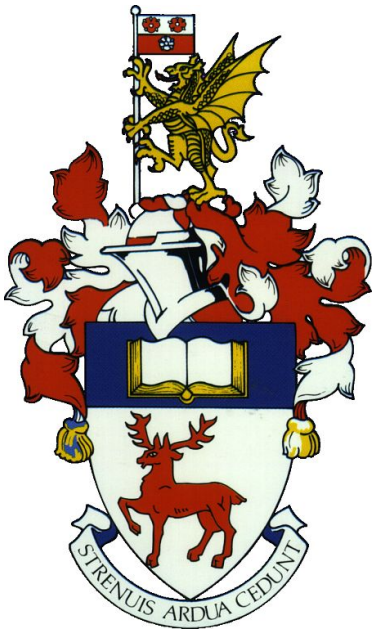
Engineering Glycosylation in HIV-1 Vaccine Design

by

Alessio D'addabbo

ORCID ID: 0000-0003-2542-6465

November 2023



Thesis for the degree of

Doctor of Philosophy

Faculty of Environmental and Life Sciences

School of Biological Sciences

University of Southampton

To my family and friends

Declaration of Authorship

Print name: **Alessio D'addabbo**

Title of thesis: **Engineering glycosylation in HIV-1 vaccine design**

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research. I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University.
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
3. Where I have consulted the published work of others, this is always clearly attributed.
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
5. I have acknowledged all main sources of help.
6. Where the thesis is based on the work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
7. Parts of this work have been contributed as:

Chapter 2: Seabright, G. E., Cottrell C. A., Gils M. J., D'addabbo A., Harvey D. J., Behrens A. J., Allen J. D., Watanabe Y., Scaringi N., Polveroni T. M., Maker A., Vasiljevic S., Val N., Sanders R. W., Ward A. B., Crispin M. Networks of HIV-1 Envelope Glycans Maintain Antibody Epitopes in the Face of Glycan Additions and Deletions. Structure. volume 28, issue 8, 897-909 (2020).

Chapter 3: Crooks E. T., Almanza F., D'addabbo A., Duggan E., Zhang J., Wagh K., Mou H., Allen J. D., Thomas A., Osawa K., Korber B. T., Tsybovsky Y., Cale E., Nolan J., Crispin M., Verkoczy L. K., Binley J. M., Engineering well-expressed, V2-immunofocusing HIV-1 envelope glycoprotein membrane trimers for use in heterologous prime-boost vaccine regimens. PLoS Pathogens. volume 17, issue 10, 1009807 (2021).

Signature: Date 27/03/2023

Acknowledgments

This thesis marks the culmination of 9 years at the University of Southampton since my arrival in 2014 as an undergraduate. As such, there are many people I would like to thank for their support throughout the course of my studies here. First and foremost, my family. Ines, Will, Gaia and Maurizio. Without you I would've never been able to reach this point. Hannah I would like to thank you as we went through so much in these last few years especially that your direct support was invaluable and any words fall short of the gratitue I owe you.

I would like to thank my supervisor, Max Crispin for giving me this opportunity as well as his academic and professional guidance. The same goes for Dr. James Binley and his group with whom it was a pleasure and privilege to work alongside. To my colleagues past and present: John (with whom I've worked alongside since my master's), Joel, Maddy, Abi, Himanshi, Yasunori, Gemma, and many others; thank you all for your help and for making this the best PhD experience possible. I also wish current newcomers Wenwen and Jacob the best of luck in their posts and hope they enjoy it as much as I did. It is my hope that this thesis a) might actually be read by someone and b) perhaps be a useful point of reference to any future members of the group, so if you're reading this then you have my gratitude as well.

Last, but certainly not least, I would like to thank some of my friends whose support was equally important. Norman and Hannah (also a Crispin group alumni), with whom I lived with for 3 years, thank you for contributing to so many good memories and being great housemates and friends. I wish you both all the very best for the future. Hamza and Ben: I really valued your friendship throughout a period of a lot of change and hope that even if my future takes me away from Southampton that we all still see each other. To my friends from undergraduate and home: Sam S, Andrew, Francois, Dave, Lewis, Will, Toby, Jake, James and Sam G; thank you all for your friendship and believe me when I say that no one is more surprised than I at actually finishing my studies here. All good things come to an end I suppose, see you around!

Table of Contents

List of Tables

List of Figures

Definitions and acronyms

- Ab: Antibody
- AIDS: Acquired immunodeficiency syndrome
- AmBic: Ammonium Bicarbonate
- API: Atmospheric pressure ionization
- B4GalNT3: β1,4-N-acetylgalactosaminyltransferase-III
- B4GalT1: β1,4-galactosyltransferase-I
- bnAb: Broadly-neutralizing antibody
- BST2: Bone marrow stromal cell antigen 2
- C1-5: Constant region 1-5
- CCR5: C-C chemokine receptor type 5
- CCS: collision cross section
- CD4: Cluster of differentiation 4
- CD4bs: CD4 binding site
- CD8: Cluster of differentiation 8
- CDR: Complementarity determining region
- cGMP: Current good manufacturing practices
- CHO: Chinese hamster ovary cells
- CLR: C-type lectin
- CMV: Cytomegalovirus
- CT: Cytoplasmic tail
- CTL: Cytotoxic T-lymphocytes
- CXCR4: CXC chemokine receptor type 4
- DC: Dendritic cells
- DC: Direct current

- DSC: Differential scanning calorimetry
- DTT: Dithiothrietol
- ELISA: Enzyme-linked immunosorbent assay
- Endo F3: Endoglycosidase F3
- Endo H: Endoglycosidase H
- eOD: Engineered outer domain
- ER: Endoplasmic reticulum
- ESCRT: Endosomal sorting complexes required for transport
- ESI: Electrospray ionization
- Fab: Fragment antigen binding
- FP: Fusion peptide
- FT: Fourier transform
- FUT6: α1,3-fucosyltransferase 6
- FUT9: α1,3-fucosyltransferase 9
- GalNAc: N-acetylgalactosamine
- GC: Gas chromatography
- GE: Glycoengineering
- gl-bnAb: Germline encoded bnAb
- GlcNAc: N-acetylglucosamine
- GnT1: N-acetylaminoglucosaminyltransferase 1
- HAART: Highly active antiretroviral therapy
- HCD: Higher energy collision dissociation
- HCDR3: Heavy chain complementarity determining region 3
- HEK: Human embryonic kidney cells
- HexNAc: N-acetylhexosamine
- HILIC: Hydrophilic interaction liquid chromatography
- HIV-1/2: Human immunodeficiency virus type 1/2

- HR1/2: Heptad repeat helices 1/2
- IFN: Interferon
- IgG: Immunoglobulin G
- IM-MS: Ion-mobility mass-spectrometry
- IMP: Intrinsic mannose patch
- IRF3: Interferon regulatory factor 3
- KI/KO: Knock-in/out
- LC: Liquid chromatography
- LC-MS: Liquid-chromatography mass-spectrometry
- m/z: mass to charge ratio
- MA: Matrix domain
- MALDI: Matrix-assisted laser desorption ionization
- MGL: Macrophage galactose-type lectin
- MHC: Major histocompatibility complex
- MPER: Membrane proximal external region
- MS/MS: Tandem mass-spectrometry
- nAb: Neutralizing antibody
- NeuAc: Sialic acid/N-acetyleuraminic acid
- NeuGc: N-glycolylneuraminic acid
- NF-κB: Nuclear factor kappa-light-chain-enhancer of activated B cells
- NGAF3: Mixture of endoF3, neuraminidase, hexoseaminidase, galactosidase
- NHP: Non-human primate
- non-nAbs: Non-neutralizing antibodies
- PAMP: Pathogen-associated molecular patterns
- PBMC: Peripheral blood mononuclear cells
- PNGaseF: Peptide N-glycosidase F
- PNGS: Potential N-linked glycosylation site

- ppGalNAcT: polypeptide N-acetylgalactosaminetransferase
- PrEP: Pre-exposure prophylaxis
- PRR: Pattern recognition receptor
- PTM: Post-translational modification
- PV: Pseudovirions/pseudoviruses
- RF: Radio frequency
- RT: Reverse transcriptase
- SDS-PAGE: Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
- SEC: Size-exclusion chromatography
- SHIV: Simian-human immunodeficiency virus
- SHPB: Serial heterologous prime-boost
- SIV: Simian immunodeficiency virus
- SIVcpz: Simian immunodeficiency virus in chimpanzees
- SIVsm: Simian immunodeficiency virus in sooty mangabeys
- SNFG: Symbol nomenclature for glycans
- ST6Gal1: β-Galactoside α2,6-sialyltransferase-I
- T/F: Transmitted/founder
- T_H: T helper cells
- T_M: T memory cells
- TAMP: Trimer-associated mannose patch
- TIC: Total ion chromatogram
- TLR: Toll like receptor
- TM: Transmembrane domain
- TMD: Transmembrane domain
- TSG101: protein tumour susceptibility gene 101
- UCA: Unmutated common ancestor
- UPLC: Ultra-performance liquid chromatography

- V1-V5: Variable loop regions 1-5
- VLP: Viral-like particles
- VNTR: variable number of tandem repeat region
- XIC: Extracted ion chromatogram

Abstract

The development of a protective HIV-1 vaccine remains a major challenge to the scientific community. The nature of the virus necessitates an immune response that is able to overcome the significant genetic diversity of HIV-1. The elicitation of broadly neutralizing antibodies offer a route to overcome this diversity. As the only viral antigen expressed on the virion surface, all bnAbs target the HIV-1 Envelope spike, and as such all candidate immunogens are based around this protein. The envelope glycoprotein is one of the most densely glycosylated proteins in nature and this dense glycan shield provides protection to conserved regions of the protein from the immune system and this is reflected by the conservation of key glycan attachment sequons despite a background of formidable sequence diversity. Broadly neutralizing antibodies exploit this conservation to allow for their broad recognition through recognizing both the protein and the glycans of Env.

As of yet, recombinant immunogens have been unable to elicit bnAbs in animal models and humans. This is due to two main facets. One is that the pathway of somatic hypermutations that are reuqired to generate bnAbs are rare, and in order to guide a glycan-binding antibody to maturation there are self-reactivity controls that need to be overcome which particularly hinder glycan-binding bnAb development. Also, off-target responses distract the immune system and result from key differences in Env glycosylation between the virus and recombinant proteins. Engineering the glycosylation of Env, therefore, represents a potential mechanism to improve recombinant immunogens towards eliciting bnAbs. There are numerous intrinsic and extrinsic properties that influence Env glycosylation, including structural constraints and enzymatic availability, that can be harnessed to alter Env glycosylation. In this thesis, multiple approaches are investigated with the goal of editing the Env glycan shield towards one that is favourable for the elicitation of bnAbs. Due to the heterogeneity of glycosylation, bespoke workflows are needed to analyse glycosylation, which is required to validate the effectiveness of different engineering approaches. To achieve this,

a methodology based around liquid chromatography-mass spectrometry (LC-MS) was used to comprehensively determine the site-specific glycosylation of Env.

First, the impact of glycan additions and deletions on the overall processing of the glycan shield was determined. Such additions and deletions are commonplace in immunogens aimed at eliciting bnAbs. This revealed that additions/deletions are well tolerated but their induction influences the processing of neighbouring glycan sites. These observations were consistent across multiple Env strains and immunogen platforms. Next, the impact of altering the availability of glycan processing enzymes on Env glycosylation was explored, demonstrating that co-expressing key enzymes in the pathway was successful in engineering the glycan shield, but these effects were unpredictable. Both of these approaches resulted in widespread alterations in the glycan shield. To limit glycan engineering to specific epitopes, the targeted epitopes of bnAbs were exploited to control glycosylation by co-transfecting Env with bnAbs, which in turn alter the availability of Env to glycan processing enzymes during glycosylation in the ER/Golgi. Finally, to alter the glycosylation of Env towards a more viral-like configuration, the gene codon usage of recombinant Env was altered to that of the native virus, which contrasted the codon optimized variants typically in use. This resulted in a decreased rate of translation, which, in turn, altered the glycan shield of recombinant Env to a more viral like state.

This thesis demonstrates the wide arsenal of approaches that can be used to change Env immunogen glycosylation, however it is difficult to predict the outcomes of such engineering without comprehensively studying the resultant immunogens. By investigating a broad range of approaches and reporting their successes and caveats it is possible for these methods to be integrated into immunogen design approaches with specific epitopes in mind. These tools may prove valuable in the design of an effective HIV-1 vaccine with an appropriate glycan profile.

1 Introduction

1.1 The human immunodeficiency virus (HIV)

HIV is the cause of acquired immunodeficiency syndrome (AIDS) and is widely acknowledged as one of the most widespread viruses across the globe. Phylogenetic and epidemiologic studies suggest that HIV originated from Sub-Saharan Africa between the years 1920-1940 where it evolved from the closely related simian immunodeficiency virus (SIV) [1]. Since then, the virus has imposed a worldwide reach and, despite most incidents still occurring within the Sub-Saharan region, continues to spread in Europe, Asia, and both North and South America with an estimated 36.9 million people living with HIV globally according to UNAIDS [2].

The identification of HIV occurred in January 1981 when Michael S. Gottlieb and colleagues observed an immunologic correlation in previously healthy homosexual men in which they expressed symptoms of fevers, weight loss, and vulnerability to rare opportunistic infections such as cytomegalovirus (CMV), pneumocystis pneumonia and Kaposi's sarcoma, conditions thought only to affect patients with a compromised or suppressed immune system.[3]

These initial findings were published in June 1981 before following it up with a more detailed report in December which found cluster of differentiation 4 positive (CD4+) helper cells virtually absent from the blood of patients; the paper also made the suggestion that a sexually transmitted disease between these homosexual patients may be the factor which causes the immunosuppression that the aforementioned opportunistic pathogens and cancers exploit [4, 5].

Thanks to intense medical research into HIV and AIDS since this identification, the prognosis of HIV has evolved from a terminal condition to a manageable condition with little effect on life span in areas with access to proper treatment. The main reason for this is the widespread usage of antiretroviral therapies [6]. The antiretroviral drugs used today are a combination

of inhibitors of various functions of the HIV virus and is known as highly active antiretroviral therapy (HAART). The development of pre-exposure prophylaxis (PrEP) methods has also allowed for the prevention of viral dissemination into non-infected people that may have come into contact with the virus. These drugs have been shown to significantly reduce mortality in patients but it is important to note that these measures only mitigate the effects of the virus. Of those previously mentioned 36.9 million people living with HIV, UNAIDS estimates that around 21.7 million were accessing antiretroviral therapy [2].

Whilst this is an impressive milestone and a credit to the research that has been undertaken, there remains much more work to be done to curb the spread. In order to reverse this global epidemic, the development of a preventative vaccine is of paramount importance, as once an infection is established the genetic variability of HIV is currently too great for a curative vaccine [7].

1.1.1 HIV classification

HIV belongs to the Lentivirus genus, family *Retroviridae*, within the subfamily *Orthoretroviridae*. It is subdivided into two biologically distinct groups, HIV-1 and HIV-2. It is believed that HIV-1 evolved from SIVs in Central African chimpanzees (SIVcpz) whilst HIV-2 evolved from West African sooty mangabeys (SIVsm) by zoonotic transmission following blood to blood contact or mucous membrane exposure to infected bodily fluids, the likely cause of which would be during bushmeat hunting [8, 9]. HIV-1 and HIV-2 differ symptomatically in that the latter has a lower virulence, transmission rate, and slower progression [10, 11].

Despite previously being thought to be confined to the western regions of Africa, HIV-2 has also begun to spread with cases being reported across Europe, India and the United States [10]. Despite this, HIV-1 accounts for the majority of cases worldwide and will therefore be the focus of this work. HIV-1 is divided further into groups M, N, O, and P of which M, dubbed “major”, comprises the majority of infections [12, 13, 14]. Within group M exist subtypes, or clades, categorised from A-K, the most prevalent of which being subtype C

which accounts for almost 50% of all global HIV-1 infections and is predominant in countries with over 80% of all global infections such as the southern regions of Africa [15].

1.1.2 HIV structure and genome

The HIV genome is comprised of two single, positive-stranded RNA molecules enclosed within the capsid core of the virion, which itself is around 100-120 nm in size [16]. The main components of this genome are the *gag, pol,* and *env* genes. The *gag* gene encodes the structural components of the virus vital for retaining the integrity of the virus: the viral matrix, the capsid, and the nucleocapsid. The *pol* gene encodes the viral enzymes protease, reverse transcriptase (RT), and integrase.

Because HIV translates its genes as polyproteins, the HIV protease is used to cleave the *pol* encoded polyproteins into individual functional proteins. RT contains two domains involved in the formation of double stranded HIV DNA: a polymerase domain and an RNAse H domain. The polymerase domain is responsible for generating DNA from the HIV RNA and the RNAse H domain is responsible for the following degradation of the positive-stranded RNA molecule to allow for the formation of double stranded HIV DNA by the polymerase domain [17]. The integrase enzyme is then responsible for incorporating this DNA into the host cell genome. The *env* gene encodes the surface transmembrane protein Env, which is embedded into the lipid bilayer of the host cell following viral budding and release. It first encodes the proprotein gp160 which, following protease cleavage by the ubiquitous human protein furin, forms a trimer of gp120-gp41 heterodimers. During the course of its translation Env undergoes heavy N-linked glycosylation which comprises around half of the mass of gp120 and gp41 [18]. The function of Env is to mediate viral entry into host cells and is highly antigenic. It is the sole target on the virion for antibodies and therefore is the focus of vaccine design. Figure 1.1 demonstrates the structure and genomic arrangement of HIV-1.

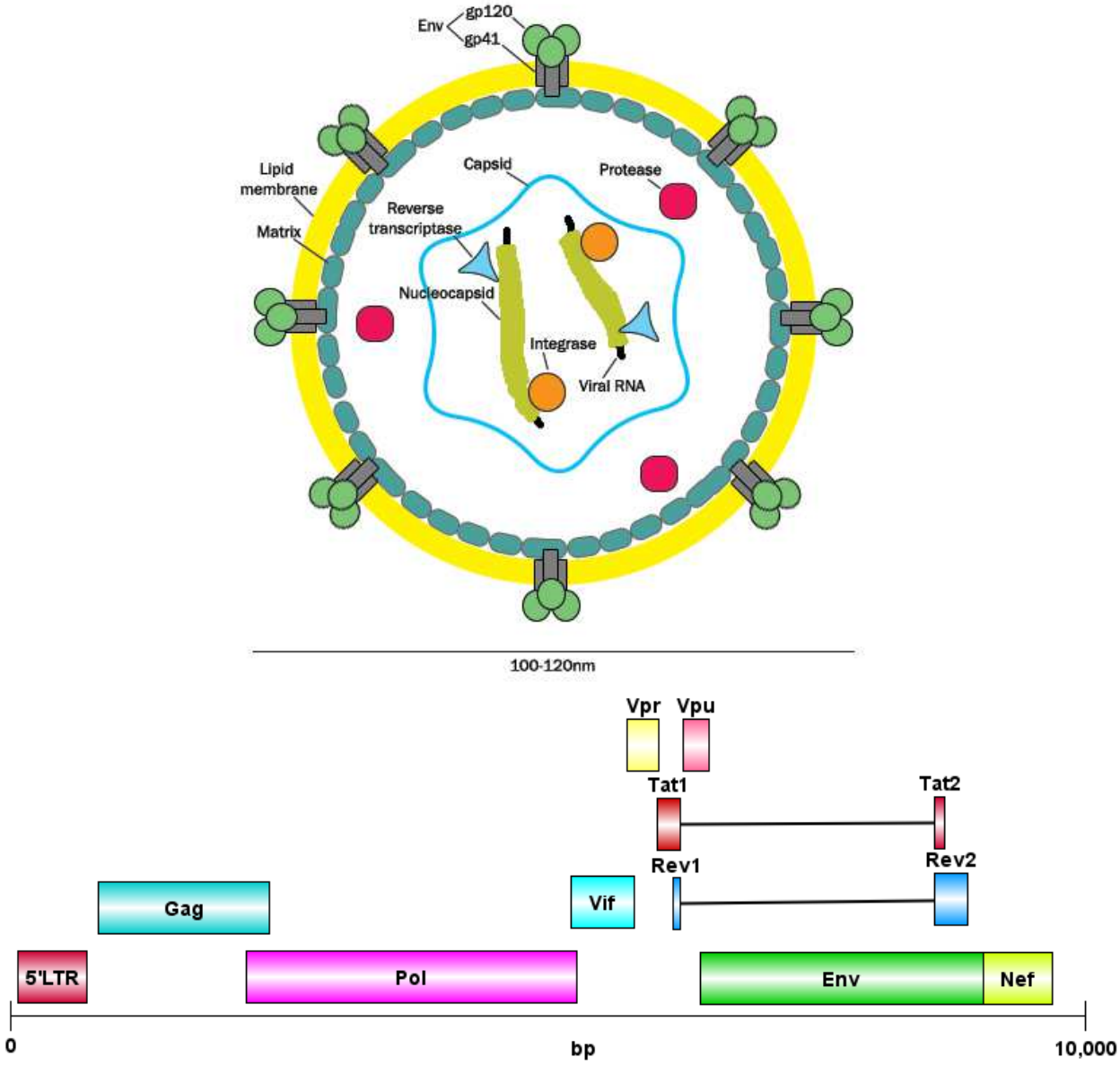


Figure 1.1: **Structure and genomic arrangement of HIV-1**
(Upper) Virions are formed from budding of particles from a host cell membrane around 100-120 nm in diameter [16]. Within, are two copies of the positive-sense ssRNA genome enclosed in nucleocapsid proteins within the cone-shaped capsid that also contains the enzymes integrase and reverse transcriptase. This structure is reinforced by a surrounding matrix which ensures structural integrity of the virion, with the HIV protease in-between the capsid and the matrix. Each HIV-1 virion contains around 14 Env spikes, which is made up of a trimer of heterodimers of gp120 and gp41 [19].
(Lower) The HIV-1 single-stranded RNA genome of around 10,000 kb is made up of several open reading frames (ORFs). The *gag* gene (blue) encodes the matrix (MA), capsid (C) and nucleocapsid (NC) proteins. The *pol* gene (magenta) encodes the viral protease (PR), reverse transcriptase (RT), and integrase (IN) proteins. PR also cleaves these aforementioned proteins. The *env* gene (green) encodes the gp120 and gp41 envelope glycoproteins. The regulatory proteins Tat and Rev, and the accessory proteins Vpu, Vpr, Vif and Nef are shown. Sequence information was obtained using NCBI Genome Workbench software version 2.13.0.

1.1.3 HIV entry and viral life cycle

The primary target for Env, and therefore HIV entry, is the CD4 receptor, which is found on CD4+ lymphocytes (aka T-helper or T_H cells), macrophages, and dendritic cells. Viral entry also requires co-receptor stimulation from either C-C chemokine receptor type 5 (CCR5) or CXC-chemokine receptor type 4 (CXCR4), with the former mediating the majority of co-receptor interactions that do not elicit structural changes in gp120 or gp41 themselves but bring Env closer to the membrane and CD4 to facilitate this process [20]. The gp120 subunits mediate binding to CD4 receptors which, in turn, elicit conformational changes in Env that drives co-receptor binding which is partly mediated by the variable loop 2 and 3 (V2, V3) of Env as well as fusion peptide insertion of gp41 (Figure 1.5) [21]. Fusion peptide insertion initiates the process of membrane fusion and is followed by the formation of a six-helix bundle made up of two helices of each subunit. The viral capsid is now released and uncoated in the host cell. In the cytoplasm, the viral RT produces double-stranded DNA from the viral RNA as previously described. The viral DNA forms part of the pre-integration complex, which also consists of integrase, protease, matrix, and Vpr [22]. This complex translocates into the nucleus and integrates the viral DNA into the host cell genome by the action of integrase. Once integrated, the host cell mediates the transcription and translation of the viral DNA which allows for the production of new virions. Once fully assembled, viral release occurs by budding from the host cell membrane. The viral protein Vpu aids viral release by antagonising the glycoprotein bone marrow stromal cell antigen 2 (BST2) which inhibits viral release by tethering viral particles to the cell membrane [23]. Protease cleavage at sites in Gag and Gag-Pol polyproteins lead to the production of a mature condensed viral core and hence mature viral particles [24]. In an alternate series of events, HIV also possesses the ability to replicate inactively within T memory (T_M) cells and thus persist in patients irrespective of antiretroviral treatments [25].

1.1.4 Immune response to HIV

The first challenge that HIV encounters upon infection is the innate immune response. Pattern recognition receptors (PRRs), such as Toll like receptors (TLRs), on host cells are able to respond to viral pathogen-associated molecular patterns (PAMPs). This leads to activation of signalling cascades that stimulate the nuclear translocation of interferon regulatory factor 3 (IRF3) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB), ultimately resulting in the production of type I interferons (IFN) and thereby the expression of antiviral and production of pro-inflammatory cytokines to create an environment to suppress infection and promote the process of adaptive immunity [26]. HIV limits these pathways by interfering with TLR signalling, inhibiting NF-κB by the action of accessory protein Vpu, and through the shielding of viral nucleic acids within the cytoplasm until it is translocated into the nucleus [27, 28].

Following sexual exposure, one of the initial targets in HIV infection are dendritic cells (DCs), specifically Langerhans cells which are present both in the skin and mucosal epithelium [29]. As the HIV Env glycoprotein is heavily glycosylated, it is recognised by PRRs on these cells known as C-type lectin receptors (CLRs). CLRs are able to specifically bind and recognise glycan moieties on Env through their carbohydrate recognition motifs, a conserved feature amongst this group of receptors and the tool in which these receptors mediate their affinities [30].

The CLRs expressed on DCs include: Mannose receptor, DC-SIGN, Macrophage galactose-type lectin (MGL), dectin-1, and langerin. Their functions are to mediate pathogen recognition and antigen internalisation for antigen presentation [30]. Of these receptors, DC-SIGN and langerin recognise and bind to the abundant, oligomannose glycans on Env. As DC-SIGN is expressed on DCs and macrophages, and langerin on Langerhans cells, the outcome of Env binding is cell-context dependent. HIV binding to DC-SIGN is also believed to result in host infection and transmission to T-cells [31]. Conversely, langerin is believed to pre-

vent HIV transmission through viral internalisation and degradation within Birbeck granules within Langerhans cells [32].

The interactions between dendritic cells and HIV mediate both the adaptive immune response and DC-mediated transmission of HIV into target cells. There are two ways in which DCs allow for the propagation of HIV, namely trans-infection and cis-infection. During trans-infection, the virus is located within endosomes which concentrate at contact zones between DCs and CD4+ T-cells (known as the virological synapse) [33]. Viral transfer occurs via the transfer of these infectious vesicles across the synapse. Cis-infection is instead characterised by the infection of dendritic cells and the CD4 and co-receptor mediated dissemination of progeny virions to CD4+ T-cells [34].

As HIV-1 viremia peaks, the initial CD8+ T-cell responses appear and peak within around two weeks before contracting in line with falling virus load [35, 36]. The selective pressure that CD8+ T-cells impose on HIV-1 combined with its intrinsic mutability causes the selection of viral escape mutants within 1-2 months after peak viremia and throughout infection [37, 38, 39, 40]. Figure 1.2 illustrates the balance between viral load and CD4+ T-cell count throughout the course of an infection. The HIV-1 protein Nef also contributes to escape from both adaptive and innate immunity through downregulation of major histocompatibility (MHC) class I molecules (alongside the HIV proteins vpu and tat) which interact with the CD8 receptor. Nef is also involved in downregulating HLA-A and HLA-B expression on the surface of HIV-1 infected cells, which in turn enables it to evade cytotoxic T-lymphocytes (CTL) responses from both the adaptive immune system (CD8+ T-cells) as well as the innate immune system (natural killer cells) respectively [41, 42].

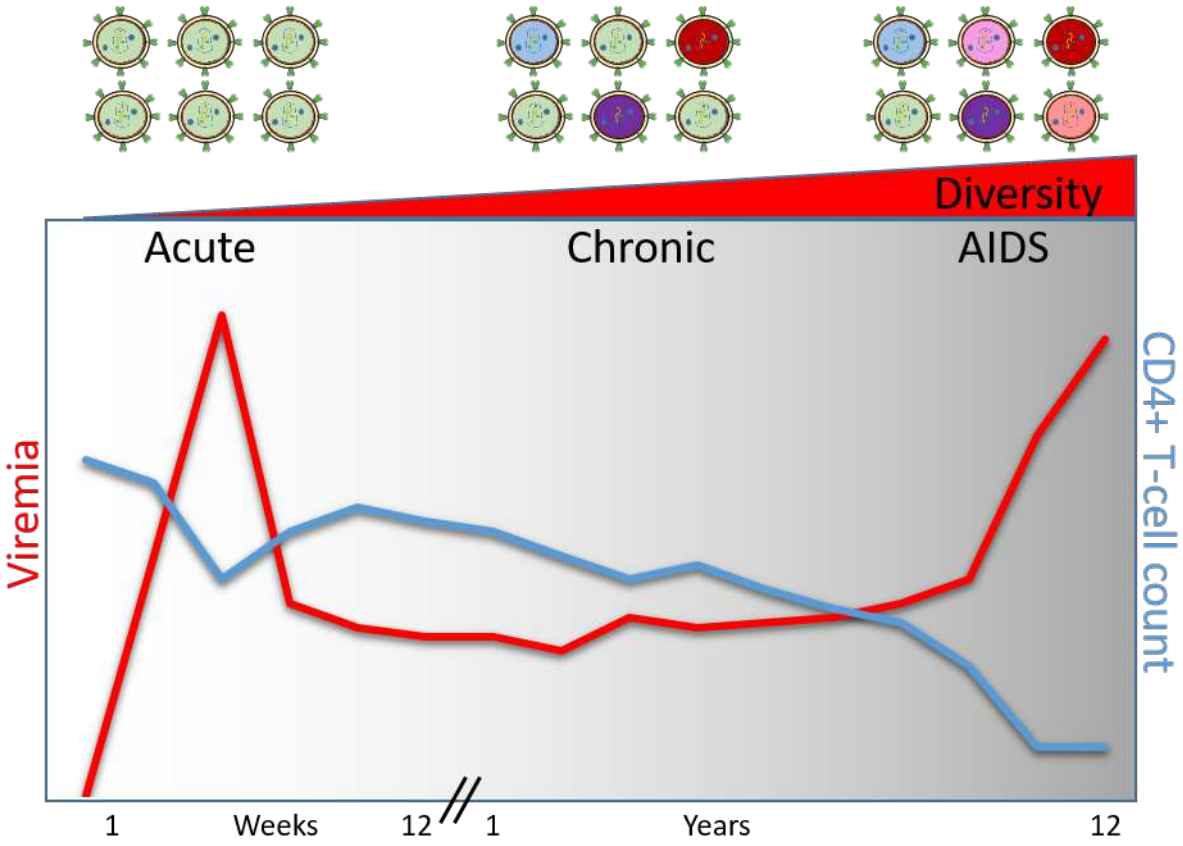


Figure 1.2: **HIV-1 viremia and CD4+ T-cell count throughout infection**
Plasma viremia and CD4+ T-cell counts throughout the average course of HIV-1 infection. As HIV-1 viremia first peaks, initial CD8+ T-cell responses appear which coincides with falling virus load. The selective pressure CD8+ T-cells exert on HIV-1 facilitates the formation of viral escape mutants which eventually causes a resurgence of viral load as the disease progresses to AIDS. Adapted from [36].

B-cell responses commence within a week of detectable levels of HIV-1 and are detectable in the form of antigen-antibody complexes [43]. Then, antibody responses initially targeting gp41 arise within two weeks of infection whilst anti-gp120 antibodies appear four weeks post-infection, mainly targeting the V3 loop [44]. These initial antibody responses are non-neutralising however, which is explained by the lack of the effect of these antibodies on viremia or selective pressure on Env [43, 44]. A neutralising antibody (nAb) response follows 3-5 months post-infection and have been found to target exposed variable regions such as the V1/V2 loop and the constant C3-V4 region [45, 46, 47, 48]. Of particular note with some early nAb responses is that they have been shown to target exposed protein epitopes

lacking in glycans (otherwise known as "glycan holes") [49]. This concept is elaborated in section 1.4.2.2.

In the virus, these holes form immunodominant epitopes which confer a strong selective pressure on HIV; however, the variable nature of the glycan shield allows for relatively simple viral escape by mutations that can cause a change in glycan position to mask these epitopes at a low fitness cost, which would render any antibody against this epitope ineffective. The evasion of the immune response by the glycan shield is discussed in section 1.2.8.

Following viral escape, other antibodies develop to target escape variants thus engaging a game of cat and mouse between HIV escape mutants and the humoral immune response. The interplay between antibodies and HIV-1 neutralization has produced a form of classification whereby HIV-1 isolates are ranked on their conformation and ability to be neutralized. These neutralization tiers correspond to the frequency of Env in a closed (Tier-2 or 3), open (Tier-1A) or intermediate (Tier-1B) conformation [50]. As the Env trimer conformation opens, more epitopes are exposed and is therefore more liable to neutralization by antibodies; The most common circulating conformation in a patient is Tier-2, a predominantly closed Env trimer configuration and one that most circulating antibodies are unable to neutralize [50]. In summary, both the innate and humoral immune response is insufficient in dealing with HIV-1 before the transition to a latent stage and eventually AIDS.

1.1.5 Broadly neutralizing antibodies

In the 1980s, the first anti-HIV antibodies were isolated from infected patient sera [51]. When these sera were tested as therapies these antibodies proved highly ineffective as as they were unable to cope against the high mutability of Env [52, 53]. Crucially, studies involving infected patient sera demonstrated that a small subset of HIV-1+ patients developed antibodies that could target shared epitopes across antigenically different strains [54, 55]. This led to the concept of broadly neutralising antibodies.

Broadly neutralising antibodies (bnAbs) are a class of antibody that possess the ability to

neutralize multiple, genetically divergent strains of HIV-1 by targeting conserved epitopes that can persist following mutation. This ability is known as breadth, and whilst bnAbs can evolve within patients, this usually occurs in around 20% of people following two or more years of infection [56]. This breadth localises to areas with a requirement for sequence homology in order to achieve its intrinsic functions, for example the CD4 binding site (CD4bs) [57]. This was a significant finding as if these antibodies could bind and neutralize multiple Env strains, the ability to induce them in a vaccine could prove as an effective prophylactic measure against HIV.

The first generation of bnAbs were isolated from the screening of many antibodies (Abs) cloned from anti-Env memory B-cells, and showed promise when used in challenge studies using HIV in humanized mouse models [58, 59] and simian-human immunodeficiency virus (SHIV) challenge in non-human primates (NHPs) [60, 61, 62]. Innovations in high-throughput screening of individual B-cells pioneered by the Nussenzweig group yielded a new generation of more potent bnAbs that could be selected for their ability to bind Tier-2 Env strains prior to their characterization [63, 64, 65]. Despite this, the challenge remained in finding a candidate for a vaccine that could elicit the production of these antibodies, as until this point they could only be achieved after infection. In order to understand how to achieve bnAb elicitation, knowledge of the fine structure and biosynthesis of Env is necessary in order to contextualize it as an antigen and vaccine candidate.

1.2 The Viral Envelope Glycoprotein, Env

1.2.1 Env structure and function

The trimeric envelope glycoprotein Env is composed of a dimer of gp120 and gp41 trimers. The former is highly variable and contains five variable loop domains as well as five constant regions; the latter is more conserved and its structure consists of a fusion peptide, heptad repeat helices 1 and 2 (HR1/2), a membrane proximal external region (MPER), a transmem-

brane domain (TM), and a cytoplasmic tail (CT). Figure 1.3 shows the structure and domain architecture of Env. Despite difficulties in determining structure, including but not limited to: the high level of glycosylation, the gp120-gp41 interactions, and the nature of variable loops; studies have successfully been published showing X-ray and cryo-EM structures of Env. The first structures to be elucidated often consisted of monomeric gp120 subunits binding the CD4 co-receptor [66] but; with the passing of time and the improvement of techniques, multiple groups have since published higher resolution structures that interrogated the finer structural and mechanistic aspects of Env in its native form and when binding CD4 and its co-receptors, most recently in 2019, where the laboratory of Professor Bing Chen published a 3.9 Å resolution structure of full length gp120 in complex with CD4 and CCR5 [20, 67, 68]. From this structure they were able to observe the interactions between the V3 loop of gp120 and the chemokine binding pocket of CCR5, and the CCR5-CD4 interactions that help bring gp120 closer to the membrane. Whilst CCR5 itself doesn't facilitate any structural changes in gp120 or gp41, it likely stabilises and anchors the CD4-induced conformation of Env in proximity of the membrane to help kickstart the structural changes that lead to viral entry.

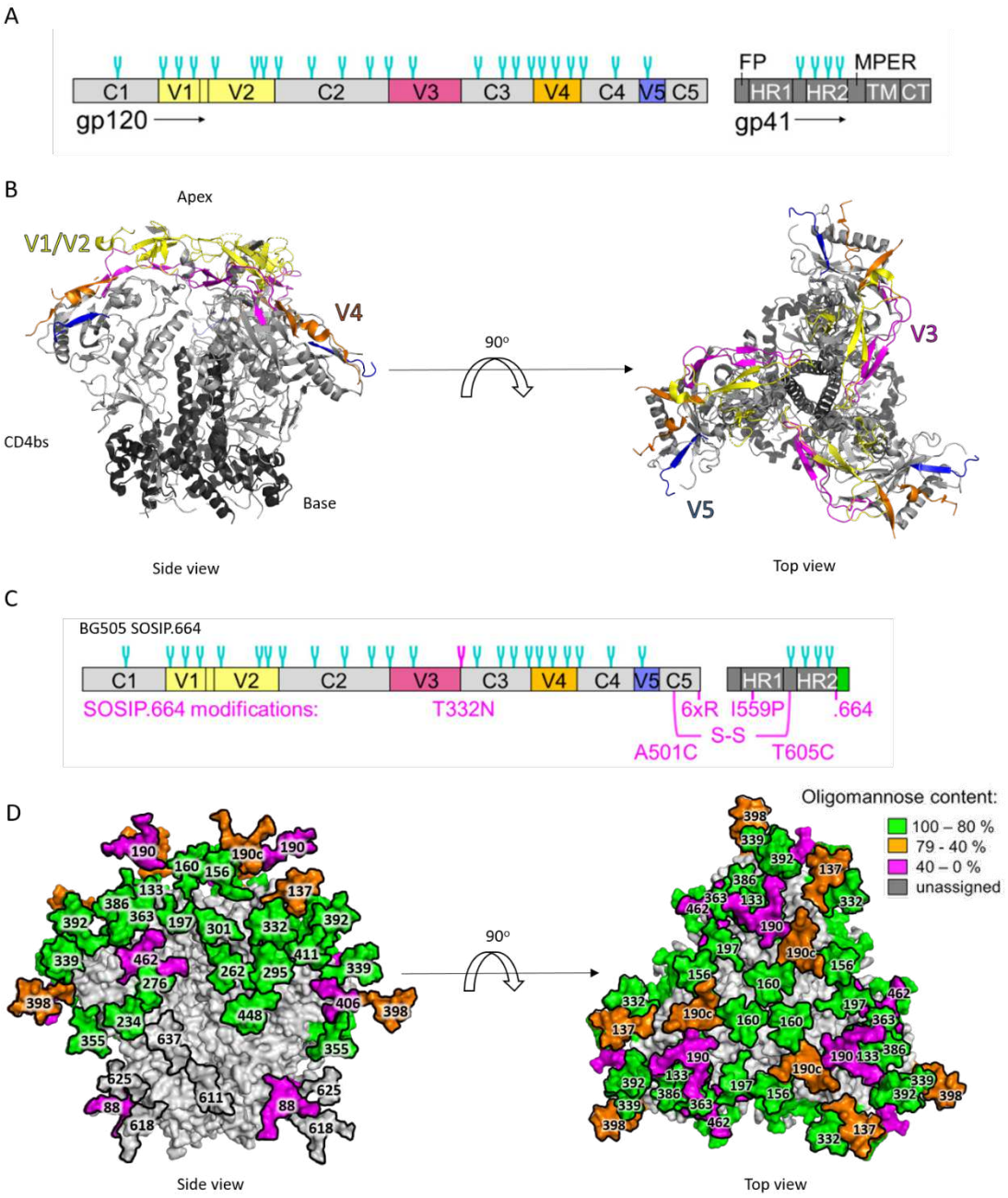


Figure 1.3: **Structure and domain representation of Env and the BG505 SOSIP.664 trimer**
(A) Primary structure of HIV-1 gp120 and gp41; the constant regions (C1-C5) are coloured in grey whilst the variable regions (V1-V5) are yellow, magenta, orange, and navy respectively. gp41 domains are dark grey; FP (fusion peptide), HR1 (heptad repeat 1), HR2 (heptad repeat 2), MPER (membrane proximal external region), TM (transmembrane domain), CT (cytoplasmic tail). **(B)** Structure of the BG505 SOSIP.664 trimer in its closed pre-fusion state. Model made in PyMOL using PDB ID: 58VL [68]. **(C)** Primary structure of BG505 SOSIP.664 including PNGSs and modifications. **(D)** Model of a fully glycosylated BG505 SOSIP.664 trimer, adapted from Behrens et al., 2016 and Behrens et al., 2017 [69, 70]. Model was made in PyMOL and Coot using PBD ID: 5ACO [71]

Whilst co-receptor binding does not induce any structural changes in the structure of Env; CD4 binding induces large structural changes as well as significant changes in entropy that allude to a "considerable conformational flexibility within gp120" [72]. Changes include the formation of a co-receptor binding site made up of a four stranded antiparallel β-sheet, otherwise known as the bridging sheet, and the V3 loop [73]. Together with co-receptor binding which causes Env to move closer to the membrane, the insertion of the N-terminal fusion peptide of gp41 into the target cell membrane is triggered, shortly followed by the rearrangements of the HR1/2 helices forming an intermediate within gp41 before forming a six-helix bundle that drives the fusion of the two membranes and viral entry [74]. Figure 1.4 shows a model whereby Env-mediated membrane fusion occurs.

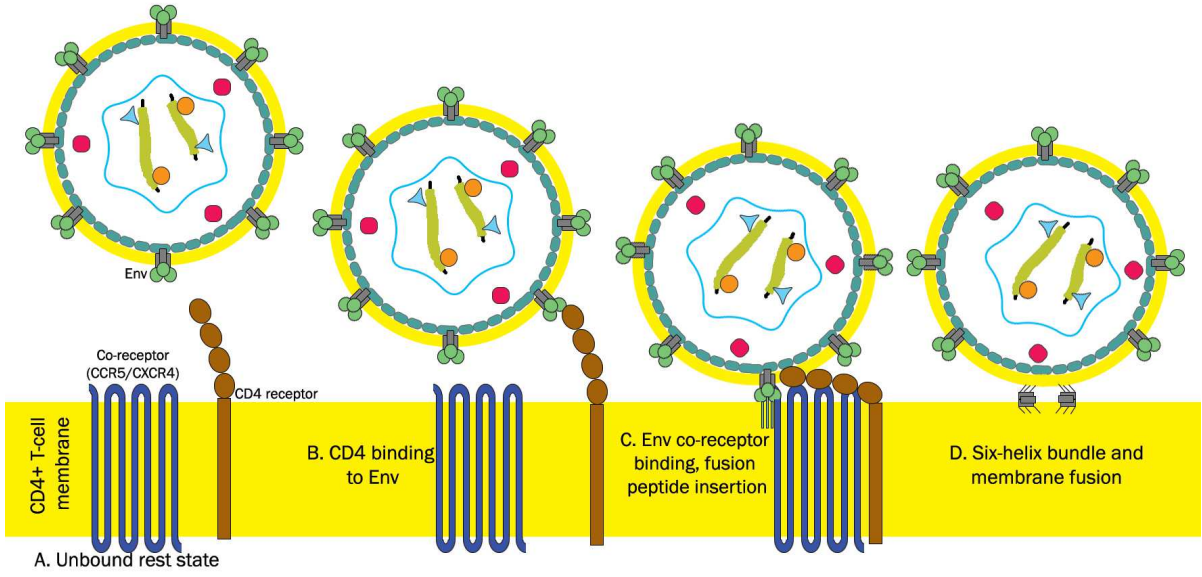


Figure 1.4: **Model of HIV-1 membrane fusion**
(A) The HIV-1 Envelope trimer is anchored in the viral membrane and mediates fusion with the cell membrane of CD4+ T-cells. In the unbound state, Env is in a closed conformation. **(B)** Upon CD4 binding with a gp120 subunit, conformational changes occur that form and stabilise the co-receptor binding site. **(C)** The Env trimer is pulled by CD4 bending to the host cell membrane where it comes into contact with a co-receptor (either CCR5 or CXCR4). Co-receptor binding initiates insertion of the fusion peptide of gp41 into the host cell membrane. **(D)** The gp41 subunits shed their interactions with gp120 and form a 6 helix bundle, made up of two helices of each subunit. Membrane fusion and subsequent viral entry into the host cell occurs.

1.2.2 Env biosynthesis

During translation, gp160 is targeted to the endoplasmic reticulum (ER) membrane by a signal peptide located on its N-terminus which is co-translationally removed by signal peptidases, leaving the transmembrane domain (TMD) of gp41 anchored in the lumen of the ER. This TMD contains a stop-transfer sequence that prevents it from its full release into the ER lumen [75]. In the ER, the monomeric components of gp160 assemble into its active trimeric structure, as well as some non-functional dimers and tetramers [76].

Env is also subject to extensive glycosylation concomitantly with translation, which adds many N-linked glycans to both gp120 and 41 subunits at Asn-X-Ser/Thr sequons, whereby X is any amino acid except proline [77]. The position of N-glycan sites are determined by the protein sequence and, by extension, the HIV-1 genome; however, the structure of the N-glycan itself is also influenced by the host cell machinery and their relevant enzymes, which is cell-context dependent [78].

Mucin-type O-linked glycans are attached to serine, threonine, and tyrosine residues by the action of polypeptide N-acetylgalactosaminetransferases (ppGalNAcT),and form what is known as a "Tn antigen" [79, 80]. Unlike N-glycosylation, O-linked glycans do not possess a sequon for their attachment, but their presence can be detected in mucin-like domains through the examination of levels of serine, threonine and proline residues [80, 81, 82]. The presence of O-linked glycans on Env is a contested topic and is addressed further in section 1.2.7. In either case, as Env contains several glycan attachemnt sites and its production is derived from the infected host cell, an explanation of mammalian N-linked glycosylation is necessary to understand what differentiates this protein and its glycans from otherwise 'self' glycoproteins. This is discussed in Section 1.2.3.

The mature and active viral spike is generated following cleavage of gp160 to generate gp41 and gp120 subunits (otherwise known as gp140) by the host cell protease, furin, at the C-terminal consensus REKR sequence [83]. It is believed that furin is primarily localised in the

trans-Golgi network, though it is also able to localise to the cell membrane in some instances [84, 85].

Once cleaved and in its proper form, gp140 is transported to the plasma membrane where it is associated with Gag and integrated into new virions [86]. The mechanism by which gp140 is incorporated is still a matter for debate, with four differing models proposed; the passive model which proposes that new virions bud from the membrane and carry with it any Env on the membrane at the given location [87]. The evidence to support this claim is that assembly of a virion with a foreign viral glycoprotein occurs readily [88, 89].

The second, direct Gag-Env interaction model proposes that the motifs within the matrix (MA) domain of Gag and the cytoplasmic tail of gp41 contain interaction domains that influence Env incorporation into virions [89]. The third model, known as the Gag-Env co-targeting model suggests that incorporation is enhanced by the targeting of Gag and Env to lipid raft-like regions of the plasma membrane [90]. Lastly, the indirect Gag-Env interaction model postulates that an adaptor protein binds both Gag and Env to facilitate the formation of viral particles [91]. All models excluding the passive model propose the involvement of the MA domain and the cytoplasmic tail of gp41 in this process [75, 86, 87, 89, 91].

Following incorporation and assembly, new virions must undergo a fission event at the membrane in order to be released from the cell surface; for this to occur, the cellular endosomal sorting complexes required for transport (ESCRT) machinery is hijacked by Gag [86]. The mechanism by which Gag hijacks the ESCRT machinery, and that by which the ESCRT machinery drives membrane fission is yet to be clearly defined. Some of the first observations on viral release found that deletion of the late domain p6 on the C-terminus of Gag resulted in blocking of virion release and accumulation of particles attached to the cell membrane [92]. Subsequent analysis of this domain mapped the release function of Gag to a conserved P(S/T)AP motif close to the N-terminus of the p6 domain [93].

