable and well-formed modified trimer and was shown to have the highest melting temperature, which has been linked to trimers with optimal antigenic and immunogenic properties due to their fully closed, prefusion conformations (Guttman *et al.*, 2015; Julien *et al.*, 2015). The neutralisation of CAP255.GT2 by CAP255 early intermediate antibodies suggests that this trimer could be further modified to enable binding by V3-glycan UCAs. This indicated that CAP255.GT2 would be the best candidate to advance for further development as a germline-targeting immunogen.

The three putative CAP314 bNAb-initiating WT Envs did not express as stable well-formed trimers. This could be due to the shorter V1 loop of the CAP314 trimers negatively influencing

the stability of these trimers. This hypothesis is supported by the decrease in the expression and stabilisation that was seen with the CAP255.GT4 trimer, where the entire V1 loop was removed, compared to the CAP255.WT trimer, with a V1 loop length of 24 residues, which yielded high quantities of quality trimer. CAP314.GTL, which had the longest V1 loop, was the only CAP314 Env that expressed a high enough yield to be tested in downstream binding assays. It is also important to note that CAP314 did not naturally contain any of the stabilising mutations that naturally occur in CAP255 and improve the expression and stability of the trimer. This indicates that sequence variability also plays a role in the non-optimal expression of a trimer and that some Envs possess sequence encoded characteristics that make them inherently better at forming Env trimers. Sequence analysis can only inform on putative bNAb-initiating viruses up to a point, screening and down-stream analysis in the lab is needed to confirm if the selected sequences are correct. Introduction of the CAP255.GT2 mutations (Steichen *et al.*, 2019) (**Table 3.1**) into CAP314 WT Envs did not result in any change in their expression. However, CAP314.GTS2 did demonstrate the most well defined trimer peak of all purified CAP314 trimers and resulted in a conformationally correct trimer. This could potentially be the result of the CAP255.GT2 stabilising mutations (Steichen *et al.*, 2019) that were introduced.

To date little has been published with regards to V3-glycan immunogens derived from subtype C virus Envs. This study resulted in the development of a stable subtype C Env, CAP255.GT2, that is able to interact with early intermediate antibodies of a V3-glycan directed bNAb lineage. As subtype C is the most dominant HIV-1 subtype worldwide, a trimer that is representative of this subtype and is able to elicit bNAbs may be advantageous. For these reasons a stable subtype C immunogen should be used in conjunction with immunogens from other subtypes such as BG505 (subtype A), and B41 (subtype B) (Sanders *et al.*, 2013; Pugach *et al.*, 2015). Using germline-targeting Envs from a participant such as CAP314 who developed broad lineages towards two key epitopes on the HIV-1 Env it could be possible to induce bNAb precursor antibodies to both V3-glycan and CD4bs. These Envs could be further stabilised and designed to represent features of both bNAb-initiating Envs of the co-circulating lineages through the addition of mutations. Through this study we were able to identify a stable subtype C Env on which future trimer design and immunogen studies can be based.

5 Conclusion

The aim of this project was to identify and characterise bNAb-initiating Envs from three South African HIV-1 infected individuals (CAP255, CAP314 and CAP177) that developed V3-glycan directed bNAb responses. Like other V3 bNAbs, the bNAbs isolated from participants CAP255 and CAP314 made use of IGHV4 germline genes. Thus future identification of conserved functional motifs in the IGVH4 region is important for development of a successful germline-targeting immunogen. Individuals infected with subtype C HIV-1 have been shown to predominantly elicit bNAb responses directed to the V3-glycan site, making this an attractive vaccine design strategy. Here, a total of 12 subtype C potential germline-targeting trimers were designed based on putative bNAb-initiating viruses, most with additional modifications to improve binding affinity and increase stabilisation. These trimers were expressed and characterised with the use of biophysical, binding and functional assays. UCA and intermediate antibodies tested against the trimers did not bind or neutralise, highlighting the difficulty of identifying bNAb-initiating Envs due to low affinity UCA interactions. This indicates the need for further modifications to increase valency and enhance binding to UCAs and further investigation of putative bNAb-initiating Envs. The analysis of *env* deep sequencing data also enabled the identification of a putative bNAb-initiating Env from the CAP314 CD4bs lineage which possessed a glycan hole previously implicated in enhanced UCA binding. Ultimately, this study has identified a subtype C trimer, CAP255.GT2, with the potential to be used as a priming immunogen and showed how SOSIP trimer stabilisation approaches cannot be applied universally to all *env* sequences due to the high degree of inherent variability observed between Envs.

References

Allan, J. S. *et al.* (1985) 'Major glycoprotein antigens that induce antibodies in AIDS patients are ncoded by HTLV-III', *Science*, 228(4703), pp. 1091–1094. doi: 10.1126/science.2986290.

Anthony, C. *et al.* (2017) 'Cooperation between strain-specific and broadly neutralizing responses limited viral escape and prolonged the exposure of the broadly neutralizing epitope', *Journal of Virology*, 91(18), pp. 1–16. doi: 10.1128/JVI.00828-17.

Arage, G. *et al.* (2014) 'Adherence to antiretroviral therapy and its associated factors among children at South Wollo Zone Hospitals, Northeast Ethiopia: a cross-sectional study', *BMC Public Health*, 14(1), p. 365. doi: 10.1186/1471-2458-14-365.

Auvert, B. *et al.* (2005) 'Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 trial', *PLoS Medicine*, 2(11), p. e298. doi: 10.1371/journal.pmed.0020298.

AVERT (2018) *Prevention of mother-to-child transmission (PMTCT) of HIV*. Available at: https://www.avert.org/professionals/hiv-programming/prevention/prevention-mother-child#footnote1_74puao0. (Accessed: 10 November 2021)

Balla-Jhagjhoorsingh, S. S. *et al.* (2013) 'The N276 glycosylation site is required for HIV-1 neutralization by the CD4 binding site specific HJ16 monoclonal antibody', *PLoS ONE*, 8(7), p. e68863. doi: 10.1371/journal.pone.0068863.

Barré-Sinoussi, F. *et al.* (1983) 'Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)', *Science*, 220(4599), pp. 868–871. doi: 10.1126/science.6189183.