In HIV-1, this late domain was found to interact with the protein tumour susceptibility gene 101 (TSG101), which forms part of ESCRT-I [94, 95, 96]. The ESCRT machinery itself is

composed of several multi-protein complexes (around 20) and has crucial roles in mediating a variety of membrane-scission processes within the cell [86, 97]. Ubiquitinylation of proteins is often a requirement for interaction with ESCRT as many of its constituents contain ubiquitin-binding domains, and a link has since been established between Gag ubiquitinylation, viral late domains and ESCRT recruitment [98].

1.2.3 Mammalian N-linked glycosylation

N-linked glycans are made up of monosaccharide carbohydrate units and are categorised by the number of carbons. As constituents of N-glycans, monosaccharides are typically represented as 5-6 carbon rings held together by an oxygen atom. These ring structures are known as furanose (5) or pyranose (6). The composition and arrangement of side groups on these rings define their nomenclature, the latter additionally defining isomers of monosaccharides. 6 carbon monosaccharides are also known as hexose sugars which comprise mannose, glucose and galactose. Figure 1.5A demonstrates the structures of common monosaccharides presented on N-linked glycans.

Glucose and mannose only differ by the orientation of the hydroxyl group at the position 2 carbon. Monosaccharides may also carry additional modifications such as the addition of N-acetylamine groups to the position 2 carbons of glucose or galactose to make N-acetylglucosamine (GlcNAc) or N-acetylgalactosamine (GalNAc). These sugars can be referred to generally as N-acetylhexosamine (HexNAc) groups. Glycosidic bonds between monosaccharides are created through a condensation reaction between hydroxyl groups on the carbon rings. The precise location on the carbon rings is important and is represented in glycosidic bond nomenclature. For example, the linkage between mannose and N-acetylglucosamine can occur between the hydroxyl group at carbon 1 of the mannose and the hydroxyl group at position 4 on the GlcNAc to make a β1-4 glycosidic bond, as seen in figure 1.5B. To simplify the reporting of these structures, the symbol nomenclature for

glycans (SNFG) [99], as demonstrated in Figure 1.5A and C, allow for a simple way to condense information about specific structures and linkages.

Other modifications include the addition of a negatively charged side chain found in sialic acid (NeuAc) residues. In addition to the composition and arrangement of the monosaccharides the linkages between monosaccharide units can also have implications with regards to enzyme specificity. Many enzymes are specific to certain monosaccharides and bonds and mammalian cells additionally have enzymes with overlapping specificities that are only expressed in certain cell subtypes, some of which are summarised in Figure 1.5D.

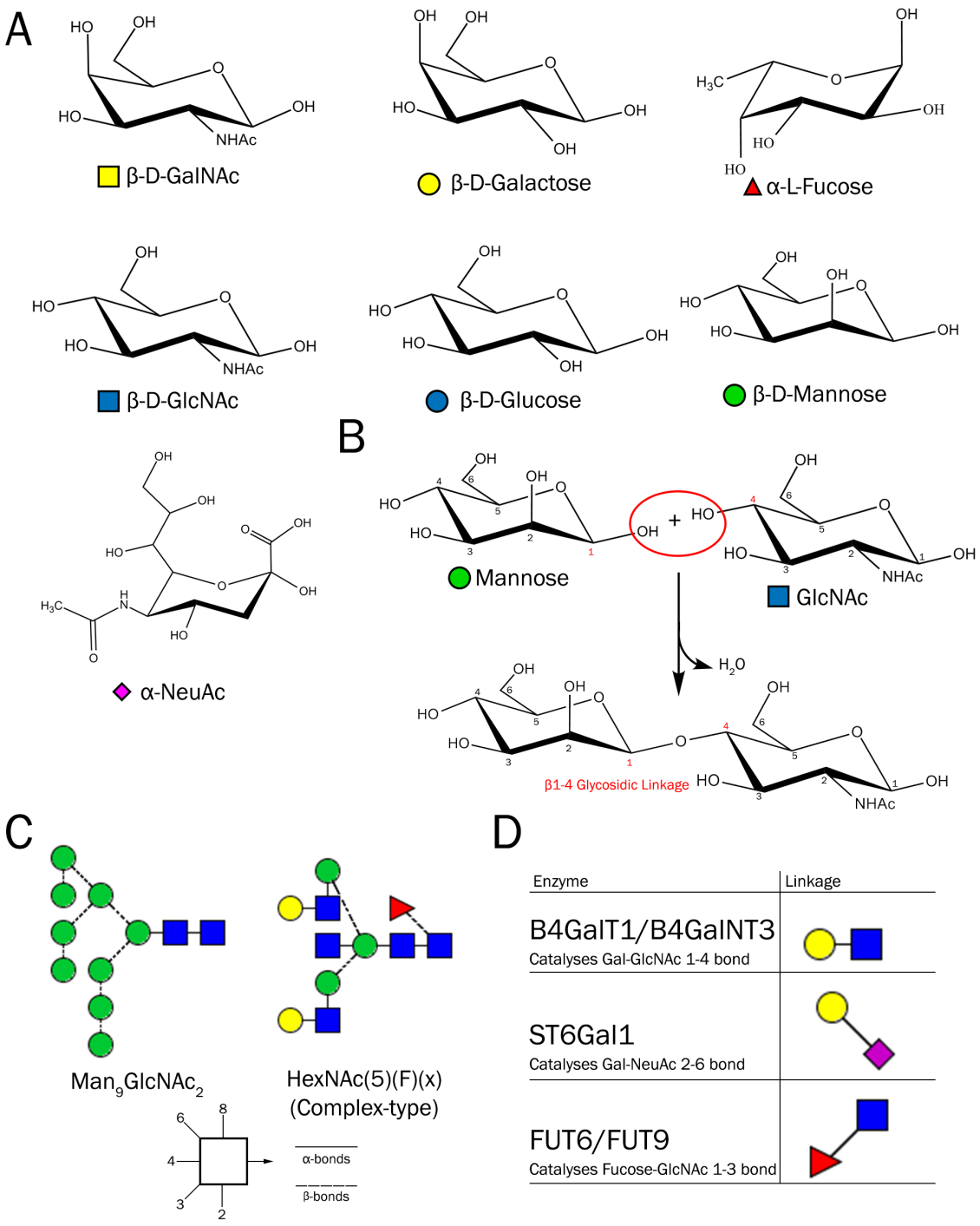


Figure 1.5: **Monosaccharides, glycosidic bonds and how they are represented**
(A) Common monosaccharide structures frequently found on N-linked glycans. Chemical structures are shown above the names of the corresponding monosaccharides and SNFG symbols [99].
(B) Formation of a glycosidic bond between a mannose and N-acetylglucosamine monosaccharide.
(C) An example oligomannose (left) and complex-type (right) glycan displayed with SNFG symbols and Oxford bond angles [99].
(D) Examples of overlapping glycosidases and their enzymatic specificities.

The mammalian N-linked glycosylation pathway (as shown in Figure 1.6) begins with en bloc transfer of $Glc_3Man_9GlcNAc_2$ from dolichol pyrophosphate to the amide group of an asparagine within Asn-X-Ser/Thr sequons at potential N-glycosylation sites (PNGSs) by the action of oligosaccharyl transferase. ER resident α-glycosidases I and II cleave the glucose residues to leave $GlcMan_9GlcNAc_2$ which, following protein folding and interactions with calnexin and calreticulin, has a further mannose residue removed by ER α-mannosidase and is then allowed to leave the ER.

If the protein is recognised as misfolded in the calnexin/calreticulin cycle however, it is re-glucosylated by UDP-glycose glucosyltransferase and subject to further rounds of folding aided by these chaperones. Alternatively, the protein is degraded via the ER degradation pathway. Gp160 folding within the ER has been found to be much slower than other proteins due to extensive isomerisation of disulphide bonds that ensure its release from the ER in the correct native form; this process is suggested to be intertwined with the cleavage of its signal sequence which itself downregulates protein folding [100, 101].

Once a glycoprotein is correctly folded and its glycans trimmed to $Man_9GlcNAc_2$, it is able to move through the Golgi network, where the N-linked glycans are further trimmed by ER and Golgi α-mannosidases. One of the many peculiarities of Env that distinguishes it from other glycoproteins observed in humans is the limited N-glycan trimming at certain sites which gives rise to the $Man_{5-9}GlcNAc_2$ series, otherwise known as oligomannose glycans (Figure 1.6B) [102]. The presence of these species and their implications will be discussed in section 1.2.7.

The glycans that are processed by ER and Golgi α-mannosidases are then able to undergo a variety of modifications and further processing by Golgi glycosyltransferases and glycosidases that can give rise to complex- and hybrid-glycans. As previously mentioned, these processes and the various enzymes involved are cell context dependent and as such can give rise to microheterogeneity; whereby the same glycan site may possess a different glycoform.

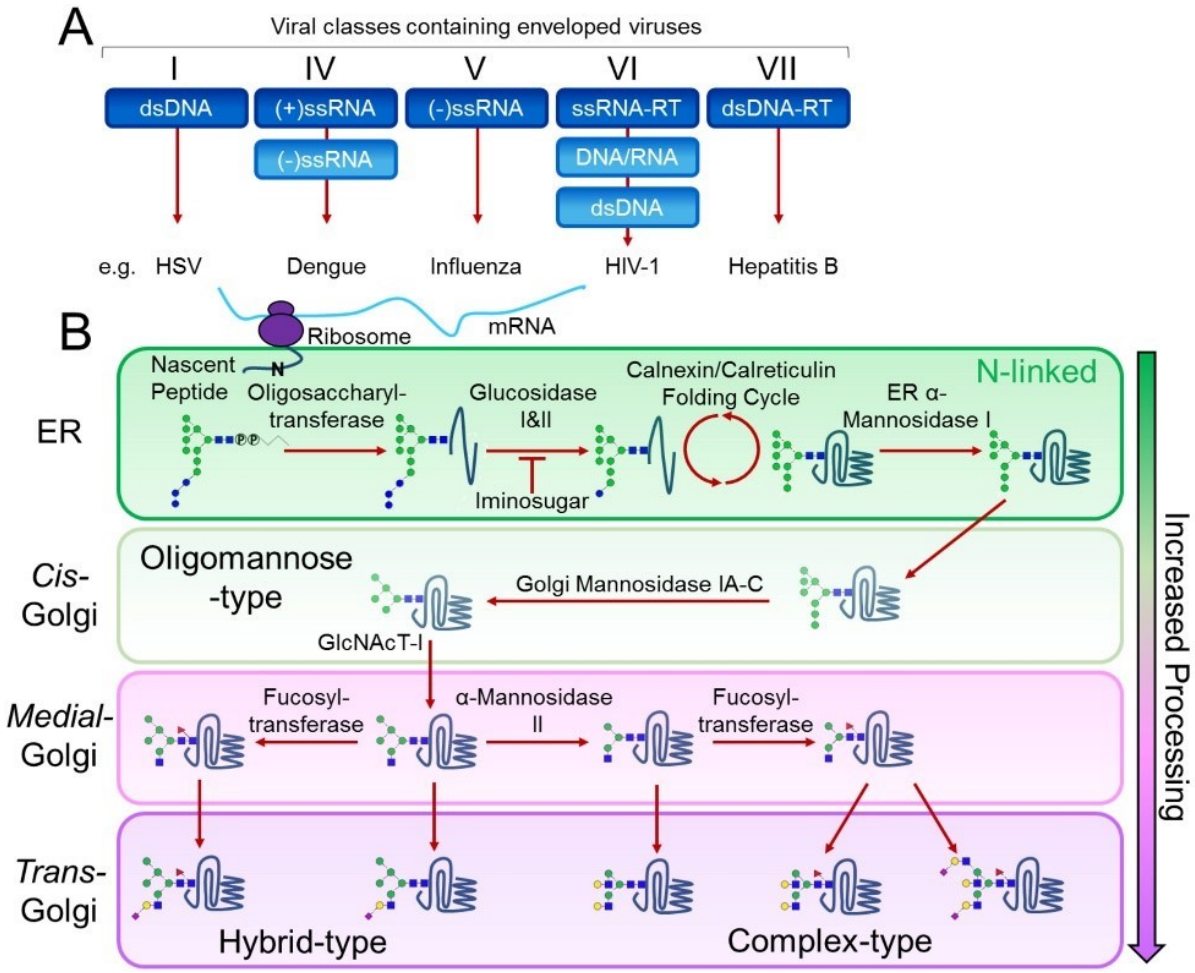


Figure 1.6: **The Mammalian N-linked Pathway**

(A) Viral classes that contain enveloped viruses including examples.

(B) The co-translational process of N-glycosylation begins with en bloc transfer of $Glcn_3Man_9GlcNAc_2$ from dolichol pyrophosphate to the amide group of an asparagine within Asn-X-Ser/Thr sequons at potential N-glycosylation sites (PNGSs) of nascent peptides by the action of oligosaccharyl transferase. ER resident α-glycosidases I and II cleave the glucose residues to leave $GlcMan_9GlcNAcn_2$ which then enters the calnexin/calreticulin folding cycle. ER and Golgi mannosidases then cleave mannose residues down to $Man_5GlcNAc_2$. GlcNAc transferase I then initiates the first branch of the N-glycan through the addition of a GlcNAc residue to the D1 arm of $Man_5GlcNAc_2$. Fucosyl-trasnferases can then act on $GlcNAcMan_5GlcNAc_2$ to create a range of hybrid type glycans. Alternatively, α-mannosidase II removes the two outermost mannose residues to then allow the action of various fucosyl, galactosyl, and sialyl-transferases to act on the glycan and produce a wide range of complex-type glycans. Glycans symbols are shown in accordance with the Consortium for Functional Glycomics symbolic nomenclature and Oxford system linkages [99], as per the key (top right). Figure adapted from Watanabe et al., 2019 [80].

1.2.4 Analysis of N-linked glycans

Analysis of N-linked glycans by mass spectrometry or ultra-performance liquid chromatography (UPLC) is multifaceted and allows for the analysis of released glycans as well as glycopeptides. Released glycan analysis comprises the analysis of released N-linked glycans following the treatment of Env with peptide N-glycosidase F (PNGase F). Released glycans can then be attached to a fluorophore, such as procainamide, to be analysed using UPLC. Using Endoglycosidase H (Endo H), a glycosidase that cleaves the β1-4 bond between the two GlcNAc residues on all oligomannose and some hybrid-type N-glycans in these analyses allows for quantification and comparison of oligomannose and hybrid-type glycans as the spectra compared to PNGase F only treated N-glycans will show a depletion of oligomannose-type glycoforms.

Qualitative mass spectrometric analysis of released glycans can also be performed via direct infusion of unlabelled, released glycans into the mass spectrometer. Using ion-mobility mass spectrometry and tandem mass spectrometry (IM-MS, MS/MS), it is possible to obtain a qualitative glycomic profile of all released glycans from a given Env and fragment specific sugars to determine structures. The ion-mobility cell within mass spectrometers makes this possible by allowing the separation of glycan regioisomers (same mass-to-charge ratio; aka m/z, different structure) by their shape, or collision cross section (CCS) [103]. Ions are separated by CCS when they travel through an ion mobility cell, a vacuum chamber in which a weak voltage wave carries ions against a counter current of an inert gas, for example nitrogen or helium. The speed at which they travel is a function of their three-dimensional shapes and therefore two ions of the same m/z can be separated if one is more compact than the other, as the more compact ion will travel more efficiently [103].

Glycopeptide analysis allows for quantitative site-specific analysis of N-linked glycans. First, the Env protein to be analysed is denatured before being treated with a range of proteolytic enzymes to maximise coverage of the protein. These enzymes include trypsin, which cleaves

at positive residues (Arg, His, Lys); chymotrypsin, which cleaves at aromatic residues (Phe, Tyr, Trp), or α-lytic protease (Thr, Ala, Ser Val). This glycopeptide mixture is then enriched to concentrate the sample and remove salts and enzymes before being subjected to liquid-chromatography mass-spectrometry (LC-MS). Glycopeptide data is searched against a standard library of known N-linked glycans. Glycopeptide fragmentation data is then individually evaluated for each glycopeptide and scored as a true-positive upon observation of correct fragment ions, the chromatographic areas for each true-positive are then compared to quanify microheterogeneity at each site [104] Glycopeptide analysis comprises the majority of the results contained in this thesis. As such, it is important to explain the process of acquiring and analysing glycopeptide data through LC-MS.

1.2.5 Mass Spectrometry

Mass spectrometry is a biophysical technique that allows for the measurement of the mass-to-charge (m/z) ratio of analyte ions. Mass spectrometers produce gas-phase ions from the user's analyte which enters an ion source. Unresolved ions are then accelerated to a mass analyser where ions are separated by their m/z. Results are then output as a spectra of signal intensity of ions against m/z. This section will provide a description of the type of mass spectrometry that will be undertaken for the analysis of digested glycopeptides, and the theory that underpins this process. A schematic of the Orbitrap Eclipse™ mass spectrometry instrument used for site-specific glycan analysis is outlined below in Figure 1.7 [105].

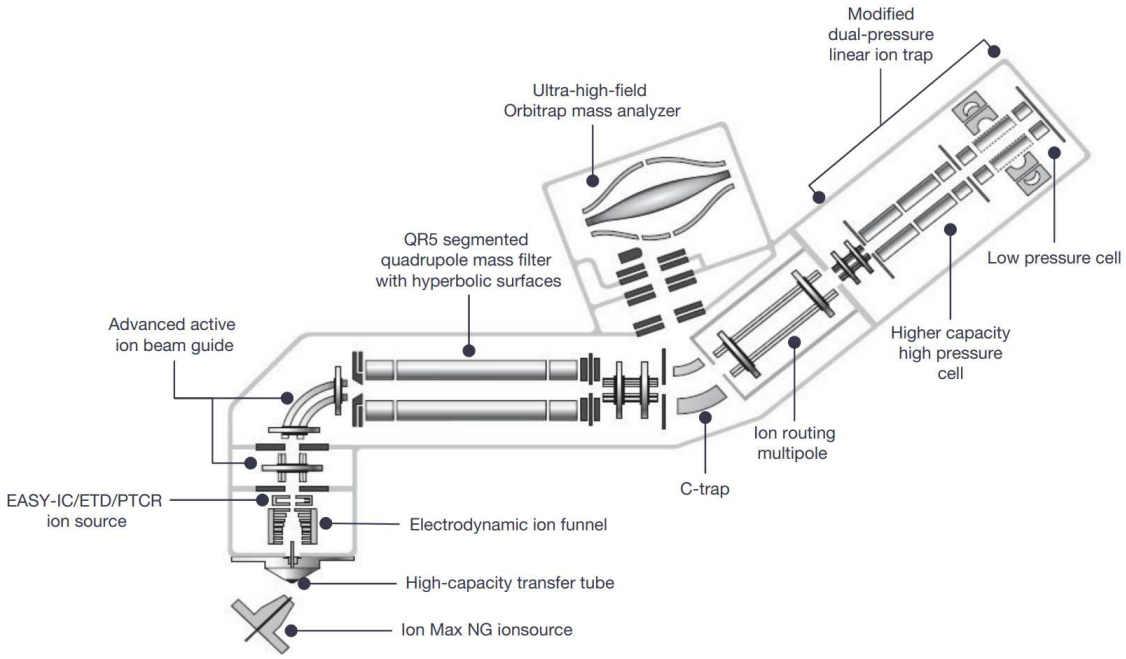


Figure 1.7: **Schematic of the Thermo Fisher Orbitrap Eclipse™ mass spectrometer** Schematic reproduced from [105].

1.2.5.1 Principles and components of a mass spectrometer

Sample inlet

The major components that comprise a mass spectrometer serve to ionise, separate and detect a given sample. Firstly, a sample introduction system is required to allow for the introduction of a sample to the ion source in a suitable quantity. Whilst mass spectrometers are able to analyse samples containing complex mixtures, they are somewhat limited in their ability to characterise these mixtures from one another. This produces a requirement for the prior separation of the components within a mixture before ionisation and analysis, which is often performed using either a gas or liquid chromatography (GC/LC) step before introduction [103]. The Thermo Scientific™ Orbitrap Fusion and Eclipse™ mass spectrometers we utilise uses a nano-LC system capable of sub ul/min flow rates as opposed to traditional LC systems that operate at 0.1-1 ml/min.

This type of chromatography is known as reversed-phase LC, as the stationary phase is a

nonpolar organic compound on a solid support (as opposed to a polar stationary phase used in normal-phase chromatography). The rationale for the use of an organic stationary phase is that organic compounds on an aqueous mobile phase will preferentially adsorb onto this stationary phase [103]. To elute these compounds, a mobile phase is used that progressively transitions from a polar, aqueous substance to an organic one. In our case, a gradient between aqueous formic acid (0.1%) and acetonitrile is used. The order in which parts of a sample are eluted is a function of their preference of either the mobile or stationary phase, with hydrophilic compounds eluting earliest in accordance with the highest concentration of polar mobile phase.

Ion source

Following separation by LC, the sample reaches the ion source. Ion sources serve two functions: to create either positive or negative ions from the sample and to transfer these ions to the mass analyser [103]. Ion sources have different ways to ionise samples such as electric fields, photons, ions, and electrons. These sources will generate ions not only of the sample, but any residual or background molecules present also. The formation of ions can occur either within a vacuum chamber (in vacuo) of the mass spectrometer, for example matrix-assisted laser desorption ionisation (MALDI); or outside of the instrument, collectively known as atmospheric pressure ionisation (API) methods.

Electrospray ionisation (ESI)

An example of an API method and the one that our mass spectrometers employ is electrospray ionisation. In ESI, the analyte flows through a narrow steel tube which applies a high voltage, causing the removal of ions with a polarity opposite to the applied voltage. This reduces our analyte to ions of only one charge type [103]. Adjacent to this tube is another, through which a flow of nitrogen gas nebulises the sample eluate. Once nebulised, there is additional flow of heated nitrogen gas that acts to dry the charged droplets, progressively

releasing ions into a gaseous phase. One of the reasons for the use of aqueous formic acid during the LC phase is that it also supplies protons for this process of ionisation. As they evaporate, these charged, nebulised droplets are passed into a vacuum chamber that connects the source to the mass analyser. Vacuum chambers are required to prevent the loss of ions with unwanted ions and neutral gases.

Mass analysers

The role of mass analysers is to separate ions on the basis of their m/z ratios, so that they may be recorded alongside their intensities by the detector. There are different forms of mass analysers, and any given instrument may employ one or more mass analysers (MS/MS) to improve analysis. For example, the Orbitrap Fusion™ mass spectrometer possesses a quadrupole and orbitrap mass analyser.

Quadrupole

Quadrupole mass analysers consist of four parallel metal rods that are aligned in a square. Opposite pairs of rods are connected, and each pair will oscillate between a positive and negative direct current (DC), with an additional radio frequency (RF) voltage. The process of mass separation occurs on the basis that ions oscillate upon entering the field produced by the DC and RF voltages. A given combination of DC and RF voltages will only allow ions of one m/z to have a stable trajectory through the quadrupole to the detector, a process known as bounded oscillation [103]. Other ions of differing m/z values will have an unstable trajectory and will discharge onto the quadrupole rods (known as unbounded oscillation). Scanning of different DC and RF voltages enables the progressive separation of ions across a range of masses which will, in turn, yield a spectra of different m/z values. The ability for quadrupoles to scan and filter out ions by their m/z is advantageous to the user, as once a m/z value that corresponds to an analyte is determined, the voltages of the quadrupole can be set so that only analyte ions of a specific m/z will pass, allowing it to act as a mass

filter. This process is known as selected ion monitoring and is particularly useful for MS/MS experiments that use more than one mass analyser.

Orbitrap and tandem mass spectrometry (MS/MS)

The second mass analyser present in the Orbitrap Fusion/Eclipse™ is the orbitrap. The orbitrap is an electrostatic ion trap made up of an inner electrode and an outer electrode which is split. Prior to entry into the orbitrap, ions are collected in a component known as the C-trap which then pulses ions into the orbitrap at high speeds. Within the cell, ions follow a trajectory comprised of two motions: a rotary motion around the inner motion and an axial oscillatory motion along the length of the inner electrode [103]. The frequency of this axial oscillation is proportional to the m/z of the ion; and is therefore recorded and interpreted into a spectrum using a Fourier transform (FT) function.

The combination of both orbitrap and quadrupole allow for the selection of analyte ions for their analysis, and subsequent further fragmentation which often occurs in a second collision cell for the analysis of constituent fragment ions, this is known as tandem mass spectrometry (MS/MS). MS/MS offers the ability to gain further structural information about specific ions. The first stage is to determine the m/z of ions within the sample that may correspond to peptides and glycopeptides. Peptides and glycopeptides digested with trypsin and other similar proteases will have an m/z between 500-2000; as such, the quadrupole's bounded oscillation is adjusted to accommodate ions within this range for the first scan of ions. This is known as the MS1. Additional filters are then applied to ions that are recorded within this range to determine which precursor ions that pass through this range are likely to be peptides and glycopeptides and therefore appropriate choices for further fragmentation and an MS2 scan.

There is a requirement for these additional filters because of the ongoing LC gradient does not permit for continual MS2 scans, as the instrument is unable to collect data for precursor ions during this time. These additional filters therefore serve as a means to fragment ions

in the MS1 that only correspond to potential peptides and glycopeptides, the criteria for fragmentation include presence of a monoisotopic peak, a minimum threshold of intensity of 2e4 (manufacturer recommended) and have a charge state of 2-5; which the MS is able to calculate at very rapid speeds.

After performing MS1, the instrument then selects ions in the MS1 spectra within these filters to fragment. The quadrupole voltages will be changed to allow the passage of only one particular ion. This ion is then fragmented by higher energy collision dissociation (HCD). Once the instrument has acquired a sufficient quantity of ions, known as the AGC target; or a certain time has elapsed (the injection time), these fragments are then subjected to separation based on their m/z values in the orbitrap in a second round of MS, known as the MS2. The quadrupole will then select for the next ion to isolate. This processed repeats for a specified amount of time (3 seconds) before reverting to the acquisition of the MS1.

Ion detection

Once analyte and fragment ions have been generated, they must be detected in order to discern and store their signals. Depending on the type of mass analysers used, the detection system may differ. Quadrupole instruments detect mass using ion currents. When ions arrive at the point of collection, they encounter a surface from which electrons are released. These electrons are then accelerated against a second surface whose contact causes the release of more electrons.

This process is repeated many times, resulting in a growing cascade of electrons which produces an ion current. This current is then amplified and recorded by a computer. The orbitrap uses what is known as an image current for detection. Ion detection is based on the induction of an electric current when ions pass close to a metal surface. Thus, when ions follow their own unique trajectory within the orbitrap it generates a frequency that is proportional to its m/z value [103]. This data is then analysed computationally to determine m/z of the constituent analyte ions and their fragments.

Data analysis

Mass spectrometric data for a given sample is stored as a .raw file which contains the total ion chromatogram of a given sample which inside contains both MS1 and MS2 data. It is represented as a chromatogram whereby the x-axis represents the retention time (from the previous LC step), and the y-axis represents the signal intensity of all the detected ions, this is referred to as the total ion chromatogram (TIC). This file must be processed in order to obtain useful information on the sample. There are many different software suites (both 1st and 3rd party) that this can be performed on. Our laboratory uses Byonic™ and Byologic© software (Protein Metrics Inc.).

Byonic™ acts as a search engine for our MS data. It takes the .raw file as input alongside the protein sequence, enzyme-specific cleavage sites (C-terminus of Arg, Lys, and C-terminus of Tyr, Trp, Phe for trypsin and chymotrypsin, respectively), a glycan library as well as other parameters such as tolerance for missed cleavages and other peptide modifications. It then simulates the digestion of the proteins and searches for the calculated masses of peptide fragments with the attached glycans from the library within the MS1 spectra obtained during the LC-MS run. Upon finding a matching ion mass it will extract the associated MS2 fragmentation spectra for that particular m/z from the entire dataset. The software will then assign the fragment ions to those predicted based upon the simulated precursor ion and assign a score to the peptide based on the quality of its fragmentation in the MS2 spectra. The main ions that contribute to the score are b and y ions as they occur between peptide residues.

Byologic© software uses the identifications output by Byonic™ and the .raw file as input to display the individual information on each identified glycopeptide from both the MS1 and MS2 spectra. This software allows for the quantification of the MS data by extracting information from the .raw file related to the list of detected peptide and glycopeptide ions from Byonic™. Whilst the TIC represents all of the ions detected over a given time, Byologic© filters this chromatogram to display the particular ions that correspond to the identifications

in Byonic™ and the time point at which they were detected. This is known as an extracted ion chromatogram (XIC). By integrating the XIC peak, we can determine the total amount of a particular ion. Byologic™ also further displays the XIC plots, ion fragmentation and isotope plots of each peptide and assigned glycan. Based on the quality of these parameters and the score, each identified peptide is then manually evaluated and categorised as either a true- or false-positive. The criteria for validation as a true-positive is as follows:

- Absolute minimum ion intensity score of 30 (The higher the quality of data the higher this value can be set).
- Identification of peptide b and y ions and oxonium ions from HCD fragmentation (as per Table 1.1).
- The peptide does not contain more than one N-linked glycan site.

The intensities of the extracted ion chromatograms (XIC) for all charge states for each glycopeptide with the same amino acid sequence classified as true-positive are then calculated as a percentage to determine the percentage of each glycoform at a given site. The way in which this data is then categorised according to different glycans can at times differ, but for the most part in this thesis it is split into oligomannose, Hybrid (sometimes combined with oligomannose to make the 'High mannose' category), Complex-type (classified according to the number of HexNAc monosaccharides and the presence/absence of fucose), and core glycans. In some instances, glycan sites may not be resolved, which can be attributed to a number of reasons. One of the most predominant reasons is that Envelope glycan sites are clustered closely to one another, and this creates difficulties in acquiring data on singly occupied glycopeptides if cleavage sites for the used enzymes are not situated between sites. A way to overcome this is to use several enzymes in separate digests in order to increase the likelihood of acquiring data on a particular site. In some instances, data on a particular glycopeptide digest may be acquired in a particular Envelope trimer or replicate which is not in other versions or replicates of the same protein or strain. This can be partly be at-

tributed to the ability for enzymes to occasionally cleave at non-canonical peptide sequences, or alternatively technical limitations in data acquisition. A complete list of the entire glycan library used during analysis and the resultant glycan classification can be found in Appendix C. Mass spectrometric analysis of viral Env has typically been hampered by the low quantities of virally expressed Env. The production of recombinant variants of Env in greater quantities allowed for the study of the glycosylation of Env. The creation of these proteins is discussed in the following section.

Oxonium ion (m/z)	**Description[a]**
127.06	Hex - $2H_2O$
147.07	Fuc
163.06	Hex
168.09	HexNAc - $2H_2O$
186.09	HexNAc - H_2O
204.09	HexNAc
274.09	NeuAc - H_2O
292.08	NeuAc
366.14	HexHexNAc
528.19	HexNAcHex
657.23	NeuAcHexHexNAc

Table 1.1: **Characteristic oxonium ions from HCD fragmentation of glycopeptides.** a. Hex, Hexose (e.g. mannose or galactose); HexNAc, N-Acetylhexosamine (e.g. N-Acetylglucosamine); Fuc, fucose; NeuAc, N-Acetylneuraminic acid (e.g. sialic acid); - H_2O, loss of water. Adapted from [106]

1.2.6 Env immunogens and the SOSIP platform

Structural studies and vaccination attempts on viral Env have proved challenging due to the low numbers of viral spikes per virion and the difficulty associated with purifying and stabilising membrane proteins; in addition, the flexibility of glycans that coat this protein hinder structural determination. This led researchers to try create Env constructs to mimic the wild type in both protein and glycan structure, whilst improving the production process and facilitating further research. It is now accepted that whilst you can create structures

that resemble the protein component of Env, mimicking the glycan composition of this glycoprotein is far more challenging and many constructs have been unable to do so, as the quaternary structural constraints of a functional Env trimer determine the glycan content, an aspect many constructs were unable to replicate [107].

Currently, one of the most studied trimer mimics and a relatively recent discovery is the SOSIP format that contains a number of mutations and modifications to produce a stable, soluble mimic that can be, under current good manufacturing practices (cGMP), be produced in gram quantities [108]. The most used variant of SOSIP is BG505 SOSIP.664 which is based on a clade A transmitted/founder strain, BG505. An artificial disulphide bond has also been introduced between gp120 and the gp41 ectodomain (SOS), the latter also being truncated at position 664 to exclude the MPER and the C-terminal region [109]. In addition, the stability was improved through a isoleucine to proline (IP) mutation within HR1 of gp41 [110, 111].

BG505 SOSIP.664 has also had a glycan site at position N332 knocked in (T332N) for antibody epitopes relying on this glycan. The furin cleavage site is also mutated from its original REKR sequence to a hexa-arginine (R6) sequence which is more prone to cleavage [112]. Lastly, the sequence has been codon-optimised for to allow for easier translation, as the wild-type sequence encoding Env, and other HIV-1 proteins, are known to impede the process of translation [113].

Structural studies on BG505 SOSIP.664 have provided many structural and mechanistic insights into the HIV-1 envelope trimer that would have otherwise been unattainable from wild-type Env. SOSIP displays similar antigenicity to wild-type Env which has allowed for studies of bnAb epitopes not only in the context of the BG505 SOSIP trimers but also SOSIP trimers of other HIV genes, such as B41 and AMC008, both of which are clade B envs, and ZM197M, a clade C env. When the structure of these different SOSIPs were determined by cryo-EM, it was found that they often did not adopt a fully closed, native trimeric structure [114].

Much like Env, the SOSIP trimers were found to be flexible and able to fluctuate between closed and more open conformations and sometimes sampling the fully open CD4-bound conformation [114]. This conformational flexibility in turn allowed for the exposure of non-neutralising antibody epitopes (non-nAbs) which, compared to bnAb epitopes, can cause an immunodominance that hinders the induction of bnAbs.

The mechanism by which they are proposed to cause this immunodominance is that higher affinity non-nAbs enter germinal centres and block antigen binding to B-cell receptors of lower affinity that have specificity for bnAb epitopes [115, 116]. The rationale was then established that limiting conformational flexibility and non-nAb epitopes would likely improve the immunogenicity of the SOSIP platform and their likelihood of inducing a bnAb response. Thus, a series of stabilising substitutions were designed to improve the SOSIP platform. To improve the gp41 pre-fusion structure, a double I535M and L543Q/N mutant was created to form SOSIP.v3.1 and v3.2 respectively; the highly non-nAb immunogenic V3 region was given an A316W substitution to enhance the hydrophobic packing and thereby decrease the propensity of this region to flip out [114].

Previous studies on env structure in the presence of a viral entry inhibitor VIR165 found that two mutations, at position 64 and 66 of gp120 conferred both resistance and dependence on this entry inhibitor, in absence of VIR165, the CD4 bound conformation was found to be impeded, whilst in its presence however, trimers were able to undergo CD4 induced conformational changes [117]. When investigated, it was proposed that these residues could be part of a switch that transduces CD4 induced conformational changes from the CD4 binding site to gp41, and that perhaps the changes in structure or charge from these substitutions impede this process.

Enzyme-linked immunosorbent assay (ELISA) studies indicated that these substitutions reduced or even eliminated CD4 induced non neutralising epitopes to AMC008 and BG505 SOSIPs, and thermostability increased also as measured by differential scanning calorimetry (DSC). Based on the already existing v3 platform, the A316W substitution was combined

with either E64K or H66R substitutions to create SOSIP.v4.1 and v4.2 respectively. When compared to wild-type SOSIP.664, each mutant from each clade showed large decreases or abolishment of non-nAbs against V3 and CD4-induced epitopes [114].

Env is covered in a dense array of glycans (Figure 1.3D) that are critical to its own structural and antigenic properties. Many crystal and cryo-EM structures of natively glycosylated Env trimers have been published including that of BG505 SOSIP.664 at 3.5Å resolution and have been crucial to the understanding of the glycan network around this protein [118]. Despite this, cryo-EM and X-ray crystallography only give an averaged snapshot of glycan states at individual sites and therefore do not capture potential microheterogeneity. Through the use of mass spectrometry however, it is possible to perform analysis of N-linked glycans at a site-specific level, and with the development of SOSIPs the ability to get a site-specific glycan profile of a native-like Envelope was at last achieved.

1.2.7 The Env glycan shield

Within a single gp120-gp41 protomer, there are between 25 and 30 N-linked glycans 18. The role of viral glycans is split between interactions with host cell receptors for viral entry and masking of the underlying protein surface from immune recognition; of which HIV-1 uses the latter with most other viruses using glycans for receptor interactions [80, 119]. Due to the sheer abundance and proximity of the N-linked glycans on the Env surface, steric constraints within the glycan shield that prevent glycosidases such as ER and Golgi α-mannosidases from performing their actions, thereby resulting in the presence of oligomannose-type glycans ($Man_{9-5}GlcNAc_2$) that are less varied than typical host cell N-glycans.

The glycan shield constantly adapts to conceal epitopes from newly produced nAbs, and although the N-linked glycans they derive are host cell-context dependent, the inability of the aforementioned glycosidases to trim these glycans result in the clustering of oligomannose-type glycans in what is known as the “intrinsic mannose patch” (IMP) [107, 120]. The IMP is a conserved feature of gp120 across all clades as well as longitudinally during an infection

[102, 120, 121, 122, 123]. Studies on native-like trimers and virion-derived Env show a higher quantity of oligomannose-type glycans than on gp120 monomers which alluded to a presence of a trimer-associated mannose patch (TAMP), located around the apex of the trimer and protomer interface including, but not limited to, the N156, N160, N197, N262, N276, N301, and N637 sites [102, 120, 123, 124, 125].

The formation of these predominantly oligomannose-type sites seems to be heavily influenced on the folding and conformation of these trimers. Whilst native-like and cleaved HIV-1 trimers are dominated by oligomannose-type glycans; uncleaved, non-native trimers adopt an open and irregular structure without sufficient quantities of furin, rendering its co-expression alongside any Env construct a prerequisite for the formation of functional trimers [107]. Consequently, the quaternary structure of Env is able to control the degree of glycan processing as the previously inaccessible sites are now able to be efficiently processed, resulting in a greater abundance in complex and hybrid-type structures, thus allowing for the examination of N-glycosylation as a means to roughly distinguish its quaternary structure [107]. In contrast to gp120, gp41 N-linked glycans are more processed and thus are composed of more complex-type glycans [107].

Whilst oligomannose-type glycans roughly account for half of the glycan types present on Env trimers even when expressed in chinese hamster ovary and human embryonic kidney (55% in CHO, 56% in HEK), complex-type glycans present greater variability in their composition across different cell expression systems. For example, the gp41 subunit from CHO-produced BG505 SOSIP.664 is considerably more sialylated than gp41 from HEK-produced BG505 SOSIP.664 [107].

Sialic acids are terminal modifications found on complex-type glycans that are known to have effects on Env and the immune system 109. Pritchard et al., characterised complex glycans of virally derived gp120 produced in peripheral blood mononuclear cells (PBMCs) and compared them to HEK-produced Env and found that PBMC gp120 complex glycans primarily contained α2,6-linked sialic acids whereas HEK-produced Env contained exclu-

sively α2,3-linked sialic acids [126]. In HIV infection, N-linked glycans are known to have roles in complement-dependent enhancement of HIV-1 dissemination through binding of the complement proteins C3b, C4b, and Factor H, which was proposed to confer protection of HIV-1 from complement-mediated lysis [126, 127].

It is also proposed that non-nAbs raised during acute infection can work in tandem with complement to aid HIV-1 dissemination, as they are able to interact and exert either inhibitory or enhancing effects with either complement or Fc receptors [128]. The effects that sialic acids confer appear to be linkage-specific. For example, CD22 is an α2,6-linkage specific receptor that is found on mature B-cell surfaces responsible for regulating immune responses, it does not bind at all with α2,3-linked sialic acids [129]. This point becomes particularly interesting when the previous differences between PBMC-derived and HEK-derived Envs are considered, and provides a point of consideration of sialic acid linkages in immunogen design, which could impact bnAb epitopes that include complex glycans [126].

The presence and function of O-linked glycosylation on Env is a contested topic. O-linked glycans are commonly found at repeated stretches of peptides rich in proline, serine and threonine, known as variable number of tandem repeat (VNTR) regions [130]. In 1994, Bernstein et al., suggested their presence after observing a decrease in molecular weight of vaccina virus recombinant Env glycoproteins following treatment with O-glycosidase and subsequent analysis of radiolabelled carbohydrates by paper chromatography [131].

More evidence of O-linked carbohydrates on Env published by Hansen et al., found that antibodies with epitopes against GalNAcα1-O-Ser/Thr (Tn antigen) have a weak neutralising effect against HIV-1 [132]. Despite this, they also found that abrogation of the regions within the V3 loop with O-glycosylation sequons had no effect on anti-Tn antibody sensitivity 116. It has been reported that recombinant Envs contain O-glycan sites at positions T499 and T606 [123, 133, 134, 135, 136]. Studies on viral Env are however, mixed. Yang et al., report the presence of an O-linked Gal1GalNAc1 at position T499 [137]. In direct contradiction, Stansell et al., found no detectable O-glycans on virion-derived gp120, suggesting that the

transmembrane area of gp41 may hinder glycosyltransferase accessibility [136].

Behrens et al., performed quantitative mass-spectrometric analysis in search of O-glycan occupancy and found that the T499 O-linked site showed around 7% occupancy for gp120 monomers, but also found minimal (0.2%) occupanccy for BG505 SOSIP.664 trimers, with similar findings (1%) for pseudotrimers [123]. The T499 site is located in close proximity of the C-terminus of gp120, and the near presence of gp41 is believed to create a local steric environment which constrains glycosyltransferase activity [123]. In addition, the structures of BG505 SOSIP.664 and the membrane-associated JR-FL trimer show the T499 site to be buried in the presence of gp41, thereby meaning an O-glycan could not be accommodated within the gp120-gp41 interface [67, 138, 123].

1.2.8 Immune evasion and the glycan shield

The glycans that envelop the underlying protein surface of Env are constantly changing shape to escape the actions of the immune system, as changes in DNA sequences encoding glycosylation sites are able to be deleted or shifted due to the intrinsic, error prone nature of HIV RT. However, a balance must be struck between functionality and immune escape, which therefore limits the quantity and locations of PNGSs, evidenced by the fact that number of sites on Env has remained somewhat constant over 30+ years of evolution [139]. Despite the relative conservation of N-linked glycosylation sequons across Envs of different isolates and clades, sites on gp120 are known to shift during infection [140, 141]. These shifts in glycan sites can be caused by changes in the DNA. For example, a mutation from an NxTxT to NxNxT can move the position of a PNGS up by two residues, which can be accomplished by a change in a single nucleotide. This is due to N, S, and T sharing extremely similar triplet codes that allows for the change from N to T to occur by exchanging the ACT triplet (T) to AAT (N) [142].

A particular site that has been identified as subject to this shift in sequon is the N332 site [49]. It was observed that in some patients with a HIV-1 clade C virus, there was a glycan

at position N334 instead. Over the course of the infection, the site would shift between N332 and N334, each time facilitating viral escape from neutralization [49]. Env sequence analysis from the Los Alamos database reveals that N332 occurs in 73% of non-redundant sequences whilst N334 appears in 20% [142, 143]. This suggests that the preservation of a glycan within this region is of somewhat importance as they are present in a total of 93% of viruses [143]. Interestingly, the bnAbs PGT121-128, which include N332 within its epitope, are able to promiscuously bind N334 in some cases [49, 144]. The ability for these bnAbs to overcome escape mutations through the use of their broad epitopes provide a good rationale for immunogen vaccine design aimed at raising these antibodies.

1.3 Targeting the Glycan Shield

1.3.1 Broadly neutralising antibodies: structure and features

Broadly neutralising antibodies are unique in the sense that they are able to target epitopes composed of both protein and glycan components, in doing so overcoming self-tolerance checkpoints that would otherwise eliminate their production before it would even begin. In order to achieve this, bnAbs possess multiple uncommon characteristics that enable them to be effective, including but not limited to: long complementarity determining region (CDR) H3 loops, domain swapping, significant somatic hypermutation, and uncommon post-translational modifications (PTMs) [142]. A good example to describe these structural intricacies is the bnAb 2G12, the first bnAb identified to bind to the glycan shield [145]. It has an unique domain-swapped structure in which the VH domain from each antigen-binding fragments (Fab) within the immunoglobulin G (IgG) swap with the neighbouring Fab to create an uncharacteristically linear IgG structure with two adjacent Fab fragments that pack tightly on one another [142, 146].

Wild-type 2G12 is extensively somatically mutated, containing 38 and 16 somatic amino acid mutations in the heavy and light chains respectively; it should be noted however that only

5 to 7 of these wild-type mutations are required to be introduced within the heavy chains to induce Fab dimerization [142, 147]. Unlike other bnAbs that interact with both protein and glycans, 2G12 is proposed to interact with between 2 and 4 oligomannose-type sugars on Env, particularly with glycans at positions 262, 295, 332, 339, 386, 392, and 448 [140]. Another bnAb, PGT145, is an apex-targeting broadly neutralising antibody that engages all three gp120 protomers simultaneously with a CDRH3 region arranged in an anionic beta hairpin structure that penetrates the glycan shield at the three-fold axis to contact all three gp120 protomers [138]. One of the tyrosine residues within the CDRH3 region of PGT145, Y100, is known to be sulphated and it is thought that this sulphate may interact with the N160 glycan of the first gp120 protomer [138]. In summary, broadly neutralising antibodies use uncommon characteristics and extensive mutations in order to target multiple glycans on the HIV-1 envelope shield whilst still ensuring tolerance to self-glycans.

1.3.2 Broadly neutralising antibody epitopes on the glycan shield

The number of bnAbs that are isolated has steadily been increasing year on year. The bnAb database, bnaber.org, reports 90 unique antibodies and 112 PDB structures have been submitted as of July 2019 [148]. Whilst the glycan shield has constantly evolved against the immune system and antibody neutralization, it does possess glycan epitopes that remain relatively constant. These regions include the CD4 binding site, glycans located on the V1, V2, and V3 loops, the mannose patch, the MPER on gp41, and an extended region of residues between gp120 and gp41 [149, 150]. Of these, the mannose patch is one of the most conserved locations across strains and glycans within this region form epitopes for many bnAbs, especially the central N332 glycan which forms part of what is dubbed as the “supersite of immune vulnerability” [151].

The ability to determine complexes of bnAbs and Env has been greatly aided by X-ray crystallography and cryo-EM, allowing for the visualisation of the previously mentioned unique characteristics that allow for broad neutralization [67, 138]. These techniques however rely

on the averaging of the microheterogeneity within a given site in order to produce a consensus structure which may not accurately reflect the finer details and interactions between glycan and antibody [152]. A way to overcome this problem is to use site-specific analysis of digested Env glycoproteins which would allow for the mapping of antibody structures onto the most common glycoform on a given site.

For the most part, there is good agreement between site-specific analysis and structural studies of glycan epitopes; however, some incongruities remain between the two. Both Pancera et al., and Vidya et al., report that bnAbs PG16 and PG9, two clonally related apex-binding antibodies, preferentially bind to α2,6-sialylated glycans at N156 [153, 154]. Yet studies by Struwe et al., and Cao et al., that involved site specific analysis of both BG505 SOSIP trimers and virally-derived Env found this site to be predominantly occupied by oligomannose-type glycoforms [104, 155]. Such inconsistencies highlight a need for further structural clarity of the finer structure of the glycan shield and the interactions that govern these epitopes as well as refined and standardised procedures between researchers to prepare and analyse samples.

1.3.3 Potential for therapeutic use

The application of bnAbs in HIV-1 treatment could prove to be useful years to come. The goal of immunogen design is to be able to induce production of bnAbs against HIV-1 to confer protection and memory against HIV-1. This is not the only application however, as administration of bnAbs against HIV-1 to infected individuals could prove equally effective. When infected humanised mice were administered with three bnAbs: 3BNC117, a CD4bs targeting antibody; PG16, a V1/V2 loop region targeting antibody and, 10-1074, a V3 stem targeting antibody, a large reduction in viremia and HIV-1 DNA was observed [156].

When rhesus monkeys infected with a SHIV (simian-human immunodeficiency virus) were administered with either PGT121 or a cocktail of bnAbs (PGT121, 3BNC117, b12), viremia declined quickly to undetectable levels [157]. Similar SHIV studies on macaques using CD4bs targeting bnAbs 3BNC117 and 10-1074 also showed strong declines in viremia and viral load

to undetectable levels [158]. When the bnAb 3BNC117 was tested in phase 1 clinical trials on HIV-1 infected and non-infected individuals, it displayed encouraging efficacy, reducing viremia significantly over a 28 day period and lowering viral load by 0.8-2.5 log10 from a single 30mg kg-1 infusion [159]. As an alternative to HIV-1 vaccine development, it is hoped that bnAbs could be combined with other drugs to activate latent cells and subsequently eradicate them.