Becker, N. *et al.* (2020) 'Individual, household, and community level barriers to ART adherence among women in rural Eswatini', *PLoS ONE*, 15(4), p. e0231952. doi: 10.1371/journal.pone.0231952.

Bhiman, J. N. *et al.* (2015) 'Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies', *Nature Medicine*, 21, p. 1332. Available at: https://doi.org/10.1038/nm.3963.

Binley, J. M. *et al.* (2002) 'Enhancing the proteolytic maturation of human immunodeficiency

virus type 1 Envelope glycoproteins', *Journal of Virology*, 76(6), pp. 2606–2616. doi: 10.1128/JVI.76.6.2606-2616.2002.

Blumenthal, R. *et al.* (2012) 'HIV entry and envelope glycoprotein-mediated fusion', *Journal of Biological Chemistry*, 287(49), pp. 40841–40849. doi: 10.1074/jbc.R112.406272.

Bonsignori, M. *et al.* (2017) 'Staged induction of HIV-1 glycan-dependent broadly neutralizing antibodies', *Science Translational Medicine*, 9(381). doi: 10.1126/scitranslmed.aai7514.

Briggs, J. A. G. and Kräusslich, H.-G. (2011) 'The molecular architecture of HIV', *Journal of Molecular Biology*, 410(4), pp. 491–500. doi: 10.1016/j.jmb.2011.04.021.

Brochet, X. *et al.* (2008) 'IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis', *Nucleic Acids Research*, 36(Web Server), pp. W503–W508. doi: 10.1093/nar/gkn316.

Burton, D. R. *et al.* (2004) 'HIV vaccine design and the neutralizing antibody problem', *Nature Immunology*, 5(3), pp. 233–236. doi: 10.1038/ni0304-233.

Canducci, F. *et al.* (2009) 'Dynamic features of the selective pressure on the human immunodeficiency virus type 1 (HIV-1) gp120 CD4-binding site in a group of long term non progressor (LTNP) subjects', *Retrovirology*, 6(1), p. 4. doi: 10.1186/1742-4690-6-4.

Cardoso, R. M. F. *et al.* (2005) 'Broadly neutralizing anti-HIV Antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41', *Immunity*, 22(2), pp. 163–173. doi: 10.1016/j.immuni.2004.12.011.

Cenci, A. *et al.* (2014) 'Molecular characterization of HIV-1 subtype C gp-120 regions potentially involved in virus adaptive mechanisms', *PLoS ONE*, 9(4), p. e95183. doi: 10.1371/journal.pone.0095183.

Choe, H. *et al.* (1996) 'The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates.', *Cell*, 85(7), pp. 1135–48. doi: 10.1016/s0092-8674(00)81313-6.

Cohen, K. *et al.* (2014) 'Early preservation of CXCR5+ PD-1+ helper T cells and B cell activation predict the breadth of neutralizing antibody responses in chronic HIV-1 infection', *Journal of Virology*, 88(22), pp. 13310–13321. doi: 10.1128/JVI.02186-14.

Corey, L. *et al.* (2021) 'Two randomized trials of neutralizing antibodies to prevent HIV-1

acquisition', *New England Journal of Medicine*, 384(11), pp. 1003–1014. doi: 10.1056/NEJMoa2031738.

Cortez, V. *et al.* (2012) 'HIV-1 superinfection in women broadens and strengthens the neutralizing antibody response', *PLoS Pathogens*, 8(3), p. e1002611. doi: 10.1371/journal.ppat.1002611.

Corti, D. and Lanzavecchia, A. (2013) 'Broadly neutralizing antiviral antibodies', *Annual Review of Immunology*, 31(1), pp. 705–742. doi: 10.1146/annurev-immunol-032712-095916.

Costin, J. M. (2007) 'Cytopathic mechanisms of HIV-1', *Virology Journal*, 4(1), p. 100. doi: 10.1186/1743-422X-4-100.

Doria-Rose, N. A. *et al.* (2014) 'Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies', *Nature*, 509(7498), pp. 55–62. doi: 10.1038/nature13036.

Edelman, G. M. and Benacerraf, B. (1962) 'On structural and functional relations between antibodies and proteins of the gamma-system', *Proceedings of the National Academy of Sciences*, 48(6), pp. 1035–1042. doi: 10.1073/pnas.48.6.1035.

Endalamaw, A. *et al.* (2018) 'Adherence to highly active antiretroviral therapy among children in Ethiopia: A systematic review and meta-analysis', *AIDS and Behavior*, 22(8), pp. 2513–2523. doi: 10.1007/s10461-018-2152-z.

Engler, K. *et al.* (2018) 'Barriers to antiretroviral therapy adherence in developed countries: a qualitative synthesis to develop a conceptual framework for a new patient-reported outcome measure', *AIDS Care*, 30(sup1), pp. 17–28. doi: 10.1080/09540121.2018.1469725.

Erikson, J. *et al.* (1981) 'Assignment of the genes for human λ immunoglobulin chains to chromosome 22', *Nature*, pp. 173–175. doi: 10.1038/294173a0.

Escolano, A. *et al.* (2019) 'Immunization expands B cells specific to HIV-1 V3 glycan in mice and macaques', *Nature*, 570(7762), pp. 468–473. doi: 10.1038/s41586-019-1250-z.

Falkowska, E. *et al.* (2014) 'Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved Envelope trimers', *Immunity*, 40(5), pp. 657–668. doi: 10.1016/j.immuni.2014.04.009.

Fauci, A. S. *et al.* (1984) 'Acquired immunodeficiency syndrome: Epidemiologic, clinical,

immunologic, and therapeutic considerations', *Annals of Internal Medicine*, 100(1), pp. 92–106. doi: 10.7326/0003-4819-100-1-92.

Fauci, A. S. (2003) 'HIV and AIDS: 20 years of science', *Nature Medicine*, 9(7), pp. 839–843. doi: 10.1038/nm0703-839.

Feng, Y. *et al.* (1996) 'HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor', *Science*, 272(5263), pp. 872–877. doi: 10.1126/science.272.5263.872.

Flaherty, D. (ed.) (2012) 'Antibody diversity', in *Immunology for Pharmacy*. 1st edn. Saint Louis: Elsevier, pp. 79–86. doi: 10.1016/B978-0-323-06947-2.10010-0.