1.4 HIV-1 Vaccine design: strategies and challenges

1.4.1 HIV-1 vaccine trials

Owing to the previously mentioned abilities for HIV-1 to evade immune surveillance, it is unsurprising that efforts to generate effective vaccines have proved unsuccessful thus far. The first generation of vaccine strategies that have undergone clinical trials, some using gp120 as a target, and proved unsuccessful at eliciting protective antibodies perhaps owing to the pervasiveness of exposed non-neutralising epitopes [160, 161].

The results of these trials fed into current design strategies and allowed for the understanding that B-cell response targeting vaccines require functional epitopes that will be presented on the viral envelope spike. Other strategies for vaccination included the use of T-cell response targeting vaccines, whose responses could be targeted at more conserved features of the virus. A downside to this platform is that they would only be able to target antigens that had already been processed for presentation by MHC class I molecules following initial viral infection. Vaccine trials aimed at stimulating CD8+ T-cell responses were also unsuccessful in stimulating protection and were prematurely halted due to safety issues [162, 163].

B-cell vaccines, which would in contrast be able to target HIV-1 before infection, have been thus far broadly unsuccessful. Of the first 9 HIV-1 vaccine efficacy trials conducted thus far, only one showed some degree of efficacy (RV144/NCT00223080) [164, 165]. Despite this, subsequent trials (HVTN702/NCT02968849 and HVTN705/NCT0306029) to improve

on this did not yield any increased efficacy [166, 167, 168]. Vaccine trials evaluating safety and efficacy of newer platforms such as BG505 SOSIP and engineered outer domain (eOD) await publishing and peer-review [169, 170], yet preliminary phase 1 results from the latter report that 97% of participants developed VRC01-class IgG B-cells, which are precursors to bnAb induction against the CD4bs and would represent a significant milestone in the effort to generate an effective vaccine. Table 1.2, adapted from [171], summarises the first 9 HIV-1 vaccine trials and their results.

Trial	Start	End	Vaccine	Location	Result	References
VAX004 (NCT00002441)	1999	Jan 2000	Bivalent clade B gp120 in alum	US, EU	No efficacy	[161, 172, 173]
VAX003 (NCT00006327)	Mar '99	Aug '00	Bivalent CRF_01AE/B in gp120 in alum	Thailand	No efficacy	[174, 160, 175]
HVTN502 (NCT00095576)	Nov '04	Sep '09	Adenovirus type 5 clade B gag/pol/nef	South Africa	No efficacy	[163, 176, 177] [178, 179]
RV144 (NCT00223080)	Sep '05	Apr '09	AVLAC with gag/pro/env; bivalent CRF_01AE/B gp120 in alum	Thailand	60% at 12mo 31% at 42mo	[164, 165, 180] [181, 182, 183] [184, 185, 186] [187, 188, 189]
HVTN505 (NCT00865566)	May '09	Oct '17	DNAs with clade B gag/pol/env and DNAs with with clade A, B, C Envs; adenovirus type 5 with gag/pol and clade A, B, C Envs	US	No efficacy	[190, 191, 192] [193, 194, 195] [196, 197]
HVTN703/ HPTN081 (NCT02568215)	May '16	Mar '21	Antibody Mediated Protection (AMP) Trial of VRC01 neutralizing antibody infusion IV	Sub-Saharan Africa	No overall efficacy	[198, 199, 200]
HVTN704/ HPTN085 (NCT02716675)	Apr '16	Dec '20	Antibody Mediated Protection (AMP) Trial of VRC01 neutralizing antibody infusion IV	North and South America, Switzerland	No overall efficacy	[198, 199, 200]
HVTN702 Uhambo (NCT02968849)	Oct '16	Sep '21	AVLAC-C with gag/pol/Env; bivalent gp120s in MF59	Sub-Saharan Africa	No efficacy	[166, 201]
HVTN705 Imbokodo (NCT03060629)	Nov '17	Aug '21	Ad26, 4 valent T-cell mosaic genes, boost with clade C gp140 Env	Sub-Saharan Africa	No efficacy	[202]

Table 1.2: **Initial HIV-1 vaccine efficacy trials. Adapted from [171]**

The various trial attempts highlight a broad approach in trying to elicit protection against HIV-1 infection. Unfortunately, only the RV144 trial appeared to show any effect. Studies suggested that antibodies against the V1/V2 region correlated with protection from trial participants [165]. In this regard, the low trial efficacy could be explained by the understanding that the V1/V2 epitope is likely best elicited from an intact trimer rather than a single gp120; as the peptides and glycans are in close proximity to one another at the trimer apex and there is a lower chance of non-nAb elicitation. The majority of these trials used gp120 as the immunogen and none of these trials were able to stimulate bnAbs, and at best stimulated nAbs. The issue with nAb elicitation owes to the high antigenic diversity of HIV-1. As previously discussed, this allows HIV-1 to confer escape mutations to escape immune pressure. Highly variable regions of Env such as the V1/V2 region are particularly prone to this phenomenon [203]. Antigenic diversity notwithstanding, the prevalence of immunodominant regions on gp120 is also a significant factor that contributes to the lack of bnAb responses. As this protein is endogenously presented as a trimer, its presentation as a monomer results in the exposure of the inner domain of gp120. This domain is made up of the parts that drive the structural rearrangements prior to viral entry via gp41. This region constitutes a large protein surface that is not occluded by N-linked glycans, and as such would drive a strong antibody response to this region that would in turn, repress a more sensitive bnAb response. Furthermore, in the case of an infection any such antibodies would only instigate a non-nAb response as these epitopes would be occluded in a fully-assembled and functional Env trimer.

1.4.2 Issues with HIV-1 vaccine design strategies

1.4.2.1 Issues with native-like trimer vaccines

It was thought that the new generation of native-like vaccine candidates that better mimic the antigenicity of Env would yield better results, owing to their similarities in structure and glycosylation. Despite this, immunogenicity studies using native-like trimers have only been

able to elicit autologous Tier-2 nAbs or heterologous Tier-1 nAbs [114, 204, 205, 206, 207, 208, 209]. Understandably, if it takes an extended amount of time with a HIV-1 infection to generate bnAbs, then it is rational to assume that a single vaccination even with a native-like trimer will not be enough to generate a bnAb response. Part of the reason bnAb epitopes are conserved is that their conservation is related to the poor immunogenicity of these epitopes. Other issues with these constructs pertain to immunodominant epitopes from the exposed base of soluble trimers, which would otherwise be inaccessible on the virus [210]. To overcome these issues, engineering strategies to either increase the immunogenicity of bnAb epitopes, decrease immunogenicity of unwanted epitopes or present specific epitopes have been undertaken, creating tailored versions of Env that present only specific epitopes or filling in potentially un-wanted immunogenic epitopes with glycan sites. Holes in the glycan shield may arise from HIV strain-specific differences in sequence which result in areas of the protein not being shielded from the immune system by a glycan or at times due to glycan under-occupancy, whereby a PNGS is not filled by a glycan. On a related note, whilst native-like trimers exhibit great fidelity in their glycosylation compared to the virus, key differences exist: the first, as previosuly mentioned, pertains to areas of the protein that are left exposed due to incomplete PNGS glycosylation. The second issue relates to differences in the fine glycosylation between immunogen and viral Env. In both 293F and CHO-derived immunogens there are higher levels of oligomannose-type glycans outside of the sterically occluded regions such as the mannose patch and the trimer apex [104]. In contrast, site-specific glycan analysis of viral-derived Env show near complete PNGS occupancy and increased levels of processing. These differences likely mediate the hotspots of immunogenicity and affinities to bnAb epitopes that may be presented on the Envelope trimer.

1.4.2.2 Holes in the glycan shield

Glycan holes are described as the absence of an N-linked glycan within the Env trimer that exposes the underlying protein surface, either though deletion of the sequon encoding for N-glycosylation or through steric hindrance from neighbouring glycans that results in a given site being unoccupied. The exposure of the underlying protein surface, whilst rare, has the potential to induce nAbs or non-nAbs against these regions, in a way which may negatively impact induction of bnAbs [114, 211].

Many nAbs directed at glycan holes from transmitted/founder (T/F) strains were found to influence the filling of these glycan holes within 2 years, partly due to escape mutants filling these spaces to avoid targeting from the immune system [211]. Wagh et al., subsequently analysed the kinetics of the neutralising response in the context of holes in the glycan shield which yielded several hypotheses into potential mechanisms: firstly, the majority of T/F glycan holes are likely targeted by autologous nAbs; subsequent viral escape from these responses likely resulted from the filling of glycan holes subject to nAb targeting during early infection; thirdly, after the filling of most holes present in the T/F Env, the breadth of neutralization began to rise; lastly, this delay in acquisition of neutralization breadth associated with glycan holes could be caused by the impediment of bnAb lineages by dominant strain-specific responses, similar to germinal centre domination of non-nAbs [115, 116, 211, 212].

With the correlation between neutralization breadth and glycan holes in mind, Ringe et al., performed a study looking into the redirection of antibody responses upon opening and closing of glycan holes, the goal of the study was to see if the nAb response to SOSIP trimers could be manipulated and consequently aid future immunogen design [213]. In doing so, they identified the three most immunogenic nAb epitope clusters on the BG505 SOSIP.v4 trimer: the first, known as the 241/289 glycan hole is created by an absence of glycans at N241 and N289, and can be blocked by a glycan knock-in at either position; the second, known as the C3/465 site is a stretch of amino acids in the C3 region and can be blocked by a glycan knock-in at position 356 or 465; the third, and least immunogenic epitope is located

in the V1 region near residues 133 and 136 [213, 214].

The study established that by inserting sequons for N-linked glycans at these locations it was possible to abrogate immunogenicity of the 241/289 glycan hole. Furthermore, by deleting glycan sequences at positions 197, 234, 276, 332, and 355, they found that immunogenic neo-epitopes could be created, in doing so opening new holes in the glycan shield and redirecting the nAb response to these newly created epitopes [213]. The ability to manipulate antibody responses to specific parts of the glycan shield may be relevant for rational immunogen design in guiding responses to broadly neutralising epitopes.

1.4.2.3 Beyond native-like trimers: germline targeting

The reason for further engineering efforts is that it is thought that bnAb generation will now likely require a vaccination regimen that includes differently modified immunogens in order to replicate the bnAb development pathway and guide a nascent B-cell response to maturity. In a HIV-1 infection, bnAb-producing B-cells were found to have B-cell receptor characteristics that are associated with poly- or autoreactivity [215, 216]. Such B-cells constitute an majority of the population of early B-cells but are mostly deleted upon maturity [217]. In addition, such B-cells need to evolve rare characteristics such as long heavy chain complementarity determining region 3 (HCDR3) regions in order to bind certain conserved epitopes develop bnAbs [218, 219, 220].

A successful HIV-1 vaccination regimen will likely combine have to guide bnAb development from a naive B-cell using immunogens specifically designed to optimise binding at different stages. This would involve an initial 'priming' immunogen designed to engage precursor lineages and subsequent 'boost' immunogens that help develop the epitope response towards a bnAb [221]. The name of this concept is germline targeting as it encapsulates the concept of taking the naive germline B-cell of a desired antibody through to a mature B-cell using a course of tailored immunogens.

Heterologous prime-boost vaccine regimens have been trialled recently for safety and efficacy

humans in NHPs. They were found to induce robust anti-Env responses in both groups and in the NHP study where challenge studies using heterologous SHIV strains were conducted, the vaccine regimen conferred 67% protection against SHIV infection [170, 202]. The results of this trial provide a robust proof-of-concept in prime-boost regimens. The results from the eOD-GT8 trials additionally provide hope in achieving the first 'priming' step [169, 170]. However, in order to sequentially guide bnAb development, the design of all immunogens within a regimen must be carefully considered.

1.5 Scope of Thesis

The purpose of this thesis is to explore ways in which to improve HIV-1 Envelope immunogens through the engineering of their glycosylation. In chapter 3, we explore the effects of addition and removal of glycan site PNGSs on the soluble BG505 SOSIP.664 Envelope trimer. The purposes of this chapter is twofold.

Firstly, we wanted to investigate the tolerance of this envelope trimer to addition of glycan PNGSs. Studies using this trimer prior to viral challenge in macaques highlighted a nAb response to epitopes at positions 241 and 289 [222]. In other HIV-1 strains, these areas are filled in by N-linked glycans and as such these areas are sometimes known as the 241/289 glycan hole. We therefore decided to test the tolerance of BG505 SOSIP.664 to addition of PNGSs to both of these sites on their own and in combination (BG505 SOSIP.664 +N241, +N289, +N241 +N289) and see the ramifications of these additions on these newly added sites as well as the other glycan sites. Ideally, this study will show that it is able to fill in exposed areas on the protein surface with PNGSs without aversely affecting the glycosylation of neighboring sites. It is possible however that adding in these glycan sites may have effects on neighboring sites.

The second aspect of this chapter pertains to the effects of removing PNGSs from BG505 SOSIP. Vaccine candidates based on BG505 SOSIP have already been developed and partly tested in which they have removed the N276 glycan site [223, 224]. The purpose of this

strategy is to better stimulate germline bnAb precursors against the CD4bs, as it has been shown that early stage bnAbs targeting this area do not bind well around this glycan, but evolve to accomodate it over time. As such, such strategies may prove useful in prime-boost vaccine regimens where initial germline bnAbs are elicited to an intentional and highly immunogenic glycan hole which is then filled in to guide bnAb development. The purpose of our PNGS removal strategy in this chapter was to investigate the feasibility of this strategy with several glycan sites and see what the effects are on any neighboring glycan sites. If successful, not only could this strategy could be widely applied to bnAb epitopes across the trimer to improve their immunogenicity; but also in filling in unwanted, immunodominant glycan holes.

Chapter 4 also takes a twofold approach to engineering immunogen glycosylation. The first follows on from Chapter 3 in investigating the addition and removal of PNGSs as well as other amino acid mutations as a strategy to improve the immunogenicity of a different membrane-bound trimeric immunogen presented on viral-like particles (VLPs), JR-FL SOS. In the first half of the chapter, engineering efforts were made to improve two separate and distinct epitope regions. One of the epitope regions in this trimer already exhibited sensitivity to antibodies targeting it whereas the other did not. Both epitope regions had been recognized as exhibiting some level of plasticity in the way that antibodies bind to it, which suggested an ability to increase sensitivity to this area through glycan additions, deletions, and other repairing mutations to boost glycan occupancy in certain sites.

The second half of Chapter 4 explored the use of glycosyltransferase co-transfection to achieve the same goal. It has been reported that some bnAbs exhibit preferential binding to glycans on Env with certain characteristics such as enhanced sialylation for example [225]. The effects of a panel of co-translated glycosyltransferases were tested on VLP-presented trimers as well as corresponding gp120s. In addition, a combination of glycosyltransferase co-transfection and post-purification treatment with neuraminidases, hexoseaminidases and galactosidases were undertaken to see which sites the enzymatic co-transfection map to on the trimeric

constructs. In doing so, we aimed to observe which sites were liable to enzymatic action through the modification of the glycans as well as which ones were accessible to processing. As the action of enzymes on Envelope glycosylation are only limited by the accessibility of their substrates, we sought to investigate ways in which site-specific changes in glycosylation may be imparted.

In Chapter 5, we sought to investigate ways to induce specific changes in site-specific glycosylation. Current methods to alter glycosylation such as those explored in Chapters 3 and 4 lack specificity as they can either have knock-on ramifications on allosteric glycan sites or in the case of enzymatic co-expression, affect all available substrates equally. Similarly, the choice of modified cell lines lacking in particular enzymes involved in the glycosylation pathway (e.g. GnT I -/- cells) or the use of enzymes inhibitory to the pathway (e.g. kifunensine) will exhibit similarly global effects on Env glycosylation. Aside from potential drawbacks on envelope structural integrity arising from these effects, it is understood that changes to Env glycosylation that can improve bnAb binding to certain regions may also simultaneously abrogate bnAb binding elsewhere. As such, this chapter aimed to explore the ability to impart specific changes to Env glycosylation whilst maintaining overall integrity of glycosylation with regards to the wild-type through co-expressing BG505 SOSIP trimers with a panel of bnAbs that target distinct epitope regions.

Chapter 6 sought to explore the differences in codon usage between the viral-expressed Env and current immunogens. Viral Env translation has been noted for its use of rarer codons which in turn, match to less abundant tRNAs. Current immunogens such as SOSIP however require the use of the most abundant codons in order to match the levels of protein production necessary for commercial use. If we consider that immunogen glycosylation is less processed and less occupied than viral glycosylation, these differences in codon optimization could be an attributing factor as the less optimised viral sequences might contribute to a prolonged translational period which can allow for greater time for both glycan attchment and processing. In this regard, Chapter 6 contains glycan analysis of 4 different strains of

Env on a SOSIP v4 backbone with codon optimized variants (SOSIPs) and variants with the original viral de-optimized sequences (SLOSIPs). The purpose of this chapter was to explore the effects of sequence optimization on glycosylation and to see if any of these effects were common across the different strains.

2 Materials and methods

2.1 Materials and methods used in Chapter 3

2.1.1 Protein expression and purification

BG505 SOSIP.664 and its knock-in/knock-out (KI/KO) variant proteins were transiently co-expressed in HEK 293F cells with a Furin expression plasmid at a ratio of (4:1). BG505 SOSIP proteins were purified using either 2G12, PGT145, or PGT151 affinity chromatography, as described in [109].

Briefly, transfection supernatants were vacuum filtered through 0.2 μm filters (Merck) and then passed (0.5–1 mL/min flow rate) over the column. The columns (Econo-Column Chromatography Columns, Bio-Rad) were made from CNBr-activated Sepharose 4B beads (GE Healthcare) coupled to the bnAb. A 0.5 M NaCl, 20 mM Tris, pH 8.0 buffer was used for column equilibration and washing. Bound Env proteins were eluted using 3 M MgCl2. The eluted proteins were immediately buffer exchanged into 75 mM NaCl, 10 mM Tris, pH 8.0. Comparisons were only drawn between BG505 SOSIP proteins purified using the same antibody.

2.1.2 ELISA

BG505 SOSIP.664 proteins were incubated in 96 well assay plates (Corning) at a concentration of 10 μg/mL in PBS overnight at 4 ^{O}C. Plates were washed with PBS 0.5% Tween (v/v) and blocked for 1 hour at room temperature with 5% milk in PBS 0.5% Tween. Following another wash step, the primary antibody was incubated (1:2 dilution series with a starting concentration of 20 μg/mL) in PBS for another hour at room temperature. Plates were once again washed and an anti-human IgG-horseradish peroxidase conjugant (Abcam) secondary antibody was added at a 1:2000 dilution in PBS. Plates were washed and TMB substrate

solution (Thermo Fisher Scientific) was added. The reaction was stopped after 5 minutes with sulphuric acid and the OD at 450 nm was measured.

2.1.3 UPLC: Sample preparation

N-linked glycans were released from BG505 SOSIP proteins by in-gel digestion with PNGase F (New England Biolabs). 30 μg BG505 SOSIP was resolved by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and stained with Coomassie Blue. Following destaining, the protein band was excised and washed alternatively with acetonitrile and water. Gel bands were then incubated with PNGase F for 16h at 37^{O}C. Released glycans were eluted from the gel with water and dried in a SpeedVac concentrator.

Released glycans were fluorescently labeled with procainamide (Abcam). Dried glycans were resuspended in 30 μL water before addition of 80 μL labeling mixture: 110 mg/mL procainamide, 60 mg/mL sodium cyanoborohydride in a solution of 70% dimethyl sulfoxide, 30% acetic acid. Samples were incubated at 65^{O}C for 4 h. Labeled glycans were purified using Spe-ed Amide-2 cartridges (Applied Separations).

2.1.4 UPLC: Data acquisition and analysis

Fluorescently labeled glycans were separated by Hydrophilic interaction liquid chromatography (HILIC-UPLC) on a Waters ACQUITY H-Class instrument using a 2.1mm $\times$ 100mm Glycan BEH Amide Column (1.7μm particle size; Waters). The following gradient was run: time=0min (t=0): 22% A, 78% B (flow rate of 0.5mL/min); t=38.5: 44.1% A, 55.9% B; t=39.5: 100% A, 0% B (0.25 mL/min); t=44.5: 100% A, 0% B; t=46.5: 22% A, 78% B (0.5 mL/min), where solvent A was 50mM ammonium formate, pH 4.4, and solvent B was acetonitrile. Fluorescence was measured using an excitation wavelength of 310nm and a detection wavelength of 370nm.

Quantification of oligomannose-type glycans was measured by digestion with Endo H, which cleaves oligomannose- (and hybrid-) type glycans, but not complex-type (New England Bio-

labs). Labeled glycans were resuspended in water and digested with Endo H for 16h at 37^{O}C. Digested glycans were cleaned using a PVDF protein-binding membrane plate (Merck Millipore) prior to HILIC-UPLC analysis as above. The abundance of oligomannose-type glycans was calculated, as a relative percentage, by integration of the HILIC-UPLC chromatograms before and after Endo H digestion, following normalization.

2.1.5 LC-MS Sample preparation

BG505 SOSIP proteins (100-150 μg each) were buffer exchanged using Vivaspin 100 kDa columns, denatured, reduced, and alkylated by sequential 1 h incubations at room temperature (RT) in the following solutions: 50 mM Tris/HCl, pH 8.0 buffer containing 6 M urea and 5 mM dithiothreitol (DTT), followed by the addition of 20 mM iodacetamide (IAA) for a further 1h at RT in the dark, and then additional DTT (20 mM), to eliminate residual IAA. The proteins were then buffer-exchanged into 50 mM Tris/HCl, pH 8.0 using Vivaspin 3 kDa columns and aliquots were digested with trypsin or chymotrypsin (Mass Spectrometry Grade, Promega) at a ratio of 1:30 (w/w) for 16 h at 37^{O}C. The reactions were dried and glycopeptides were extracted using C18 Zip-tip (Merck Millipore) following the manufacturer's protocol. Briefly, tips were equilibrated by alternating in acetonitrile and 0.1% trifluoracetic acid. The reaction mixture was loaded on to the tip and eluted with 50% acetonitrile, 0.1% trifluoracetic acid.

2.1.6 LC-MS: Data acquisition and analysis

Eluted glycopeptides were dried again and re-suspended in 0.1% formic acid prior to mass spectrometry analysis. An aliquot of glycopeptides 10-20 μL was analyzed by LC-MS with an Easy-nLC 1200 system coupled to an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) using higher energy collisional dissociation (HCD) fragmentation. Peptides were separated using an EasySpray PepMap RSLC C18 column (75 μm x 75 cm) with a 275 minute linear gradient consisting of 0%–32% acetonitrile in 0.1% formic acid over 240 minutes

followed by 35 minutes of 80% acetonitrile in 0.1% formic acid. The flow rate was set to 200 nL/min. The spray voltage was set to 2.8 kV and the temperature of the heated capillary was set to 275^{O}C. HCD collision energy was set to 50%, appropriate for fragmentation of glycopeptide ions. Glycopeptide fragmentation data were extracted from the raw file using Byonic™ (Version 2.7) and Byologic™ software (Version 2.3; Protein Metrics Inc.). The glycopeptide fragmentation data were evaluated manually for each glycopeptide; the peptide was scored as true-positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the glycan identified. The relative abundance of each glycan at each site was calculated using the extracted ion chromatograms for true-positive peptides. Oligomannose-type glycans are categorized according to the number of mannose residues present, hybrid-type glycans are categorized according to the presence/absence of fucose, and complex-type glycans are categorized according to the number of processed antennas and the presence/absence of fucose.

2.1.7 LC-MS: Relative quantification of glycan types

In order to acquire glycopeptide data for BG505 SOSIP sites with 1 PNGS/peptide, 100-150 μg BG505 SOSIP proteins were first digested with Endo H to cleave oligomannose- and hybrid-type glycans (high mannose), leaving a single GlcNAc residue at the corresponding site. The reaction mixture was then dried and resuspended in a mixture containing 50 mM ammonium bicarbonate and PNGase F using only ^{18}O-labeled water (Sigma-Aldrich) throughout. This second reaction cleaves the remaining complex-type glycans, leaving the GlcNAc residues from high-mannose glycans intact. The use of H_2^2O in this reaction enables complex glycan sites to be differentiated from unoccupied glycan sites as the hydrolysis of the glycosidic bond by PNGase F leaves an ^{18}O isotope on the resulting aspartic acid residue. The resultant peptides were purified by C18 ZipTip, as outlined above, and subjected to LC-MS in a similar manner to before; but using a lower HCD energy of 27% as glycan fragmentation was not required.

2.1.8 Model construction

A model was made by docking the gp120 and gp41 domains from the BG505 SOSIP.664 structure (PDB 5ACO [71]). Glycans were modeled on manually using Coot based on the most abundant glycoform for each site as per site-specific glycan data. Predominantly unoccupied glycan sites were filled in with a $Man_5GlcNAc_2$ glycan structure. Majority oligomannose-type glycan sites from relative data were filled in with a $Man_9GlcNAc_2$ glycan structure whilst complex-type sites were filled in with a fucosylated bi-antennary glycan structure.

2.2 Materials and methods used in chapter 4

2.2.1 Protein production and purification (From the laboratory of Dr. James Binley, San Diego Biomedical Research Institute)

VLPs were produced by co-transfecting human embryonic kidney 293T cells (ATCC) with an Env-expressing plasmid (typically pCAGGS JR-FL gp160Δ CT SOS E168K and mutants thereof), pMuLV Gag and pMV-Rev 0932 using polyethyleneimine (PEI Max, Polysciences, Inc.). Two days later, supernatants were collected, precleared by low speed centrifugation, filtered, and pelleted at 50,000 x g in a Sorvall SS34 rotor. To remove residual medium, VLP pellets were diluted with 1ml of PBS, then re-centrifuged at 15,000 rpm and resuspended in PBS at 1,000 x the original concentration. JR-FL gp120s were produced by transfecting human embryonic kidney 293F cells with PEI Max and purified via N-terminal 6x His-tag using a 5ml HisTrap column (GE). Columns were first equilibrated and washed using 50mM NaPO4 0.3M NaCl 10mM imidazole buffer at a flow rate of 5 ml/min. Loading of the column with the supernatant was performed at a flow rate of 1 ml/min. Proteins were eluted in 50 mM NaPO4 0.3 M NaCl 150 mM imidazole. Following elution, proteins were concentrated using 50 kDa molecular weight cut-off vivaspin columns (Sartorious) and buffer exchanged into PBS.

2.2.2 Glycosyltransferase co-transfection (From the laboratory of Dr. James Binley, San Diego Biomedical Research Institute)

Glycosyltransferase plasmids were co-transfected at a ratio of 1% total transfected DNA. The exception to this was B4GALT1 + ST6GAL1 co-expression, where B4GALT1 was transfected at 1% and the sialyltransferase at 2.5% total transfected DNA. For increasing galactosylation and increasing sialylation, 5 mM D-galactose (Sigma) was added to the medium prior, during and post-transfection in conjunction with co-transfection of glalactosyltransferases. the glycosyltransferase plasmids used for co-transfection were pEE6.4 for B4GALT1 and B4GALNT3, pEE14.4 for ST6GAL1, FUT6 and FUT9.

2.2.3 Reduction, alkylation, and digestion of JR-FL gp120s and VLPs

JR-FL gp120 monomer (100-150 μg) was denatured for 1h in 50 mM Tris/HCl, pH 8.0 containing 6M urea and 5 mM dithiothreitol (DTT). Next, Env proteins were reduced and alkylated with 20mM iodoacetamide (IAA) for 1h in the dark, followed by a 1h incubation with 20mM DTT to eliminate residual IAA. Alkylated Env proteins were buffer exchanged into 50mM Tris/HCl, pH 8.0 using Vivaspin columns (3 kDa). Aliquots were digested separately overnight using trypsin and chymotrypsin (Mass Spectrometry Grade, Promega). The next day, peptides were dried and extracted using C18 Zip-tip (Merck Millipore).

JR-FL VLPs were processed in the same way, except that were initially buffer exchanged into 50mM Tris HCl 0.1% Triton X-100 (w/w) to disperse lipids. To identify the glycome of the trimers that complexed with MAb CH01, VLPs were mixed with excess CH01 and incubated for 1h at 37^{O}C. Screw cap spin columns were incubated with protein A–agarose for 10 minutes to allow for spin column resin equilibration before washing with gentle Ag-Ab binding buffer (Thermo Fisher Scientific). VLP-CH01 complexes were then applied to the

spin columns and left to incubate for 30 minutes. Columns were washed twice with gentle Ag-Ab binding buffer prior to elution in 100–200 μL gentle Ag-Ab elution buffer (Thermo Fisher Scientific). Eluted VLP-CH01 mixtures were then buffer exchanged into 100μL 50mM Tris/HCl pH 8.0 for subsequent reduction and alkylation.

2.2.4 LC-MS: Settings and analysis

Peptides were dried again, re-suspended in 0.1% formic acid and analyzed by nanoLC-ESI MS with an Ultimate 3000 HPLC (Thermo Fisher Scientific) system coupled to an Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific) using stepped higher energy collision-induced dissociation (HCD) fragmentation. Peptides were separated using an EasySpray PepMap RSLC C18 column (75 μm $\times$ 75 cm). A trapping column (PepMap 100 C18 3μM 75μM x 2cm) was used in line with the LC prior to separation with the analytical column. LC conditions were as follows: 280 minute linear gradient consisting of 4–32% acetonitrile in 0.1% formic acid over 260 minutes, followed by 20 minutes of alternating 76% acetonitrile in 0.1% formic acid and 4% acetonitrile in 0.1% formic acid to ensure all the sample elutes from the column. The flow rate was set to 300nL/min. The spray voltage was set to 2.7 kV and the temperature of the heated capillary was set to 40^{o}C. The ion transfer tube temperature was set to 275^{o}C. The scan range was 3751500 m/z. Stepped HCD collision energy was set to 15%, 25% and 45% and the MS2 for each energy was combined. Precursor and fragment detection were performed with an Orbitrap at a resolution MS1 = 120,000, MS2 = 30,000. The AGC target for MS1 was set to standard and injection time set to auto which involves the system setting the two parameters to maximize sensitivity while maintaining cycle time.

2.2.5 Site-specific glycan classification

Glycopeptide fragmentation data were extracted from the raw file using Byos (Version 3.5; Protein Metrics Inc.). Glycopeptides were evaluated in reference to UniProtKB Q6BC19

(ectodomain of JR-FL gp160Δ CT). All samples carry additional mutations SOS (A501C, T605C) and E168K and N189A, along with other mutations, as denoted in the.txt files found in the MassIVE database (MSV000088108). Data were evaluated manually for each glycopeptide. A peptide was scored as true-positive when the correct b and y fragment ions were observed, along with oxonium ions corresponding to the glycan identified. The MS data was searched using the Protein Metrics "N-glycan 309 mammalian no sodium" library with sulfated glycans added manually. All charge states for a single glycopeptide were summed. The precursor mass tolerance was set at 4 ppm and 10 ppm for fragments. A 1% false discovery rate (FDR) was applied. Glycans were categorized according to the composition detected.

HexNAc(2)Hex(10+) was defined as M9Glc, HexNAc(2)Hex(93) was classified as M9 to M3. Any of these structures containing a fucose were categorized as FM (fucosylated mannose). HexNAc(3)Hex(56)X was classified as Hybrid with HexNAc(3)Hex(5–6)Fuc(1)X classified as Fhybrid. Complex glycans were classified according to the number of HexNAc subunits and the presence or absence of fucosylation. As this fragmentation method does not provide linkage information, compositional isomers are grouped, so, for example, a triantennary glycan contains HexNAc(5) but so does a biantennary glycans with a bisect. Core glycans refer to truncated structures smaller than M3. M9Glc- M4 were classified as oligomannose-type glycans. Glycans containing at least one sialic acid were categorized as NeuAc and at least one fucose residue in the "fucose" category.

Glycans were categorized into IDs ranging from 1 (M9Glc) to 19 (HexNAc(6+)(F)(x)). These values were multiplied by the percentage of the corresponding glycan divided by the total glycan percentage excluding unoccupied and core glycans to give a score that pertains to the most prevalent glycan at a given site. Arithmetic score changes were then calculated from the subtraction of these scores from one sample against others as specified.

2.2.6 Model construction

The model representation of the JR-FL SOS E168K+N189A trimer was constructed using SWISS-MODEL based on an existing structure of the 426c DS-SOSIP D3 trimer (pdb: 6MYY).Glycans were modelled on to this structure based on the most abundant glycoform identified from site-specific glycan analysis using WinCoot version 0.9.4.1 and ChimeraX version 1.2.5. For sites which were not identified, a $Man_9GlcNAc_2$ glycan was modelled. Conditional color formatting was used to illustrate the predominant glycoforms of modeled glycans, as follows: green (oligomannose), orange (mixed) and magenta (complex).

2.3 Materials and methods used in chapter 5

2.3.1 Protein expression and purification

HEK293F cells were maintained at a density of 1e6 - 3e6 cells per mL at 37 degrees with 8% CO2 and 125rpm shaking. Plasmids encoding BG505 SOSIP.664 containing a C-terminal 6xHis-tag were transiently co-transfected with a Furin-encoding plasmid (4:1) in HEK293F cells using PEI MAX transfection reagent. Plasmids encoding bnAbs or Trastuzumab were used in a 2:1 LC:HC ratio. The cells were transfected at a density of 1e6 cells per ml and incubated for 6 days at 37 degrees with 8% CO2 and 125rpm shaking.

WT BG505 SOSIP.664 trimers were purified using a 5ml HisTrap column (Cytiva). Columns were first equilibrated and washed using 50mM NaPO4 0.3 M NaCl 10 mM imidazole buffer at a flow rate of 5 ml/min. Loading of the column with the supernatant was performed at a flow rate of 1 ml/min. Proteins were eluted in 50 mM NaPO4 0.3 M NaCl 300 mM imidazole. Following elution, Env constructs were concentrated using 50 kDa molecular weight cut-off vivaspin columns (Sartorious) and buffer exchanged into PBS. Env trimers were further purified by size exclusion chromatography using a Superdex 200 column (GE Healthcare).

For constructs that were co-purified by VRC01, VRC02, or PGT151; separate transfections of the aforementioned bnAbs were first purified with a Protein A column (GE Healthcare).

Column equilibration and washing was carried out with a 20 mM NaPO4, pH 7.5 buffer, and 0.1 M citric acid, pH 3.0 was used for elution. Antibodies were concentrated using 50 kDa molecular weight cut-off vivaspin columns (Sartorious) and buffer exchanged into PBS prior to further purification via size exclusion chromatography using a Superdex 200 column (GE Healthcare). 2 mg antibody was then incubated in HEK293F filtered supernatant containing BG505 SOSIP.664 for 24h at 2^oC. SOSIP-Ab complexes were then co-purified and buffer exchanged into PBS with the same protein-A procedure used to initially purify the antibodies without SEC.

bnAb co-expressed SOSIPs were co-purified with a protein A column whilst sCD4 and Trastuzumab co-expressed SOSIPs were purified by means of HisTrap and SEC columns as previously outlined. The ratios of plasmids used, and the number of biological replicates expressed for each co-expression is outlined in Table 2.1.

Expressed plasmids	**No. Biological replicates**
BG505 SOSIP.664 6xHis 4:1 Furin	3
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 PGT145	1
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 PG16	1
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 PGT121	1
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 PGT135	1
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 2G12	1
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 PGT151	1
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 VRC01	6
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 VRC02	3
(BG505 SOSIP.664 6xHis 4:1 Furin) 2:1 VRC01	3
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:2 VRC01	3
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:2 Trastuzumab	3
(BG505 SOSIP.664 6xHis 4:1 Furin) 1:1 sCD4	3
(BG505 SOSIP.664 6xHis 4:1 Furin) VRC01p	3
(BG505 SOSIP.664 6xHis 4:1 Furin) VRC02p	3
(BG505 SOSIP.664 6xHis 4:1 Furin) PGT151p	1

Table 2.1: **Table of co-expressed proteins and the number of biological replicates expressed**

2.3.1.1 LC-MS: Sample preparation

50ug protein aliquots were resolved by SDS–PAGE using NuPAGE™ 4-12% Bis-Tris gels (Thermo Fisher Scientific) and stained with Coomassie Blue. SDS PAGE settings were 125V and 100 mA for 45 minutes. Following destaining, excised SDS-PAGE gel bands corresponding to gp140 were incubated in 500 uL acetonitrile until opaque. The acetonitrile was then removed and replaced with 50 uL 10 mM Dithiothrietol (DTT) in 100 mM Ammonium Bicarbonate (AmBic). The bands were incubated in the DTT solution for 30 minutes at 56^{o}C. The gel pieces were cooled down and washed with 500 uL acetonitrile. Gel pieces were then incubated in the dark for 20 minutes in 50 uL 55 mM Iodoacetamide (IAA) in 100 mM AmBic. Bands were then washed once with 500 uL 100 mM AmBic and acetonitrile.

Washed bands were then incubated on ice with either α-lytic protease, trypsin or chymotrypsin (13 ng/uL) in a solution of 10 mM AmBic 10% acetonitrile. Further enzymatic solution was added and left to incubate on ice for another 90 minutes ensuring the gel bands were fully submerged. 10-20 uL 10 mM AmBic was added and then this mixture was incubated overnight at 37^{o}C. Following digestion, the gel pieces were spun down and the supernatant was set aside in a fresh tube. The individual supernatants were dried and glycopeptides were extracted using C18 Zip-tip (Merck Millipore). Tips were equilibrated by alternating in acetonitrile and 0.1% trifluoracetic acid. The reaction mixture was loaded on to the tip and eluted with 50% acetonitrile, 0.1% trifluoracetic acid.

2.3.2 LC-MS: Settings and analysis

Eluted glycopeptides were dried again and re-suspended in 0.1% formic acid prior to mass spectrometry analysis. Glycopeptides were analyzed by LC-MS with an Easy-nLC 1200 system coupled to an Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific) using stepped higher energy collisional dissociation (HCD) fragmentation (15, 25, 45%). Peptides were separated using an EasySpray PepMap RSLC C18 column (75 μm x 75 cm) with a

275 minute linear gradient consisting of 0%–32% acetonitrile in 0.1% formic acid over 240 minutes followed by 35 minutes of 80% acetonitrile in 0.1% formic acid. The flow rate was set to 200 nL/min. The ion transfer tube temperature was set to 275^{o}C. The scan range was 400 - 1600 m/z. The spray voltage was set to 2.8 kV and the temperature of the heated capillary was set to 40^{o}C. Precursor and fragment detection were performed using Orbitrap at a resolution MS1 = 100,000. MS2 = 30,000. The AGC target for MS1 = 4e5 and MS2 = 5e4 and injection time: MS1 = 50 ms MS2 = 54 ms.

Glycopeptide fragmentation data were extracted from the raw file using Byos (Protein Metrics Inc.). The following parameters were used for data searches in Byonic: The precursor mass tolerance was set at 4 ppm and 10 ppm for fragments. Peptide modifications included in the search include: Cys carbamidomethyl, Met oxidation, Glu $\rightarrow$ pyroGlu, Gln $\rightarrow$ pyroGln and N deamidation. For each protease digest a separate search node was used with digestion parameters appropriate for each protease (Trypsin: RK, Chymotrypsin: YFW, α-lytic protease: TASV) using a semi-specific search with 2 missed cleavages. A 1% false discovery rate (FDR) was applied. Digests of the same biological repilicate were combined into a single file for downstream analysis. All charge states for a single glycopeptide were summed. The glycopeptide fragmentation data were evaluated manually for each glycopeptide; the peptide was scored as true-positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the glycan identified.

The relative amounts (determined by comparing the XIC of each glycopeptide, summing charge states) of each glycan at each site as well as the unoccupied proportion were determined by comparing the extracted chromatographic areas for different glycotypes with an identical peptide sequence. Glycans were categorized according to the composition detected. HexNAc(2), Hex(9-3) was classified as M9 to M3. Any of these compositions that were detected with a fucose are classified as FM. HexNAc(3)Hex(5-6)Neu5Ac(0-4) was classified as Hybrid with HexNAc(3)Hex(5-6)Fuc(1)NeuAc(0-1) classified as Fhybrid. Complex-type glycans were classified according to the number of processed antenna and

fucosylation. Complex-type glycans are categorized as HexNAc(3)(X), HexNAc(3)(F)(X), HexNAc(4)(X), HexNAc(4)(F)(X), HexNAc(5)(X), HexNAc(5)(F)(X), HexNAc(6+)(X) and HexNAc(6+)(F)(X). Core glycans are any glycan smaller than HexNAc(2)Hex(3).

2.3.3 Model construction of BG505 SOSIP trimer

The model representation of the BG505 SOSIP trimer was constructed using PDB ID: 6V0R [210] . Glycans were modelled on to this structure based on the most abundant glycoform identified from site-specific glycan analysis using WinCoot version 0.9.4.1 and ChimeraX version 1.2.5. For sites which were not identified, a $Man_9GlcNAc_2$ glycan was modelled. Conditional color formatting was used to illustrate the predominant glycoforms of modeled glycans, as follows: green (oligomannose), orange (mixed) and magenta (complex). For percentage change model representations the key provided in the figure indicated percentage point increases (red) or decreases (blue) in the abundance of high-mannose glycoforms.

2.4 Materials and methods used in Chapter 6

2.4.1 Design of SOSIP and SLOSIP constructs

Envelope nucleotide sequences were obtained from NCBI GenBank for four strains originating from different HIV-1 clades: BG505 (Clade A, GenBank: MW650637.1), B41 (Clade B, GenBank: EU576114.1), ZM197M (Clade C, GenBank: DQ388515.1), X1193 (Clade G, EU885761.1) [226, 227, 228, 229]. These native Env sequences were combined with mutations that make up SOSIP v4.1 trimers (E64K, A316W) in addition to the mutations included in BG505 SOSIP [114, 109]. The plasmids that comprise the de-optimized, native *env* sequences, are referred to as SLOSIPs; whereas the plasmids that comprise the same protein sequence but makes use of the most optimized codons as per Thermo Fisher GeneArt are referred to as SOSIPs.

2.4.2 Protein expression and purification

HEK293F cells were maintained at a density of 1e6 - 3e6 cells per mL at 37 degrees with 8% CO2 and 125rpm shaking. Plasmids encoding SOSIP or SLOSIP v4.1 trimers containing a C-terminal 6xHis-tag were transiently co-transfected with a Furin-encoding plasmid (4:1) in HEK293F cells using PEI MAX transfection reagent. Trimers were purified using a 5ml HisTrap column (Cytiva). Columns were first equilibrated and washed using 50mM NaPO4 0.3 M NaCl 10 mM imidazole buffer at a flow rate of 5 ml/min. Loading of the column with the supernatant was performed at a flow rate of 1 ml/min. Proteins were eluted in 50 mM NaPO4 0.3 M NaCl 300 mM imidazole. Following elution, Env constructs were concentrated using 50 kDa molecular weight cut-off vivaspin columns (Sartorious) and buffer exchanged into PBS. Env trimers were further purified by size exclusion chromatography using a Superdex 200 column (GE Healthcare).

2.4.3 LC-MS: Sample preparation

50ug protein aliquots were resolved by SDS–PAGE using NuPAGE™ 4-12% Bis-Tris gels (Thermo Fisher Scientific) and stained with Coomassie Blue. SDS PAGE settings were 125V and 100 mA for 45 minutes. Following destaining, excised SDS-PAGE gel bands corresponding to gp140 were incubated in 500 uL acetonitrile until opaque. The acetonitrile was then removed and replaced with 50 uL 10 mM Dithiothrietol (DTT) in 100 mM Ammonium Bicarbonate (AmBic). The bands were incubated in the DTT solution for 30 minutes at 56^{o}C. The gel pieces were cooled down and washed with 500 uL acetonitrile. Gel pieces were then incubated in the dark for 20 minutes in 50 uL 55 mM Iodoacetamide (IAA) in 100 mM AmBic. Bands were then washed once with 500 uL 100 mM AmBic and acetonitrile.
Washed bands were then incubated on ice with either α-lytic protease, trypsin or chymotrypsin (13 ng/uL) in a solution of 10 mM AmBic 10% acetonitrile. Further enzymatic

solution was added and left to incubate on ice for another 90 minutes ensuring the gel bands were fully submerged. 10-20 uL 10 mM AmBic was added and then this mixture was incubated overnight at 37^{o}C. Following digestion, the gel pieces were spun down and the supernatant was set aside in a fresh tube. The individual supernatants were dried and glycopeptides were extracted using C18 Zip-tip (Merck Millipore). Tips were equilibrated by alternating in acetonitrile and 0.1% trifluoracetic acid. The reaction mixture was loaded on to the tip and eluted with 50% acetonitrile, 0.1% trifluoracetic acid.

2.4.4 LC-MS: Settings and analysis

Eluted glycopeptides were dried again and re-suspended in 0.1% formic acid prior to mass spectrometry analysis. Glycopeptides were analyzed by LC-MS with an Easy-nLC 1200 system coupled to an Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific) using stepped higher energy collisional dissociation (HCD) fragmentation (15, 25, 45%). Peptides were separated using an EasySpray PepMap RSLC C18 column (75 μm x 75 cm) with a 275 minute linear gradient consisting of 0%–32% acetonitrile in 0.1% formic acid over 240 minutes followed by 35 minutes of 80% acetonitrile in 0.1% formic acid. The flow rate was set to 200 nL/min. The ion transfer tube temperature was set to 275^{o}C. The scan range was 400 - 1600 m/z. The spray voltage was set to 2.8 kV and the temperature of the heated capillary was set to 40^{o}C. Precursor and fragment detection were performed using Orbitrap at a resolution MS1 = 100,000. MS2 = 30,000. The AGC target for MS1 = 4e5 and MS2 = 5e4 and injection time: MS1 = 50 ms MS2 = 54 ms.

Glycopeptide fragmentation data were extracted from the raw file using Byos (Protein Metrics Inc.). The following parameters were used for data searches in Byonic: The precursor mass tolerance was set at 4 ppm and 10 ppm for fragments. Peptide modifications included in the search include: Cys carbamidomethyl, Met oxidation, Glu $\rightarrow$ pyroGlu, Gln $\rightarrow$ pyroGln and N deamidation. For each protease digest a separate search node was used with digestion parameters appropriate for each protease (Trypsin: RK, Chymotrypsin: YFW, α-lytic pro-

tease: TASV) using a semi-specific search with 2 missed cleavages. A 1% false discovery rate (FDR) was applied. Digests of the same biological repilicate were combined into a single file for downstream analysis. All charge states for a single glycopeptide were summed. The glycopeptide fragmentation data were evaluated manually for each glycopeptide; the peptide was scored as true-positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the glycan identified.

The relative amounts (determined by comparing the XIC of each glycopeptide, summing charge states) of each glycan at each site as well as the unoccupied proportion were determined by comparing the extracted chromatographic areas for different glycotypes with an identical peptide sequence. Glycans were categorized according to the composition detected. HexNAc(2), Hex(9-3) was classified as M9 to M3. Any of these compositions that were detected with a fucose are classified as FM. HexNAc(3)Hex(5-6)Neu5Ac(0-4) was classified as Hybrid with HexNAc(3)Hex(5-6)Fuc(1)NeuAc(0-1) classified as Fhybrid. Complex-type glycans were classified according to the number of processed antenna and fucosylation. Complex-type glycans are categorized as HexNAc(3)(X), HexNAc(3)(F)(X), HexNAc(4)(X), HexNAc(4)(F)(X), HexNAc(5)(X), HexNAc(5)(F)(X), HexNAc(6+)(X) and HexNAc(6+)(F)(X). Core glycans are any glycan smaller than HexNAc(2)Hex(3).

2.4.5 Model construction of SOSIP/SLOSIP trimers

The model representation of the BG505, B41, ZM197M and X1193 trimers were constructed using SWISS-MODEL based on an existing structure of 8EU8 for BG505, B41, ZM197M and X1193 trimers which were constructed separately from the same structure [230]. Glycans were modelled on to this structure based on the most abundant glycan category identified from site-specific glycan analysis using WinCoot version 0.9.4.1 and ChimeraX version 1.2.5. For sites which were predominantly oligomannose-type, a $Man_9GlcNAc_2$ glycan was modelled. For mixed sites, a $Man_5GlcNAc_2$ glycan was modelled. For complex-type sites, a bi-antennary complex-type glycan was modelled. For sites which could not be derived or

those that were not identified, a $Man_5GlcNAc_2$ glycan was modelled. Conditional color formatting was used to illustrate the predominant glycoforms of modeled glycans, as follows: green (oligomannose), orange (mixed) and magenta (complex), under-occupied (cyan), not derived (n.d.; grey).

3 Mutations affecting the glycan shield

The glycans on the HIV-1 envelope spike form epitopes for broadly neutralising antibodies. The breadth at which they are able to neutralize is of note when considering that the processing of these glycans can be substantially influenced depending on the presence or absence of neighbouring glycan attachment sequons on this highly mutable protein. This chapter addresses the extent to which mutations which add or delete these glycan sites within a network of glycans comprising an antibody epitope impact on their processing, as well as the impact on antibody binding.

3.1 Contributions

Chapter 3 contains a comparison of several BG505 Env constructs whose residues were individually introduced to add or delete glycan PNGSs. The results from this chapter were obtained in collaboration with Dr. Gemma Seabright with whom we collectively obtained the glycan and ELISA data. The samples analysed are as listed in the following table:

	SOSIPs analysed
1	BG505 SOSIP.664
2	BG505 SOSIP.664 + N295A
3	BG505 SOSIP.664 + N332T
4	BG505 SOSIP.664 + N339A
5	BG505 SOSIP.664 + N386A
6	BG505 SOSIP.664 + N392A
7	BG505 SOSIP.664 + N411A
8	BG505 SOSIP.664 + N448A
9	BG505 SOSIP.664 + N241
10	BG505 SOSIP.664 + N289
11	BG505 SOSIP.664 + N241 + N289
12	BG505 SOSIP v5 + N289
13	BG505 SOSIP v5 + N241 + N289

Table 3.1: **Table of SOSIPs analysed in chapter 2**

3.2 Impact of point mutations on glycan processing

The HIV-1 envelope spike is the sole protein expressed on the exterior of the virion and is the only target for a bnAb response. This makes it an attractive target for vaccine design whose strategies involve the reproduction of recombinant trimers to mimic the native Env trimer for its bnAb epitopes and immunogens specifically designed to target the germline-encoded bnAb precursor (gl-bnAb), otherwise known as the B cell receptor [231, 232].

Whilst the purpose of the glycan shield of Env is to protect the underlying protein surface from neutralising Abs, bnAbs have evolved to neutralize epitopes consisting either partially or entirely of N-linked glycans [143, 153, 233, 234, 235, 236, 237, 238, 239, 240]. These bnAbs recognise glycans at four distinct regions of Env: The V1/V2 loops at the trimer apex, the area surrounding the CD4bs, the gp120-gp41 interface, and the cluster of oligomannose-type glycans on the outer domain of gp120 [152]. Many bnAbs have been found to target the N332 glycan central to this cluster of oligomannose glycans which has led to it being dubbed as the 'supersite of immune vulnerability' [151].

One of the most widely studied Env mimics are the BG505 SOSIP.664 trimers due to their native-like structure and antigenicity; however, the transmitted/founder virus is inherently lacking in the conserved N332 site which was introduced into BG505 SOSIP.664 by means of a T332N point mutation [109]. In addition to this site, glycans N241 and N289 which are respectively present in 97% and 72% of HIV-1 isolates are also missing from the BG505 sequence. The presence of breaches within the glycan shield has drawn a lot of discussion of late due to their associated roles in initiating immunodominant nAb responses that direct away from responses against broadly neutralising epitopes [214, 241, 242, 243, 207, 244].

When rabbits were immunised with BG505 SOSIP trimers, the antibody responses they elicited were found to target the N241/289 holes by means of autologous nAbs [207, 245]. Filling in these breaches by means of glycan sequon insertions block nAb responses directed at these positions [245]. Similar observations occurred with autologous nAbs against holes in

positions 130, 197, and 465 in trimer immunogenicity studies of different isolates [214, 242, 243, 207, 244]. It is possible to redirect these autologous nAb responses in vivo by filling in these holes and creating alternate ones elsewhere on the trimer [213]. Such findings reflect phenomena already observed within the context of a natural infection, as certain glycans are able to shift their sequon positions in face of the selective pressures imposed on them by nAbs [49, 141, 211, 246]. The N332 glycan has been observed to shift to N334 and back in response to nAbs during infection [49].

It is understood that breaches in the glycan shield of Env offer immunodominant distractions that enable the elicitation of autologous nAbs; the extent to which this distraction detracts from bnAb development is yet to be characterised. Previous studies have shown that that the level of glycan shield integrity (i.e. lack of breaches) in transmitted/founder strains correlate with breadth of neutralization in infected individuals [211]. In future, it may be possible that immunogens may be engineered to remove these glycan holes in order to redirect the antibody response to broadly neutralising glycan epitopes [213, 245].

A potential drawback to this approach however relates to the elicitation of gl-bnAbs at their precursor B-cells. In order to be effective, an immunogen must be able to engage the gl-bnAb before affinity maturation of the bnAb. However, gl-bnAbs have low affinity with Env which is driven by their inability to accommodate conserved N-linked glycans [115, 247, 248, 249, 248, 250]. An alternative approach to eliciting bnAbs would be to initially prime with a glycan-depleted immunogen that can engage the gl-bnAb and then boost with immunogens that have those initially missing glycans in order to stimulate the development of neutralization breadth [251, 252, 223, 253, 232].

The density of the glycans presented on the HIV-1 envelope impacts the amount of glycosylation processing some of these glycans can undergo, which in turn influences the presentation of epitopes, as some sites will be unable to be processed by host cell α-mannosidases [70]. Links between glycan density and the glycans displayed on env have been previously established [254, 255]. Studies on recombinant, monomeric gp120 and trimeric immunogens have

revealed the presence of a population of under processed oligomannose-type glycans known as the intrinsic mannose patch [102, 120, 134, 256].

In gp120 mutant studies where single glycan sites from the IMP were removed, decreases in oligomannose-type glycans in the surrounding sites were found to be greater than expected as the glycan site deletions made the others nearby more susceptible to processing 188. Env trimers, particularly those being touted as possible immunogens, display a native-like conformation that mimics the structure of Env.

This structure imparts additional steric restrictions compared to gp120 due to additional glycan and protein components from the additional protomers and gives rise to further under-processed clusters of oligomannose-type glycans, known as the trimer-associated mannose patch [155, 255]. Thus, whilst oligomannose-type glycans constitute a conserved feature of the glycan shield of Env and a target for bnAbs, their presence is influenced by their local environment and can be subject to glycosylation processing if the environment is suitable.

With these factors in mind, this section aimed to establish the impact of glycan site additions and deletions on the glycosylation of a native-like envelope trimer, with particular interest in this effect at bnAb binding sites. Site-specific analysis of trimeric BG505 SOSIP.664 showed that addition and deletion of glycan sites not only influenced local (sites proximal to addition or deletion) glycosylation but also the processing in sites further away from the site of addition or deletion. This may allude to a form of glycan dependency upon certain sites for some bnAbs that rely on glycans for their epitopes.

3.3 Results

Glycan deletion increases mannose processing throughout the trimer

The BG505 SOSIP.664 trimer sequence was used to make the following trimer mutants with individual glycan deletions: N295A, N332T, N339A, N386A, N392A, N411A, and N448A. These mutations were introduced at the consensus sequence N-x-S/T (where x is any amino

acid except proline. These sites were chosen as they either directly form part of the 2G12 epitope (N295, N332, N339, N392) or are located within the surrounding glycan network (N411, N448, N386) (Figure 3.1B). As a benchmark for the glycan analysis on these deletion mutants, we performed glycopeptide analysis on three biological replicates of BG505 SOSIP.664 to validate that observed differenecs are attributed to the mutations (Figure 3.1A).

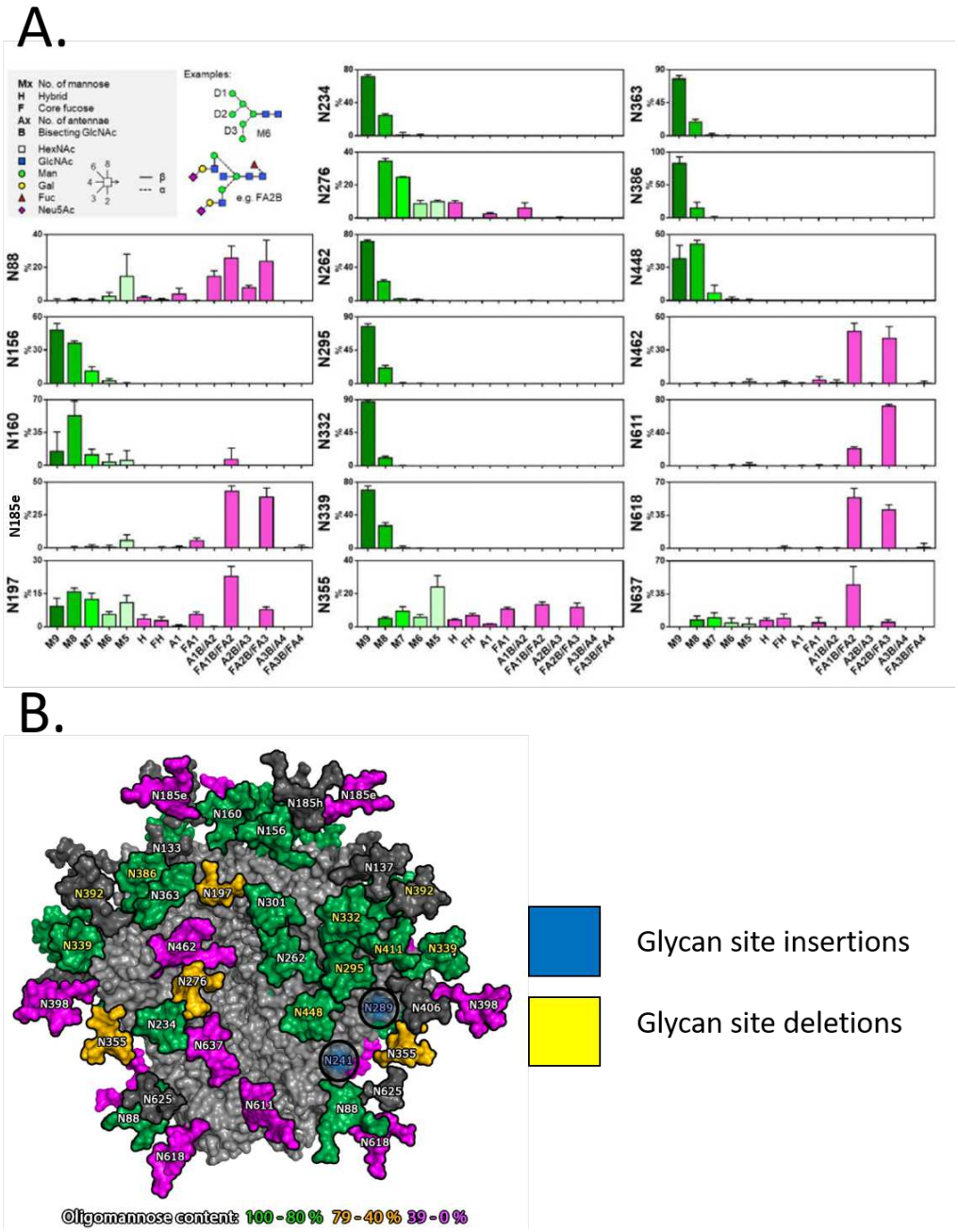


Figure 3.1: **Site-specific glycopeptide analysis of PGT145 purified BG505 SOSIP.664 and schematic of glycan site insertions and deletions surrounding 2G12 epitope.**

(A) Site-specific quantification of three biological replicates of BG505 SOSIP.664. bars represent standard error of mean. HexNAc = N-acetylhexosamine, GlcNAc = N-acetylglucosamine, Man = mannose, Gal = galactose, Fuc = fucose, Neu5Ac = sialic acid.

(B) Schematic of the BG505 SOSIP.664 trimer with glycans coloured according to their oligomannose content, derived from **(A)**. Trimers were reduced, alkylated, and digested with trypsin or chymotrypsin before analysis by LC-ESI MS. Sites highlighted in yellow represent those that were removed in trimer mutants, whilst those in blue represent sites added in trimer mutants. Figure adapted from [257].

Glycan site deletion in the aforementioned mutants resulted in a general increase in glycan processing at the adjacent sites. This was most prominent in the N332T, N411A, and N448A mutants which surround the N295 site. Loss of any of these three glycans resulted in a 61, 80 and 78 percentage point (p.p.) reduction in $Man_9GlcNAc_2$ glycans respectively, which was mainly compensated for with an increase in $Man_{5-8}GlcNAc_2$ structures. A reciprocal, albeit lesser, effect was observed in the N295A mutant whose absence increased processing at N332 and N448 sites (-18 and -37 p.p. $Man_9GlcNAc_2$ respectively, Figure 3.2). Elsewhere, the N386A mutant caused a 42 p.p. reduction in $Man_9GlcNAc_2$ at the neighbouring N363 site, and the N392A mutant decreased Man_9GlcNA_2 by 30 p.p. at the N339 site and 35 p.p. at the N386 site (Figure 3.2).

Increased processing at sites were also observed at sites further away from the site of the glycan deletion. The N332T mutant for example increased processing at the N448 and N339 sites (-30, -35 p.p., respectively; Figure 3.2). Once again this effect was reciprocated a the N448A mutant increased N332 processing (-21 p.p., Figure 3.2). Elsewhere, the N234 site which forms a cluster of glycans around the CD4 binding site further away from the IMP showed increased processing following deletion of N386, N392, or N448 sites (Figure 3.2). These results relay a varied impact of glycan deletions on the processing state of both neighbouring glycans and those further afield, which in turn may be related to underlying glycan-glycan interactions [258].

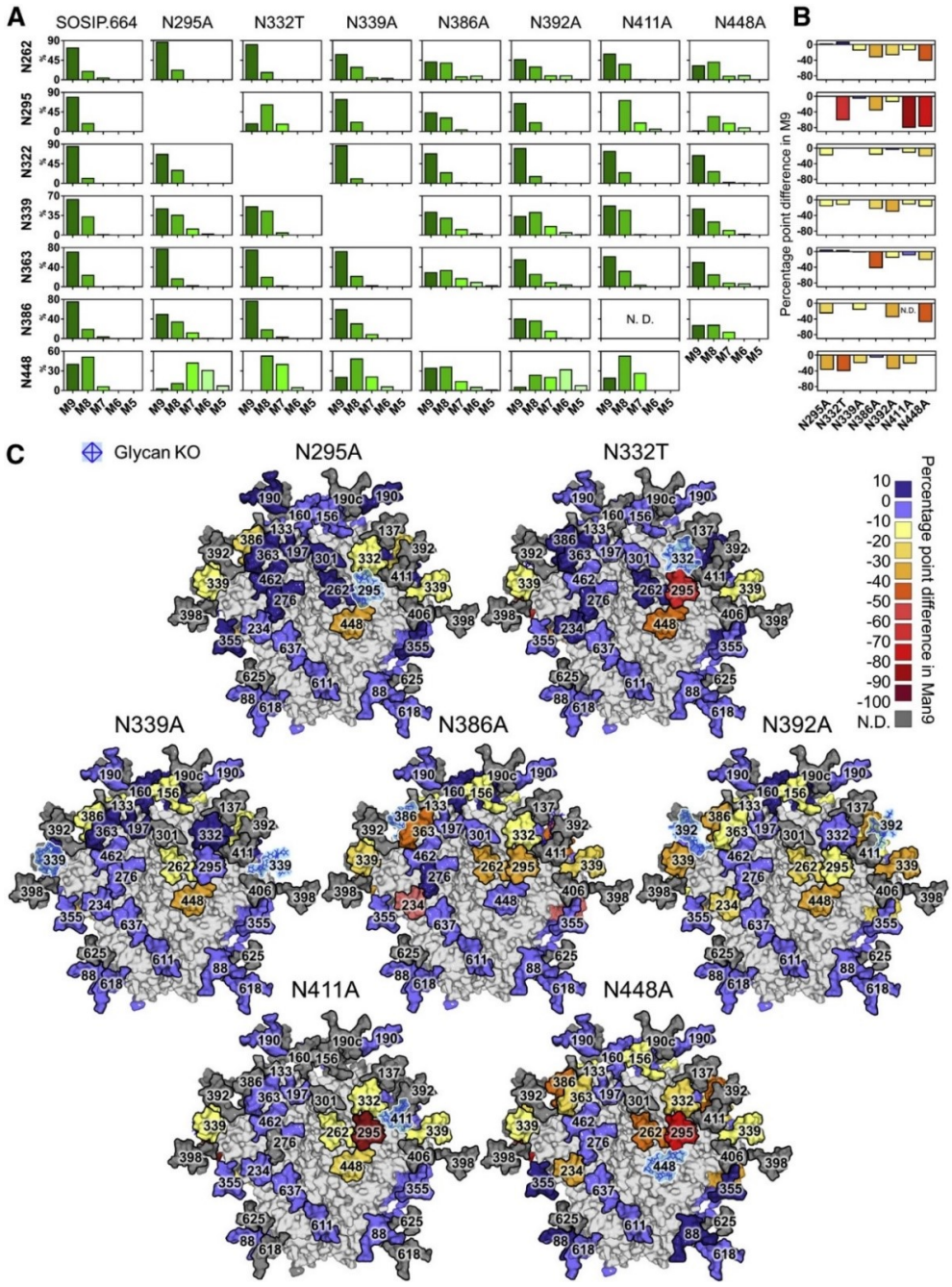


Figure 3.2: **Glycan deletion increases mannose trimming throughout the trimer.** **(A)** Relative quantification of IMP sites from BG505 SOSIP.664 and the N295A, N332T, N339A, N386A, N392A, N411A, and N448A glycan knockouts. M9 = Man9 (dark green) to M5 = Man5 (pale green). **(B)** The percentage point difference in the abundance of Man9 at IMP sites in the glycan knockouts, compared with BG505 SOSIP.664. Decreases in the abundance of Man9 are colored as per the key in **(C)**. **(C)** Heatmap demonstrating the percentage point difference in the abundance of Man9 at each site in the glycan knockouts, compared with BG505 SOSIP.664. Differences are calculated as follows: (% Man9 in knockout % Man9 in BG505 SOSIP.664). KO, knockout; N.D., not determined. Figure reproduced from [257]. Mutant datasets are from one biological replicate subtracted from the mean WT dataset.

Glycan addition confers differential effects on glycan processing

The BG505 SOSIP trimer was used to make three mutants carrying glycan knock-in mutations at positions 241, 289, and both simultaneously. The +N241 mutant was found to restrict glycan processing at the neighbouring N448 site which displayed a +21 p.p. increase in $Man_9GlcNAc_2$ glycans (Figure 3.3). This finding was in agreement with the idea that glycan density limits mannosidase processing [70].

In contrast, the +N289 mutant resulted in an increase in glycan processing at glycans proximal to the introduced site and further away throughout the trimer. Nearby N262, N295, and N332 sites displayed a decrease in $Mann_9GlcNAcn_2$ glycans (-37, -20, -28 p.p. decrease, respectively; Figure 3.3). N363 and N234 sites were similarly affected despite their distance from the knock-in. The +N241, N289 double glycan mutant was found to restrict processing at N448 (+12 p.p. $Man_9GlcNAc_2$ glycans) and also increase N262, N295, N332, and N363 processing (Figure 3.3).

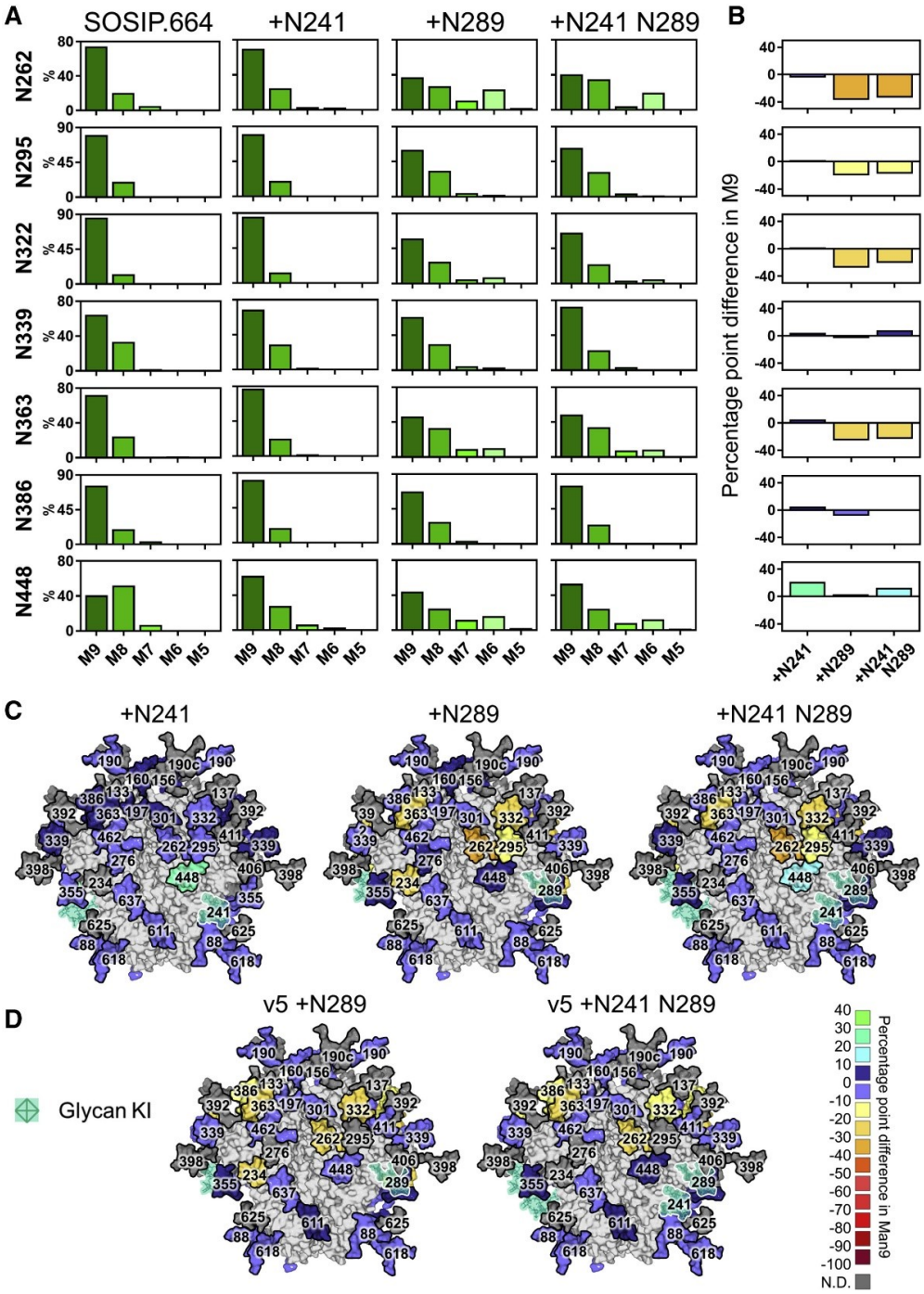


Figure 3.3: **Differential effect of glycan additions on glycosylation processing**
(A) Relative quantification of IMP sites from BG505 SOSIP.664 and the +N241, +N289, and +N241 N289 glycan knock-ins. M9 = Man9 (dark green) to M5 = Man5 (pale green).
(B) The percentage point difference in the abundance of Man9 in the glycan knock-ins, compared with BG505 SOSIP.664. Graphs are colored according to the key in **(D)**.
(C) Heatmap demonstrating the percentage point difference in the abundance of Man9 at each site in the glycan knock-ins, compared with BG505 SOSIP.664: (% Man9 in knock-in - % Man9 in BG505 SOSIP.664).
(D) Percentage point difference in the abundance of Man9 at each site in the SOSIP.v5 glycan knock-ins compared with BG505 SOSIP.v5. KI, knock-in. Adapted from [257]. Mutant datasets are from one biological replicate subtracted from the mean WT dataset.

The increased glycan processing associated with the +N289 mutation was surprising considering that increased glycan density is often associated with the opposite process. It was thought that as the BG505 SOSIP.664 protein generally lacks this site it may be due to an inability to accommodate a glycan at this position, and that its presence may be conferring conformational changes to the protein that may facilitate mannosidase trimming. If true, then it is possible that certain glycan epitopes may be adversely affected. To address this, we assessed binding of these mutants to a panel of antibodies targeting distinct epitopes on env by ELISA and found the mutants to be antigenically similar to the unmutated BG505 SOSIP.664 protein (Figure 3.4B).

In addition, we sought to determine the impact of these mutants on a hyperstabilised variant of the BG505 SOSIP.664 protein, SOSIP.v5 [259]. Glycopeptide analysis of this protein and its mutants confirmed that N289 addition coincided with increased glycan processing, albeit to a lesser extent than what was observed on the SOSIP.664 protein (Figure 3.3D). The SOSIP.v5 double mutant also showed similar changes to the SOSIP.664 double mutant, but also to a smaller extent.

It was thought that the observed decreases in oligomannose-type glycans in the trimer mutants could be attributed to the added sites comprising predominantly complex-type glycans. It was not possible to determine the precise composition of the N241 or N249 sites as they co-occupy digested peptides with the N234 and N295 sites, respectively. As a means to broadly determine their composition, we subjected glycopeptide fragments of the trimer mutants to sequential digests with Endo H (cleaves oligomannose-type glycans) and PNGase F to cleave complex-type glycans 140. In all mutants, the N241 and N289 sites were found to be predominantly occupied with oligomannose-type glycans, which confirms that decreases in oligomannose-type structures are attributed to increased glycan processing elsewhere on the trimer (Figure 3.4A).

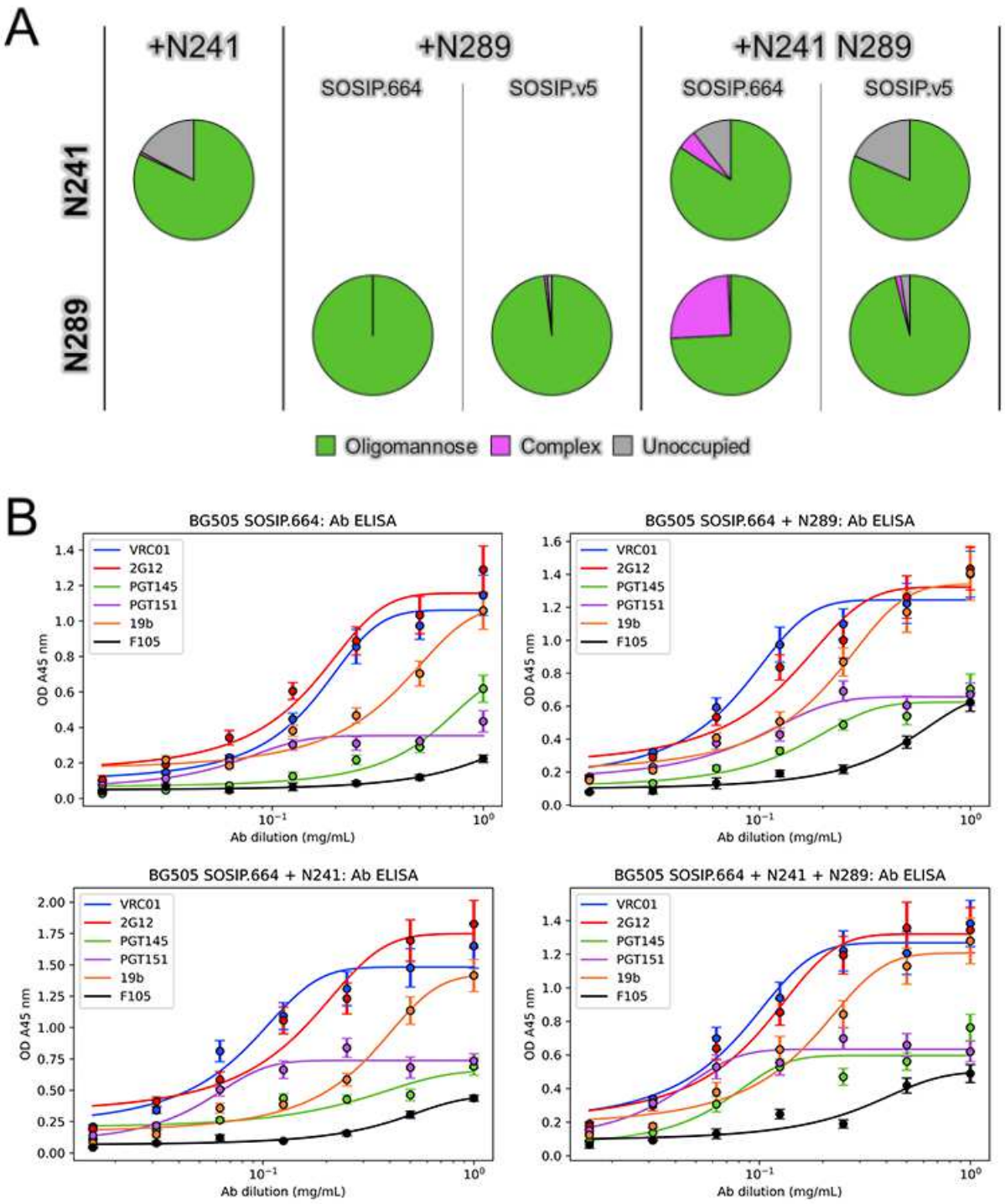


Figure 3.4: **Glycan analysis of +N241 and +N289 sites**
(A) Relative quantification of glycan-types occupying the N241 and N289 sites in the SOSIP.664 and SOSIP.v5 glycan addition mutants. Glycopeptides were digested with Endo H to cleave oligomannose-type glycans, leaving behind a GlcNAc residue. The remaining complex-type glycans were then digested with PNGase F in the presence of ^{18}O- labeled water, to convert the Asn to an ^{18}O-labeled Asp, prior to LC-MS analysis. Datasets are from one biological replicate for each SOSIP and SOSIP.v5 mutant.
(B) ELISA binding of SOSIP.664 mutants to a panel of env-targeting neutralizing antibodies to confirm antigenic similarity to wild-type BG505 SOSIP.664. Adapted from [257]. Data and SEM error bars from three technical repeats.

3.4 Discussion

In the context of a natural HIV infection, the high density of glycans on Env suggests an advantage in using host-cell derived glycans to evade immune surveillance [260]. However, the glycan shield does not evolve to a maximum number. Instead, the shield must adapt, by means of glycan rearrangements, for effective immune evasion [49, 141, 211, 246].

Using the BG505 SOSIP.664 env trimer as a model system, this section aimed to assess the impact of glycan deletions or additions in the form of point mutations on the processing state of glycan sites proximal and further away from the point of mutation. With regards to vaccination strategies, the development of neutralization breadth may be aided through strategies involving immunogens with 'filled-in' glycan holes or using immunogens with deliberate holes in certain positions on the protein [251, 252, 223, 253, 232]. This is of particular interest as bnAbs may evolve to recognise these structures [49].

The 2G12 antibody directly targets four oligomannose-type glycans at the IMP, and it is thought that the surrounding glycans also play a role in binding. By deleting glycan sites either directly contributing to its epitope or those surrounding it, previous studies have shown that such interventions may disrupt antibody binding and neutralization [69, 140, 245]. This may partly be explained by the idea that if a glycan site is removed, mannosidases may no longer be sterically hindered from accessing neighbouring glycan sites. This would be reflected in the composition of glycans at these neighbouring sites in the form of reduced oligomannose-type glycans.

Compared to viral-derived env, analysis of soluble BG505 SOSIP.664 trimers yields a slightly higher level of oligomannose-type glycans [104, 261]. Whilst it could be argued that mutations that increase the processing of SOSIP.664 could be beneficially used to better sync the glycan profile to that of viral env, sites that are predominantly oligomannose-type tend to be conserved between the two [104, 261]. The integrity of the mannose patch is associated with correctly folded trimers in the native-like conformation, and substantial derivations in

abundance of oligomannose glycans at particular sites can indicate non native-like conformations [107]. Therefore, the location of glycan site mutations are of significant importance as they may inadvertently affect the conformation of the glycoprotein [256, 262, 263, 264]. In response to this, all trimers used in these experiments were purified using a column loaded with a quaternary structure-dependent antibody, PGT145 or PGT151, as a means of reducing the amount that misfolded proteins may contribute to the observed data [234, 239]. All mutants displayed oligomannose-type glycan profiles that reflect a native-like configuration [69, 123]. The +N289 mutant was the only trimer mutant that exhibited a larger than expected increase in glycosylation processing in both SOSIP.664 trimers and SOSIP.v5, suggesting that the mutation may be adversely affecting the conformation of the trimer. The mutation to mutant involved replacing a proline at position 291. This conformational issue could betherefore attributed to glycan-glycan interactions as well as changes in the underlying protein ructure due to the known role of proline as a helix-breaker.

Understanding the role of glycans and how they affect one another will be an important part of understanding how immune selective pressure may influence viral mutations and, in turn, bnAb epitope presentation. With regards to immunogen design, the removal or introduction of a glycan site can have significant consequences in terms of epitope presentation and immunogenicity. The processing state of glycans in some instances is crucial for bnAb epitope formation [151, 265]. In other instances, the presence of glycans independent of their processing state is enough. In accordance with this view, both deletion of glycan sites surrounding the 2G12 epitope and addition of glycan sites proximal to this network increased the glycan processing of the epitope, however the only mutations that significantly reduced 2G12 binding were the deletion mutants [257]. As a result, the network of glycans surrounding this epitope likely plays a role in maintaining the epitope by providing support for the glycans that bind 2G12 and are likely not restricted to a particular glycan type.

These glycan network effects seemingly apply to antibodies targeting other parts of the HIV-1 Env. Previous studies on apex-targeting bnAbs PG9 and PG16 report a preference

of these antibodies for sialylated glycans at N156 [266, 236, 237]. Conversely, analysis of N156 containing glycopeptides in this section and in literature have shown glycans at this site to be predominantly oligomannose-type [69, 123, 155, 261]. This plasticity that bnAbs display in their tolerance of different glycans demonstrates that the loss of binding upon the deletion of a site is symptomatic of their roles in stabilising glycan epitopes as well as partaking directly in them. It is as of yet unknown if a successful HIV-1 immunogen must display particular glycan epitopes in a vaccine setting. Further understanding of the roles glycans have in the shaping of bnAb epitopes and the factors that govern the glycosylation of this protein will aid the development of env immunogens.

4 Approaches to engineer antibody epitopes

The ability to add or remove glycan PNGSs on Env trimers offers an editing method which has the potential to fine tune the immunogenicity of Env constructs by either exposing or masking certain epitopes. Simply adding or removing glycan sites however may not be enough to drive bnAb binding, and thus other engineering techniques may be necessary in order to create Env constructs with more immunogenic glycan epitopes. This chapter explores the use of enzymatic co-expression as well as targeted changes to the primary sequence in the form of additions and removals of PNGSs and other residues as a means of improving the immunogenicity of a membrane-bound, VLP-expressed trimer to certain antibody epitopes.

4.1 Contributions

Chapter 4 contains a comparison of several VLP or gp120-expressed JR-FL Env trimers. External collaborators from the laboratory of Dr. James Binley (San Diego Biomedical Research institute, USA) supplied the following samples for analysis:

	VLPs analysed
1	JR-FL SOS E168K + N189A Parent (x2; 2020/2021)
2	JR-FL SOS E168K + N189A + S158T
3	JR-FL SOS E168K + N189A + S364T
4	JR-FL SOS E168K + N189A + D197N
5	JR-FL SOS E168K + N189A + D197N + S199T
6	JR-FL SOS E168K + N189A + T49N
7	JR-FL SOS E168K + N189A + N611Q
8	JR-FL SOS E168K + N189A + T49N + N611Q
9	JR-FL SOS E168K + N189A + N138A + N141A
10	JR-FL SOS E168K + N189A + N138A + N141A - CH01 purified
11	JR-FL SOS E168K + N189A + B4GalT1 + ST6Gal1
12	JR-FL SOS E168K + N189A NGAF3-treated
	gp120s analysed
13	JR-FL gp120 WT
14	JR-FL gp120 + FUT6
15	JR-FL gp120 + FUT9
16	JR-FL gp120 + B4GalNT3
17	JR-FL gp120 + B4GalT1

Table 4.1: **Table of JR-FL VLPs and gp120s analysed**

4.2 Engineering well-expressed, V2-immunofocusing HIV-1 envelope glycoprotein membrane-bound trimers for use in heterologous prime-boost vaccine regimens

HIV-1 vaccine immunofocusing strategies have the potential to induce broadly reactive nAbs. Here, we engineered a panel of diverse, membrane-resident native HIV-1 trimers vulnerable to two broad targets of neutralizing antibodies (nAbs), the V2 apex and fusion peptide (FP). Selection criteria included i) high expression and ii) infectious function, so that trimer neutralization sensitivity can be profiled in pseudovirus assays. The expression of these trimers was improved firstly by truncating gp41 and introducing a gp120-gp41 SOS disulfide to prevent gp120 shedding. "Repairs" were made to fill glycan holes and other strain-specific aberrations. To try to increase V2-sensitivity, clashing glycan sites were removed and the V2 loop's C-strand was modified. Glycopeptide analysis of JR-FL trimers revealed near

complete sequon occupation and that filling the N197 glycan hole was well-tolerated. In contrast, sequon optimisation and inserting/removing other glycans in some cases had local and global "ripple" effects on glycan maturation and sequon occupation in the gp120 outer domain and gp41. V2 mAb CH01 selectively bound trimers with small oligomannose glycans near the base of the V1 loop, thereby avoiding clashes. Knocking in a N49 glycan perturbs gp41 glycans via a distal glycan network effect, increasing FP nAb sensitivity - and sometimes improving expression. Overall, we identified 7 diverse trimers with a range of sensitivities to two targets that should enable rigorous testing of immunofocusing vaccine concepts.

4.2.1 Introduction

Current HIV-1 vaccine candidates are able to induce high titres of autologous tier-2 nAbs [242, 207, 267, 268, 222]. Cross-neutralization of divergent HIV-1 strains however, remains unpredictable [204, 269]. This is partly due to the idea that nAbs in general target strain-specific gaps of exposed Env protein surface, known as 'glycan holes' [214, 242, 245, 222]. In contrast, bnAb targets are frequently protected by adjacent glycans on the HIV-1 Env that nascent nAbs must either evolve to avoid or additionally bind [118, 225, 254].

Over the last decade many bnAbs have been isolated from HIV-1-infected donors, the majority of which target 5 conserved epitope clusters on Env: the V2 apex, V3 glycan, CD4bs, gp120-gp41 fusion peptide (FP) and the gp41 MPER [57] Recent animal vaccine studies have shown that nAbs with some level of breadth can be induced [270, 271, 272], however further efforts are needed to improve vaccine-induced nAb breadth, titre and consistency.

There are currently three main ideas regarding improving the quality of vaccine-induced nAbs [273, 274, 275, 276]. The first is to use vaccines to trigger the expansion of bnAb precursor cells (unmutated common ancestors, UCAs) [277, 278, 279, 280, 281]. Vaccine-mediated triggering [115, 282, 250, 253, 252, 283, 284] may be improved through the removal of clashing glycans [243, 267, 270, 223, 285], reducing glycan size [266, 225, 286, 287, 288], or by priming with core epitopes or scaffolds [269, 271, 272, 279, 284, 289, 290, 291, 292,

293]. Ideally, priming with a vaccine creates an initial diverse pool of antibodies that can be 'filtered' by using carefully selected boost vaccines that would promote bnAb development [269, 285, 293]. A second approach would involve replicating bnAb development, using patient-derived Env clones in vaccines to guide bnAb lineages to breadth [277, 253, 254, 294, 295, 296, 297, 298, 299, 300, 301]. A third strategy would be to 'immunofocus' nAbs, using trimers with high sensitivities to desired areas [243, 244, 267, 269, 270, 223, 285, 289].

Repeated immunisations with the same Env trimer may cause nAbs to overly focus on distinct features of the vaccine strain, e.g. glycan holes. Trimer variants may be needed to drive nAb cross-reactivity. Polyvalent mixtures of vaccines have been attempted to limited success [244], perhaps because the most sensitive trimer in the mixture dominates Ab responses. Furthermore, whilst polyvalent mixtures are intended to promote cross-reactivity they likely do not provide a direction for the development of the nAb.

A second format is homologous prime-boosts. It would perhaps begin with modified trimers of the same strain. This has been attempted with some success, typically by priming with trimers in which glycans surrounding the target are removed and boosting with trimers in which these glycans are reinstated [243, 253, 302, 267, 270, 285, 303].

A third format, serial heterologous prime-boosts (SHPB), uses trimers from different strains in each shot with the intention of eliminating strain-specific autologous nAbs [207, 244, 304, 305, 306, 307]. Success may hinge on whether nascent nAb-expressing memory B cells sufficiently cross-react with boosting trimers to keep them 'on track'.

Here, we sought to assemble a panel of diverse trimers expressed on VLPs to simultaneously immunofocus both the V2 apex and FP epitopes. These two epitopes were chosen for several reasons. First, since the two sites do not overlap, the probability of inducing bnAbs is doubled. Second, both sites can accommodate multiple nAb binding modes [244, 269, 290]. This 'plasticity' may increase the frequency of compatible germline Ab precursors. FP nAbs have common in-repertoire features, and can be induced in many species, including mice [269, 271, 308]. Some V2 nAbs (e.g. CH01, VRC38.01) also exhibit sufficiently common features,

and thus do not depend heavily on rare V(D)J recombination and/or hypermutation-driven events occurring during their formation, thus making them plausible vaccine prototypes [280, 281, 308].

V2 and FP nAbs both bind epitopes comprising of protein and glycan contacts. V2 bnAbs bind the N160 glycan and the neighbouring basic C strand of a 5-strand β-barrel at the trimer apex [236, 280, 309]. However, the binding may be regulated by protecting V1/V2 glycans and long V1/V2 loops [266, 153, 244, 254, 310, 223, 303]. FP bnAbs, exemplified by VRC34, recognise the N-terminal fusion peptide, along with the proximal N88 glycan of gp120, but clash with gp41's N611 glycan [269, 290, 291, 311, 312]. The preferential binding of CH01 and VRC34 to trimers that are produced in N-acetylglucosaminyltransferase 1 (GnT1-) deficient cells, in which glycan maturation is blocked, suggests that both nAbs contact the stems of their respective glycans at positions N160 and N188, respectively.

4.2.2 Results

Here, glycopeptide analysis was conducted on a series of trimers to immunofocus both V2 and FP epitopes. The initial identification stage originally sought out V2-sensitive trimers from which the JR-FL strain was chosen due to its good expression and because E168K + N189A mutations were shown to introduce V2 sensitivity into the JR-FL strain [225, 235], by increasing C strand charge and eliminating an overlapping sequon. Demonstrating the feasibility of knocking V2 sensitivity into group 2 strains. Afterwards, a series of mutations was carried out to induce FP-nAb sensitivity into these trimers. The analysed samples are listed in Table 4.1.

4.2.2.1 JR-FL modifications

We first modified our prototype vaccine strain, JR-FL, hoping to improve V2 and FP sensitivity. The E168K + N189A mutant is sensitive to VRC38, PG9 and PGT145 and is partially sensitive to CH01. V2 sensitivity might be improved by removing local clashing

glycans and/or by increasing strand C's basic charge. These modifications were made in the JR-FL SOS gp160ΔCT background.

We next investigated the effects of removing the overlapping V2 sequons N188 and N189 alone or together. Unlike the E168K, E168K + N188A and E168K + N188A + N189A mutants, E168K + N189A was modestly CH01-sensitive. We therefore used JR-FL E168K + N189A as a "parent" clone to overlay further mutations.

V2 nAb sensitivity typically depends on glycans N156 and N160 and basic (K/R) residues in strand C [244, 303, 309, 236, 235] . K168 and K171 are V2 nAb "anchor" residues [280, 309]; K/R residues at 166, 169 and 170 also contribute to strand C's positive charge . Our JR-FL E168K + N189A mutant contains these anchor residues and a net charge of +2, which is somewhat lower than many V2-sensitive strains that exhibit K/R at position 169 of strand C.

The JR-FL S158T and S364T modifications were induced separately to see if NxT mutations improved glycan occupancy at the N156 and N362 sites. the D197N +/- S199T modifications sought to successfully knock-in the N197 glycan however due to its partial occupancy across different strains the latter S199T mutation sought to use the NxT sequon to achieve higher occupancy and V2-nAb sensitivity [242]. The T49N/N611Q modifications sought to increase FP sensitivity to the VRC34 antibody. Modeling places the N49 glycan in between the N276 and N637 glycan sites (Figure 4.1A). Its proximity to N637 led us to wonder if it might impact the local glycan network that regulates sensitivity to FP nAbs like VRC34 that contact N88 glycan and clash with N611 glycans. Single and double mutant variants were analysed to ascertain the impact on glycosylation of these three sites as the reduction in N88 glycan maturation may be associated with improved VRC34 sensitivity. The N138A + N141A mutations were associated with better V2 sensitivity and this trimer was analysed on its own and after an additional CH01 complexing to observe differences between the endogenously produced VLP and those bound by the antibody.

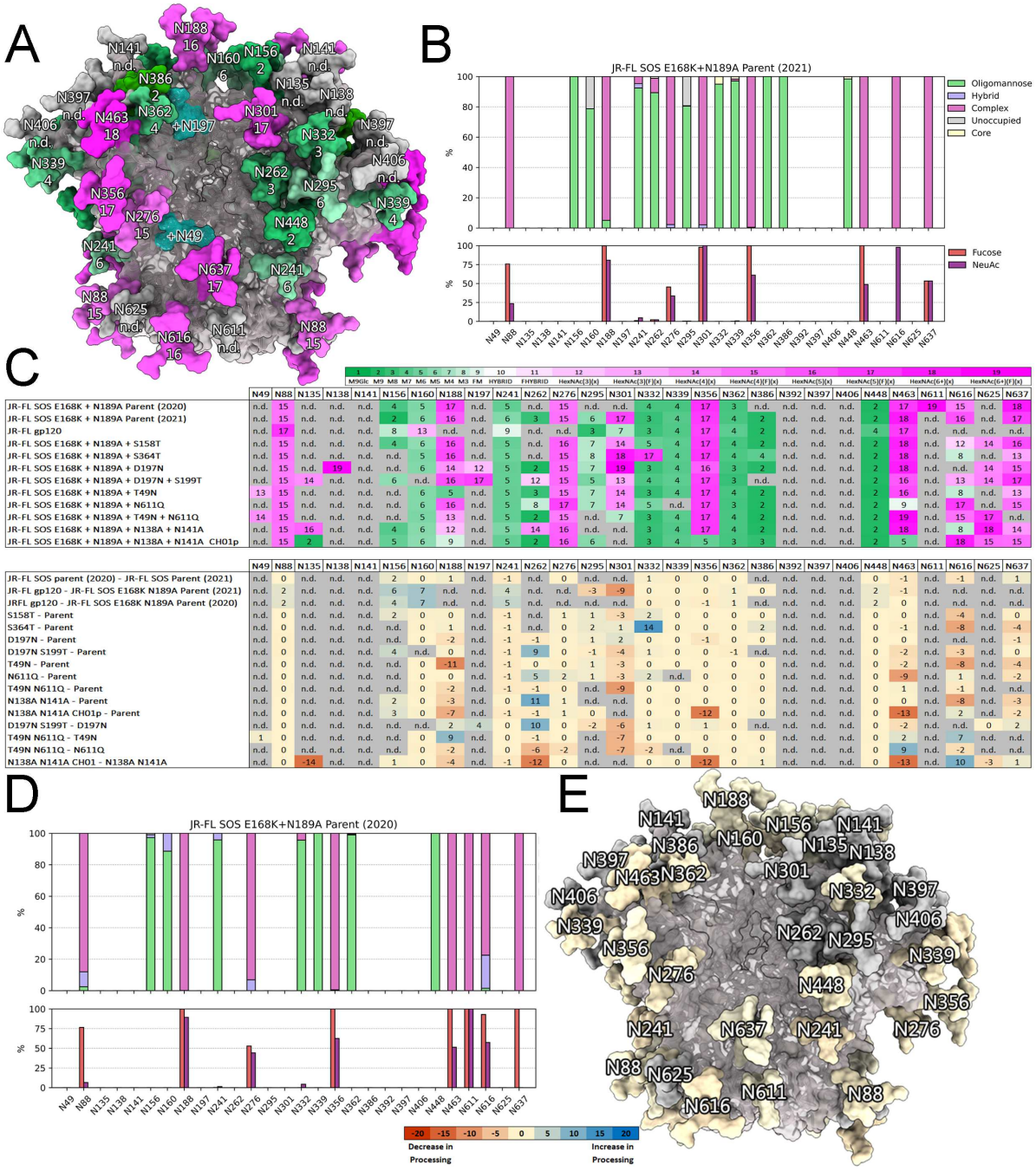


Figure 4.1: **Effects of mutants on JR-FL membrane trimer glycan maturation and occupation.**
(A) JR-FL trimer model (PDB ID: 6MYY) Glycans are colored according to a score that corresponds to the average maturation state (Key below **(B)**). Untrimmed oligomannose glycans are dark green. Glycans in grey were not resolved (n.d.). Inserted glycans at positions N49 and N197 are modeled as blue translucent masses. **(B)** Site-specific glycosylation of the 2021 JR-FL Parent trimer. The upper graph shows the site-specific proportion of oligomannose, hybrid, complex, core glycans and under-occupied glycopeptides. The lower shows site-specific levels of fucose and sialic acid (NeuAc). Data is from one biological replicate. **(C)** *(Upper)* site-specific glycan score for each analysed sample. Numbers correspond to the average maturation state for each site. *(Lower)* Changes in glycan scores at each position between sample pairs. A negative score implies a shift to less mature glycan and vice versa. Data are only shown at positions where a glycan was detected in more than 10% of the equivalent peptides of both samples in each pair. **(D)** Site-specific glycosylation of the 2020 JR-FL Parent trimer. Data is from one biological replicate. **(E)** Change in glycan score of the 2020 JR-FL Parent versus the 2021 Parent trimer as per **(C)** mapped onto the JR-FL trimer.