Forthal, D. N. (2014) 'Functions of antibodies.', *Microbiology spectrum*, 2(4), pp. 1–17. doi: 10.1128/microbiolspec.AID-0019-2014

Freedberg, K. A. *et al.* (2015) 'The HIV cure research agenda: The role of mathematical modelling and cost-effectiveness analysis.', *Journal of virus eradication*, 1(4), pp. 245–249. doi: 10.1016/S2055-6640(20)30929-8

Freund, N. T. *et al.* (2017) 'Coexistence of potent HIV-1 broadly neutralizing antibodies and antibody-sensitive viruses in a viremic controller', *Science Translational Medicine*, 9(373), pp. 1–14. doi: 10.1126/scitranslmed.aal2144.

Gallo, R. C. *et al.* (1984) 'Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS', *Science*, 224(4648), pp. 500–503. doi: 10.1126/science.6200936.

Garces, F. *et al.* (2014) 'Structural evolution of glycan recognition by a family of potent HIV antibodies', *Cell*, 159(1), pp. 69–79. doi: 10.1016/j.cell.2014.09.009.

Gray, E. S. *et al.* (2011a) 'Isolation of a monoclonal antibody that targets the alpha-2 helix of gp120 and represents the initial autologous neutralizing-antibody response in an HIV-1 subtype C-infected individual', *Journal of Virology*, 85(15), pp. 7719–7729. doi: 10.1128/JVI.00563-11.

Gray, E. S. *et al.* (2011b) 'The neutralization breadth of HIV-1 develops incrementally over four years and is associated with CD4+ T cell decline and high viral load during acute infection',

Journal of Virology, 85(10), pp. 4828–4840. doi: 10.1128/JVI.00198-11.

Guenaga, J. *et al.* (2015) 'Well-ordered trimeric HIV-1 subtype B and C soluble spike mimetics generated by negative selection display native-like properties', *PLoS Pathogens*, 11(1), pp. 1–16. doi: 10.1371/journal.ppat.1004570.

Gulick, R. M. *et al.* (1997) 'Treatment with Indinavir, Zidovudine, and Lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy', *New England Journal of Medicine*, 337(11), pp. 734–739. doi: 10.1056/NEJM199709113371102.

Gupta, R. and Brunak, S. (2001) 'Prediction of glycosylation across the human proteome and the correlation to protein function', in *Biocomputing 2002*, pp. 310–322. doi: 10.1142/9789812799623_0029.

Gürtler, L. G. *et al.* (1994) 'A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon', *Journal of Virology*, 68(3), pp. 1581–1585. doi: 10.1128/jvi.68.3.1581-1585.1994.

Guttman, M. *et al.* (2015) 'Antibody potency relates to the ability to recognize the closed, pre-fusion form of HIV Env', *Nature Communications*, 6(1), p. 6144. doi: 10.1038/ncomms7144.

Hemelaar, J. (2012) 'The origin and diversity of the HIV-1 pandemic', *Trends in Molecular Medicine*, pp. 182–192. doi: 10.1016/j.molmed.2011.12.001.

Hong, H.-W. *et al.* (2013) 'Induced degradation of tat by nucleocapsid (NC) via the proteasome pathway and its effect on HIV transcription', *Viruses*, 5(4), pp. 1143–1152. doi: 10.3390/v5041143.

Hu, W.-S. and Hughes, S. H. (2012) 'HIV-1 reverse transcription', *Cold Spring Harbor Perspectives in Medicine*, 2(10), pp. a006882–a006882. doi: 10.1101/cshperspect.a006882.

Huang, J. *et al.* (2012) 'Broad and potent neutralization of HIV-1 by a gp41-specific human antibody', *Nature*, 491(7424), pp. 406–412. doi: 10.1038/nature11544.

Huang, J. *et al.* (2014) 'Broad and potent HIV-1 neutralization by a human antibody that binds the gp41–gp120 interface', *Nature*, 515(7525), pp. 138–142. doi: 10.1038/nature13601.

IAVI (2021) *Media roundup: News coverage highlights preliminary data from IAVI G001 clinical trial*. Available at: https://www.iavi.org/news-resources/features/media-roundup-news-

coverage-highlights-preliminary-data-from-iavi-g001-clinical-trial. (Accessed: 6 March 2022).

Jacob, R. A. *et al.* (2015) 'Anti-V3/glycan and anti-MPER neutralizing antibodies, but not anti-V2/glycan site antibodies, are strongly associated with greater anti-HIV-1 neutralization breadth and potency', *Journal of Virology*, 89(10), pp. 5264–5275. doi: 10.1128/JVI.00129-15.

Jardine, J. G. *et al.* (2016) 'HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen', *Science*, 351(6280), pp. 1458–1463. doi: 10.1126/science.aad9195.

Jönsson, U. *et al.* (1991) 'Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology.', *BioTechniques*, 11(5), pp. 620–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1804254.

Joseph, A. M. *et al.* (2003) 'Human immunodeficiency virus-1 Nef protein interacts with Tat and enhances HIV-1 gene expression', *FEBS Letters*, 548(1–3), pp. 37–42. doi: 10.1016/S0014-5793(03)00725-7.

Julien, J.-P. *et al.* (2013) 'Broadly neutralizing antibody PGT121 allosterically modulates CD4 binding via recognition of the HIV-1 gp120 V3 base and multiple surrounding glycans', *PLoS Pathogens*, 9(5), p. e1003342. doi: 10.1371/journal.ppat.1003342.

Julien, J.-P. *et al.* (2013) 'Crystal structure of a soluble cleaved HIV-1 Envelope trimer', *Science*, 342(6165), pp. 1477–1483. doi: 10.1126/science.1245625.

Julien, J.-P. *et al.* (2015) 'Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens', *Proceedings of the National Academy of Sciences*, 112(38), pp. 11947–11952. doi: 10.1073/pnas.1507793112.

Katoh, K. and Standley, D. M. (2013) 'MAFFT multiple sequence alignment software version 7: improvements in performance and usability.', *Molecular biology and evolution*, 30(4), pp. 772–80. doi: 10.1093/molbev/mst010.

Keele, B. F. *et al.* (2006) 'Chimpanzee reservoirs of pandemic and nonpandemic HIV-1', *Science*, 313(5786), pp. 523–526. doi: 10.1126/science.1126531.

van den Kerkhof, T. L. G. M. *et al.* (2016) 'HIV-1 escapes from N332-directed antibody neutralization in an elite neutralizer by envelope glycoprotein elongation and introduction of

unusual disulfide bonds', *Retrovirology*, 13(1), p. 48. doi: 10.1186/s12977-016-0279-4.

Klein, F. *et al.* (2013) 'Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization', *Cell*, 153(1), pp. 126–138. doi: 10.1016/j.cell.2013.03.018.

Kong, L. *et al.* (2013) 'Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120', *Nature Structural & Molecular Biology*, 20(7), pp. 796–803. doi: 10.1038/nsmb.2594.

Korber, B. *et al.* (2000) 'Timing the ancestor of the HIV-1 pandemic strains', *Science*, 288(5472), pp. 1789–1796. doi: 10.1126/science.288.5472.1789.

Krachmarov, C. P. *et al.* (2006) 'Factors determining the breadth and potency of neutralization by V3-specific human monoclonal antibodies derived from subjects infected with clade A or clade B strains of human immunodeficiency virus type 1', *Journal of Virology*, 80(14), pp. 7127–7135. doi: 10.1128/JVI.02619-05.

Krebs, S. J. *et al.* (2019) 'Longitudinal analysis reveals early development of three MPER-directed neutralizing antibody lineages from an HIV-1-infected individual', *Immunity*, 50(3), pp. 677-691.e13. doi: 10.1016/j.immuni.2019.02.008.

Kumar, R. *et al.* (2011) 'Improving immunogenicity of HIV-1 envelope gp120 by glycan removal and immune complex formation', *Vaccine*, 29(48), pp. 9064–9074. doi: 10.1016/j.vaccine.2011.09.057.

Kwong, P. D. *et al.* (2002) 'HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites', *Nature*, 420(6916), pp. 678–682. doi: 10.1038/nature01188.

Kwong, P. D. and Mascola, J. R. (2018) 'HIV-1 vaccines based on antibody identification, B cell ontogeny, and epitope structure', *Immunity*, 48(5), pp. 855–871. doi: 10.1016/j.immuni.2018.04.029.

LaBranche, C. C. *et al.* (2019) 'Correction: HIV-1 envelope glycan modifications that permit neutralization by germline-reverted VRC01-class broadly neutralizing antibodies', *PLOS Pathogens*, 15(3), p. e1007646. doi: 10.1371/journal.ppat.1007646.

Landais, E. *et al.* (2016) 'Broadly neutralizing antibody responses in a large longitudinal sub-saharan HIV primary infection cohort', *PLOS Pathogens*, 12(1), p. e1005369. doi: 10.1371/journal.ppat.1005369.

Landais, E. and Moore, P. L. (2018) 'Development of broadly neutralizing antibodies in HIV-1 infected elite neutralizers', *Retrovirology*, 15(1), pp. 1–14. doi: 10.1186/s12977-018-0443-0.

Larsson, A. (2014) 'AliView: a fast and lightweight alignment viewer and editor for large datasets', *Bioinformatics*, 30(22), pp. 3276–3278. doi: 10.1093/bioinformatics/btu531.

Lefranc, M.-P. (2001a) 'Nomenclature of the human immunoglobulin heavy (IGH) genes', *Experimental and Clinical Immunogenetics*, 18(2), pp. 100–116. doi: 10.1159/000049189.

Lefranc, M.-P. (2001b) 'Nomenclature of the human immunoglobulin kappa (IGK) genes', *Experimental and Clinical Immunogenetics*, 18(3), pp. 161–174. doi: 10.1159/000049195.

De Leys, R. *et al.* (1990) 'Isolation and partial characterization of an unusual human immunodeficiency retrovirus from two persons of west-central African origin', *Journal of Virology*, 64(3), pp. 1207–1216. doi: 10.1128/jvi.64.3.1207-1216.1990.

Li, G. and De Clercq, E. (2016) 'HIV genome-wide protein associations: a Review of 30 years of research', *Microbiology and Molecular Biology Reviews*, 80(3), pp. 679–731. doi: 10.1128/MMBR.00065-15.

Li, H. *et al.* (2008) 'Identification of an N-Linked glycosylation in the C4 region of HIV-1 Envelope gp120 that is critical for recognition of neighboring CD4 T cell epitopes', *The Journal of Immunology*, 180(6), pp. 4011–4021. doi: 10.4049/jimmunol.180.6.4011.

Liao, H.-X. *et al.* (2013) 'Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus', *Nature*, 496(7446), pp. 469–476. doi: 10.1038/nature12053.

MacLeod, D. T. *et al.* (2016) 'Early antibody lineage diversification and independent limb maturation lead to broad HIV-1 neutralization targeting the Env high-mannose patch', *Immunity*, 44(5), pp. 1215–1226. doi: 10.1016/j.immuni.2016.04.016.

Martinez, V. M. *et al.* (2021) 'A fractional calculus model for HIV dynamics: real data, parameter estimation and computational strategies', *Chaos, Solitons & Fractals*, 152(August), p. 111398. doi: 10.1016/j.chaos.2021.111398.

Mauclère, P. *et al.* (1997) 'Serological and virological characterization of HIV-1 group O infection in Cameroon', *AIDS*, 11(4), pp. 445–453. doi: 10.1097/00002030-199704000-00007.

McCoy, L. E. *et al.* (2016) 'Holes in the glycan shield of the native HIV Envelope are a target of trimer-elicited neutralizing antibodies', *Cell Reports*, 16(9), pp. 2327–2338. doi: 10.1016/j.celrep.2016.07.074.

McGuire, A. T. *et al.* (2013) 'Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies', *Journal of Experimental Medicine*, 210(4), pp. 655–663. doi: 10.1084/jem.20122824.

McLellan, J. S. *et al.* (2011) 'Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9', *Nature*, 480(7377), pp. 336–343. doi: 10.1038/nature10696.