4.2.2.2 Effects of mutations on glycan maturation and occupation

To understand the basis of the effects of these mutants on the JR-FL SOS E168K + N189A trimer, we assessed glycan occupation and maturation of these mutants by glycopeptide LC-MS. Each glycan type was assigned a 'glycan score' ranging from 1 to 19, depending on the average maturation state. The least processed oligomannose glycan in our quantified series, M9Glc, is given a score of 1, while the most highly branched and fucosylated complex glycan HexNAc(6+)(F)(x) has a score of 19. Glycan scores of the most recent parental E168K + N189A trimer is modeled in Fig 4.1A and the score change between it and the older parent trimer is modeled in Fig 4.1E. The scores and diversity at each site of all mutants are summarized in Fig 4.1C (upper). The nature of glycans at each site generally match a previous report that categorized JR-FL pseudovirus (PV) Env glycans by another method [261], although the N160 and N386 glycans were mostly high mannose in our hands but were complex by the other method.

We next evaluated glycan score differences at each site in pairs of samples. Score changes were recorded in a table (Fig 4.1C) for sites that were over 10% occupied by glycans (excluding core glycans, i.e., truncated glycan structures smaller than M3) in both samples. Score differences for each pair are modeled below. Glycan under-occupancy and core glycans are shown in the glycosylation charts such.

We first compared two preparations of JR-FL SOS E168K + N189A VLP trimers ('parent') analyzed at different times to gauge sample and assay variation (Fig 4.1B vs D). Minor differences in glycan maturation were observed, e.g., at position N156. under-occupancy was rare and varied between samples, occurring at positions N156 (0.87%) and N362 (0.40%) in one sample and at N160 (21.24%) and N295 (19.24%) in the other. Glycan core was found occasionally (around 5% or less) at 4 sites in one sample, but not at all in the others. Several sequons could not be assigned a glycan (e.g., N135 and N138), as their proximity made it difficult to isolate glycopeptides with only one glycan.

A comparison of 'parent' trimers and monomeric JR-FL gp120 revealed that glycan types

(i.e., high mannose or complex) were similar at many positions (Fig 4.3C, D). However, gp120 monomer glycans were more differentiated at positions N88, N156, N160 and N241, perhaps reflecting the greater access to glycan processing enzymes (Fig 4.3B). Similar observations were made in a comparison of BG505 SOSIP and gp120 [123]. Conversely, glycans N295 and N301 were less mature, which, as for SOSIP, may be because the relatively fast Env production rate reduces contact time with glycan processing enzymes (Fig 4.3A).

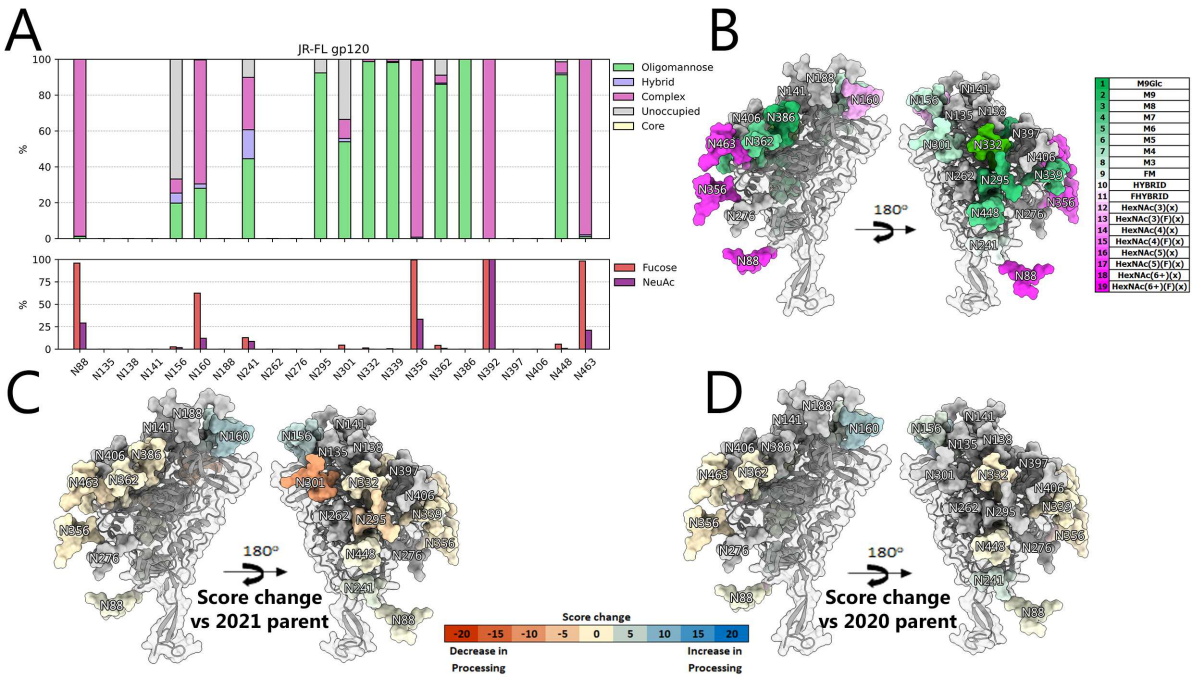


Figure 4.2: **JR-FL gp120 glycosylation.**
(A) Site-specific glycosylation of JR-FL gp120. The upper graph shows the site-specific proportion of oligomannose, hybrid, complex, core glycans and under-occupied glycopeptides. The lower shows site-specific levels of fucose and sialic acid (NeuAc). Data is from one biological replicate. **(B)** JR-FL gp120 model (PDB ID: 6MYY) Glycans are colored according to a score that corresponds to the average maturation state (Key alongside **(B)**). Untrimmed oligomannose glycans are dark green. Glycans in grey were not resolved (n.d.). **(C-D)** site-specific glycan score change for JR-FL gp120 compared to both parent trimers. Glycans are coloured according to changes in glycan scores at each position between sample pairs. A negative score implies a shift to less mature glycan and vice versa. Data are only shown at positions where a glycan was detected in more than 10% of the equivalent peptides of both samples in each pair.

Glycan under-occupancy was common in gp120: 8 out of 14 sequons were partially unoccupied(Fig 4.2A). Of these, N156 was most frequently skipped (66.7%), followed by N301 (33.5%) and N362 (8.9%). As noted above, N156 and N362 were also occasionally skipped in membrane-bound trimers, albeit to a far lesser extent. Both sequons have a serine at the 3rd position, i.e., NxS.

Since NxT is a better substrate for glycan transfer [313, 314], we made S158T and S364T mutants. Glycopeptide LC-MS revealed that S158T mutant glycosylation was broadly similar to the parent, with score changes less than 5 (Fig 4.1C). Differences in both directions were observed, e.g., at N156 and N301. Modest under-occupancy and core glycans at N160 could be a direct consequence of the adjacent S158T mutation, as in BG505 SOSIP [313].

Skipping also occurred at N339 (Fig 4.3A). S364T dramatically increased N332 glycan differentiation and caused skipping at N339 and N637 (Fig 4.3D). Some of the above effects were distal from the two mutation sites, suggesting allosteric 'glycan network' effects (Fig 4.3F). In both cases, gp41 sites N616 and N637 glycans were less mature (Fig 4.3F). Collectively, our findings suggest that these two mutants disturbed conserved serine residues, resulting in modest structural changes.

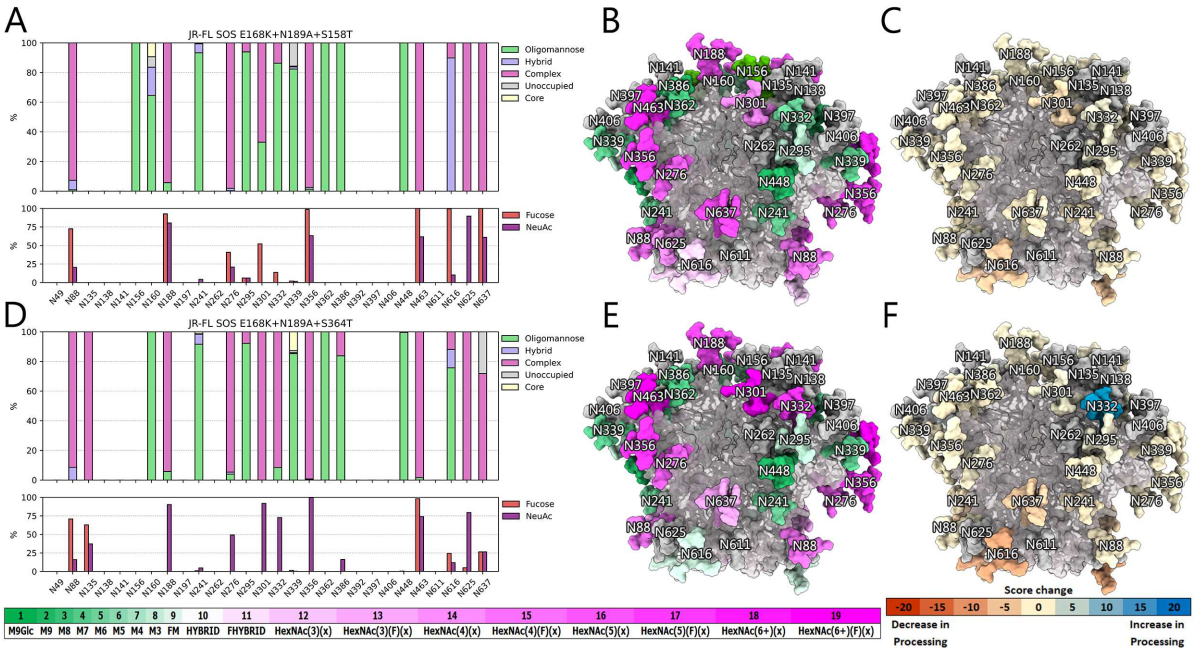


Figure 4.3: **JR-FL SOS S158T/S364T glycosylation.**
(A, D) Site-specific glycosylation of JR-FL SOS E168K+N198A S158T/S364T mutants. The upper graph shows the site-specific proportion of oligomannose, hybrid, complex, core glycans and under-occupied glycopeptides. The lower shows site-specific levels of fucose and sialic acid (NeuAc). Data is from one biological replicate. **(B, E)** JR-FL SOS model (PDB ID: 6MYY) Glycans are colored according to a score that corresponds to the average maturation state (Key alongside **(B)**). Untrimmed oligomannose glycans are dark green. Glycans in grey were not resolved (n.d.). **(C, F)** Site-specific glycan score change for JR-FL SOS mutants compared to the latest parent trimer. Glycans are coloured according to changes in glycan scores at each position between sample pairs. A negative score implies a shift to less mature glycan and vice versa. Data are only shown at positions where a glycan was detected in more than 10% of the equivalent peptides of both samples in each pair.

The D197N mutant completely knocked in the N197 glycan (Fig 4.1C, 4.4A). N301 maturation was modestly increased, perhaps due to its proximity (Fig 4.4C). N160 and N637

sequons were both fully occupied, unlike the S158T and S364T mutants, respectively. However, like S158T and S364T, some skipping occurred at N339 (Fig 4.4A). Overall, D197N was well-tolerated compared to the S158T and S364T (Fig 4.4E).

A sequon-optimised D197N + S199T mutant was inferior. It only filled the N197 site to around 90% efficiency and caused glycan holes elsewhere, most notably at N463 (∼91% skipped) and N295 (∼36% skipped) (Fig 4.4B). N262 partly toggled to complex, whereas N301 became immature. Overall, D197N + S199T was poorly tolerated, like S158T and S364T, affecting distal glycans in a global "ripple" effect, further cautioning against mutations at conserved positions.

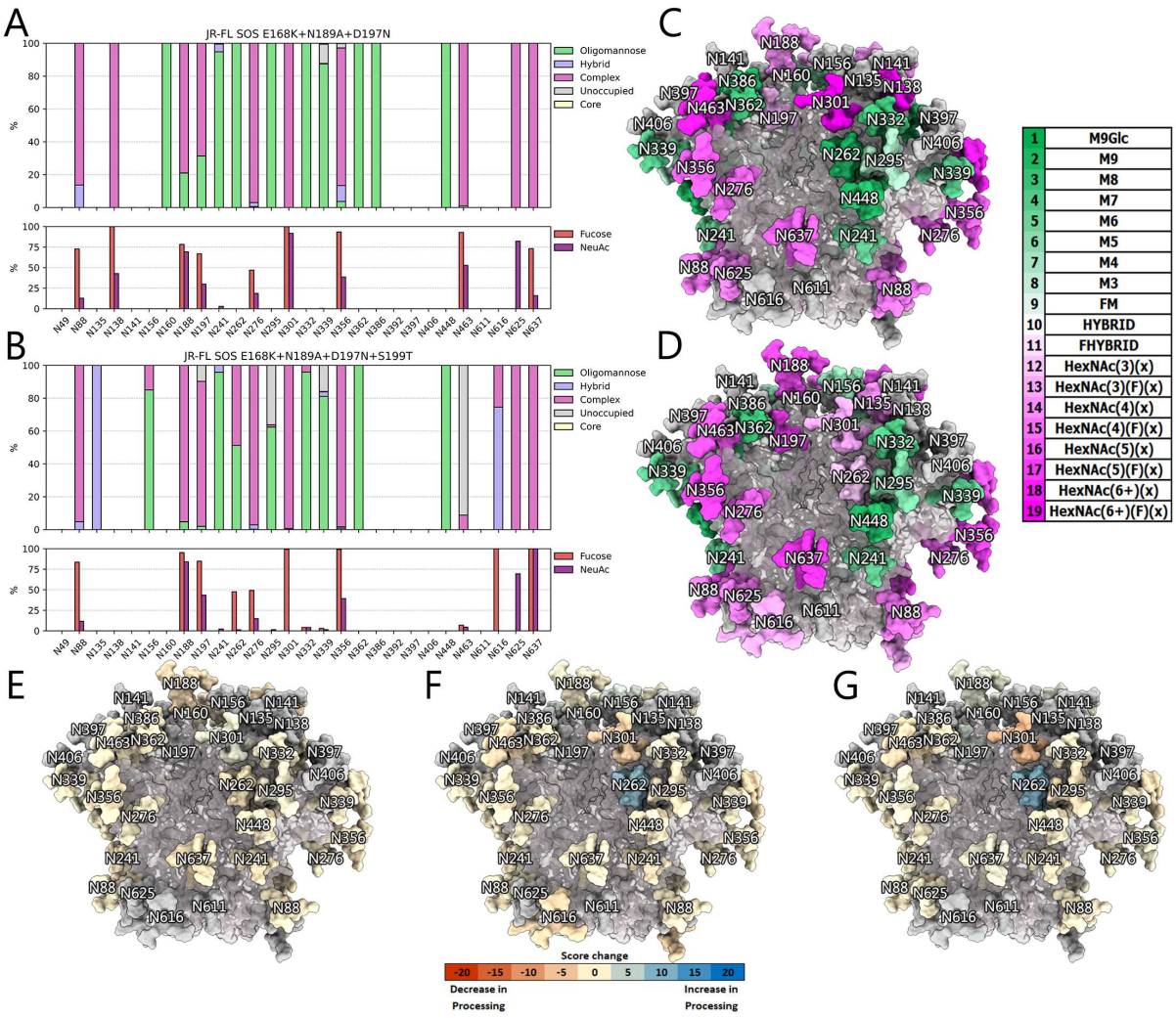


Figure 4.4: **JR-FL SOS D197N +/- S199T glycosylation.**
(A-B) Site-specific glycosylation of JR-FL SOS E168K+N189A D197N +/- S199T. The upper graph shows the site-specific proportion of oligomannose, hybrid, complex, core glycans and under-occupied glycopeptides. The lower shows site-specific levels of fucose and sialic acid (NeuAc). Data is from one biological replicate. **(C-D)** JR-FL SOS model (PDB ID: 6MYY) Glycans are colored according to a score that corresponds to the average maturation state (Key alongside **(B)**). Untrimmed oligomannose glycans are dark green. Glycans in grey were not resolved (n.d.). **(E-G)** site-specific glycan score change for JR-FL SOS compared to the most recent parent trimer (E-F; E: D197N vs Parent, F: D197N+S199T vs Parent). **(G)** illustrates D197N+S199T vs D197N. Glycans are coloured according to changes in glycan scores at each position between sample pairs. A negative score implies a shift to less mature glycan and vice versa. Data are only shown at positions where a glycan was detected in more than 10% of the equivalent peptides of both samples in each pair.

T49N successfully added a complex glycan that led to decreased maturation at positions N188, N301, N616, and N637 (Fig 4.5A), presumably due to overcrowding. This contrasted sharply with the mild effects of D197N. We could not obtain glycopeptide data for N611 that might have given insights into how the N49 glycan improves VRC34 sensitivity. While N611 is not close to the N49 (Fig 4.5D), it is possible that smaller glycans at the other gp41 sites provide space for the N611 glycan to move aside for VRC34 binding. Our model suggests that some effects are localized whilst others (e.g., N188) are distal, suggesting a global conformational change, consistent with partial V3 MAb sensitivity.

N611Q led to increased N262 glycan maturation, decreased N463 glycan maturation and skipping at N188 and N339 (Fig 4.5B). T49N + N611Q led to reduced N301 maturation and skipping at N339 and N463 (Fig 4.5B, E). Reduced N301 maturation of the single T49N and N611Q mutants appeared to be amplified in the double mutant. However, other effects in the single mutants were absent in the double mutant, suggesting that N49 knock in partially compensates for N611 glycan loss (Fig 4.5C). Comparing the double mutant to its component single mutants again highlighted differences at N188, N262, N301, N463 and N616 (Fig 4.5C, J, K), although the patterns did not resemble those above, implying unpredictable and subtle effects on trimer folding.

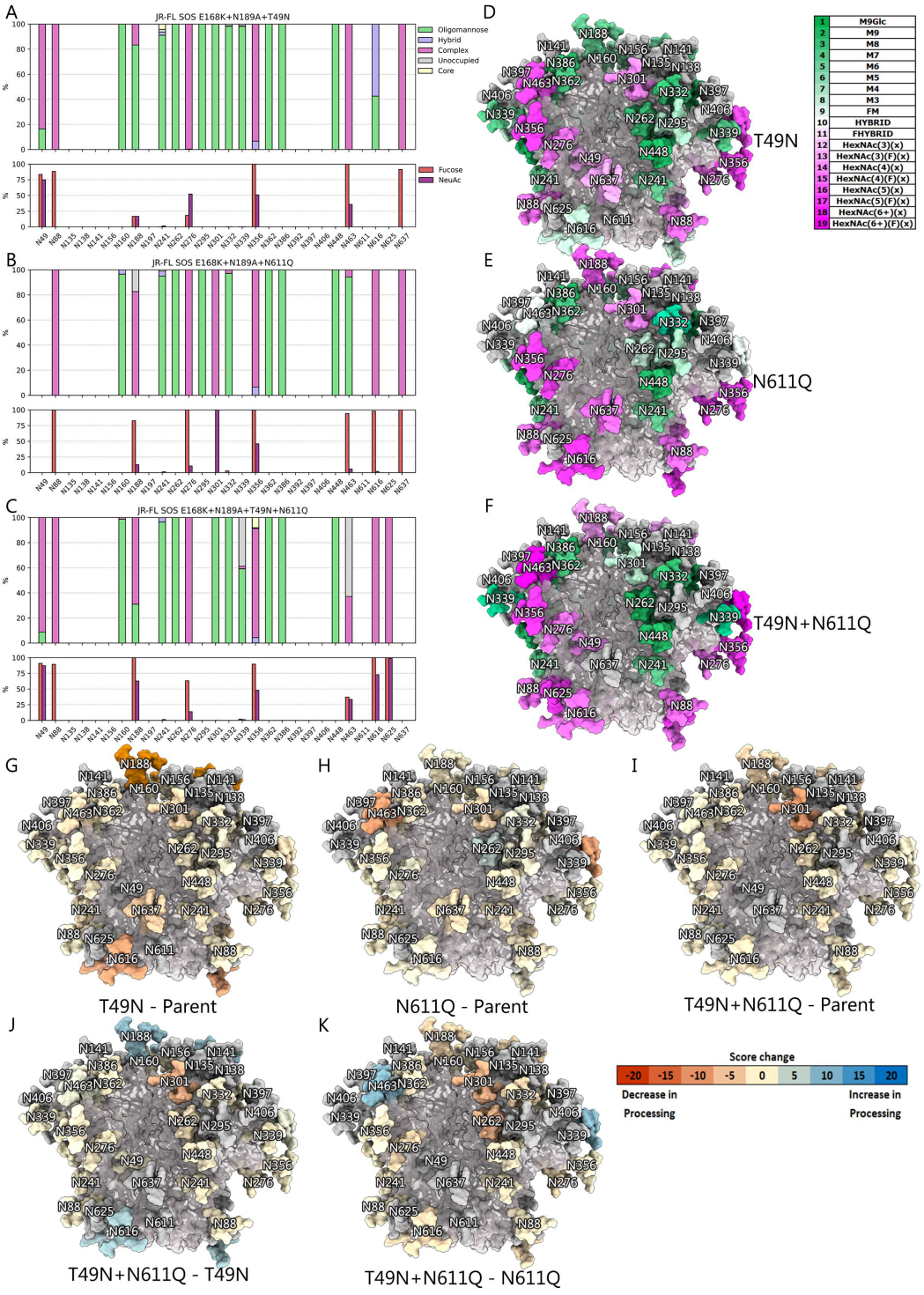


Figure 4.5: **JR-FL SOS T49N/N611Q glycosylation.**
(A-C) Site-specific glycosylation of JR-FL SOS E168K+N189A T49N/N611Q/T49N+N611Q. The upper graph shows the site-specific proportion of oligomannose, hybrid, complex, core glycans and under-occupied glycopeptides. The lower shows site-specific levels of fucose and sialic acid (NeuAc). Data is from one biological replicate. **(D-F)** JR-FL SOS model (PDB ID: 6MYY) Glycans are colored according to a score that corresponds to the average maturation state. Untrimmed oligomannose glycans are dark green. Glycans in grey were not resolved (n.d.). **(G-K)** site-specific glycan score change for JR-FL SOS compared to the latest parent trimer. Glycans are coloured according to changes in glycan scores at each position between sample pairs. A negative score implies a shift to less mature glycan and vice versa. Data are only shown at positions where a glycan was detected in more than 10% of the equivalent peptides of both samples in each pair.

Analysis of N138A + N141A revealed that the N135 glycan is complex in the absence of these neighboring glycans (Fig 4.6A, C). Significant skipping at positions N262 and N295 was observed, along with some at N339 (Fig 4.6A). The small amount of glycan detected at position N262 was far more mature than on the parent (Fig 4.6A).Since this glycan is structurally important, its absence could cause some misfolding [67]. Glycan maturation differences were also observed at N156, N188, N616 and N637.

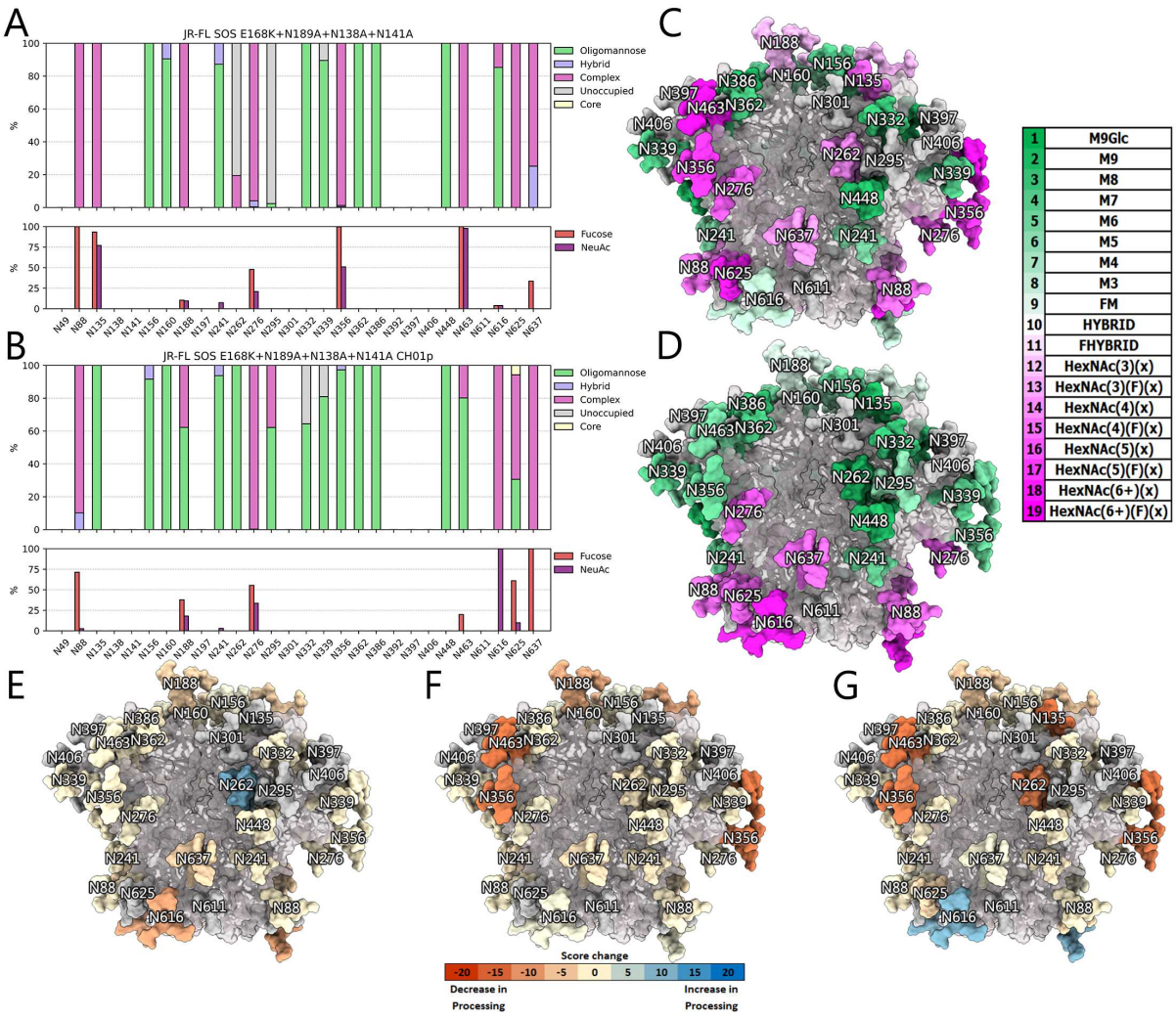


Figure 4.6: **JR-FL SOS N138A+N141A +/- CH01 glycosylation and parent strain comparison.**
(A-B) Site-specific glycosylation of JR-FL SOS E168K+N189A N138A+N141A and CH01p VLP trimers. The upper graph shows the site-specific proportion of oligomannose, hybrid, complex, core glycans and under-occupied glycopeptides. The lower shows site-specific levels of fucose and sialic acid (NeuAc). Data is from one biological replicate. **(C-D)** JR-FL SOS model (PDB ID: 6MYY) Glycans are colored according to a score that corresponds to the average maturation state. Untrimmed oligomannose glycans are dark green. Glycans in grey were not resolved (n.d.). **(E-G)** site-specific glycan score change for JR-FL SOS N138A+N141A +/- CH01 compared to the most recent parent trimer **(E-F)** and between CH01p and non-CH01p VLPs **(G)**. Glycans are coloured according to changes in glycan scores at each position between sample pairs. A negative score implies a shift to less mature glycan and vice versa. Data are only shown at positions where a glycan was detected in more than 10% of the equivalent peptides of both samples in each pair.

N138A + N141A membrane-bound trimers complexed with CH01 exhibited radical differences at some positions: a shift to high mannose at N135 and N188 may help to minimize clashes at the binding site (Fig 4.6B, D) [225]. However, N356 and N463 glycans were also less mature, despite being distal from the CH01 epitope. Intriguingly, N262 and N295 glycans were efficiently detected in CH01-bound sample, even though both were largely skipped in the unbound sample. Conversely, N332 exhibited more skipping and N616 was more complex in the CH01-bound sample. Thus, it appears that rare glycan species were found in CH01-trimer complexes that were not detected in uncomplexed trimers. Since CH01 neutralizes the N138A + N141A mutant to a maximum of only around 75%, it appears to bind only to trimers where N135 and N188 glycan clashes are minimal. This trimer population carries other unusual glycans at distal sites that may further improve CH01 binding or are inextricably linked with high mannose glycans at the epitope.

Overall, this data reveals that outer domain glycans (N156-N339) are prone to maturation changes, while inner domain glycans N88, N356-N448 are largely static. Under-occupancy is also more common at some outer domain glycan sites, particularly N339. Importantly, our findings provide some guidance regarding how well our engineering efforts are tolerated.

4.2.3 Discussion

Here, we demonstrate engineering efforts to generate V2- and FP-sensitive trimers, with view to the potential use of the modified JR-FL trimers as boosting immunogens to help induce bnAbs. Collaborator efforts showed that gp41 tail truncation and SOS mutation consistently improve expression [315]. The optimisation process can be accelerated by initially correcting all *"errors of nature"*, by filling glycan holes, and resolving insertions, deletions and overlapping glycan sequons. Although these and other repair strategies occasionally improved expression or infectivity, we did not discover a rigid strategy for engineering desirable membrane-bound trimers. Depending on the Env strain, membrane-bound trimers can exist in a variety of conformational states, which has implications for the epitopes that a candidate

would display, and therefore the types of antibodies that it could induce. This suggests that trimers of certain strains, depending on their state, may be more disposed to these types of engineering strategies. the JR-FL strain in this instance has proved to be a good choice for use as a membrane-bound candidate as the trimer retains its robust V3-resistance in face of multiple engineering strategies [315].

Most repair strategies that are helpful for SOSIPs [279, 266, 316] proved to be unhelpful for membrane-bound trimers. This difference is underlined by the distinct limiting factors of membrane-bound trimers (e.g. expression) versus SOSIPs, which is reflected in the fact that the different strains make better prototypes between membrane-bound and soluble trimers e.g., JR-FL and BG505, respectively. The fact that codon optimisation, modified signal peptides [317], different expression plasmids and vectors were unhelpful in elevating expression of membrane-bound material could suggest that the expression bottleneck is not at the level of transcription, but is instead dictated by protein folding. It may be that optimising the aforementioned factors can have positive effects in increasing the overall quantity of produced trimers, but in doing so produces misfolded trimers whose conformational states are inappropriate for their purpose. Strategies to avoid producing misfolded trimers such as identifying and replacing residues in the interprotomer interface between the gp120/41 monomers or by covalently or chemically linking the protomers together could yield unwanted knock-on effects on trimer glycosylation, and may not be reproducible across Env strains [109, 318, 319].

In our analysis, glycan under-occupancy of these membrane-bound constructs was found to be more limited when compared to soluble Env glycan analysis [104, 261, 155, 313, 320]. Additionally, NxT sequon optimisation efforts yielded ambiguous results. For the N156 glycan site, the S158T mutation was well tolerated and the site remained 100% oligomannose-type however this was already observed in one of the parent trimers (2021) and the parent trimer analysed in 2020 was also above 90% oligomannose-type also (Fig 4.1B, D). This was also the case for the S364T mutant. The D197N +/- S199T mutants showed that this strategy does

not always work however, as the D197N mutation was well tolerated and the glycan site was successfully inserted into the JR-FL SOS backbone with 100% occupancy whereas the double D197N + S199T mutant reported a much lower subset of oligomannose-type glycosylation as well as a small proportion of glycan under-occupancy (Fig 4.4A, B). Again, this may be due to the flexibility of membrane-bound trimers, versus rigid SOSIP, such that mutations can have more dramatic consequences for membrane-bound trimers [254, 155, 255, 122, 257]. In our case it appears that membrane trimer mutants are only successful when they introduce a more conserved alternative residue at a given position. The varied effects of glycan toggling on proximal and distal glycans provides reason for caution; for example, a glycan hole that increases V2 apex sensitivity may only be effective if it does not open up other unwanted glycan holes. On the other hand, if "off-target" holes differ between successive vaccine shots, this could limit the problem [213].

Our collaborators found that V1V2 base glycans had the biggest impact on V2 nAb sensitivity. The greater effect of N135 on JR-FL compared to N138 or N141 is consistent with its closer physical proximity to the V1V2 apex, which is why the former two glycan sequons were swapped out for alanine residues. Furthermore, the N130 glycan is absent in V2-sensitive group 1 strains and its removal from group 2 strains also improved their sensitivities. However, some V1V2 glycans can affect expression. For JR-FL, knocking in the N197 glycan did not appreciably impact either expression or V2 sensitivity, although N197 toggling can impact the V2 sensitivity of other strains [321].

Given the increased activity of CH01 on GnT1- PVs, it was no surprise that CH01 selectively bound to small high mannose glycans at N135 to avoid clashes. This selective glycovariant binding has been reported previously [313]. What we did not expect is for glycans at other sites to also be far less mature. The idea that CH01 bound an early "high mannose" trimer glycoform can be ruled out, as not all sequons were affected. For example, the N616 glycan was more mature in CH01 complexes. It could be that glycan maturation at different sites is co-dependent via a glycan network. If so, how can we reconcile the rarity or total absence

of glycoforms at several sites in the corresponding uncomplexed parent? We assume that all glycoforms are equally infectious, but this may not be the case. Indeed, a significant fraction of JR-FL trimers remains uncleaved, while our glycopeptide analysis provides data on the total Env of VLPs, regardless of processing. Further analysis of MAb-complexed trimers may provide insights.

To conclude, this project tested various immunogen design strategies in membrane-bound trimers presented on VLPs. Whilst a number of these strategies could be imparted onto the protein with minimal impact on glycosylation, others had more widespread and at times unpredictable consequences. This highlights a challenge in immunogen design between soluble and membrane-bound material in that a strategy suitable for the former may prove unsuitable for the latter. The fact that more intrusive strategies such as knock-ins and knock-outs of glycan sequons could be achieved without averse effect however is positive as it provides a potential avenue for creating bespoke immunogens in prime-boost regimens, which may require such changes between regimens to induce desired neutralising antibodies.

4.3 Co-translational glycoengineering of HIV-1 Env

4.3.1 Introduction

The extensive glycosylation of the HIV-1 envelope protein combined with its intrinsic heterogeneity constitutes a considerable barrier to the induction of a bnAb response. As a follow-on to a previous study investigating the use of glycoengineering (GE) tools in HIV-1 env constructs on sensitivity to bnAbs, the effects of GE on Env and gp120 glycosylation were investigated to observe how changes in the fine processing of these constructs relate to previous findings on bnAb affinity [225].

To date, there are a handful of methods that allow for the alteration of glycans on a protein. These include the use of glycosyltransferase inhibitors such as kifunensine, which prevents

the trimming of $Man_9GlcNAc_2$ structures; use of cell lines with knocked-out glycosyltransferases; use of enzymes in-vitro, and co-transfection of glycosyltransferase plasmids with the target glycoprotein. Some of these methods have been previously used in studies to modify the glycans on HIV-1 env, and whilst they have been found to increase sensitivity to some bnAbs, they often come at a cost in the form of reduced affinity to other antibodies [225].

In the previous study by Crooks et al., pseudovirions (PVs) bearing Env constructs that were co-transfected with several glycosyltransferases were screened against a panel of bnAbs and their germline precursors. Binding of bnAbs under the effects of GE tools varied but some, namely V2-apex specific antibodies PG9 and CAP256.09, were 30-fold more potent against PVs co-transfected with β1,4-N-acetylgalactosaminyltransferase-III (B4GalNT3) and β-Galactoside α2,6-sialyltransferase-I (ST6Gal1) [225]. Which is consistent with previous studies that PG9 and CAP256.09 engage with α2,6-sialic acid structures [266, 310]. Also of note was that co-transfection of Env with β1,4-galactosyltransferase-I (B4GalT1) increased sensitivity to endo-H, through the partial replacement of complex with hybrid-type glycans, thereby reducing PV sensitivity to PGT151, whose epitope has a preference for multiantennary glycans [225].

The effect of the enzymes B4GalT1 and ST6Gal1, which were co-transfected prior to post-purification treatment with Endo F3 mixed with a neuraminidase, hexoseaminidase and galacotosidase (referred to as NGAF3) on the glycan shield was studied in comparison to JR-FL VLP-expressed trimers soley co-expressed with B4GalT1 and ST6Gal1. Clade B Gp120 JR-FL proteins were also co-transfected with α1,3-fucosyltransferase 6 and 9 (FUT6, FUT9), B4GalT1, and B4GalNT3. The aim of this study was to probe into the effects of these enzymes on env constructs at a site-specific level to ascertain the efficacy of this chosen panel of glycosyltransferases and glycosidases. It is thought that these GE tools could potentially aid immunogen design in generating preferred epitopes to guide bnAb elicitation.

4.3.2 Results

4.3.2.1 Site-specific analysis of JR-FL SOS viral-like particles (VLPs)

Aliquots of purified JR-FL constructs were reduced and alkylated, followed by subsequent enzymatic digestion with chymotrypsin. The digested glycopeptides were analysed by LC-MS to determine the N-glycan compositions at a site-specific level.

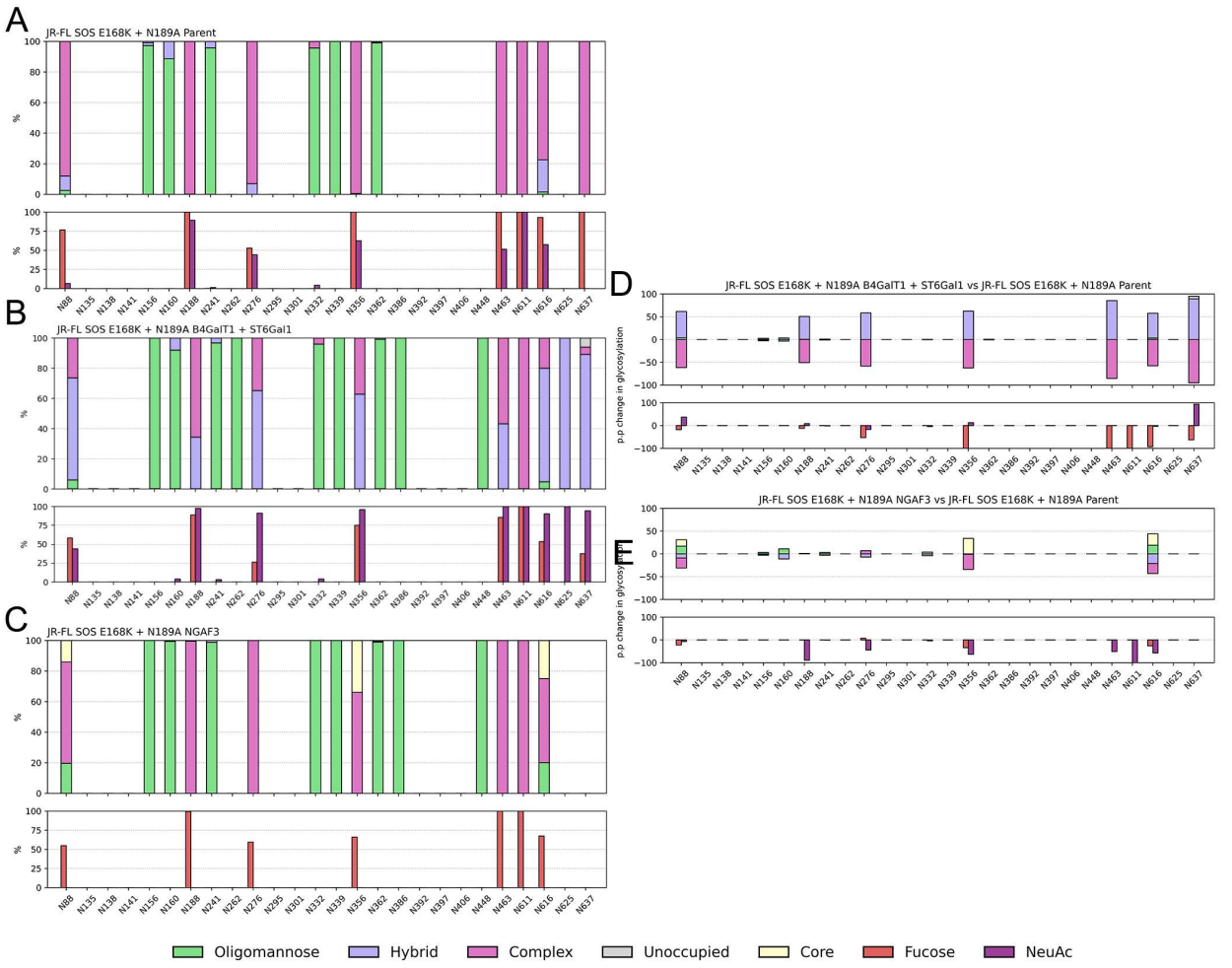


Figure 4.7: **Site-specific analysis of JR-FL SOS E168K + N189A VLP-derived gp140s**

(A-C) Site specific analysis of JR-FL SOS gp140 trimers, as determined by LC-MS. The intensity of different glycoforms at a single N-linked glycan site were compared and glycopeptides were categorised according to the number of mannose residues (green), hybrids that possess one unprocessed arm and one processed and also the number of processed branches and the presence/absence of fucose for complex-type glycans (magenta). Each dataset is from one biological replicate.

(D-E) Arithmetic percentage point (p.p.) difference between JR-FL SOS trimer co-expressed with GE enzymes and control (Parent). The proportion of un-occupied asparagines at each site, or sites in occupied with only one N-acetylglucosamine residue are coloured grey. PNGs numbered according to HXB2 alignment.

When the B4GalT1 and ST6Gal1 treated construct is compared to the parent strain, we can observe increased hybrid-type structures with concomitant decrease of antennary structures; in particular tri- and tetra-antennary structures (Figure 4.7, A-B). These changes are most evident at sites N88, N187, N356, N463, N616, and N637. When differences in hybrid-type glycans identified were compared between the control and the B4GalT1/ST6Gal1-treated sample, no unique hybrid-type glycans were identified (Figure 4.8).

Between the two strains, increased sialylation in the form of terminal N-acetylneuraminic acid (sialic acid, NeuAc) residues was observed throughout the population of complex- and hybrid-type structures, indicative of the synergistic effect of B4GalT1 and ST6Gal1 in adding terminal galactose and NeuAc residues (Figure 4.9). This was most prevalent at sites N88, N356, N463, N616, N637.

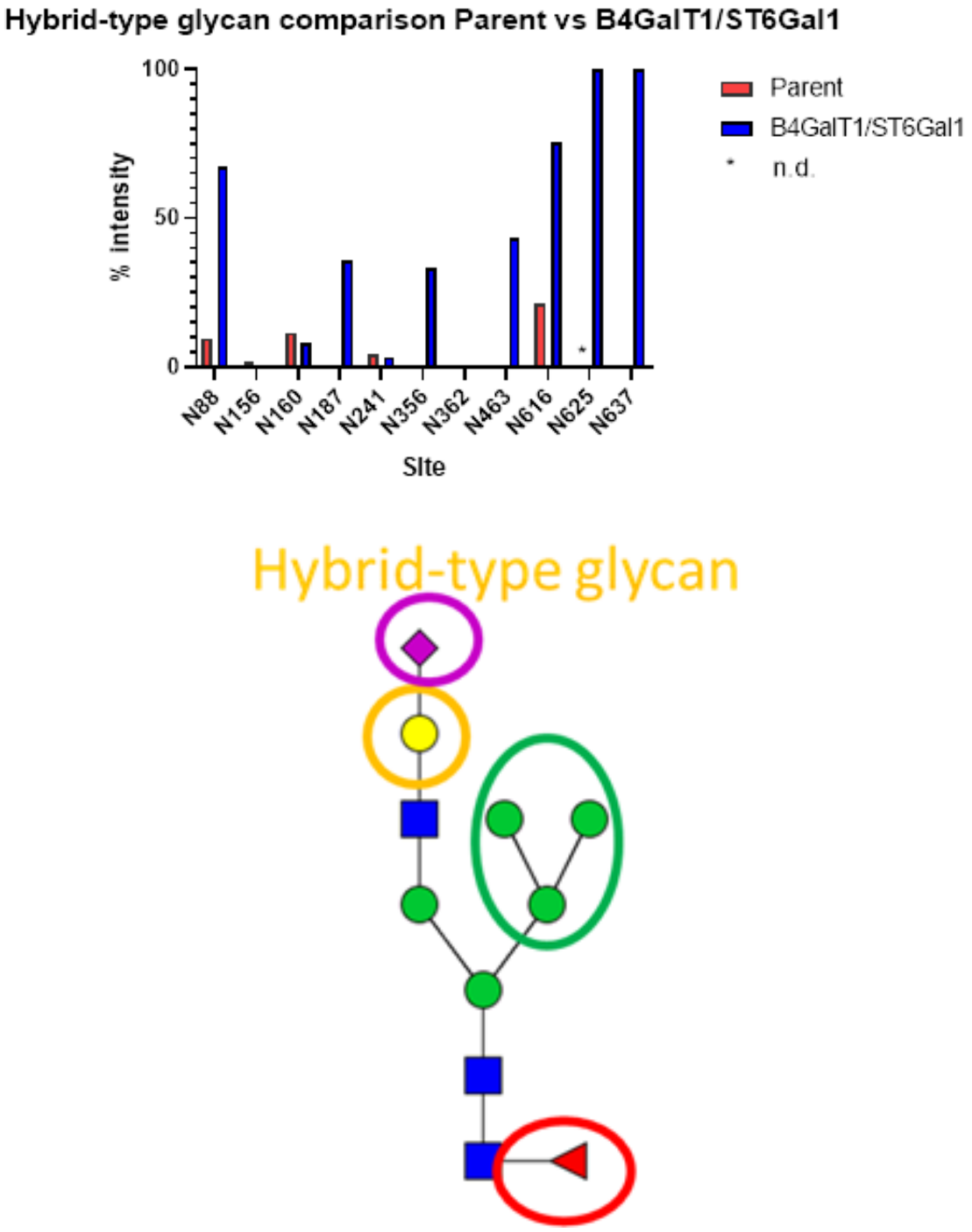


Figure 4.8: **Hybrid-type glycan comparison between JR-FL SOS E168K N189A +/- B4GalT1 + ST6Gal1**
Site-specific glycan analysis of JR-FL SOS E168K + N189A gp140 trimers, determined by LC-MS. All hybrid-type glycans were compared across both trimers and percentage intensities were compared for each site containing hybrid-type glycans. PNGSs numbered according to HXB2 alignment. Hybrid type glycans are identified as having one glycan arm consisting of only mannose residues and the other containing additional GlcNAc, galactose, and/or sialic acid residues. The first, asparagine-linked GlcNAc residue may also be fucosylated.