Medina-Ramírez, M. *et al.* (2017) 'Design and crystal structure of a native-like HIV-1 envelope trimer that engages multiple broadly neutralizing antibody precursors in vivo', *The Journal of Experimental Medicine*, 214(9), pp. 2573–2590. doi: 10.1084/jem.20161160.

Merk, A. and Subramaniam, S. (2013) 'HIV-1 envelope glycoprotein structure', *Current Opinion in Structural Biology*, 23(2), pp. 268–276. doi: 10.1016/j.sbi.2013.03.007.

Montefiori, D. C. (2009) 'HIV Protocols', 485. doi: 10.1007/978-1-59745-170-3.

Moody, M. A. *et al.* (2016) 'Immune perturbations in HIV-1–infected individuals who make broadly neutralizing antibodies', *Science Immunology*, 1(1), pp. 1–22. doi: 10.1126/sciimmunol.aag0851.

Moore, P. L. *et al.* (2009) 'Limited neutralizing antibody specificities drive neutralization escape in early HIV-1 subtype C infection', *PLoS Pathogens*, 5(9), p. e1000598. doi: 10.1371/journal.ppat.1000598.

Moore, P. L. *et al.* (2012) 'Evolution of an HIV glycan–dependent broadly neutralizing antibody epitope through immune escape', *Nature Medicine*, 18(11), pp. 1688–1692. doi: 10.1038/nm.2985.

Moosa, A. *et al.* (2019) 'Long-term adherence to antiretroviral therapy in a South African adult patient cohort: a retrospective study', *BMC Infectious Diseases*, 19(1), p. 775. doi: 10.1186/s12879-019-4410-8.

Mouquet, H. *et al.* (2012) 'Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies', *Proceedings of the National Academy of Sciences*, 109(47). doi: 10.1073/pnas.1217207109.

Mouquet, H. (2014) 'Antibody B cell responses in HIV-1 infection', *Trends in Immunology*. Elsevier Ltd, 35(11), pp. 549–561. doi: 10.1016/j.it.2014.08.007.

Moyo, T. *et al.* (2020) 'Targeting the N332-supersite of the HIV-1 envelope for vaccine design', *Expert Opinion on Therapeutic Targets*. Taylor & Francis, 24(6), pp. 499–509. doi: 10.1080/14728222.2020.1752183.

Murphy, K. *et al.* (2012) *Janeway's immunobiology*. 8th ed. New York : Garland Science. Available at: http://lib.ugent.be/catalog/rug01:002038139.

Myers, G. *et al.* (1997) 'Human retroviruses and AIDS 1996. A compilation and analysis of nucleic acid and amino acid sequences', in.

NIAID (2021) *HIV vaccine candidate does not sufficiently protect women against HIV infection*. Available at: https://www.hiv.gov/blog/hiv-vaccine-candidate-does-not-sufficiently-protect-women-against-hiv-infection. (Accessed: 6 March 2022).

NIH (2017) *NIH and partners launch HIV vaccine efficacy study.* Available at: https://www.niaid.nih.gov/news-events/nih-and-partners-launch-hiv-vaccine-efficacy-study. (Accessed: 22 January 2022).

NIH (2019) *NIH and partners to launch HIV vaccine efficacy trial in the Americas and Europe.* Available at: https://www.niaid.nih.gov/news-events/nih-and-partners-launch-hiv-vaccine-efficacy-trial-americas-and-europe. (Accessed: 22 January 2022).

NIH (2020) *Experimental HIV vaccine regimen ineffective in preventing HIV.* Available at: https://www.niaid.nih.gov/news-events/experimental-hiv-vaccine-regimen-ineffective-preventing-hiv. (Accessed: 24 January 2022).

NIH (2022) *A clinical trial to evaluate the safety and immunogenicity of BG505 MD39.3, BG505 MD39.3 gp151, and BG505 MD39.3 gp151 CD4KO HIV trimer mRNA vaccines in healthy, HIV-uninfected adult participants*, *ClinicalTrials.gov*. Available at: https://clinicaltrials.gov/ct2/show/NCT05217641#wrapper. (Accessed: 4 March 2022).

Palella, F. J. *et al.* (1998) 'Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection', *New England Journal of Medicine*, 338(13), pp. 853–860. doi: 10.1056/NEJM199803263381301.

Pantophlet, R. *et al.* (2008) 'Neutralizing activity of antibodies to the V3 loop region of HIV-1 gp120 relative to their epitope fine specificity', *Virology*, 381(2), pp. 251–260. doi: 10.1016/j.virol.2008.08.032.

Parren, P. W. H. I. *et al.* (1997) 'HIV-1 antibody — debris or virion?', *Nature Medicine*, 3(4), pp. 366–367. doi: 10.1038/nm0497-366d.

Peeters, M. *et al.* (1997) 'Geographical distribution of HIV-1 group O viruses in Africa', *AIDS*, 11(4), pp. 493–498. doi: 10.1097/00002030-199704000-00013.

Pejchal, R. *et al.* (2011) 'A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield', *Science*, 334(6059), pp. 1097–1103. doi: 10.1126/science.1213256.

Pinto, D. *et al.* (2019) 'Structural basis for broad HIV-1 neutralization by the MPER-specific human broadly neutralizing antibody LN01', *Cell Host & Microbe*, 26(5), pp. 623-637.e8. doi: 10.1016/j.chom.2019.09.016.

Prabakaran, P. *et al.* (2007) 'Structure and function of the HIV Envelope glycoprotein as entry mediator, vaccine immunogen, and target for inhibitors', in *Advances in Pharmacology*. Elsevier, Inc, pp. 33–97. doi: 10.1016/S1054-3589(07)55002-7.

Price, D. A. *et al.* (1997) 'Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection', *Proceedings of the National Academy of Sciences*, 94(5), pp. 1890–1895. doi: 10.1073/pnas.94.5.1890.

Pugach, P. *et al.* (2015) 'A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene', *Journal of Virology*, 89(6), pp. 3380–3395. doi: 10.1128/JVI.03473-14.

Ratner, L. *et al.* (1985) 'Complete nucleotide sequence of the AIDS virus, HTLV-III', *Nature*, 313(6000), pp. 277–284. doi: 10.1038/313277a0.

Rerks-Ngarm, S. *et al.* (2009) 'Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand', *New England Journal of Medicine*, 361(23), pp. 2209–2220. doi: 10.1056/NEJMoa0908492.