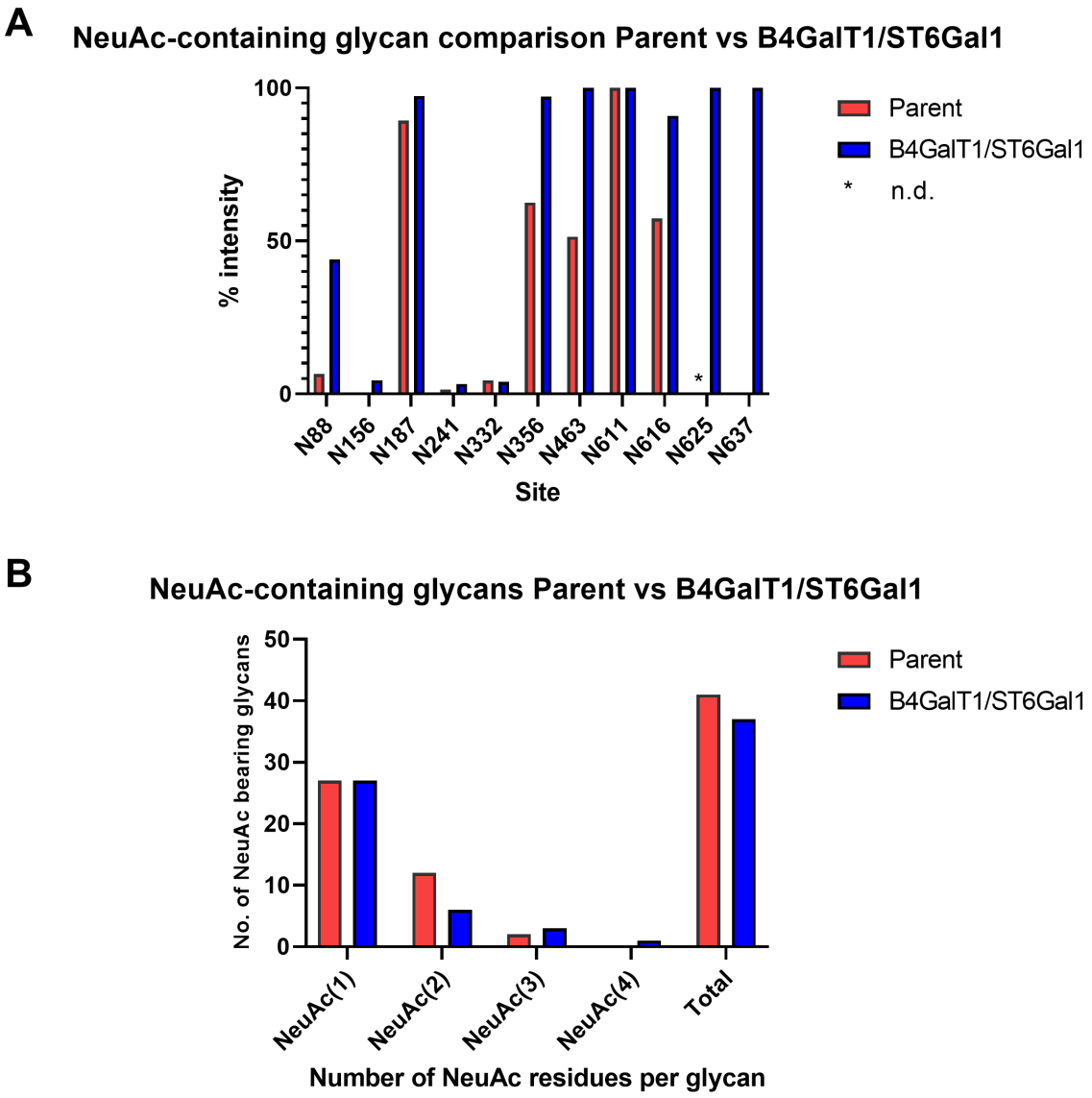


Figure 4.9: **Sialic acid (NeuAc) comparison and quantification between JR-FL SOS E168K + N189A +/- B4GalT1 & ST6Gal1**
(A) Site-specific analysis of JR-FL SOS E168K + N189A gp140 trimers. All sialic acid bearing glycans were compared for each site containing sialic acid bearing glycans. PNGSs numbered according to HXB2 alignment.
(B) Number of glycans containing between one and four NeuAc residues per glycan as well as totals for JR-FL SOS E168K + N189A +/- B4GalT1 & ST6Gal1 gp140s determined by site-specific glycan analysis.

Despite this increase, we observed less NeuAc-bearing glycans in the B4GalT1/ST6Gal1-treated sample, which suggests that there may be a glycoform focusing effect from the co-expression of these glycosyltransferases that goes beyond what is known with regards to their enzymatic specificities.

It has been reported that B4GalT1 and ST6Gal1 are able to form homo- or heterodimers between themselves, it is possible that glycosyltransferase activity may vary depending on the form of interaction these enzymes are undergoing, which itself is regulated by the Golgi environment [322]. It is thought that heterodimerisation can enhance substrate channeling through these two enzymes, which may be the cause for the aforementioned focusing effect on the glycoforms displayed on the B4GalT1/ST6Gal1-treated JR-FL trimer. The prevalence of oligomannose-type glycans remained similar between these two constructs [322].

The NGAF3-digested constructs, when compared to the control, resulted in the formation of the "stump" species, which consists of only the first GlcNAc residue which may be fucosylated. Endoglycosidase F3 (Endo F3) is known to act on bi- and triantennary complex-type glycans. This enzymatic mixture was observed to act mainly at the sites N88, N187, N356, and N616, mostly at the cost of the intended target; but additionally hybrid-type residues at these sites were almost entirely eliminated from the sample at these sites (Figure 4.7 A, C).

Importantly, Endo F3 is not known to digest hybrid-type glycans indicating that the effect of the enzymes goes beyond predicted catalytic properties. Interestingly, the glycans at the N160 site were found to be 11.34% hybrid-type structures in the control, 8.05% in the B4GalT1/ST6Gal1-treated sample, and 0.66% in the NGAF3-treated sample. Also, of note is that no stump species were detected in the latter sample at this site. This is explained by the elevation of oligomannose-type structures which means there are no substrates (and therefore no products) for Endo F3.

4.3.2.2 Site-specific glycan analysis of JR-FL gp120 by LC-MS

Aliquots of purified JR-FL gp120 constructs were reduced and alkylated, followed by subsequent enzymatic digestion with chymotrypsin. The digested glycopeptides were analysed by LC-MS to determine the N-glycan compositions at a site-specific level.

Unlike the aforementioned JR-FL constructs, gp120 are lone monomers which constitute part of an env trimer. Due to their smaller size they constitute a more malleable protein with regards to its steric access for certain glycans, which is of use in investigating the signatures that some of these enzymes may leave and comparing them to the trimeric JR-FL constructs.

FUT6 possesses the activity required for the transfer of α1,3-linked fucose to the GlcNAc of a type II lactosamine structure, which may be sialylated (Galβ1-4GlcNAcβ1-3-R or NeuAcα2-3Galβ1-4GlcNAcβ1-3-R).

When compared to the control (No treatment, Figure 4.10A), sites predominantly displaying fucosylated structures vary little in their quantity of fucosylated structures (e.g. N88, N160, N356, N463, Figure 4.10B), with subtle changes indicating a shift towards structures bearing more antennae. These changes are more apparent when looking at sites in the control that did not contain a large amount, if any, of antennary structures (e.g. N156 and N386); the latter site showing a +5.48 p.p. and a +13.76 p.p. increase in bi- and tri-fucosylated structures, respectively.

In terms of individual structures within each site, there was an observable but not marked increase in bi- and tri-fucosylated structures, which implies the action of FUT6 in the addition of further fucose residues. When the number of fucosylated glycans were compared between these two samples, there was both an increase in the total number of fucosylated glycans as well as the number of glycans with two or more fucose residues (Figure 4.11).

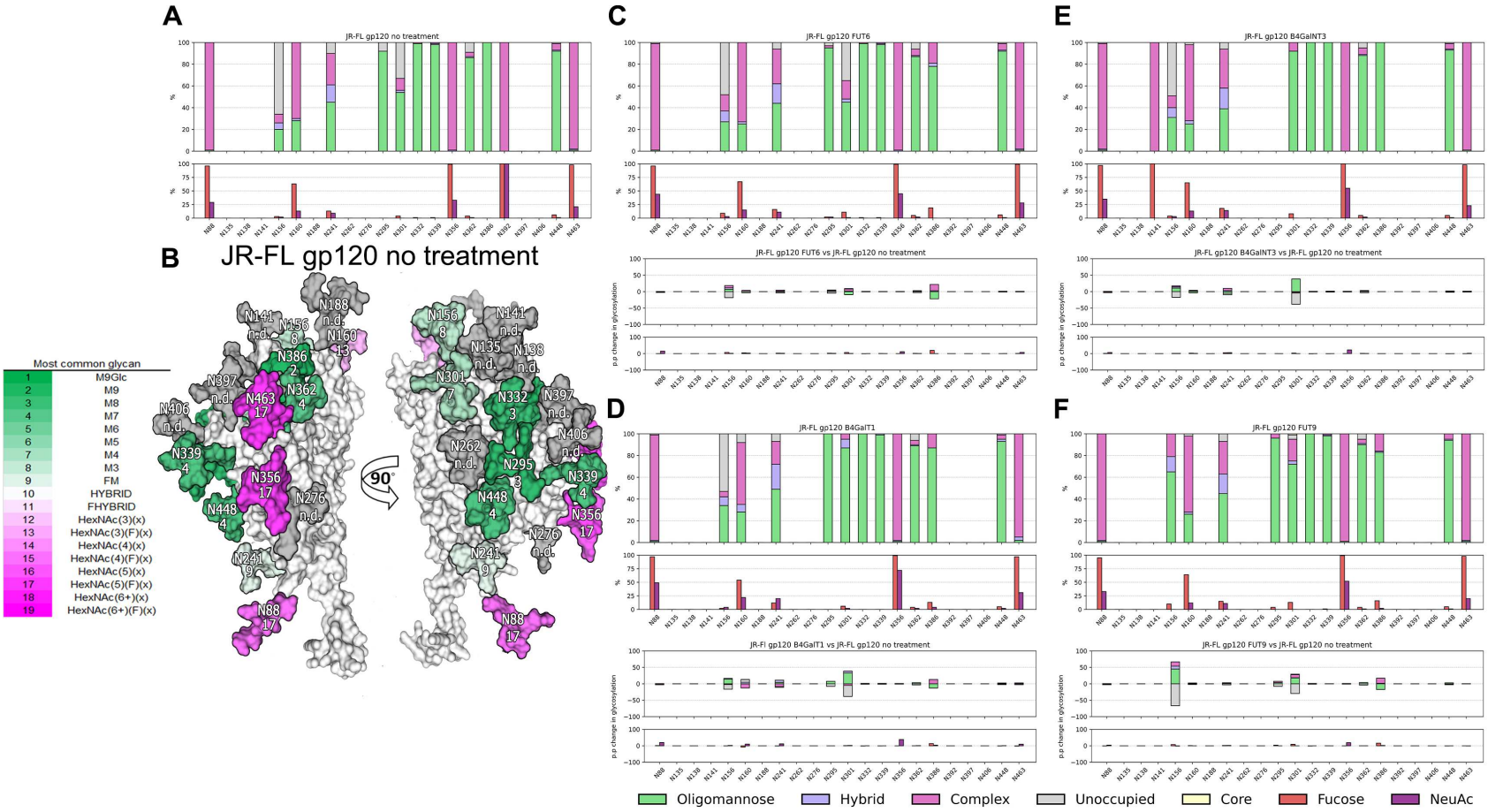


Figure 4.10: **Site-specific glycan analysis of JR-FL gp120s**

(A) Site-specific analysis of JR-FL gp120 no treatment.The upper graph shows the site-specific proportion of oligomannose, hybrid, complex, core glycans and under-occupied glycopeptides. The lower shows site-specific levels of fucose and sialic acid (NeuAc). Data is from one biological replicate. **(B)** JR-FL gp120 model (PDB ID: 6MYY). Glycans are colored according to a score that corresponds to the average maturation state (Key alongside **(B)**). Untrimmed oligomannose glycans are dark green. Glycans in grey were not resolved (n.d.).

(C-F) Upper: Site-specific analysis of JR-FL gp120 with glycosyltransferases. Lower: Arithmetic percentage point change between JR-FL gp120s treated with glycosyltransferases and JR-FL gp120 no treatment. Data is from one biological replicate.

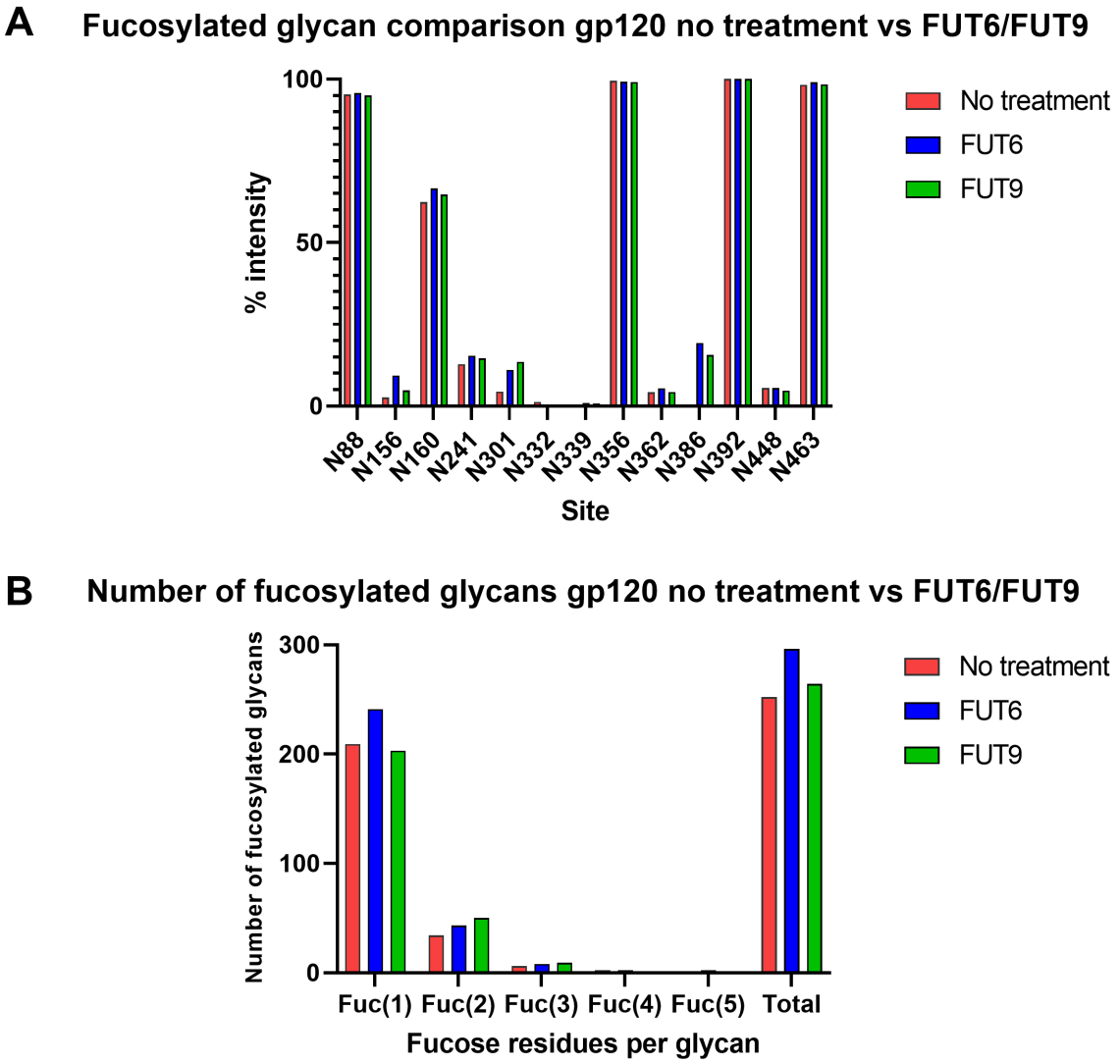


Figure 4.11: **Site-specific analysis of fucosylation between JR-FL gp120s +/- FUT6, FUT9**

(A) Site-specific analysis of JR-FL gp120s, as determined by LC-MS. All fucose bearing glycans were compared across both constructs and percentage intensity was calculated between sites where fucosylation was observed. PNGs numbered according to HXB2 alignment.

(B) Number of fucosylated glycans against fucose residues per glycan as determined by site-specific analysis.

B4GalNT3 is responsible for the transfer of GalNAc residues to α1,4-linked-GlcNAc residues synthesizing LacdiNAc structures on both N- and O-glycans. When compared to the control (Figure 4.10A), larger increases in oligomannose glycosylation were accompanied by subtle changes, generally increases, in the amount of bi- tri- and tetra-antennary complex-type glycans elsewhere (e.g. N88, N160, N241, and N356, Figure 4.10C). Increased occupancy at sites N156 and N301 were also observed with a concomitant increase in $Man_5GlcNAc_2$ (+42.61 p.p.; Figure 4.10A, C). Generally, increases in glycosylation were linked to the dominating glycan type for each site.

In terms of structures found within each site, there was no discernible differences in the quantity of HexNAc structures (GlcNAc or GalNAc) identified which would be indicative of the action of B4GalNT3 in adding GalNAc residues and the number of peptides bearing glycans with 5 or more HexNAc structures was relatively similar to that of the control.

B4GalT1 is an enzyme that is involved in the transfer of Gal residues to an acceptor GlcNAc by a β1,4-bond. When compared to the control (Figure 4.10A), total increases in hybrid-type glycans, particularly at sites N160, N241 and N301 were observed (Figure 4.10A, D). When hybrid and fucosylated hybrid-type glycans were compared between no treatment and B4GalT1, the observed elevation in hybrid type structures in each site showed a bias towards afucosylated hybrid-type glycans (Figure 4.10A, D; Figure 4.12).

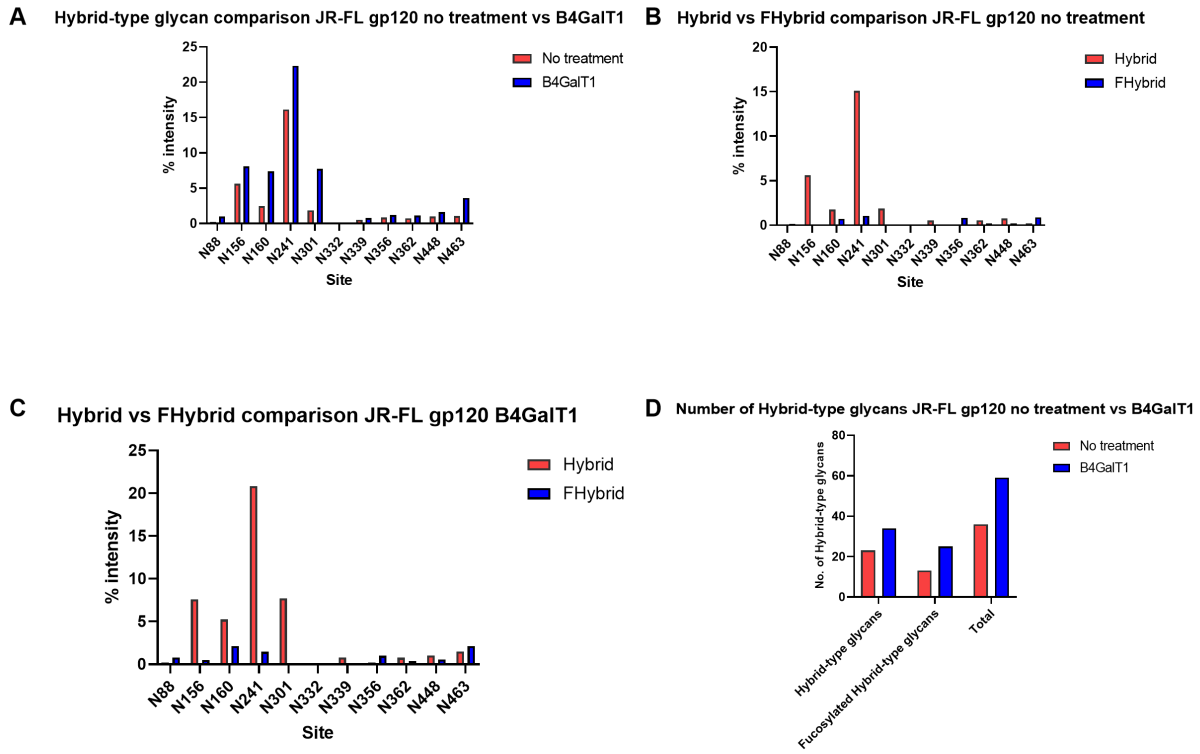


Figure 4.12: **Hybrid-type glycan comparison between JR-FL gp120 +/- B4GalT1**
(A) Site-specific analysis of JR-FL gp120 monomers, determined by LC-MS. All hybrid-type glycans were compared across both trimers and percentage intensities were compared for each site containing hybrid-type glycans. PNGSs numbered according to HXB2 alignment.
(B-C) Site-specific comparison of fucosylated and afucosylated hybrid-type structures within the gp120 monomers.
(D) Comparison of the number of hybrid-type glycans +/- fucose within the different gp120s, including total numbers.

When differences in hybrid-type glycans identified were compared between the control and B4GalT1-treated gp120s, five different glycans were identified in the B4GalT1-treated gp120s that were not in the control. It should be noted however individually none of these glycans comprised more than 0.12% intensity on their given site. These glycans are listed below (Table 4.2).

Start AA	End AA	Var. Pos. Protein	Protein Sequence	Glycan	Score	Validate	XIC Area summed	Percentage	Classification
75	98	88	W.ATHACVPTDPNPQEVVLEnVTEHF.N	HexNAc(4)Hex(6)Fuc(1)	160.17	True-positive	1.08E+07	0.037%	FHYBRID
75	98	88	W.ATHACVPTDPNPQEVVLEnVTEHF.N	HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	136.59 - 178.92	True-positive	3.20E+07	0.109%	FHYBRID
75	98	88	W.ATHACVPTDPNPQEVVLEnVTEHF.N	HexNAc(4)Hex(6)Fuc(3)	30.00 - 131.60	True-positive	3.42E+07	0.116%	FHYBRID
75	98	88	W.ATHACVPTDPNPQEVVLEnVTEHF.N	HexNAc(4)Hex(7)Fuc(2)	30.00 - 30.00	True-positive	9.00E+06	0.031%	FHYBRID
237	252	241	F.NGKGPCKnVSTVQCTH.G	HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	45.96	True-positive	4.23E+06	0.01%	FHYBRID
237	252	241	F.NGKGPCKnVSTVQCTH.G	HexNAc(4)Hex(6)Fuc(3)	30	True-positive	2.18E+06	0.01%	FHYBRID
363	375	362	F.nHSSGGDPEIVMH.S	HexNAc(4)Hex(6)	63.09	True-positive	1.70E+07	0.03%	HYBRID

Table 4.2: **Table of unique hybrid-type glycan structures found in JR-FL gp120 + B4GalT1 compared to WT**
Site-specific analysis of JR-FL gp120 co-transfected with B4GalT1 not present in JR-FL gp120 no treatment, as determined by LC-MS. All hybrid-type glycans were compared across all sites and percentage intensity was calculated against all glycans within the given construct. PNGs numbered according to HXB2 alignment.

FUT9, like FUT6, has the activity required for the transfer of α1,3-linked fucose to the GlcNAc of a type II lactosamine structure. When compared to the control (Figure 4.10A), there does not appear to be as unified a shift towards more fucosylated structures or structures bearing more antennae as seen in the FUT6-treated sample, however there is an observed increase in structures bearing two fucose residues as well as a general elevation in fucosylation when compared to the control, particularly at N301 and N386 (Figure 4.10E; Figure 4.11).

The sites containing mostly complex- or hybrid-type glycans (N88, N160, N241; Figure 4.10E) are relatively similar in composition between the control and FUT9-treated sample. Interestingly, The N356 site shows a reduction in mono- and bi-antennary fucosylated glycans (-3.87 p.p., -9.11 p.p. respectively; Figure 4.10E), which appears to be compensated by an increase in tri- and tetra-antennary fucosylated glycans (+1.72 p.p., +11.38 p.p. respectively; Figure 4.10E).

When number of fucosylated glycans were compared with the control, the FUT9-treated sample showed a greater number of fucosylated glycans as well as more glycans with two or more fucose residues (Figure 4.11). Additionally, the N386 site which in the control lacked

any complex- or hybrid-type glycan, was 17.13% complex/hybrid-type in the FUT9-treated sample, with an +11.73 p.p. increase in tri-antennary fucosylated glycans at this site (Figure 4.10A, E; Figure 4.11A).

Of note across multiple sites was a substantial drop in un-occupancy with no obvious explanation, particularly at sites N156 and N301, which showed a -13.95 p.p. and -29.46 p.p. reduction respectively compared to the control (Figure 4.10E). This was observed for all gp120 constructs except for the FUT6 treated gp120 construct.

4.3.3 Discussion

This project sought to back up previous findings that certain bnAbs were more effective against α2,6-hypersialylated glycans [225]. Indeed, in the B4GalT1 and ST6Gal1 treated sample, whilst there were less sialylated glycans in total, the overall sialylation increased, indicating that the previously observed increase in bnAb binding may be attributable to increased sialylation.

Increased sialylation could offer a benefit to immunogen design, as there is a suggestion that α2,6-hypersialylated immunogens may be ideal for boosting strategies as they better mirror peripheral blood mononuclear cell (PBMC) derived viral glycosylation [225]. It is also thought that using GE tools could allow for further investigation of the effects of variation in glycosylation in eliciting neutralising antibodies [225, 243].

The NGAF3 treated trimer was expected to yield 'stump' species as an indicator for the action and accessibility of Endo F3 on well-folded trimers. The typical cleavage targets for Endo F3 are biantennary and tetra-antennary complex glycans. Of the sites predicted to display glycans sensitive to Endo F3, only a subset exhibit the cleavage product. This is explained by both incomplete digestion and a change in glycan processing at some sites which eliminates the existance of glycans susceptible to Endo F3 cleavage. Whilst some decreases in bi- and tetra-antennary complex glycans were observed, the unexplained decrease in hybrid-type glycosylation that was observed at sites bearing the indicator 'stump' glycans require

further clarification as it goes beyond expected enzymatic properties. It is possible that the added neuraminidase, hexoseaminidase and galactosidases may be exerting a quantum shift during the process of glycosylation that may lead to the formation of alternate glycans but in order to prove this their effects will require individual characterisation. The fact that the N160 site did not exhibit any cleavage product and still showed a decrease in hybrid-type structures lends itself to this hypothesis.

Of the gp120 constructs that were co-expressed with FUT6 or FUT9 fucosyltransferase plasmids, the sites that appeared to be most susceptible to their effects were predominantly complex-type. As such, increases in fucosylation both as a measure of intensity and of total number of fucosylated glycans did not substantially deviate from that observed in the control. Importantly, as these constructs were purified from HEK 293T cells, it is known that this expression vector already endogenously produces the highly similar FUT8 α1,3-fucosyltransferase. This raises difficulties in comparing the co-expression of these two enzymes due to the additional involvement of this fucosyltransferase.

Both FUT6 and FUT9 co-expression led to elevated levels of fucosylation compared to the control. FUT6 exhibited the largest increase in total fucosylated glycans, particularly of those with only one fucose reside, whilst FUT9 was found to elevate the number of glycans with two or more fucose residues. Co-expression of both enzymes led to substantial reductions in glycan under-occupancy at N156, but also N301 for FUT9. Certain sites also showed signs of a shift that favouring the formation of additional antennae of hybrid-type glycans following FUT9 co-expression, which could indicate a synergistic effect in adding additional glycan arms for fucosylation. The N386 site was exclusively oligomannose-type in the control, yet in both FUT6 and FUT9 co-expressed constructs we observe an increase in both fucosylated complex-type glycans as well as hybrid-type glycans which also goes against our current understanding of the enzymatic properties of these fucosyltransferases. Further work is required to better characterise these enzymes in order to better understand in order to understand where they differ. Characterising the effects of FUT8, or lack thereof when

examining other fucosyltransferases, will allow for a better understanding of what separates these enzymes in their function.

The co-expression of B4GalNT3 showed the least widespread effect on the JR-FL gp120 construct in comparison to the other co-expressed enzymes. Generally, co-expression of B4GalNT3 led to small increases in glycans that were associated with the dominating glycan profile of the site. The most substantial change in this construct was a focusing of $Man_5GlcNAc_2$ glycans at N301 which came mostly at the cost of the observed under-occupancy at this site in the control. This increase was most pronounced in the B4GalNT3 co-expressed construct but was visible in all co-expressed constructs excluding FUT6.

B4GalT1 co-expression led to similar but more pronounced effects than those observed in the B4GalNT3 co-expressed constructs. There were greater increases in glycosylation of glycans pertaining to the dominant glycan type in most sites, and similar drops in glycan under-occupancy at N156 and N301. In addition, there was a notable, but not significant increase (1-5%) in hybrid-type glycans at a number of sites. Comparison of hybrid-type glycosylation in terms of both intensity and raw number of hybrid-type glycans further validated this observation. The co-expression of B4GalT1 was presumed to yield increased hybrid-type glycans that were not previously present in the control [225]; whilst hybrid-type glycosylation was increased at certain sites, there were few unique glycans that were not previously identified in the control gp120 construct. The lack of widespread changes in hybrid-type glycans in this construct suggests a limitation in the availability or accessibility of its substrate, which in turn could affect the ability of this enzyme to introduce such glycans [70, 107].

In conclusion, glycoengineering HIV-1 env constructs offers a potentially attractive avenue for both immunogen design and potential strategies for vaccines such as prime-boost regiments [253, 302]. With regards to immunogen design, better understanding of the effects that the co-expression of glycoengineering enzymes can confer on env constructs, both monomeric and trimeric, could allow for the fine tuning of desired epitopes. Alternatively, they may serve as a way of identifying undiscovered bnAbs from infected donors [225]. Understanding

how these glycans act on env and how that in turn affects the immunogenicity of candidate env immunogens will hopefully lead to a better understanding of the requirements for a broadly neutralising antibody response.

5 Epitope-directed glycan engineering

The previous chapter outlined that enzymatic interventions on Env glycans can be undertaken to widespread effect across multiple sites. Whilst these enzymatic interventions have the potential to improve the design of Env vaccine candidates, they do not present a controlled intervention on the glycan shield. In this chapter, we sought other co-translational methods by which we could make more controlled interventions on the glycan shield without affecting the entire glycan shield.

Full control of the glycan processing of glycoproteins offers a route to the development of bespoke, optimised biologicals. However, control of glycan processing has typically been restricted to global approaches where engineering glycosylation at one site necessitates the same intervention at another. While chemoenzymatic approaches can offer a route to site-specific control of glycosylation of simple targets, the site-specific control of heavily glycosylated biologics using cellular expression systems remains a challenge. Here, we exploit the specificity of broadly neutralizing antibodies (bnAbs) to site-selectively control the glycan processing. We exemplify the approach using the densely glycosylated envelope glycoprotein of HIV-1, Env. Co-expression of a recombinant, soluble form of Env with antibodies targeting epitopes including or proximal to glycans can provide a localized steric blockade to glycan processing thus punctuating Env with localized changes in glycosylation. Using this approach, it was possible to selectively restrict glycan processing at sites surrounding the CD4 without compromising the glycan processing of distant glycosylation sites. Epitope-directed control of glycan processing exploits the specificity of antibodies and provides a a method by which the site-specific modulation of glycan processing can be achieved.

5.1 Contributions

Chapter 5 involves a comparison of BG505 SOSIP.664 Env trimers that were recombinantly expressed alongside bnAb, trastuzumab (a non Env-specific antibody) and sCD4. The results from this chapter were obtained in collaboration with Dr. Gemma Seabright who aided in the sample preparation and data analysis for some of the bnAb-Env co-expressions. The samples analysed are as follows:

SOSIP Analysed	**No. Biological replicates**
BG505 SOSIP.664 WT	3
BG505 SOSIP.664 1:1 PGT145	1
BG505 SOSIP.664 1:1 PG16	1
BG505 SOSIP.664 1:1 PGT121	1
BG505 SOSIP.664 1:1 PGT135	1
BG505 SOSIP.664 1:1 2G12	1
BG505 SOSIP.664 1:1 PGT151	1
BG505 SOSIP.664 1:1 VRC01	6
BG505 SOSIP.664 1:1 VRC02	3
BG505 SOSIP.664 2:1 VRC01	3
BG505 SOSIP.664 1:2 VRC01	3
BG505 SOSIP.664 1:2 Trastuzumab	3
BG505 SOSIP.664 1:1 sCD4	3
BG505 SOSIP.664 VRC01p	3
BG505 SOSIP.664 VRC02p	3
BG505 SOSIP.664 PGT151p	1

Table 5.1: **Table of co-expressed proteins and the number of biological replicates expressed**

5.2 Introduction

Glycoproteins, such as antibodies and viral spike proteins, [323, 324], represent important biological therapeutic targets. The glycosylation of these proteins may modulate their function in vivo and their structures have the ability to dictate the type of response a particular therapeutic may cause, therefore, controlling the glycosylation of the targets is of therapeutic

interest. Currently, control of glycosylation is limited to non-specific modifications of the entire glycome, such as the choice of expression system, the use of inhibitors of glycosylation enzymes (e.g. kifunensine), the use of cell lines lacking particular glycosyltransferases (e.g. GnT I-/- cells), co-transfection with glycosyltransferase encoding plasmids, and in vitro enzymatic reactions [120, 225, 325, 326]. These glycoengineering techniques often result in universal changes to in glycan structures on proteins containing more than one PNGS. This in turn can have adverse effects on the structure and function of a glycoprotein, hence more subtle, site specific control poses an alternative approach which could allow for additional studies of the influence of individual glycan sites [263]. An alternate intervention that can be used to alter the glycosylation is through the addition or removal of PNGSs [257]. Whilst this generally leads to increased processing of nearby glycans, such mutations are not predictable and may lead to unwanted consequences with regards to glycoprotein integrity [263, 257, 155, 255, 256].

A system in which control of specific PNGSs could have strong therapeutic impact is HIV-1 Env spike protein, as the glycan shield of HIV-1 forms epitopes for broadly neutralizing antibodies. The envelope glycoprotein is a trimer of gp120-gp41 heterodimers and has up to 39 PNGSs per gp120-gp41 protomer, and up to 120 PNGSs per trimer, 12 which allow masking of strongly immunogenic protein surfaces [327, 328]. A high density of glycan sites on the trimer form clusters which impose steric constraints upon glycan processing enzymes, leading to areas of under-processed oligomannose-type glycans which are referred to as mannose patches [107]. The presence of these oligomannose-type glycans at certain PNGSs can be used to distinguish between compact, native-like trimers and more open or non-native conformations as ER- and Golgi-resident mannose trimming enzymes require an extensive accessible surface area to act, which can be inhibited by steric clashes from local protein architectures [70, 123]. The conservation of the presentation of oligomannose-type glycans across diverse HIV-1 Env strains demonstrates their importance in vaccine design and as a target for therapeutic intervention [254]. It has been shown that a large subset of HIV-1

infected patients eventually form antibodies that cross-react with the broad epitopes that consist of both glycans and protein surface, known as broadly neutralizing antibodies (bnAbs).14 A notably potent class of bnAbs specifically target the oligomannose patch, although other regions of the Env trimer can form targets for bnAbs [262, 329].

Much study has gone into the development of a hypothesized prophylactic HIV vaccine, which would induce production of bnAbs in patients. Several of the candidate immunogens revolve around the use of stabilized envelope spike constructs. One such immunogen is a solubilized Env that contains a disulfide bridge between gp120 and gp41 and an isoleucine to proline mutation, known as SOSIP [109]. It now apparent that a single native-like trimer immunization will be unable to elicit bnAbs. This is due to the extensive hypermutations required to confer the ability of an antibody to overcome the dense glycan shield of Env [111, 223], even in recombinant constructs designed specifically for this purpose [223, 206, 330, 204].

Strategies to overcome these issues include using constructs with deliberate glycan site knockouts to improve germline antibody binding [213, 224, 331], however a lack of a glycan at a particular site can prevent the binding of a bnAb, making certain PNGS knockouts undesirable [233]. Previous glycoengineering attempts on SOSIPs and similar spike constructs include kifunensine treatment, which inhibits ER α-mannosidase I, which results in the conversion of all N-glycans to $Man_9GlcNAc_2$, and the use of GnTI deficient cell lines which result in the presentation of elevated oligomannose-type glycans at all sites. Kifunensine treatment corresponded to increased neutralization by bnAbs PGT121, PGT125, PGT128 and 35O22 [225, 143], and GnT I deficient cell lines induced a similar effect for bnAbs VRC01 and PGT145 [332]. Kifunensine treament and the use of GnT I-/- cell lines, however, results in poorer neutralization by bnAbs such as PG9, PG16 and PGT151 [225, 233, 235, 239].

While these glycoengineering approaches are able to beneficially enhance the neutralization of particular bnAbs, they are global approaches to a highly complex system. In order to have the best chance of inducing bnAbs, candidate immunogens must present features which

bnAbs, or their germline precursors, preferentially bind without minimal disruption of glycan epitopes distal to the sites of interest. In this study, we sought to examine the intertwined relationship between BG505 SOSIP and the bnAbs which target Env. We co-expressed SOSIPs with a panel of bnAbs, which target epitopes at several regions across the protein, and determined changes in N-linked glycosylation using both UPLC and LC-MS.

We observed that different bnAbs induce diverse changes on the co-expressed SOSIP, many of which localized to the epitope of the co-transfected bnAb. We hypothesize that the co-expression of bnAbs with viral glycoproteins could represent a novel pathway for inducing targeted changes in glycoproteins. However, the mechanism by which this process occurs does not appear to be common amongst all antibodies investigated. Despite this, clear examples of the site-specific modulation of N-linked glycosylation can be seen, providing an avenue by which the compositions of particular N-linked glycan sites can be modified to present more desirable glycan compositions. The samples analysed for this project are outlined in table 5.1.

5.3 Results

5.3.1 Influences of binding and non-binding partner co-expression on envelope glycosylation

To determine the effects of co-expression, we first determined the site-specific glycan composition of WT BG505 SOSIP.664 trimers from three independently expressed biological replicates expressed in Human embryonic kidney (HEK) 293F cells. BG505 SOSIP trimers were purified by nickel-affinity chromatography and size-exclusion chromatography (SEC). These were digested using the proteases α-lytic protease, chymotrypsin, and trypsin in order to generate glycopeptides containing a single glycosylation sequon (N-X-S/T-X, where X-any amino acid except proline) using LC-MS. We were able to quantify high mannose and complex-type glycosylation as well as glycan un-occupancy across the majority of the

glycan sites on BG505 SOSIP. To determine if target co-expression imparted changes to glycosylation BG505 SOSIP was co-expressed with soluble CD4 (sCD4) in a 1:1 (w/w) plasmid ratio. In general, BG505 SOSIP is expressed in a 4:1 ratio with furin to enhance cleavage, so when expressed in a 1:1 ratio, BG505 SOSIP 4:1 Furin accounts for '1' and an equal amount of DNA to this combination is accounted for by the sCD4 plasmid. As a negative control, we chose to co-express BG505 SOSIP with a non-targeting antibody, Trastuzumab, in a 1:2 SOSIP:Trastuzumab ratio. Trastuzumab was chosen as it constitutes a negative control that is otherwise very similar in size and biosynthesis to the subsequent antibodies analysed in this chapter. in this case the Trastuzumab antibody expression (and all other subsequent antibodies in this chapter) itself constitues a ratio of 2:1, light:heavy chain. Trimers were purified by nickel-affinity chromatography and biological triplicate datasets were averaged and compared to the WT glycan profile by arithmetic percentage point difference (Figure 5.3).

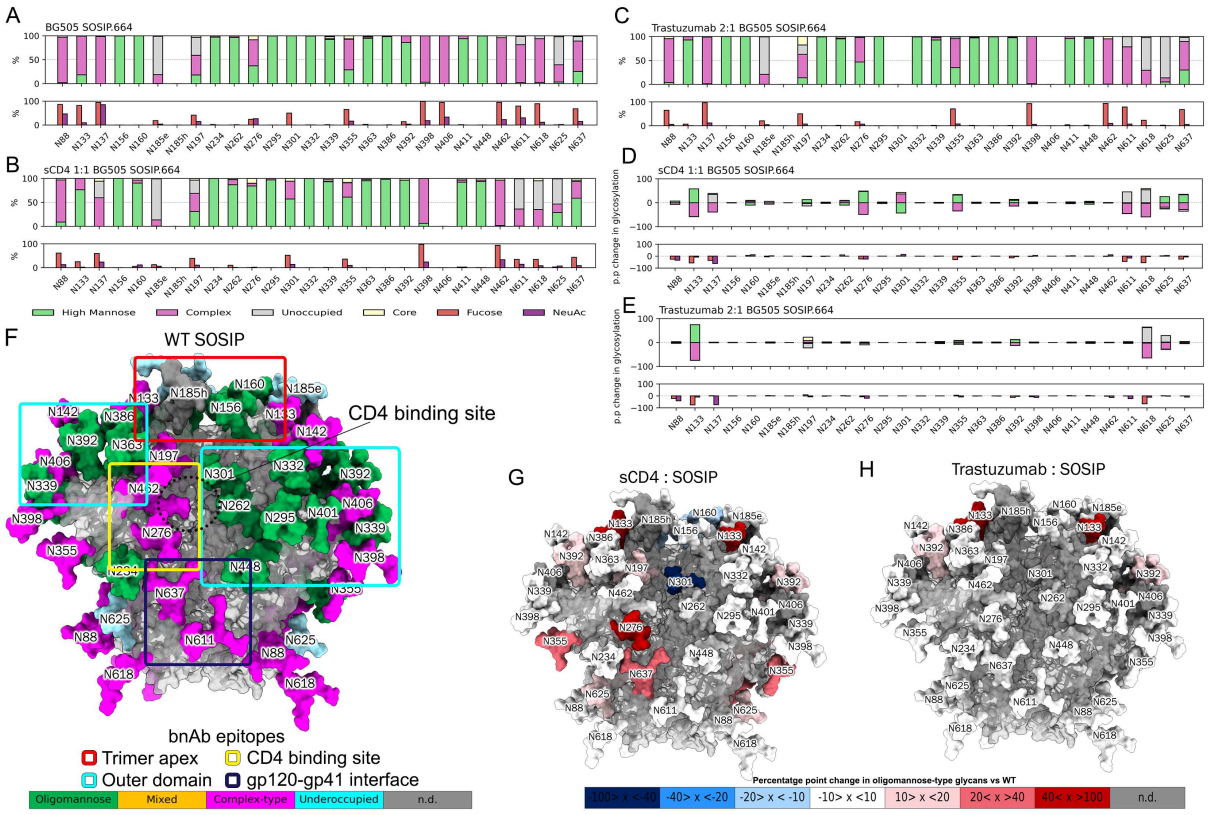


Figure 5.1: **WT BG505 SOSIP, sCD4 and Trastuzumab co-expression glycan analysis**

(A-C) Average site-specific glycan analysis of WT BG505 SOSIP, sCD4 (1:1) and Trastuzumab (2:1) co-expressions from three biological replicates. Proportions of high mannose, complex-type, core and unoccupied PNGS at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. All three datasets correspond to mean site-specific glycopeptide data from three biological replicates.

(D-E) Arithmetic percentage point difference between **(A)** and **(B-C)**.

(F) Average WT glycosylation from **(A)** plotted onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). Green sites represent >80% oligomannose, yellow (mixed): 40-79%, magenta (complex): <40%, cyan (underoccupied): >50% underoccupied, grey (n.d.).

(G-H) Arithmetic percentage point change in oligomannose-type glycans between WT and sCD4/Trastuzumab mapped onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey (n.d.).

With regards to sCD4 co-expression, it is well documented that CD4 binding to gp120 leads to structural rearrangements required for co-receptor binding and subsequent viral entry [118, 334, 66] . With this in mind the choice of envelope construct and purification method is significant, as analysis of non-trimeric material would be uninformative. The SEC traces of WT and sCD4 co-expressed BG505 SOSIP indicated a shift towards more unfolded material in the latter sample however we were still able to isolate a trimeric peak perhaps owing to enhanced structural rigidity conferred from the stabilising mutations of this construct.

Comparison of the sCD4 dataset to the WT yielded changes in glycosylation at a certain number of sites, particularly N276 and N301 which are located close to the CD4-binding site (CD4bs) (Figre 5.3B). N276 has been observed as a mixed site in this and previous analyses whilst N301 is usually found to be comprised of only high-mannose glycoforms. Here we observe a sharp +47 p.p increase in high-mannose at N276 and a -43 p.p. drop in these glycoforms at N301 (Figure 5.3D, G). Other notable differences include a sharp +55 p.p. increase in high mannose at N133 as well as increases at sites N355, N625 and N637. Glycan un-occupancy also increased at sites N611 and N618 (+46, +54 respectively)(Figure 5.3D). The glycan profile from Trastuzumab co-expression showed a high degree of fidelity with WT. The only significant changes observed from the Trastuzumab co-expressed dataset was a similar sharp increase in high-mannose at N133 (+75 p.p.) and increase in un-occupancy at N618 (+62), suggesting potential biosynthetic effects associated with co-transfection of multiple proteins(Figure 5.3E, H).

5.3.2 Glycosylation is affected by proportion of co-expressed antibody

Despite the specific changes conferred to BG505 SOSIP trimers post-sCD4 co-expression, its ability to confer structural changes to a subset of this construct makes it an unsuitable candidate for engineering its glycosylation. We therefore sought a target that might achieve this whilst maximising protein integrity. bnAbs targeting Env have been previously used to purify trimeric material based on their specificity for particular structures and epitopes [335] . Based on this we chose the bnAb VRC01, which targets an protein-based epitope around the CD4 binding site and N276 [298] , to see if its co-expression would induce similar changes to sCD4. Biological triplicate co-expressions were purified in SOSIP:VRC01 ratios of 2:1, 1:1, and 1:2 to check for the effects observed in the Trastuzumab co-expression. Trimer-bnAb complexes were purified by means of Protein-A affinity chromatography only. This allowed us to assess the glycosylation of solely co-purified BG505 SOSIP trimers. From this data we suggest that the majority of the co-purified trimers are well folded, as glycan sites that lie at the trimer apex, are predominantly high-mannose in well-folded trimeric material and show enhanced levels of complex-type glycosylation when monomeric or misfolded [123]. In the case of these trimers the N156/N160 glycans sites were overwhelmingly oligomannose-type.

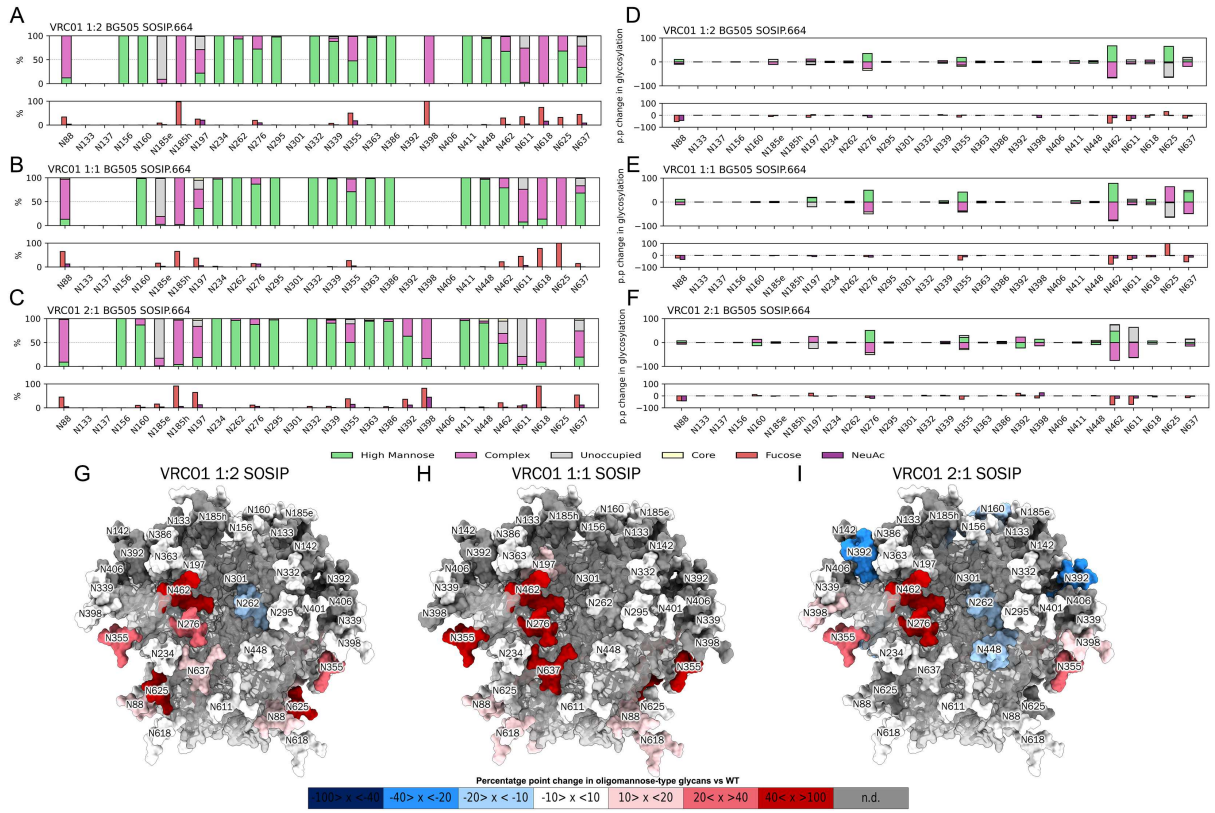


Figure 5.2: **VRC01 co-expression at different ratios**

(A-C) Site-specific glycan analysis of SOSIP co-expressed with VRC01 in a 2:1 **(A)**, 1:1 **(B)**, and a 1:2 **(C)** (w/w) ratio. Proportions of high mannose, complex-type, core and un-occupied PNGS at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. All three datasets correspond to mean site-specific glycopeptide data from three biological replicates.

(D-F) Arithmetic percentage point difference between average WT SOSIP glycosylation and VRC01 (2:1) **(D)**, VRC01 (1:1) **(E)**, VRC01 (1:2) **(F)**.

(G-I) Co-expression percentage point difference vs WT mapped onto a crystal structure of BG505 SOSIP.664: VRC01 (2:1) **(G)**, VRC01 (1:1) **(H)**, VRC01 (1:2) **(I)** (PDB ID: 4NCO [333]). Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

Glycan analysis of VRC01 co-expressed SOSIP revealed extensive increases in high-mannose glycoforms that exceeded sCD4 co-expression. At N276 $Man_9GlcNAc_2$ glycoforms, which were only previously present in sCD4 co-expression at 5% on average, exceeded 20% in all three SOSIP:VRC01 ratios (31%, 23%, 72%) (Figure 5.2A-C). Of these three ratios, only the 2:1 SOSIP:VRC01 co-expression averaged a lower proportion of high-mannose glycoforms at this site. Similarly, the adjacent N462 glycan site which is typically composed of complex-type glycans showed large increases in high-mannose glycosylation (+67, +78, +48 p.p. respectively) overwhelmingly comprised of $Man_5GlcNAc_2$ glycoforms and far exceeding the 2% high-mannose in sCD4 co-expression [69] (Figure 5.2D-F; G-I). Elsewhere, modest increases in high-mannose glycans could be observed at the nearby N197 and N637 sites, as well as more distal sites such as N355. We note that whilst elevation compared to the WT was observed in most of these sites, this effect was most pronounced in the 1:1 co-expression and we observed a drop in some high-mannose glycosylation in 2:1 and 1:2 co-expressions. With the exception of these sites, the remaining glycan sites showed very little change compared to the WT with apex and mannose patch sites maintaining the same glycosylation overall.

5.3.3 Differentiating antibody glycan biasing from co-expression

To ensure that the effects imparted from VRC01 co-expression were as a result of co-expression and not as a result of glycan biasing resulting from post-secretion binding, we decided to compare the glycan profiles of another three biological replicates of VRC01 co-expressed BG505 SOSIP (1:1) with VRC01-purified (VRC01p). WT BG505 SOSIP was incubated to completion. Separately purified VRC01 antibody was then applied to the filtered supernatant containing BG505 SOSIP and was incubated for a day prior to Protein-A purification.

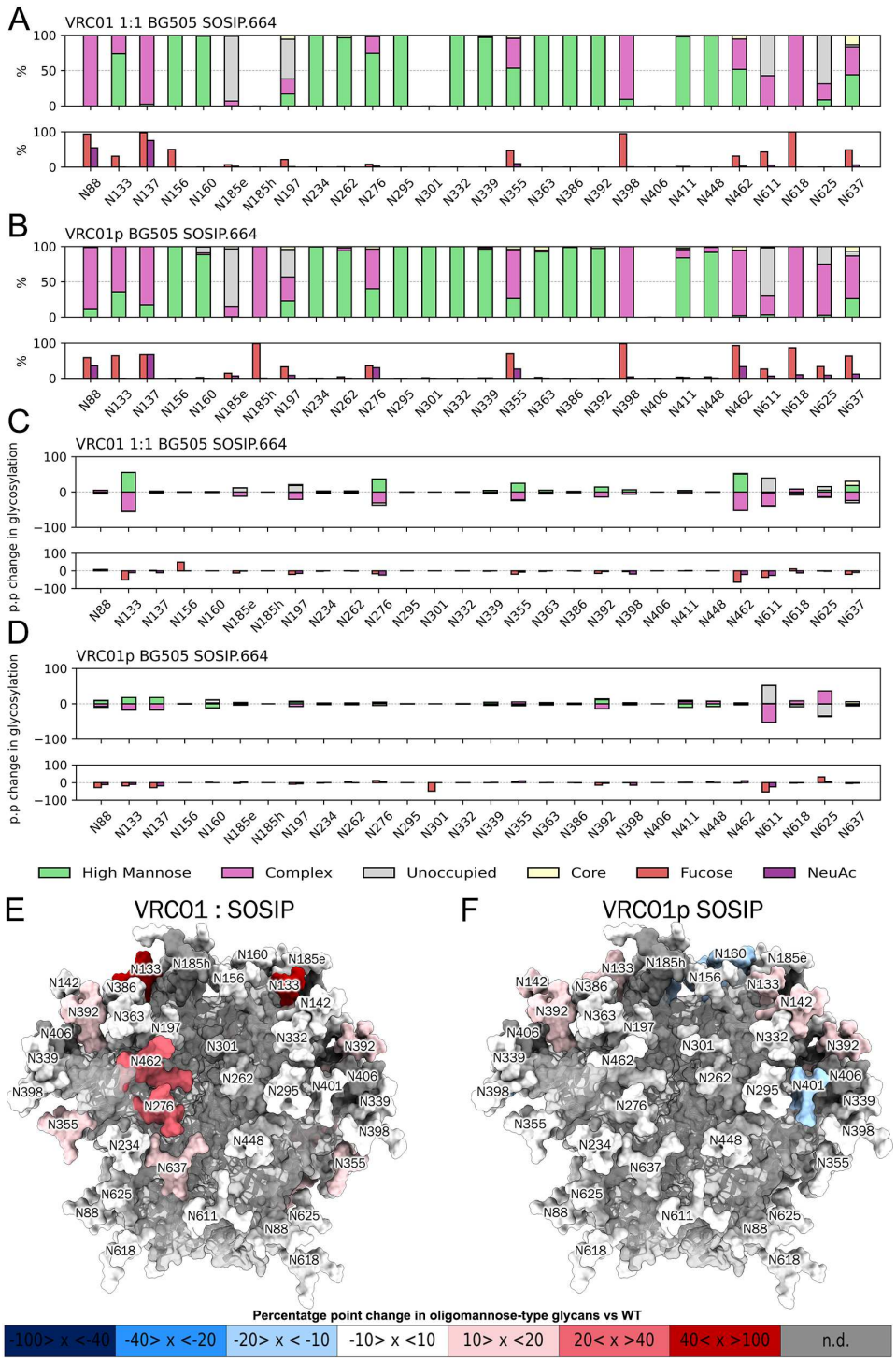


Figure 5.3: **VRC01 co-expression and VRC01-purified SOSIP glycan analysis**
(A-B) Site-specific glycan analysis of SOSIP co-expressed with VRC01 in a 1:1 ratio **(A)** WT SOSIP co-purified using VRC01 **(B)** Proportions of high mannose, complex-type, core and un-occupied PNGS at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. Both datasets correspond to mean site-specific glycopeptide data from three biological replicates.
(C-D) Arithmetic percentage point difference between average WT SOSIP glycosylation and VRC01 (1:1) **(C)**, and VRC01p **(D)**
(E-F) VRC01 co-expression/purification percentage point difference vs WT mapped onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

Comparison of the two datasets indicate that the processing effects at CD4bs glycan sites are directly attributable to VRC01 co-expression. Whilst these biological replicates of the 1:1 co-expression did not show as strong a shift to high-mannose glycosylation as the prior dataset, N276 glycosylation was once again predominantly $Man_9GlcNAc_2$ (21%) (Figure 5.3A, C). N355 and N462 also showed a lesser, but significant +25 and +51 p.p. increase in high-mannose glycoforms, respectively (Figure 5.3C, E). In comparison, VRC01p BG505 SOSIP did not show any change in high-mannose glycosylation that exceeded 20 percentage points (Figure 5.3 D, F). To assess the specificity of the SOSIP-VRC01 interaction and its effects on BG505 SOSIP trimers, we compared the glycan profiles of BG505 SOSIP co-expressed (1:1) and purified with VRC02 (VRC02p). VRC01 and VRC02 are highly related bnAbs that are derived from variants of the same IgG clone [336].

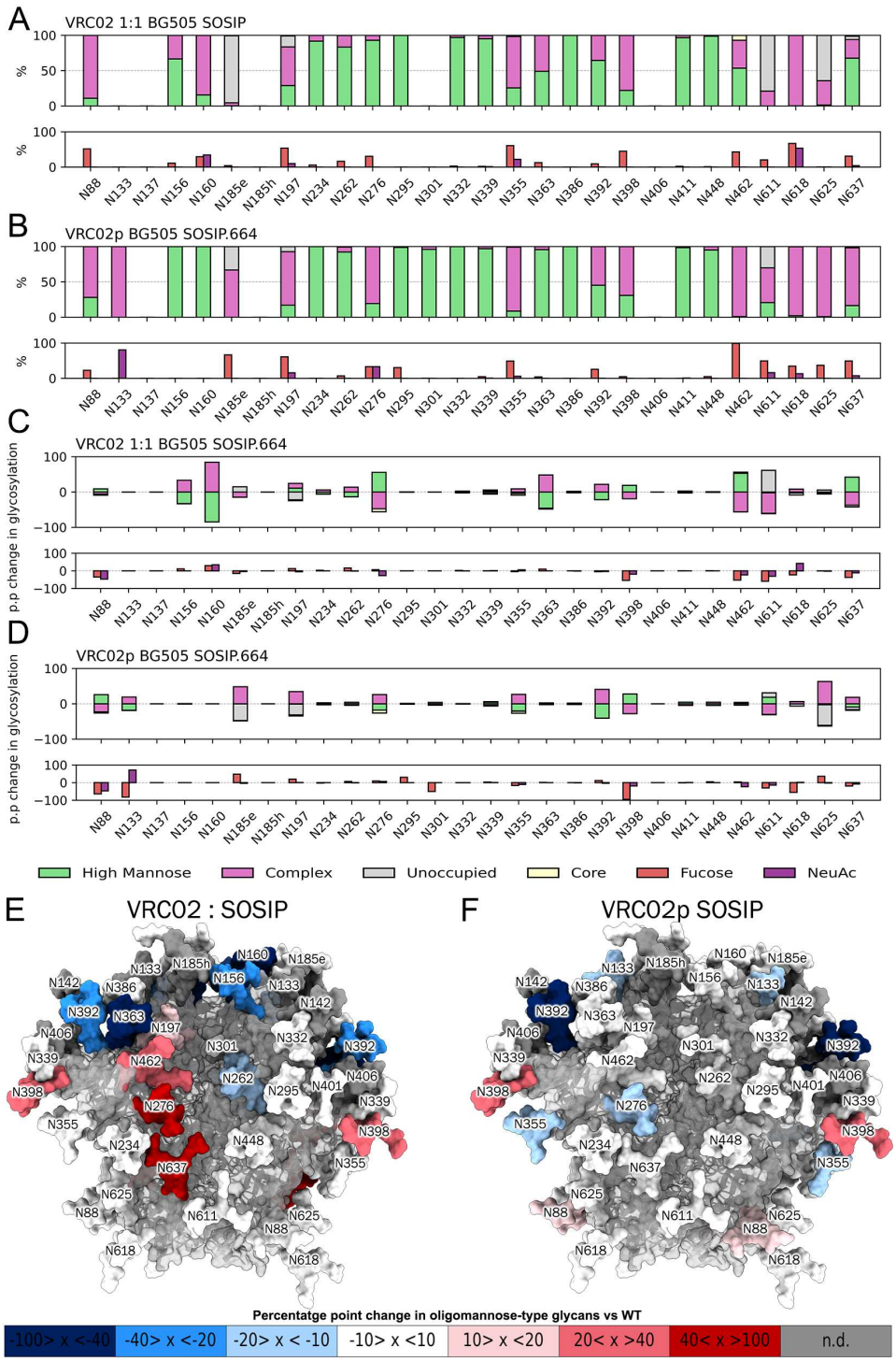


Figure 5.4: **VRC02 co-expression and VRC02-purified SOSIP glycan analysis**
(A-B) Site-specific glycan analysis of SOSIP co-expressed with VRC02 in a 1:1 ratio **(A)** WT SOSIP co-purified using VRC02 **(B)** Proportions of high mannose, complex-type, core and un-occupied PNGS at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. Both datasets correspond to mean site-specific glycopeptide data from three biological replicates.
(C-D) Arithmetic percentage point difference between average WT SOSIP glycosylation and VRC02 (1:1) **(C)**, and VRC02p **(D)**.
(E-F) VRC02 co-expression/purification percentage point difference vs WT mapped onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

Compared to VRC01p BG505 SOSIP, VRC02p SOSIP had a glycan profile with elevated levels of processing compared to WT BG505 SOSIP; however, 3/4 of the largest increases in complex-type glycosylation were at sites which reported a concomitant drop in glycan un-occupancy (N185e, N197, N625), whilst apex and mannose patch sites were largely unaffected. Conversely, VRC02 co-expression decreased high-mannose glycosylation at a number of sites such as N156 (-33 p.p.), N160 (-84 p.p.), N363 (-46 p.p.) and N392 (-22 p.p.), all of which are usually predominantly high-mannose. Similar to VRC01 co-expression however, there were increase in high-mannose glycans at N276, N462 and N637 (+56, +53, +42 p.p., respectively). This suggests that whilst these two related bnAbs likely bind overlapping epitopes, the way in which they do so and the subsequent conformations they induce differ. This, in turn, affects the glycosylation that we observe following co-expression. It is possible that VRC02 binding simulates CD4 binding which allows for greater enzymatic access to the glycan sites where we observe elevated glycan processing. VRC01 has instead been previously suggested to 'lock' Env into a closed conformation [334].

5.3.4 Co-expression of apex-targeting bnAbs

To investigate whether the effects observed for CD4bs bnAb co-transfection can be replicated with other bnAb targets, BG505 SOSIP trimers were co-expressed with the apex-targeting bnAbs PG16 and PG145. The glycan sites N156 and N160 form key contacts in both bnAb epitopes. Structural studies with PG16 and related antibodies using different constructs have reported a preference α2-6-linked sialylated hybrid-type glycans at a glycan site analogous to N156 [153, 236, 244]. It was not possible to determine the glycan content of the N156 site for the PGT145 co-expressed trimer (Figure 5.5B), but the PG16 N156 site was entirely oligomannose-type (Figure 5.5A); as was the N160 site for both trimers with the exception of some small under-occupancy in PGT145.

Given that most native-like trimers can be bound by PG16 with a high affinity, and molecular dynamics of these native-like trimers have shown the N156 and N160 sites to be resistant to processing, it is unlikely that its co-expression would therefore contribute to an increase in abundance of sialylated hybrid-type glycans at N156 [155, 69]. Mannose patch glycan sites were found to be intact and indifferent to the average WT glycosylation. Interestingly, oligomannose-type glycans were found to be enriched in both co-expressions at sites N197 (PG16: +26; PGT145 +27 p.p.) N276 (PG16: +44; PGT145 +54 p.p.); and to a lesser extent N355 and N637, similar to VRC01/02/sCD4 co-expressions (Figure 5.5C-D; E-F). In this instance, these changes were accounted for mainly by increases in $Man_5GlcNAc_2$ glycans. These changes, particularly to N276, are as significant as VRC01 co-expression, despite both bnAbs binding distal epitopes. In rationalizing these observations, the way that PG16 and PGT145 bind is important, as it appears that PGT16/PGT145 binding at the apex confers an environment that impairs, but not fully prevents, glycan processing of the N276 glycan site. This is different to VRC01 binding, as its binding likely provides a physical obstacle to glycan processing at this site, and is evident from the lack of changes to the N462 glycan site relative to the wild-type (Figure 5.5C-D; E-F).

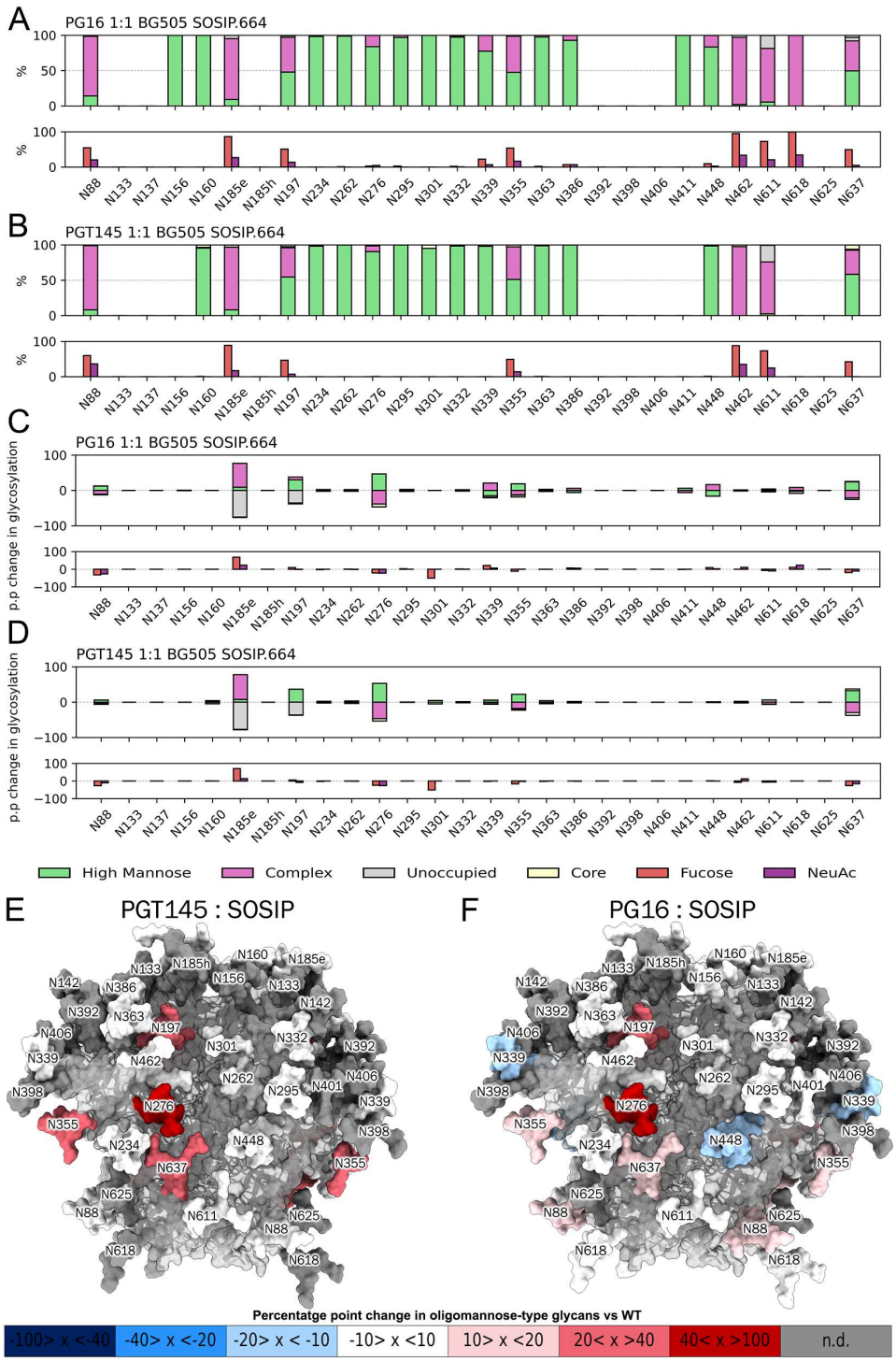


Figure 5.5: **Apex-targeting bnAb co-expression**

(A-C) SOSIP co-expressed with PG16 **(A)** and PGT145 **(B)** in a 1:1 (w/w) ratio. Proportions of high mannose, complex-type, core and un-occupied PNGS at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. Both datasets correspond to site-specific glycopeptide data from single bioligcal replicates

(C-D) Arithmetic percentage point difference between average WT SOSIP glycosylation and PG16 **(C)** and PGT145 **(D)**.

(E-F) Arithmetic percentage point change in oligomannose-type glycans between WT and the PG16/PGT145 co-expressions mapped onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

5.3.5 Co-expression of outer domain-targeting bnAbs

Three outer domain-targeting bnAbs were co-expressed with BG505 SOSIP trimers: PGT121, PGT135, and 2G12. The PGT121 bnAb recognizes the glycan sites N137, N156, N301 and N332 [239, 144, 337, 338]. It is the only bnAb in this panel that recognizes the N301 glycan, which is typically occupied by $Man_{5-8}GlcNAc_2$ structures on WT BG505 SOSIP trimers [69]. Accordingly, co-expression with PGT121 partially restricted N301 glycan processing, resulting in 25% of this PNGS containing $Man_9GlcNAc_2$ structures. There was no $Man_9GlcNAc_2$ detected at N301 in any other trimer. Sites N137 and N156 could not be determined (Figure 5.6A). Of note, the N185e site showed a marked increase in oligomannose and hybrid-type structures compared to wild-type BG505 SOSIP trimers (Figure 5.6D).

PGT135 co-expression had moderate effects on glycosylation at sites both within its epitope and further away. The PGT135 epitope contains the N295, N332, N386 and N392 glycan sites [239, 151, 265]. Additionally, it has also been reported that the presence of $Man_9GlcNAc_2$ at N392 may limit the binding of PGT135 [151, 265]. Perhaps most interestingly, there was a substantial drop in the oligomannose content at N332 by -30 percentage points (Figure 5.6C, F, I). The other sites within the epitope showed little deviation from wild-type glycosylation, including N392, which contained 69% $Man_9GlcNAc_2$, and 31% $Man_8GlcNAc_2$. Elsewhere on the trimer, the N355 glycan site showed a 23 percentage point increase in the proportion of un-occupied PNGS (Figure 5.6C, F, I). These changes came at the cost of oligomannose-type glycosylation which decreased by -16 percentage points.

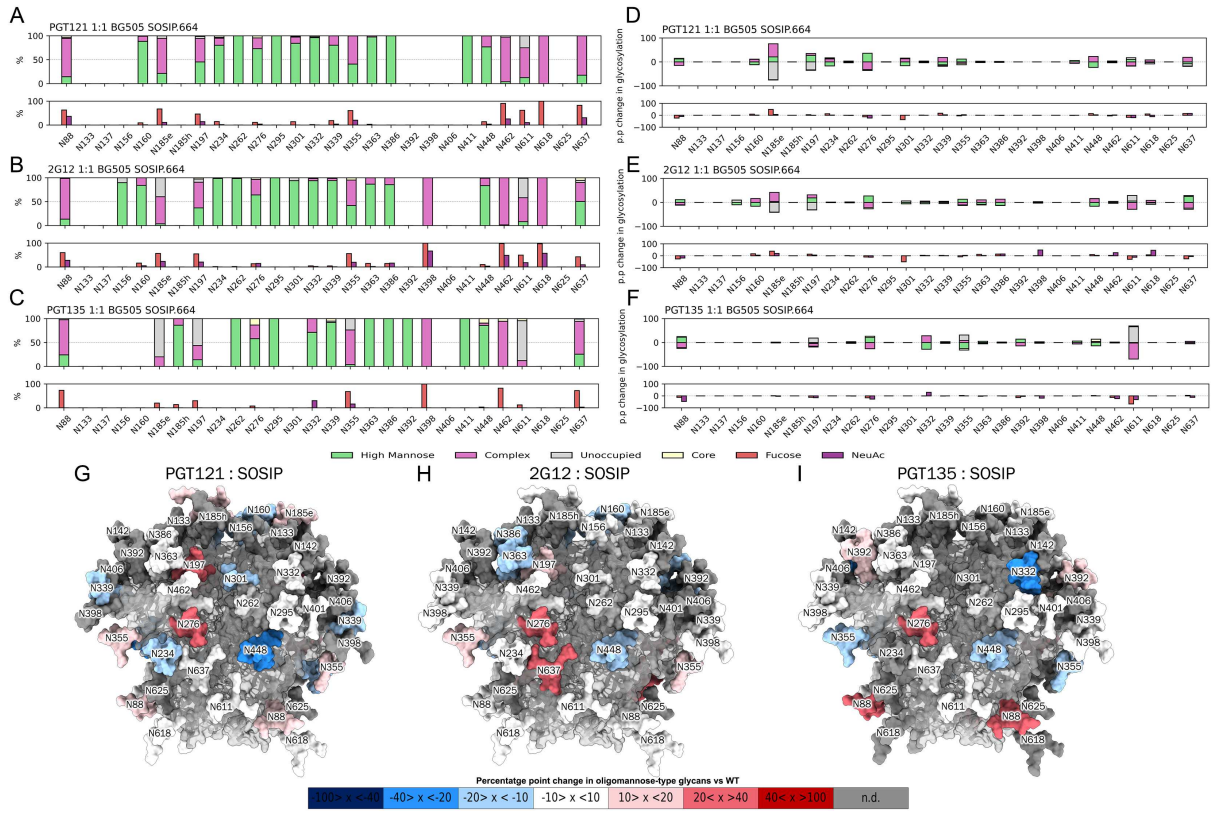


Figure 5.6: **Outer domain-targeting bnAb co-expression**
(A-C) SOSIP co-expressed with PGT121 **(A)**, PGT135 **(B)** and 2G12 **(C)** in a 1:1 (w/w) ratio. Proportions of high mannose, complex-type, core and un-occupied PNGS at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. All three datasets correspond to site-specific glycopeptide data from single biological replicates.
(D-F) Arithmetic percentage point difference between average WT SOSIP glycosylation and PGT121 **(D)**, PGT135 **(E)**, and 2G12 **(F)**.
(G-I) Arithmetic percentage point change in oligomannose-type glycans between WT and the PGT121 **(G)**, PGT135 **(H)**, and 2G12 **(I)** co-expressions mapped onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

The 2G12 epitope consists exclusively of N-linked glycans, comprising the N295, N332, N339 and N392 sites [262, 329, 140, 145, 146, 339, 340]. These sites are typically occupied by oligomannose-type glycans. Co-expression of SOSIP trimers with 2G12 did not significantly change the processing state for sites N295, N332, and N339 (Figure 5.6B, E, H). Data for N392 could not be obtained. Glycan sites N363, N386, and N448 showed small increases in glycan processing (Figure 5.6E, H). The N386 glycan site has been implicated in stabilizing the 2G12 epitope and forms part of the glycan network surrounding it. The N363 site lies

adjacent to N386 whilst N448 lies below N295. Apex sites N156 and N160 also decreased in their oligomannose content slightly (Figure 5.6E, H).

5.3.6 Co-expression of gp120-gp41 interface-targeting bnAbs

PGT151 is a bnAb that has a preference for tri- and tetra-antennary (HexNAc(5-6)(+/-F)(x)) structures at the glycan sites N611 and N637 [79, 98, 66]. Co-expression with PGT151 marginally increased the glycan occupancy of both of these sites, and an enhancement of tri- and tetra-antennary glycans was observed as well as levels of fucosylation and sialylation (Figure 5.7A).

Structural studies have suggested that the PGT151 light chain CDR loops are also in close proximity to glycan sites N262 and N448 from one protomer, and its light chain framework region with N276 from another protomer [341]. The N262 site could not be determined.

N386 in this co-expression had the lowest amount of oligomannose-type glycans in the bnAb panel (-30 p.p. vs WT; Figure 5.7C). Apex sites N156 and N160 were also subjects of significant change in their glycan content. N156 glycan occupancy decreased by -57 percentage points at the cost of oligomannose-type structures whilst N160 complex-type glycosylation increased by +39 percentage points at the cost of oligomannose-type glycosylation (Figure 5.7B).

PGT151 has a long HCDR3 which it uses to bind an interprotomer cavity at the gp120-gp41 interface [98]. This feature means that PGT151 can only bind trimers, and it is subsequently used for purifying trimers in the same way as bnAbs 2G12 and PGT145 [67]. However, if co-expression of trimers and bnAbs form complexes immediately following synthesis, it is possible that the effects of this binding in turn affects the glycosylation we observe in this co-expression. In this case, interprotomer binding may disrupt env quaternary structure at the apex and enable the enhanced processing that we observe.

To better determine this effect, BG505 SOSIP trimers purified by PGT151 were examined. Compared to PGT151 co-expression, the only large changes in glycosylation occured at gly-

can sites N611 and N625. N625 overwhelmingly increased in oligomannose-type glycans by +92 p.p. from a previous 59% under-occupancy in WT (Figure 5.7D) The change at N611 was most interesting as it showed an increase in underoccupied glycans at this site relative to WT SOSIP (Figure 5.7D, F); despite the aforementioned preference for complex-type glycans at this site. This suggests that whilst PGT151 may observe a preference for glycans at this location, it is not necessary for binding. Alternatively, its binding is dictated more by its binding to N637 than N611, whose glycosylation varies little in PGT151p BG505 SOSIP.

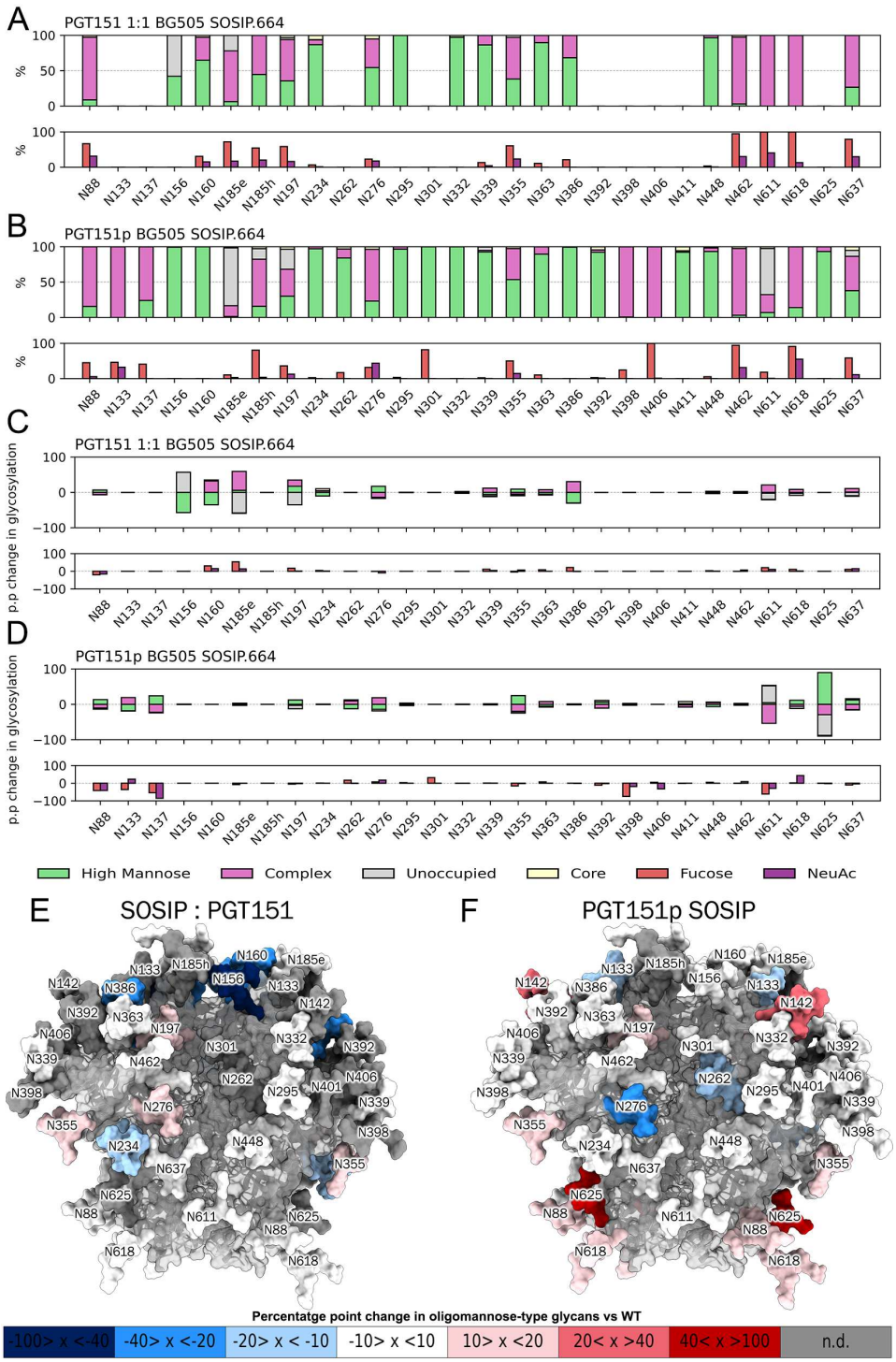


Figure 5.7: **gp120-gp41 interface-targeting bnAb co-expression**
(A-B) Site-specific glycan analysis of SOSIP co-expressed with PGT151 in a 1:1 ratio **(A)** WT SOSIP co-purified using PGT151 **(B)** Proportions of high mannose, complex-type, core and un-occupied PNGS at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. Both datasets correspond to site-specific glycopeptide data from single biological replicates.
(C-D) Arithmetic percentage point difference between average WT SOSIP glycosylation and PGT151 (1:1) **(C)**, and PGT151p **(D)**
(E-F) PGT151 co-expression/purification percentage point difference vs WT mapped onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

5.4 Discussion

The ability to selectively modify the glycan shield of HIV-1 Env spikes using antibodies that target it poses a powerful tool to alter the epitopes and immunogenicity of Env whilst maintaining the overall integrity of the glycans outside of the area being targeted by the co-expressed binding protein [120, 225, 325, 342]. As it currently stands, research in this area has made use of enzyme co-transfection, glycosylation inhibitors, or engineered cell lines; all of which impart global alterations in the glycosylation processing of an env trimer which may lead to improved binding of some bnAbs but often at the cost of others [225, 233, 143, 235]. The ability to make changes to the glycan profile of env trimers without abrogating sensitivity to bnAbs, or conversely increasing sensitivity of a trimer to a particular bnAb of interest may therefore be of benefit [225].

The dichotomy between VRC01 and VRC02 co-expression demonstrated that highly related bnAbs can cause significantly different changes in glycosylation of the same spike protein. Despite acting on the same epitope it is evident that VRC01 co-expression had a mostly restricting effect on SOSIP glycosylation whilst VRC02 caused widespread increases in SOSIP glycan processing outside of the CD4bs. When compared to the co-expression of sCD4, there were fewer widespread increases in the processing of several sites and simultaneously we also observed sharp increases in oligomannose-type glycosylation at sites N276 and N637. These molecules all have similar epitopes and yet imparted unique, albeit sometimes overlapping, changes in glycosylation when co-expressed with the similar molecules. The reasons for this can be ascribed to what impacts the binding of these molecules may have. In the case of VRC01 binding the widespread increases in oligomannose-type structures likely mean that the trimer is being maintained in a closed state, whilst the opposite is true for VRC02 and sCD4 co-expression.

Co-transfecting differing proportions of VRC01 showed overall consistent effects across the different proportions. The 1:1 co-expression on average caused the largest shift towards more

oligomannose-type structures particularly at sites N276, N462, and N637 (Figure 5.2). The changes to N276 and N462 were consistent in the 2:1 and the 1:2 co-expression but N637 glycosylation was inconsistent amongst the three subsets. Additionally, whilst these effects were maintained between two biological triplicate datasets, most importantly the ability to impart a dominant $Man_9GlcNAc_2$ glycan at N276, there was still some difference between the two (Figures 5.2 and 5.3). This highlights the inherent variance in glycosylation when producing these proteins.

Other findings with the co-expression of some bnAbs with SOSIP trimers, particularly PGT145 and 2G12, demonstrated relatively few changes from the WT trimers. This was an expected finding due to their current use in the production and purification of recombinant HIV-1 spike proteins as they are known to bind the most restricted N-glycan composition in $Man_9GlcNAc_2$. Specifically, their use in selectively binding well-folded HIV-1 spike trimers is in part due to the fact that they do not significantly alter the structure and conformation of these proteins, as confirmed by X-ray/Cryo-EM studies [339, 335, 108, 147]. On the other hand, another bnAb which has been used for the same purpose, PGT151, was shown to cause some changes in glycosylation at many sites on co-expressed SOSIP trimers. By understanding the mechanism of action of PGT151, we see that the way in which it engages Env trimers in a purification setting is likely the same interaction that drives increased glycan processing upon co-expression.

PGT151 binds a trimer-specific quaternary epitope in between the protomers. This allows a user to isolate trimers on this basis amongst a litany of potentially mis-folded (and therefore likely bearing more processed glycans) HIV-1 Env proteins. If the same mechanism of binding is applied to SOSIP trimers co-translationally, in other words as its being glycosylated, then it is possible to imagine that ER or Golgi resident glycan trimming enzymes will likely have better access to certain glycan sites in a spike protein that has an antibody bound in between its protomers. In the case of PGT151 co-expression, the SOSIP trimers are likely not assembling as tightly as they would in the absence of the antibody or, as has

been suggested for similar bnAbs, the antibody itself may be contributing to the prolonged disassembly of these trimers [343]. This is perhaps the reason why we observe high levels of processing in PGT151 co-expression but would otherwise not if purifying trimers using PGT151.

Another point of interest from our data is that in general, bnAb-SOSIP co-expression showed greater glycan occupancy at certain sites. This is particularly true for glycan sites N185e, N197 and N611. This observation was not limited to co-expressed bnAbs either as the Trastuzumab co-expression also showed increased glycan occupancy at these sites. On the other hand in some co-expressions, namely those of sCD4, PGT135, and PGT151, showed increases in the proportion of unoccupied PNGS. It isn't possible to attribute this observation to any particular effect however as many of these glycan sites can significantly vary in their occupancy without co-expression or antibody purification. The crux of this project was to prove the ability to impart site-specific changes in a select few glycan sites and this was successfully and consistently achieved using HIV-1 Env with VRC01 co-expression. This raises the possibility of the further study and use of a novel glycan engineering tool not only in HIV-1 vaccine design but with other glycoproteins and their binding partners.

6 Translational control of immunogen glycosylation

Whilst the previous chapter covered methods in which changes to site-specific glycosylation could be imparted whilst maintaining the integrity of distal glycan sites, the glycosylation of soluble Env immunogens is still in part dissimilar to that found on viral-expressed Env. The glycosylation of viral-derived and artificial soluble immunogens are characterised by areas of conserved glycosylation between the two, namely the mannose patch and the spike apex which frequently exhibit nascent oligomannose-type glycans in pre-fusion state trimers. Differences between the two frequently arise when observing glycosylation in sites outside of these epitope clusters; where in analysed viral material these sites are found to display more processed, complex-type glycans. In contrast, these sites on soluble immunogens, for example N197 or N276, display a mixture of oligomannose-type and complex-type glycans. Another observed difference is the presence of glycan underoccupancy at certain glycan sequons in soluble immunogens whereas in viral-derived material near 100% occupancy is observed at every site. In order to understand some of the factors that drive these discrepancies, this chapter aims to establish the effects of codon usage between viral-derived and artificial constructs in generating their respective glycan profiles, and if the use of rarer codons in soluble immunogens could represent a strategy to *'bridge the gap'* in glycosylation between these observed glycan profiles.

6.1 Contributions

Chapter 6 involves a comparison of Envs with optimized and unoptimized, native *env* sequences from four different clades that were produced and recombinantly purified using the SOSIP v4.1 platform. This work in this chapter was undertaken in collaboration with Dr. Joel Allen who collaborated on the design of these constructs which are detailed in Table 6.1.

SOSIP/SLOSIP Analysed	No. Biological replicates
BG505 SOSIP.v4	3
BG505 SLOSIP.v4	3
B41 SOSIP.v4	3
B41 SLOSIP.v4	3
ZM197M SOSIP.v4	3
ZM197M SLOSIP.v4	3
X1193 SOSIP.v4	3
X1193 SLOSIP.v4	3

Table 6.1: **Table of expressed proteins and the number of biological replicates expressed**

6.2 Introduction

The strategies that currently govern modern HIV-1 vaccine design strategies to induce bnAbs are based on the principle of antigenic mimicry. Vaccine candidates are designed to be as close to the viral envelope spike in structure and glycosylation as possible in order to effectively target conserved regions of the spike in face of the significant genetic diversity posed by HIV-1 [344].

What distinguishes Env from other glycoproteins is its extensive post-translational modifications; not only is it heavily decorated with N-linked glycans but it also has an intricate network of disulphide bonds [345]. The shielding effect conferred from the spatial density of N-linked glycans on the surface of Env distinguish them from endogenous N-linked glycans in that they contain high levels of oligomannose-type glycans that arise from enzymatic steric inhibition [123]. Combined with the immunogenic viral protein surface, these abnormal glycan signatures contribute to the elicitation of an ever expanding repertoire of bnAbs to act upon the virus [57]. Additionally, the success of bnAbs in macaque passive transfer studies have catalyzed their use as a passive therapy in addition to the goal of vaccine-induced bnAbs [346]. Design attempts in generating stable Env constructs to elucidate these antibodies have overcome significant engineering hurdles in generating these antigenic mimics of the viral spike.

In order to create native-like Env trimers, several mutations were introduced to create proteins with a similar antigenic profile to Env [112, 108, 347, 348, 109, 114, 306]. The range of these mutations and modifications have been reviewed extensively and these trimers are under phase I trials [108]. Whilst initial success and development of several following iterations of soluble trimers was achieved with the clade A BG505 strain this strategy has also been applied to other Env strains [349, 254].

The reliance of bnAb binding on N-linked glycans necessitated the site-specific characterization of the BG505 SOSIP.664 trimer [69, 155]. These analyses revealed several regions of the glycan shield that rely on proper folding and maturation of Env. These include the IMP, which consists of several closely packed glycan sites that all display oligomannose-type glycans and constitute targets for many bnAb strategies, usually revolving around the N332 supersite [253]. Further analyses comparing BG505 gp120 to trimeric BG505 SOSIP.664 revealed another region of oligomannose-type glycans which are only present in the trimeric form. The TAMP is focused around the apex of the trimer and in particular glycan sites N156 and N160 [70]. As many bnAbs that target the apex of Env have a dependence on oligomannose-type glycosylation at N160 it is important to ensure that this is conserved on vaccine candidates aimed at eliciting apex-binding bnAbs. The remaining glycan regions are more processed compared to the IMP and TAMP and include 4 conserved glycans on gp41. These Env glycans form a network that is resistant to change upon the deletion of a single PNGS, with studies demonstrating that only sites proximal to the deletion are impacted upon by the deletion [69, 257].

Following the initial characterizations of SOSIP proteins, several studies were undertaken to observe the glycosylation of a strain matched viral derived trimer against the SOSIP [261, 104]. Both studies highlighted similarities and differences between the glycan shields of viral Env and SOSIP trimers. Notably, the IMP and TAMP glycan sites including N156, N160 and N332 are conserved between the two. Whilst these two key bnAb epitopes show conserved glycosylation there are several other regions which show distinct glycosylation on

viral Env. The largest difference in glycosylation localized to sites which present a mixture of under processed oligomannose-type and processed complex-type glycans on SOSIP trimers such as N133, N197 and N276. On viral env these glycans are more processed. Furthermore, the processing state of complex-type glycans is different in that the amount of branching in the glycan structures is elevated on viral-derived Env when compared to SOSIP with tri- and tetra-antennary glycans common across gp120 and gp41. A possible explanation for this is that the membrane tethered viral Env is closer in proximity to the glycan processing enzymes that reside in the ER and Golgi which results in elevated processing at sites lacking steric protection from adjacent glycan sites such as N197. This idea was strengthened by a comparative glycan analysis between sequence matched recombinant SOSIPs and membrane-bound Env that showed mixed glycan sites to possess elevated processing and branching [350]. These observations could rationalize why certain bnAbs which bind complex-type glycan epitopes show preference for tri- and tetra-antennary glycans on glycan microarray studies [233].

An observation noted in [104] was that viral Env PNGSs were occupied to a far greater degree when compared to those on SOSIP. Further analysis of SOSIP trimers revealed regions of underoccupancy at sites N611 on gp41 as well as in the V1/V2 apex of gp120. These glycan holes on SOSIP trimers are thought to drive immunodominance to non-nAb epitopes when used in vaccinations whereby the glycan shield on the vaccine is compromised and the antibodies raised from the vaccine are unable to bind the autologous virus. This is exemplified in macaque immunizations whereby an EM-based epitope mapping technique revealed that macaque elicited antibodies bound to a region that should be occupied by the N611 glycan. These antibodies were incapable of neutralizing virus autologous to the same site [210]. These observations regarding changing compositions and differences in glycosylation necessitate glycan engineering techniques that could be used to modulate glycosylation to either minimize undesirable characteristics such as glycan holes or enhance characteristics such as increase the number of processed antennae displayed on complex-type glycans as a

means of improving glycan binding.

As glycan deletions and insertions only appear to affect PNGSs proximal to the modified site we sought to investigate ways to modulate the glycosylation of soluble Env constructs independently of their amino acid sequence. An advantage to this approach is that changes in amino acid composition could perturb Env quaternary structure and unintentionally expose undesirable antibody epitopes. Furthermore complications regarding full-length immunogens and their delivery and stability currently suggest that it is advantageous to remain in a soluble format. It should however be noted that RNA-based vaccines are a promising alternative to delivering full-length immunogens and that nanoparticle based immunogens are feasible and are associated with several desirable characteristics such as better lymphatic uptake [209, 351].

In this chapter, we sought to explore differences between SOSIP and viral-derived Env with which glycosylation could potentially be affected. In this regard, production of SOSIPs relies on plasmid codon optimization in order to maximise expression yields. Conversely, the virus frequently displays a bias towards rarer triplet codons encoding Env, i.e. the abundance of the tRNA for a particular amino acid codon is less common therefore reducing the rate of translation [113]. To this end, we designed SOSIP constructs from 4 distinct clades that still possess the viral *env* DNA sequence (SLOSIPs) as well as their codon optimized counterparts (SOSIPs). The reason for this study is the potential hypothesis that a slower rate of translation could in turn, provide a longer time for glycan attachment at sites that display under-occupancy or alternatively a longer time for mixed glycan sites that are more complex-type in viral Env to undergo processing, thus bridging the gap between immunogen and viral glycosylation. These strategies could be employed in immunogen design if it yields desirable features that are enhanced by this approach.

6.3 Results

6.3.1 Site-specific glycan analysis of SOSIP/SLOSIP trimers.

SOSIP and SLOSIP trimers expressed in biological triplciate batches of HEK293F cells and were purified by nickel-affinity chromatography and size-exclusion chromatography (SEC). peptides digested using α-lytic protease, trypsin, and chymotrypsin from each biological replicate were combined to make an average dataset for each SOSIP/SLOSIP trimer. The findings of each strain are hereby compared with one another.

6.3.1.1 BG505 SOSIP vs SLOSIP

Analysis of BG505 SOSIP and SLOSIP trimers revealed unexpected changes in glycosylation across certain sites. Changes in glycosylation between the two triplicate datasets were generally associated with increases in complex-type glycosylation, the largest of which was observed at sites N133, N276, N301 and N611 (Figure 6.1C). The most unexpected of these changes were the increases in complex-type glycosylation that occured at site N301, which has previously been characterized as oligomannose type in 293F, CHO and viral BG505 Env analysis [104]; in BG505 SLOSIP, this site was 51% high-mannose, representing a 41 p.p. increase in complex-type glycosylation from BG505 SOSIP. Out of the three aforementioned glycan sites which show discrepancy between immunogens and viral Env, N133 displayed the largest change with a 60 p.p. drop in high-mannose glycoforms in place of complex-type glycans. N197 showed a small increase in oligomannose-type glycosylation at the cost of complex-type glycans (20 p.p.) whereas N276 showed a larger increase in complex-type glycosylation with a concomitant decrease in oligomannose-type glycosylation (28 p.p.).