Ringe, R. P. *et al.* (2019) 'Closing and opening holes in the glycan shield of HIV-1 Envelope glycoprotein SOSIP trimers can redirect the neutralizing antibody response to the newly unmasked epitopes', *Journal of Virology*, 93(4), pp. 1–26. doi: 10.1128/JVI.01656-18.

Rodrigo, A. G. *et al.* (1997) 'Quantitation of target molecules from polymerase chain reaction-based limiting dilution assays', *AIDS Research and Human Retroviruses*, 13(9), pp. 737–742. doi: 10.1089/aid.1997.13.737.

Rognes, T. *et al.* (2016) 'VSEARCH: a versatile open source tool for metagenomics', *PeerJ*, 4(10), p. e2584. doi: 10.7717/peerj.2584.

Roskin, K. M. *et al.* (2020) 'Aberrant B cell repertoire selection associated with HIV neutralizing antibody breadth', *Nature Immunology*, 21(2), pp. 199–209. doi: 10.1038/s41590-019-0581-0.

Rusert, P. *et al.* (2016) 'Determinants of HIV-1 broadly neutralizing antibody induction', *Nature Medicine*, 22(11), pp. 1260–1267. doi: 10.1038/nm.4187.

Sacks, D. *et al.* (2019) 'Somatic hypermutation to counter a globally rare viral immunotype drove off-track antibodies in the CAP256-VRC26 HIV-1 V2-directed bNAb lineage', *PLOS Pathogens*, 15(9), p. e1008005. doi: 10.1371/journal.ppat.1008005.

Salzwedel, K. *et al.* (1999) 'A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity', *Journal of Virology*, 73(3), pp. 2469–2480. doi: 10.1128/JVI.73.3.2469-2480.1999.

SAMRC (2021) 'Efficacy results announced: global first antibody mediated prevention (AMP) trials'. Available at: https://www.samrc.ac.za/news/efficacy-results-announced-global-first-antibody-mediated-prevention-amp-trials. (Accessed: 10 March 2022)

Sanders, R. W. *et al.* (2002) 'Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1', *Journal of Virology*, 76(17), pp. 8875–8889. doi: 10.1128/JVI.76.17.8875-8889.2002.

Sanders, R. W. *et al.* (2013) 'A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-

neutralizing antibodies', *PLoS Pathogens*, 9(9), p. e1003618. doi: 10.1371/journal.ppat.1003618.

Sather, D. N. *et al.* (2009) 'Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection', *Journal of Virology*, 83(2), pp. 757–769. doi: 10.1128/JVI.02036-08.

Saunders, K. O. *et al.* (2019) 'Targeted selection of HIV-specific antibody mutations by engineering B cell maturation', *Science*, 366(6470), p. eaay7199. doi: 10.1126/science.aay7199.

Sawaya, B. E. *et al.* (2000) 'Cooperative interaction between HIV-1 regulatory proteins Tat and Vpr modulates transcription of the viral genome', *Journal of Biological Chemistry*, 275(45), pp. 35209–35214. doi: 10.1074/jbc.M005197200.

Scharf, L. *et al.* (2013) 'Structural basis for HIV-1 gp120 recognition by a germ-line version of a broadly neutralizing antibody', *Proceedings of the National Academy of Sciences*, 110(15), pp. 6049–6054. doi: 10.1073/pnas.1303682110.

Scheid, J. F. *et al.* (2009) 'Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals', *Nature*, 458(7238), pp. 636–640. doi: 10.1038/nature07930.

Scheid, J. F. *et al.* (2011) 'Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding', *Science*, 333(6049), pp. 1633–1637. doi: 10.1126/science.1207227.

Schickel, J. N. *et al.* (2017) 'Self-reactive VH4-34-expressing IgG B cells recognize commensal bacteria', *Journal of Experimental Medicine*, 214(7), pp. 1991–2003. doi: 10.1084/jem.20160201.

Schoofs, T. *et al.* (2019) 'Broad and potent neutralizing antibodies recognize the silent face of the HIV Envelope', *Immunity*, 50(6), pp. 1513-1529.e9. doi: 10.1016/j.immuni.2019.04.014.

Schorcht, A. *et al.* (2022) 'The glycan hole area of HIV-1 Envelope trimers contributes prominently to the induction of autologous neutralization', *Journal of Virology*, 96(1). doi: 10.1128/JVI.01552-21.

Schur, P. H. (1988) 'IgG subclasses. A historical perspective.', *Monographs in allergy*, 23, pp.

1–11. Available at: http://europepmc.org/abstract/MED/3290655.

Sharp, P. M. and Hahn, B. H. (2011) 'Origins of HIV and the AIDS pandemic', *Cold Spring Harbor Perspectives in Medicine*, 1(1), pp. a006841–a006841. doi: 10.1101/cshperspect.a006841.

Sheward, D. J. *et al.* (2018) 'HIV superinfection drives de novo antibody responses and not neutralization breadth', *Cell Host & Microbe*, 24(4), pp. 593-599.e3. doi: 10.1016/j.chom.2018.09.001.

Simonich, C. A. *et al.* (2016) 'HIV-1 neutralizing antibodies with limited hypermutation from an infant', *Cell*, 166(1), pp. 77–87. doi: 10.1016/j.cell.2016.05.055.

Sliepen, K. *et al.* (2015) 'Binding of inferred germline precursors of broadly neutralizing HIV-1 antibodies to native-like envelope trimers', *Virology*, 486(5), pp. 116–120. doi: 10.1016/j.virol.2015.08.002.

Sok, D., Doores, K. J., *et al.* (2014) 'Promiscuous glycan site recognition by antibodies to the high-mannose patch of gp120 broadens neutralization of HIV', *Science Translational Medicine*, 6(236), pp. 236ra63-236ra63. doi: 10.1126/scitranslmed.3008104.

Sok, D. *et al.* (2016) 'A prominent site of antibody vulnerability on HIV Envelope incorporates a motif associated with CCR5 binding and its camouflaging glycans', *Immunity*, 45(1), pp. 31–45. doi: 10.1016/j.immuni.2016.06.026.

Sok, D. and Burton, D. R. (2018) 'Recent progress in broadly neutralizing antibodies to HIV', *Nature Immunology*, 19(11), pp. 1179–1188. doi: 10.1038/s41590-018-0235-7.

Stamatatos, L. *et al.* (2017) 'Germline-targeting immunogens', *Immunological Reviews*, 275(1), pp. 203–216. doi: 10.1111/imr.12483.

Statistics South Africa (2020) *2020 Mid-year Population Estimates*.

Steichen, J. M. *et al.* (2016) 'HIV vaccine design to target germline precursors of glycan-dependent broadly neutralizing antibodies', *Immunity*, 45(3), pp. 483–496. doi: 10.1016/j.immuni.2016.08.016.

Steichen, J. M. *et al.* (2019) 'A generalized HIV vaccine design strategy for priming of broadly neutralizing antibody responses.', *Science*, 4380(October), pp. 1–23. doi: 10.1126/science.aax4380.

Stewart, J. (2012) 'Immunological principles', in Greenwood, D. et al. (eds) *Medical Microbiology*. 18th edn. Edinburgh: Elsevier, pp. 98–108. doi: 10.1016/B978-0-7020-4089-4.00023-8.

Stiegler, G. *et al.* (2001) 'A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1', *AIDS Research and Human Retroviruses*, 17(18), pp. 1757–1765. doi: 10.1089/08892220152741450.

Taswell, C. (1981) 'Limiting dilution assays for the determination of immunocompetent cell frequencies. I. Data analysis.', *The Journal of Immunology*, 126(4), pp. 1614 LP – 1619. Available at: http://www.jimmunol.org/content/126/4/1614.abstract.

Taylor, B. S. *et al.* (2008) 'The challenge of HIV-1 subtype diversity', *New England Journal of Medicine*, 358(15), pp. 1590–1602. doi: 10.1056/NEJMra0706737.

Tonegawa, S. (1983) 'Somatic generation of antibody diversity', *Nature*, 302(5909), pp. 575–581. doi: 10.1038/302575a0.

UNAIDS (2021) *2021 UNAIDS Global AIDS Update — Confronting inequalities — Lessons for pandemic responses from 40 years of AIDS.* Availible at: https://www.unaids.org/en/resources/documents/2021/2021-global-aids-update (Accessed: 13 November 2021).

Venkatesan, P. (2021) 'Preliminary phase 1 results from an HIV vaccine candidate trial', *The Lancet Microbe*, 2(3), p. e95. doi: 10.1016/S2666-5247(21)00042-2.

Veronese, F. D. *et al.* (1985) 'Characterization of gp41 as the transmembrane protein coded by the HTLV-III/LAV envelope gene', *Science*, 229(4720), pp. 1402–1405. doi: 10.1126/science.2994223.

Wain-Hobson, S. *et al.* (1985) 'Nucleotide sequence of the AIDS virus, LAV', *Cell*, 40(1), pp. 9–17. doi: 10.1016/0092-8674(85)90303-4.

Walker, B. D. and Burton, D. R. (2008) 'Toward an AIDS vaccine', *Science*, 320(5877), pp. 760–764. doi: 10.1126/science.1152622.

Walker, L. M. *et al.* (2009) 'Broad and potent neutralizing antibodies from an african donor reveal a new HIV-1 vaccine target', *Science*, 326(5950), pp. 285–289. doi:

10.1126/science.1178746.

Walker, L. M. *et al.* (2011) 'Broad neutralization coverage of HIV by multiple highly potent antibodies', *Nature*, 477(7365), pp. 466–470. doi: 10.1038/nature10373.

Wang, J. *et al.* (2008) 'The HIV-1 Vif protein mediates degradation of Vpr and reduces Vpr-induced cell cycle arrest', *DNA and Cell Biology*, 27(5), pp. 267–277. doi: 10.1089/dna.2007.0707.

Wang, W. *et al.* (2015) 'N463 glycosylation site on V5 loop of a mutant gp120 regulates the sensitivity of HIV-1 to neutralizing monoclonal antibodies VRC01/03', *Journal of Acquired Immune Deficiency Syndromes*, 69(3), pp. 270–277. doi: 10.1097/QAI.0000000000000595.

Ward, A. B. and Wilson, I. A. (2015) 'Insights into the trimeric HIV-1 envelope glycoprotein structure', *Trends in Biochemical Sciences*, 40(2), pp. 101–107. doi: 10.1016/j.tibs.2014.12.006.

Watson, C. T. *et al.* (2017) 'The individual and population genetics of antibody immunity', *Trends in Immunology*, 38(7), pp. 459–470. doi: 10.1016/j.it.2017.04.003.

Wei, X. *et al.* (2003) 'Antibody neutralization and escape by HIV-1', *Nature*, 422(6929), pp. 307–312. doi: 10.1038/nature01470.

WHO (2021) *HIV/AIDS*, *World Health Organization, Health topics*. Available at: https://www.afro.who.int/health-topics/hivaids. (Accessed: 10 November 2021).

Wibmer, C. K. *et al.* (2013) 'Viral escape from HIV-1 neutralizing antibodies drives increased plasma neutralization breadth through sequential recognition of multiple epitopes and immunotypes', *PLoS Pathogens*, 9(10), p. e1003738. doi: 10.1371/journal.ppat.1003738.

Wibmer, C. K. *et al.* (2016) 'Structure of an N276-dependent HIV-1 neutralizing antibody targeting a rare V5 glycan hole adjacent to the CD4 binding site', *Journal of Virology*, pp. 10220–10235. doi: 10.1128/JVI.01357-16.

Wibmer, C. K. *et al.* (2015) 'HIV broadly neutralizing antibody targets', *Current Opinion in HIV and AIDS*, 10(3), pp. 135–143. doi: 10.1097/COH.0000000000000153.

Wiehe, K. *et al.* (2018) 'Functional relevance of improbable antibody mutations for HIV broadly neutralizing antibody development', *Cell Host & Microbe*, 23(6), pp. 759-765.e6. doi:

10.1016/j.chom.2018.04.018.

Wilen, C. B. *et al.* (2012) 'HIV: Cell binding and entry', *Cold Spring Harbor Perspectives in Medicine*, 2(8), pp. a006866–a006866. doi: 10.1101/cshperspect.a006866.

Williams, K. L. *et al.* (2018) 'Superinfection drives HIV neutralizing antibody responses from several B cell lineages that contribute to a polyclonal repertoire', *Cell Reports*, 23(3), pp. 682–691. doi: 10.1016/j.celrep.2018.03.082.

Wu, X. *et al.* (2006) 'Neutralization escape variants of human immunodeficiency virus type 1 are transmitted from mother to infant', *Journal of Virology*, 80(2), pp. 835–844. doi: 10.1128/JVI.80.2.835-844.2006.

Wu, X. *et al.* (2010) 'Rational design of Envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1', *Science*, 329(5993), pp. 856–861. doi: 10.1126/science.1187659.

Wu, X. *et al.* (2011) 'Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing', *Science*, 333(6049), pp. 1593–1602. doi: 10.1126/science.1207532.

Wyatt, R. *et al.* (1998) 'The antigenic structure of the HIV gp120 envelope glycoprotein', *Nature*, 393(6686), pp. 705–711. doi: 10.1038/31514.

Wyatt, R. and Sodroski, J. (1998) 'The HIV-1 Envelope glycoproteins: Fusogens, antigens, and immunogens', *Science*, 280(5371), pp. 1884–1888. doi: 10.1126/science.280.5371.1884.

Yang, M. *et al.* (2017) 'Conformational heterogeneity of the HIV Envelope glycan shield', *Scientific Reports*, 7(1), p. 4435. doi: 10.1038/s41598-017-04532-9.

Zhou, T. *et al.* (2013) 'Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies', *Immunity*, 39(2), pp. 245–258. doi: 10.1016/j.immuni.2013.04.012.

Zhou, T. *et al.* (2015) 'Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors', *Cell*, 161(6), pp. 1280–1292. doi: 10.1016/j.cell.2015.05.007.

Zhou, T. *et al.* (2018) 'A neutralizing antibody recognizing primarily N-linked glycan targets the silent face of the HIV Envelope', *Immunity*, 48(3), pp. 500-513.e6. doi: 10.1016/j.immuni.2018.02.013.

Zhuang, J. *et al.* (2002) 'Human immunodeficiency virus type 1 recombination: Rate, fidelity, and putative hot spots', *Journal of Virology*, 76(22), pp. 11273–11282. doi: 10.1128/JVI.76.22.11273-11282.2002.

Appendix A: CAPRISA Letter of Permission

CAPRISA
CENTRE FOR THE AIDS PROGRAMME OF RESEARCH IN SOUTH AFRICA
Generating Knowledge - Impacting Health

Doris Duke Medical Research Institute (2nd floor), 719 Umbilo Road, Private Bag X7, Congella, 4013, Durban, South Africa
tel: +27 31 2604555 | fax: +27 31 2604549 | email: caprisa@caprisa.org | www.caprisa.org

06 April 2020

Professor Clem Penny
Chair: Human Research Ethics Committee (Medical)
Phillip V Tobias Health Sciences Building
29 Princess of Wales Terrace
Parktown
2193 Johannesburg
South Africa

Dear Prof Penny,

Re: Elizabeth Venter - MSc study "Identification of V3-Glycan-Specific bNAb-Initiating HIV-1 Envelopes to Inform the Design of Germline Targeting Immunogens"

As Head of HIV Pathogenesis and Vaccine Research at CAPRISA and as Co-Chair of the CAPRISA 002 Acute Infection Study, I write to confirm that Ms Elizabeth Venter will have access to stored samples from the CAPRISA 002 and 004 cohorts to carry out the above study for her MSc through the University of the Witwatersrand.

Participants in the CAPRISA 002 and 004 studies have consented to stored specimens being used for additional studies. The CAPRISA 002 BREC approved Specimen Storage Consent (V4.02) attached here, states that: "*We may test your cells, proteins, other chemicals in your body and genes (DNA). Some of the samples will also be tested to see how your nutritional status may be interacting with HIV infection. Your samples may be analysed in laboratories outside of South Africa*".

Ms Elizabeth Venter's study will be supervised by Prof Penny Moore, Dr Dale Kitchin and Dr Thandeka Moyo and represents an extension of our successful collaboration with her laboratory at the National Institute for Communicable Diseases, Johannesburg.

Please do not hesitate to contact me if there are any queries.

Yours sincerely,

Dr Nigel Garrett
MBBS, MRCP, MSc, PhD
Head of Pathogenesis and Vaccine Research, CAPRISA
Investigator of Record HVTN Studies eThekwini Clinical Research Site
Honorary Lecturer in Public Health, University of KwaZulu-Natal
Centre for AIDS Programme of Research in South Africa
University of KwaZulu-Natal
Tel: 031 260 4453/ 031 655 0617 Cell: 076 330 8300
Email: nigel.garrett@caprisa.org

Appendix B: Ethics Clearance Certificate

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R14/49 Ms E Venter; Drs T Moyo and D Kitchin

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M200443

NAME: (Principal Investigator)	Ms E Venter; Drs T Moyo and D Kitchin
DEPARTMENT:	School of Patholgy Department of Virology Medical School University
PROJECT TITLE:	Identification of V3-Glycan-specific broadly neutralizing antibody-initiating HIV-1 envelopes to inform the design of germline targeting immunogens
DATE CONSIDERED:	2020/04/24
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Professor P Moore
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2020/09/11

This clearance certificate is valid for 5 years from the date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the 3rd Floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to submit details to the Committee. **I agree to submit a yearly progress report**. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in **April** and will therefore reports and re-certification will be due early in the month of **April** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature Date

Appendix C: Turnitin Report

Elizabeth_Venter_MSc_Dissertation_Final-1.docx

ORIGINALITY REPORT

18%	15%	10%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	pure.uva.nl Internet Source	1%
2	open.uct.ac.za Internet Source	1%
3	Submitted to University of Witwatersrand Student Paper	1%
4	thesis.library.caltech.edu Internet Source	<1%
5	digital.lib.washington.edu Internet Source	<1%
6	hdl.handle.net Internet Source	<1%
7	dspace.cuni.cz Internet Source	<1%
8	www.frontiersin.org Internet Source	<1%
9	rupress.org Internet Source	<1%