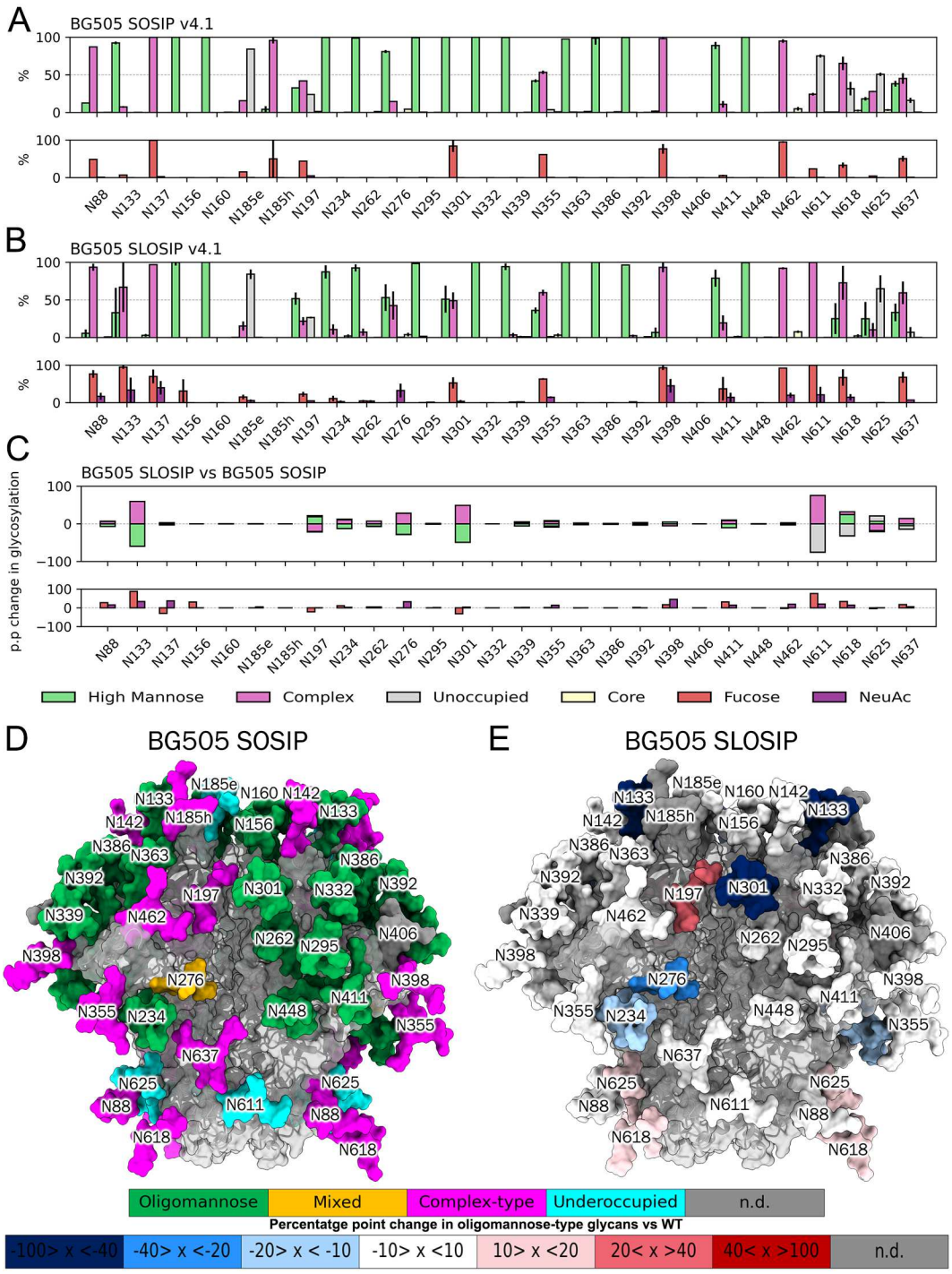


Figure 6.1: **BG505 SOSIP and SLOSIP glycan analysis**
(A-B) Site-specific glycan analysis of BG505 SOSIP v4.1 **(A)** and BG505 SLOSIP v4.1 **(B)** Proportions of high mannose, complex-type, core and unoccupied glycans at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. Error bars represent SEM values calculated from biological triplicate glycopeptide datasets.
(C) Arithmetic percentage point difference between average BG505 SLOSIP glycosylation and BG505 SOSIP.
(D-E) Average BG505 SOSIP v4.1 glycosylation **(D)** and percentage point difference in oligomannose-type glycosylation of BG505 SLOSIP vs SOSIP **(E)** mapped onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). In **(D)**, green sites represent >80% oligomannose-type glycosylation, yellow: 79-40%, magenta: <40%, cyan: >50% under-occupancy. In **(E)**, Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

Changes in glycan under-occupancy appeared to localize to the gp41 glycan sites. On average, in BG505 SOSIP, the gp41 sites N611 and N625 were predominantly under-occupied (75% and 51% respectively, Figure 6.1A) whilst N618 and N637 were 32% and 16% under-occupied, respectively. The only other sites that exhibited significant glycan under-occupancy were N185e (84%), and N197 (24%). In BG505 SLOSIP, the only sites which exhibited notable changes in glycan under-occupancy were those located on gp41. N611 was the most significant of these with a 75 p.p. decrease in under-occupancy with a concomitant rise in complex-type glycosylation (Figure 6.1C). N618 showed a smaller but also significant decrease in under-occupancy with a 32 p.p. decrease but this was reconstituted into a 25 p.p. increase in high-mannose glycoforms and a 7 p.p. increase in complex-type glycosylation. At N637 decreases in underoccupancy (-9 p.p.) and high-mannose (-5 p.p.) accounted for a 14 p.p. increase in complex-type glycosylation between BG505 SOSIP and SLOSIP. Interestingly, the N625 glycan site reported a +14 p.p. increase in under-occupancy in BG505 SLOSIP making it predominantly under-occupied. Elsewhere, the glycan underoccupancy of gp120 glycan sites N185e (0 p.p.) and N197 (+3 p.p.) appeared to be unaffected overall.

6.3.1.2 B41 SOSIP vs SLOSIP

Analysis of B41 SOSIP and SLOSIP trimers revealed more widespread changes in glycosylation than what was observed in BG505 SOSIP/SLOSIP. Out of the 29 glycan sites on this trimer, 17 sites reported a >10 p.p. change in high-mannose or complex-type glycosylation (Figure 6.2C). The majority of these changes were associated with increases in complex-type glycosylation with a few notable exceptions. The largest increases occurred at N138 (+31 p.p.), N145 (+21 p.p.), N187 (+22 p.p.), N276 (+21 p.p.), N362 (+38 p.p.), N392 (+43 p.p.), N611 (+35 p.p.), N625 (+35 p.p.) and N637 (+52 p.p.), mostly at the host of high-mannose glycoforms or the proportions of under-occupied PNGSs (Figure 6.2). The largest increases in high-mannose glycoforms occurred at sites N187 (+32 p.p.), N197 (+38 p.p.), N339 (+73 p.p.), N386 (+26 p.p.), N463 (+34 p.p.), and N625 (+35 p.p.). Once again the three sites which display elevated high-mannose glycosylation, in this case with N133 in BG505 being replaced with N138 in B41, showed decreases in N138 and N276 whereas N197 reported an increase. Some of the most interesting changes occurred at N339 and N392 which are located close to one another on the trimer (Figure 6.2D, E). In B41 SOSIP N339 was predominantly complex-type and N392 was high-mannose but in B41 SLOSIP these sites were the opposite.

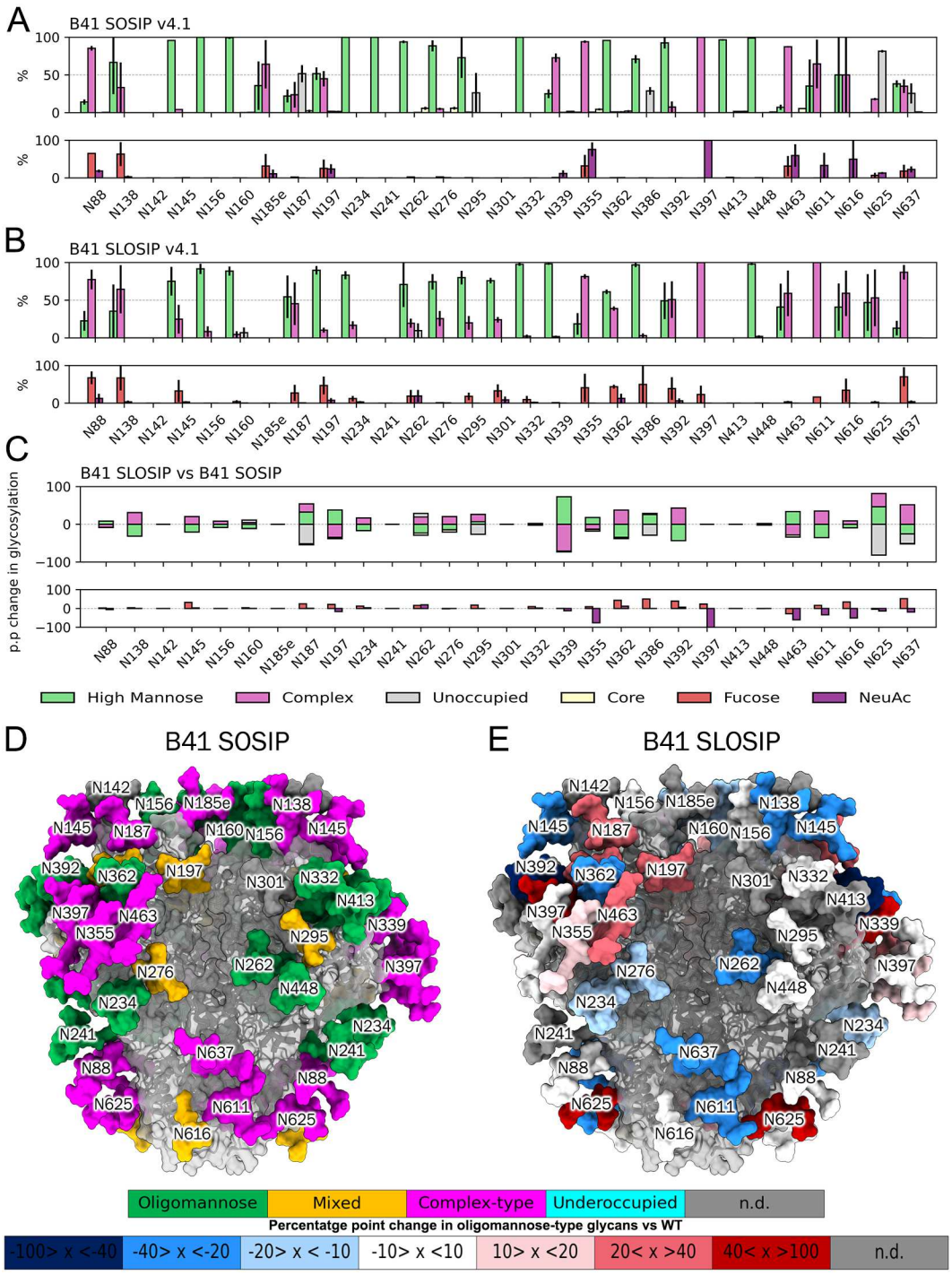


Figure 6.2: **B41 SOSIP and SLOSIP glycan analysis**

(A-B) Site-specific glycan analysis of B41 SOSIP v4.1 **(A)** and B41 SLOSIP v4.1 **(B)** Proportions of high mannose, complex-type, core and unoccupied glycans at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. Error bars represent SEM values calculated from biological triplicate glycopeptide datasets.

(C) Arithmetic percentage point difference between average B41 SLOSIP glycosylation and B41 SOSIP.

(D-E) Average BG505 SOSIP v4.1 glycosylation **(D)** and percentage point difference in oligomannose-type glycosylation of B41 SLOSIP vs SOSIP **(E)** mapped onto a crystal structure of BG505 SOSIP.664 (PDB ID: 4NCO [333]). In **(D)**, green sites represent >80% oligomannose-type glycosylation, yellow: 79-40%, magenta: <40%, cyan: >50% under-occupancy. In **(E)**, Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

Changes in glycan under-occupancy in B41 SOSIP/SLOSIP were not just limited to the gp41 glycan sites as it was in BG505. With the exception of N160 (+7 p.p.) and N262 (+9 p.p.), glycan under-occupancy was consistently lower in B41 SLOSIP than B41 SOSIP. The largest changes occured at N187 (-52 p.p.), N295 (-26 p.p.), N386 (-29 p.p.), N625 (-81 p.p.) and N637 (-26 p.p.) (Figure 6.2C). Excluding N386, all of these sites reconstituted the drop in glycan under-occupancy with increases in both high-mannose and complex-type glycans. Glycan under-occupancy in B41 SLOSIP was significantly lower than in B41 SOSIP and in this trimer we report only the N160 (7%) and N262 (10%) glycan sites as showing any glycan under-occupancy (Figure 6.2B).

6.3.1.3 ZM197M SOSIP vs SLOSIP

Analysis of ZM197M SOSIP and SLOSIP trimers also showed widespread changes in glycosylation. Out of the 31 PNGSs on these trimers, 23/31 sites showed a >30 p.p. change in either high-mannose or complex-type glycosylation (Figure 6.3C). 11 of these sites showed increases in complex-type glycosylation and 11 in high-mannose glycosylation, and one of these sites showed an increase in glycan under-occupancy. The largest increases in complex-type glycosylation were at sites N133d (+33 p.p.), N136 (+21 p.p.), N141 (+31 p.p.), N147 (+67 p.p.), N301 (+29 p.p.), and N411 (+35 p.p.). The largest increase in high-mannose glycosylation were at sites N130 (+39 p.p.), N186h (+29 p.p.), N197 (+35 p.p.), N276 (+54 p.p.), N289 (+23 p.p.), N611 (+26 p.p.) and N637 (+54 p.p.). In this case, the three sites which showed discrepancies between viral and immunogen BG505 Env (In this case N130 instead of N133), all showed increases in high-mannose glycosylation at the cost of complex type glycosylation (N130: +39, N197: +35, N276: +54 p.p.; Figure 6.3C, E). Interestingly, the adjacent glycan sites N133d - N147 which showed similar levels of mixed glycosylation in ZM197M SOSIP showed increases in complex-type glycosylation in place of high-mannose (N133d: +33, N136: +21, N141: +31, N147: +67 p.p.; Figure 6.3A-C).

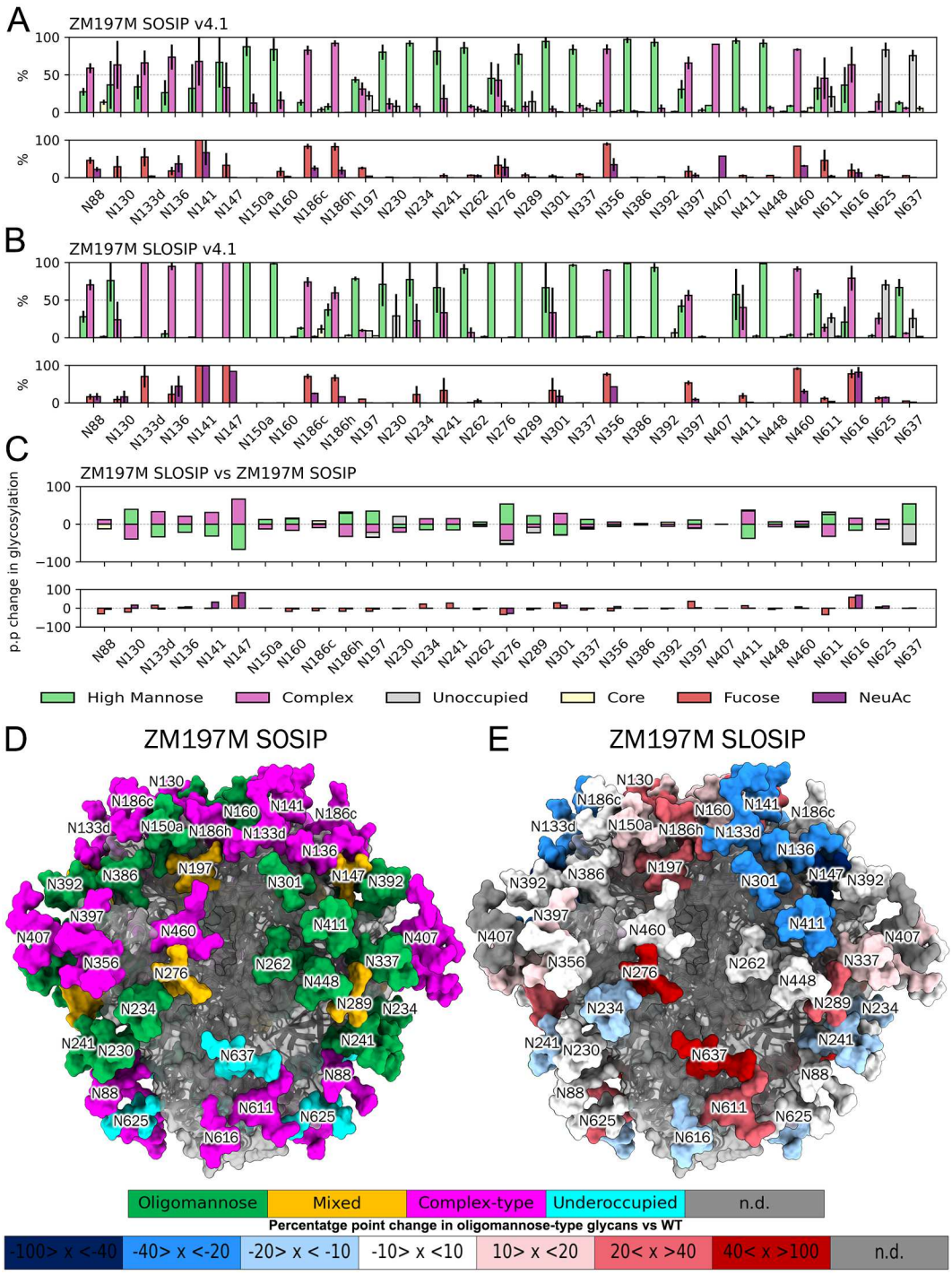


Figure 6.3: **ZM197M SOSIP and SLOSIP glycan analysis**

(A-B) Site-specific glycan analysis of ZM197M SOSIP v4.1 **(A)** and ZM197M SLOSIP v4.1 **(B)** Proportions of high mannose, complex-type, core and unoccupied glycans at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. Error bars represent SEM values calculated from biological triplicate glycopeptide datasets.

(C) Arithmetic percentage point difference between average ZM197M SLOSIP glycosylation and ZM197M SOSIP.

(D-E) Average ZM197M SOSIP v4.1 glycosylation **(D)** and percentage point difference in oligomannose-type glycosylation of ZM197M SLOSIP vs SOSIP **(E)** mapped onto a crystal structure of ZM197M SOSIP.664 (PDB ID: 4NCO [333]). In **(D)**, green sites represent >80% oligomannose-type glycosylation, yellow: 79-40%, magenta: <40%, cyan: >50% under-occupancy. In **(E)**, Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

In ZM197M SOSIP, only the N625 and N637 glycan sites were predominantly under-occupied (N625: 83%, N637: 76%; Figure 6.3A). Elsewhere on ZM197M SOSIP there was some underoccupancy at other sites, the largest of which were N197 (22%), N289 (14%) and N611 (21%). Between ZM197M SOSIP and SLOSIP, the largest changes in glycan under-occupancy localized to the sites N197 (-13 p.p.), N230 (-21 p.p.), N289 (-14 p.p.), N625 (-13 p.p.), and N637 (-50 p.p.; Figure 6.3C). With the exception of N625, decreases in under-occupancy were coupled with increases in high-mannose, and in many sites were also coupled with decreases in complex-type glycosylation as well. There was also a noticeable increase in under-occupancy at the N230 glycan site.

6.3.1.4 X1193 SOSIP vs SLOSIP

Analysis of X1193 SOSIP and SLOSIP revealed fewer changes than B41 or ZM197M SOSIP/SLOSIP. In general, changes between X1193 SOSIP and SLOSIP were associated with increases in high-mannose glycosylation. Of the 26 sites that we were able to compare between X1193 SOSIP and SLOSIP (Unable to acquire data on N137 and N406 in SOSIP/SLOSIP; N141, N144d in X1193 SOSIP), the largest increases in complex-type glycosylation occurred at the glycan sites N332 (-34 p.p.), N611 (-34 p.p.), and N616 (-40 p.p.) (Figure 6.4C, E). Of these three increases, two (N332 and N616) was in the place of high-mannose glycosylation whilst the change at N611 came in the place of glycan under-occupancy. In contrast, the largest changes in high-mannose glycosylation localized to the glycan sites N160 (+65 p.p.), N234 (+65 p.p.), N262 (+20 p.p.), N276 (+46 p.p.), N301 (+62 p.p.), N392 (+38 p.p.), N398 (+50 p.p.), N462 (+80 p.p.) (Figure 6.4C). For the most part these changes were accounted for by decreases in complex-type glycosylation.

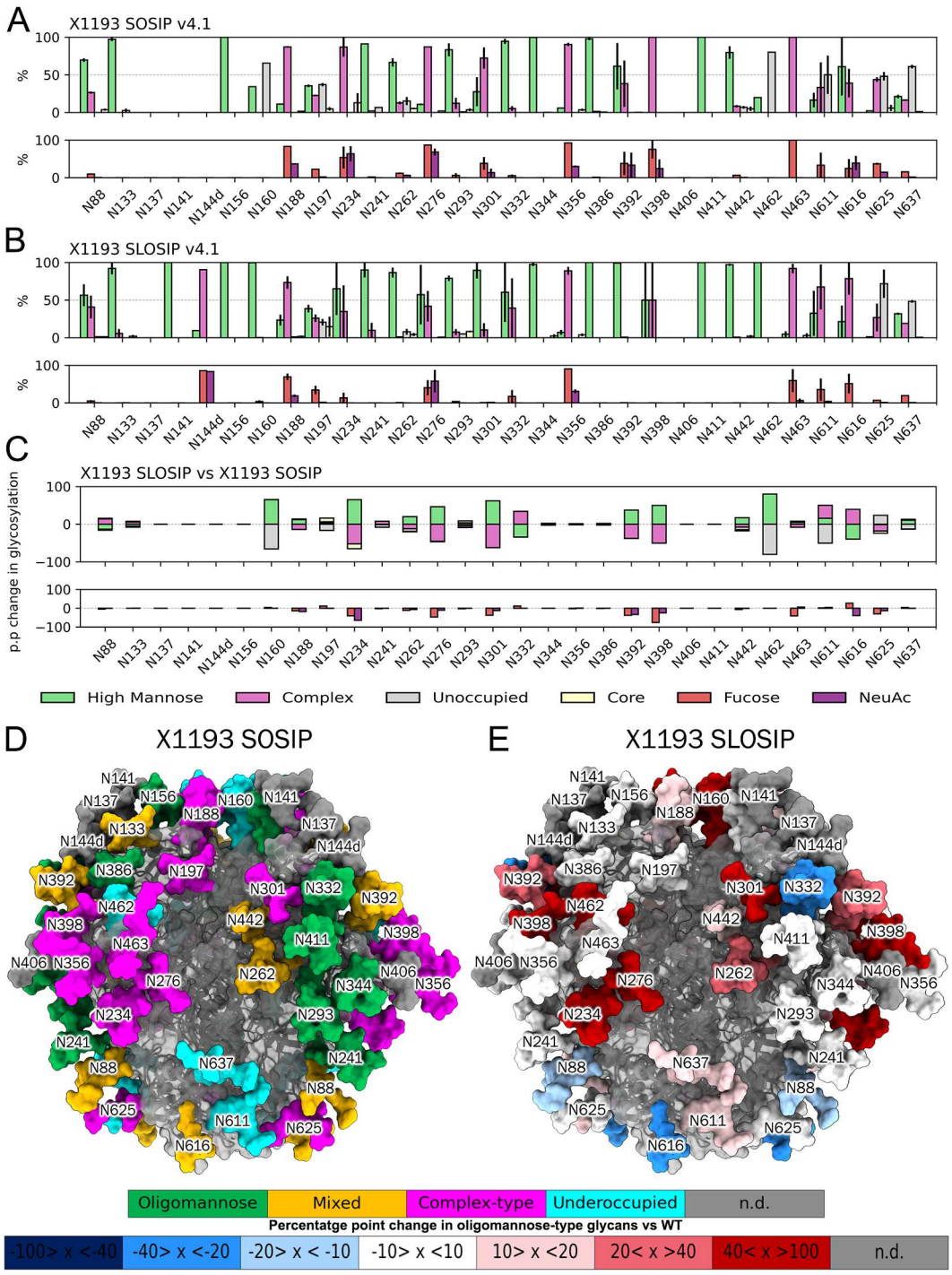


Figure 6.4: **X1193 SOSIP and SLOSIP glycan analysis**

(A-B) Site-specific glycan analysis of X1193 SOSIP v4.1 **(A)** and X1193 SLOSIP v4.1 **(B)** Proportions of high mannose, complex-type, core and unoccupied glycans at individual sites in upper chart and individual site levels of fucosylation and NeuAc glycans in lower chart. Error bars represent SEM values calculated from biological triplicate glycopeptide datasets.

(C) Arithmetic percentage point difference between average X1193 SLOSIP glycosylation and X11935 SOSIP.

(D-E) Average X1193 SOSIP v4.1 glycosylation **(D)** and percentage point difference in oligomannose-type glycosylation of X1193 SLOSIP vs SOSIP **(E)** mapped onto a crystal structure of X1193 SOSIP.664 (PDB ID: 4NCO [333]). In **(D)**, green sites represent >80% oligomannose-type glycosylation, yellow: 79-40%, magenta: <40%, cyan: >50% under-occupancy. In **(E)**, Increases in oligomannose are in red whilst decreases are in blue. Glycans which could not be determined are in grey.

The majority of changes in glycan under-occupancy between X1193 SOSIP and SLOSIP associated with decreases in glycan underoccupancy. The only increase in glycan under-occupancy above 10 p.p. was at N625 (+25; Figure 6.4C). Conversely, decreases of the same magnitude in under-occupancy occured at N160 (-60 p.p.), N197 (-16 p.p.), N462 (-80 p.p.), N611 (-50 p.p.) and N637 (-13 p.p.). With the exception of N611, these decreases were concomitant with increases in high-mannose glycosylation (Figure 6.4C, E).

6.3.2 Cross-clade observations on changes in glycosylation

Difficulties arise in making cross-clade analyses in the glycosylation of Envelope proteins due to the fact that not all glycan sites are analogous. In order to make such observations, a means of averaging the glycosylation of SOSIPs and SLOSIPs across the different glycan sites which still provides meaningful insights in the general glycosylation is necessary. Categorizing data into high-mannose, complex-type and under-occupied allows for a good summary of the changes in site-specific glycosylation of these constructs. However, as many different glycoforms make up these categories, it is possible that changes in the processing of high-mannose (**M9Glc** $\rightarrow$ **FHybrid**) and complex-type glycans (**HexNAc(3)(x)** $\rightarrow$ **HexNAc(6+)(F)(x)**) may be occuring within the bars of the charts in the previous figures. Consequently, the following results sections reformat the data as follows. For each site that was shared between the SOSIPs and SLOSIPs of each clade (glycan sites for which there was data for B41 SOSIP and SLOSIP for example), the mean for each glycan category from M9Glc to FHybrid was derived (e.g. Σ M9 / no. resolved sites). This gives us a value for the mean of the resolved levels of trimer glycosylation of each glycan category for the entire SOSIP or SLOSIP trimer which, whilst not useful for understanding site-specific changes in glycosylation can provide a general idea of the global influence of codon usage across the different clades.

6.3.2.1 Examining changes in high-mannose glycosylation

Cross clade analysis of changes in mean-resolved high-mannose levels in SOSIPs and SLOSIPs yielded variable changes between clades, albeit with some generalizable observations. With the exception of B41, Most SLOSIPs observed similar overall high-mannose glycosylation to their SOSIP counterparts (Figure 6.5A-D). ZM197M and X1193 SLOSIP was previosuly described as differing from BG505 and B41 as they showed greater levels of high-mannose glycoforms. From Figure 6.5C and D, these changes come in the form of an enrichment of $Man_9GlcNAc_2$, $Man_8GlcNAc_2$, and $Man_5GlcNAc_2$ in ZM197M SLOSIP and $Man_8GlcNAc_2$, $Man_4GlcNAc_2$, and $Man_3GlcNAc_2$ in X1193 SLOSIP. In BG505 SLOSIP (Figure 6.5A), the changes in high mannose glycoforms are not as pronounced, but instead most of the individual categories are found to be in lower quantities in the SLOSIPs compared to the SOSIPs, suggesting that these decreases may be accounted for by increases in other glycoforms such as complex-type glycans. In B41 SLOSIP (Figure 6.5B), There is a drastic drop in the $Man_9GlcNAc_2$ between B41 SOSIP and SLOSIP (26% to 5%). In its place, there is a shift towards smaller, more processed high-mannose glycoforms. Overall, there appears to be an enrichment of more processed high-mannose glycans in SLOSIPs which observe decreases and increases in the levels of high-mannose glycosylation, however these changes do not manifest themselves uniformly.

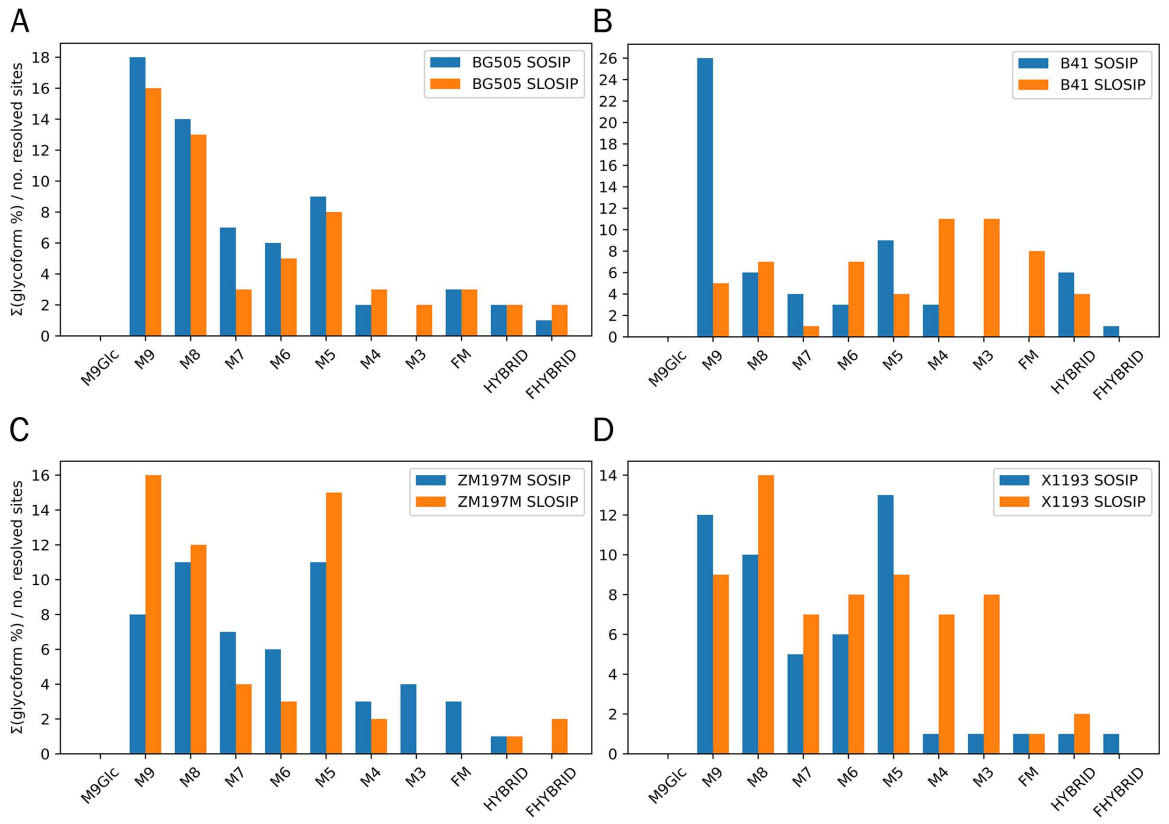


Figure 6.5: **Mean-resolved high-mannose levels in SOSIPs and SLOSIPs** **(A-D)** Mean-resolved high-mannose levels between BG505 **(A)**, B41 **(B)**, ZM197M **(C)**, and X1193 **(D)** SOSIPs and SLOSIPs.

6.3.2.2 Examining changes in complex-type glycan processing

Cross clade analysis of changes in mean-resolved complex-type glycosylation across SOSIPs and SLOSIPs also yielded different changes across clades in the levels and type of complex-type glycosylation between SOSIPs and SLOSIPs (Figure 6.6). In BG505 and ZM197M SLOSIP, there appears to be a shift towards more processed complex-type glycoforms, in particular HexNAc(5)(F)(x) (Figure 6.6A, C). In B41 SLOSIP the inverse appears to be happening, and the proportions of HexNAc(3)(x), HexNAc(3)(F)(x), and HexNAc(4)(x) are far greater than what was observed for B41 SOSIP which appears to have come at the cost of HexNAc(4)(F)(x), HexNAc(5)(x), and HexNAc(5)(F)(x) which are more dominant in B41 SOSIP (Figure 6.6C). In BG505, B41 and ZM197M SLOSIP, there are greater proportions

of the fully processed HexNAc(6)(F)(x) glycoforms than in the SOSIP counterparts. Lastly, X1193 SLOSIP shows large decreases in HexNAc(4)(F)(x) and HexNAc(5)(F)(x) glycoforms which were previously dominant with smaller increases in HexNAc(3)(x) and HexNAc(4)(x) which is perhaps rationalized by the increase high-mannose glycoforms that were previosuly described (Figure 6.6) Overall, with the exception of X1193, there appears to be a shift towards a general increase in complex-type glycosylation in SLOSIPs compared to SOSIPs but once again these changes to not coalesce to a particular glycoform.

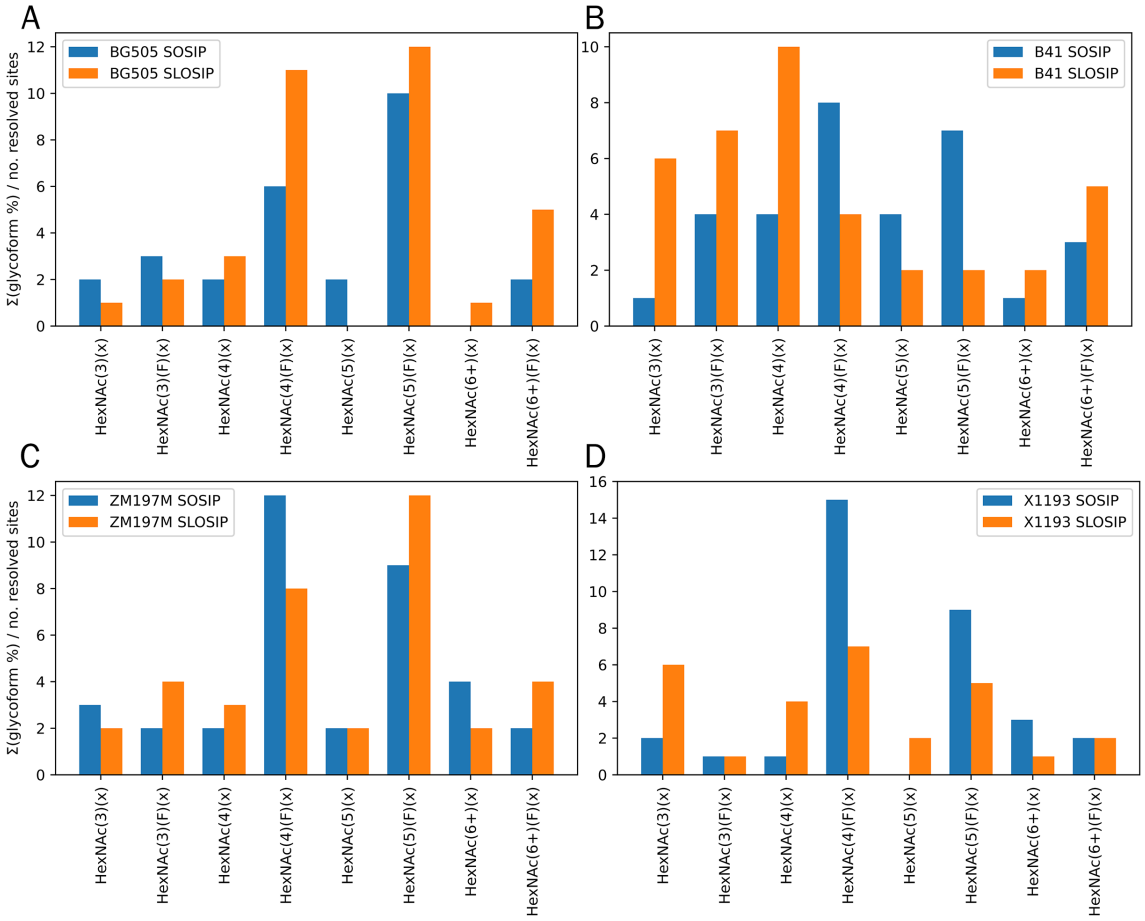


Figure 6.6: **Mean-resolved complex-type levels in SOSIPs and SLOSIPs** **(A-D)** Mean-resolved complex-type levels between BG505 **(A)**, B41 **(B)**, ZM197M **(C)**, and X1193 **(D)** SOSIPs (Blue) and SLOSIPs (Orange).

6.3.2.3 Examining changes in glycan under-occupancy

Cross clade analysis of changes in mean-resolved glycan under-occupancy was consistent in that it showed an average decrease in glycan underoccupancy in all SLOSIPs compared to SOSIPs. The largest differences in under-occupancy between SOSIPs and SLOSIPs were between the two B41 trimers, with a mean of 9% in B41 SOSIP and 1% in B41 SLOSIP (Figure 6.7). This is a substantial drop, and the difference was previously described and accounted for small levels of under-occupancy at sites N160 and N262 (Figure 6.2B). X1193 SOSIP had the highest average glycan under-occupancy at 14% but in X1193 SLOSIP this was reduced to a joint second 6%, accounted for by large decreases in glycan under-occupancy at sites N160, N462 and N611 (Figure 6.4C). ZM197M SLOSIP also displayed an average of 6% glycan under-occupancy, but had a lower level of glycan under-occupancy in ZM197M SOSIP (8%; Figures 6.7 and 6.3C). Lastly, the resolved averages of glycan under-occupancy in BG505 SOSIP (11%) was reduced to 7% in BG505 SLOSIP. Whilst this drop was bigger than that observed in ZM197M SOSIP/SLOSIP for example, the reason for BG505 SLOSIP retaining the highest levels of average under-occupancy could be explained by the increase in glycan under-occupancy at N625 as well as the retention of predominant glycan under-occupancy at N185e (Figure 6.1B, C).

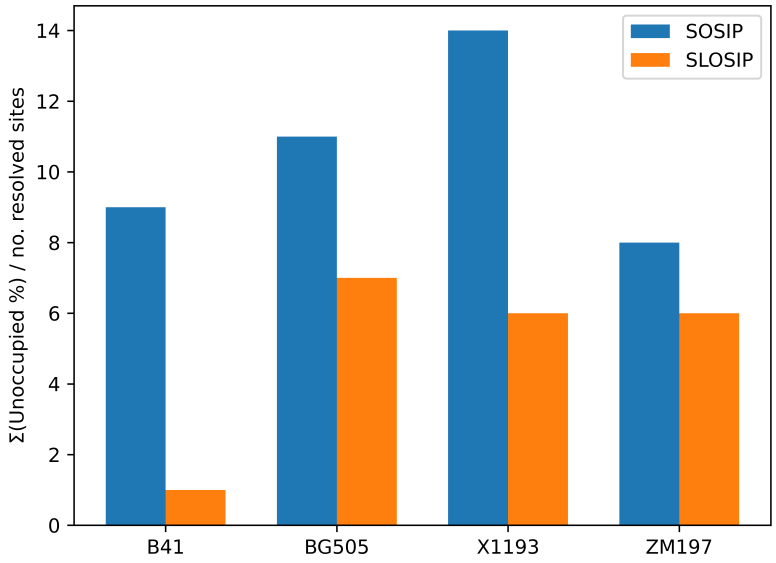


Figure 6.7: **Mean-resolved under-occupancy levels in SOSIPs and SLOSIPs (A-D)** Mean-resolved glycan under-occupancy levels between BG505, B41, ZM197M, and X1193 SOSIPs (Blue) and SLOSIPs (Orange).

6.4 Discussion

In this chapter we aimed to explore the use of rarer codons as observed in viral *env* sequences to engineer the glycan shields of four Envelope strains as a means of improving current immunogen design [113]. As studies on viral Env have highlighted differences in site-specific glycosylation and glycan under-occupancy in that viral Env is more processed and shows very little glycan under-occupancy; we hypothesised that the use of viral sequences on the SOSIP v4.1 platform might yield soluble immunogens whose glycosylation is more processed, particularly at glycan sites which exhibit mixed levels of high-mannose and complex-type glycosylation, with less glycan under-occupancy. Key glycan sites in which we hypothesized that these changes would be observed on BG505 were the N133 (and its analagous sites on B41, X1193 and ZM197M), N197 and N276. Whilst the aim of reducing glycan under-occupancy was centred around the gp41 glycan sites (N611-N637) whose prolific glycan under-occupancy attracts non-nAb responses in animal neutralization studies [210].

The most successful application of the SLOSIP constructs were in reducing glycan under-

occupancy. As shown from the mean-resolved analyses of the individual glycoforms, SLOSIP trimers displayed less glycan under-occupancy than their SOSIP trimer counterparts as a whole in all strains (Figure 6.7). The site-specific analyses of each SLOSIP trimer overwhelmingly skewed towards decreases in glycan occupancy, particularly at gp41 glycan sites. Between all strains, there were 5/16 SLOSIP gp41 sites that had a percentage point decrease in glycan under-occupancy greater than 10 p.p. and only 2/16 with inverse changes. Despite this success in replenishing glycan under-occupancy at the C-terminal sites, this effect was not consistent as outlined by the increases in glycan under-occupancy at N625 in X1193 SLOSIP (+25 p.p.) and BG505 SLOSIP (+14 p.p.; Figures 6.1 and 6.4). Increases in under-occupancy were also observed in ZM197M SLOSIP (N230: +21 p.p.) and, to a lesser extent, B41 SLOSIP (N160: +7 p.p., N262: +9 p.p.; Figures 6.2 and 6.4). B41 SLOSIP was the most successful trimer in filling in glycan under-occupancy, as the only detected glycan under-occupancy was 7% and 9% at N160 and N262, respectively.

The second aim we hoped to achieve with the SLOSIP constructs was to shift the glycosylation of these immunogens to be more akin to the virus, which aside from being nearly fully occupied also exhibits greater levels of complex-type processed glycans than in strain-matched immunogens [261, 104]. Such analyses currently exist only for B41 and BG505. Analysis of BG505 SLOSIP and SOSIP displayed increases in complex-type glycans that brought the glycan profile of BG505 SOSIP more in line with that of the analyses of viral BG505 Env, particularly at sites N133 and N276 (Figure 6.1B) [104]. The N197 glycan site, which is made up of a mixture of glycan under-occupancy, high-mannose and complex-type glycans in BG505 SOSIP, is fully occupied and is made up of 2/3 complex-type glycans and 1/3 high-mannose glycans in one of the viral analyses and almost entirely complex-type with a small fraction of glycan under-occupancy in the other[261, 104]. In our analysis of BG505 SLOSIP, there was effectively no change in the level of glycan under-occupancy at this site between BG505 SOSIP (24%) and SLOSIP (27%), but the proportions of high-mannose and complex-type glycans shifted towards high-mannose at the cost of complex type (+19

p.p., -20 p.p., respectively; Figures 6.1A-C). This was not the only unexpected change that was exhibited between the two constructs, as glycan sites which are almost entirely high-mannose in virus and immunogen such as N234 and N411 showed a notable -12 p.p. and -10 p.p. drop in high-mannose glycosylation with concomitant increases in complex-type glycosylation. The glycosylation of N411 in BG505 viral Env is somewhat disputed as one viral analysis shows it to be almost entirely complex-type whilst another shows the inverse, which perhaps suggests an ability for its glycosylation to differ; however, it should be noted that as a constituent of the intrinsic mannose patch [70, 256], this glycan ought to skew more towards high-mannose as it does per our analysis in both BG505 SOSIP and SLOSIP. Another site whose composition in viral BG505 Env analyses differs is N301. In one study, N301 is fully occupied by high-mannose glycoforms in viral Env, CHO, and HEK-produced BG505 SOSIP, and is consistent with our finding for this site in BG505 SOSIP [104]. In the other analysis of BG505, N301 was fully occupied by high-mannose glycans in BG505 SOSIP but fully complex-type in BG505 viral Env [261], whereas our analysis of BG505 SLOSIP has N301 occupied almost equally between high-mannose and complex-type glycans (Figure 6.1B). Lastly, only one of these studies were able to acquire data on some of the gp41 sites (N611 and N637) and showed them to be fully occupied by complex-type glycans [261]. BG505 SLOSIP was successful in inducing full occupancy of N611 with complex-type glycans, but at N637, despite increasing the relative quantity of complex-type glycans at the cost of high-mannose glycans and glycan under-occupancy, there was still remnants of both in BG505 SLOSIP rendering this site 60% complex-type as opposed to 100% in BG505 viral Env.

The same study also analysed B41 viral Env and B41 SOSIP.664, and compared to our analyses, provided similar observations as BG505. B41 SLOSIP was successful in eliminating in glycan under-occupancy at C-terminal sites, but not in eliminating high-mannose glycosylation as in B41 viral Env (Figure 6.2B, C) [261]. B41 SLOSIP induced more increases in complex-type glycosylation than BG505 at sites such as N187, N234, N262, N276, N295,

N362 and N392. But in many cases when these changes were compared to B41 viral env, they were either not enough to bridge the discrepancies between the high-mannose immunogen glycosylation and the complex-type viral glycosylation (for example, N187) or went beyond the glycan processing in what had been observed in B41 viral Env (for example: N234, N295, N386); and in some cases, B41 SLOSIP showed changes in glycosylation not observed in either of the previosuly analysed B41 SOSIP.664 or B41 viral Env (for example N339, Figure 6.2C) [261]. Analyses of clade C ZM197M and clade G X1193 SLOSIPs were similar in that their glycan profiles did not appear to show a pattern as to which sites exhibited increased high-mannose or complex-type glycans, with the latter SLOSIP exhibiting more of a shift to high-mannose glycans than complex-type unlike the other clades, although as both of these Envs have more PNGSs on the trimer, this could perhaps be rationalized by an enhanced steric inaccesibility to glycan trimming enzymes, particularly at sites which changed from being mostly under-occupied to high-mannose (for example, N637 in ZM197M and N160/463 in X1193; Figures 6.3 and 6.4C). In any case, without analyses of strain matched viral Envs, it is difficult to ascertain the successes of SLOSIPs in the context of these goals.

Overall, these approaches highlight an ability to draw soluble immunogens towards viral glycosylation in some sites where site-specific glycan data on strain-matched viral Env is available. The largest success of the SLOSIP constructs were their ability to effectively reduce glycan under-occupancy compared to all strain matched SOSIPs yet despite this success, they did still exhibit areas where glycan under-occupancy was either unaffected or increased, which suggests that whilst the codon usage of viral immunogens has a tangible effect on the glycan occupancy of these constructs, there are other effects at hand which require characterization. The effects of codon usage on glycan processing are a similar case. Where glycosylation data exists for both immunogen and viral Env, the SLOSIPs display a shift in their glycan profile from the former to the latter. Despite this, there are glycan sites which show either a resistance to this effect, or the direct opposite and are characterized by increases in high-mannose glycosylation. An interesting example of this phenomena

are the changes that occur at the N197 and N301 glycan sites in BG505 SLOSIP (Figure 6.1B, C). These two sites flank each other on the inner face of the trimer above the CD4bs and as per our data, the N197 glycan increases its high-mannose glycosylation at the cost of complex-type glycosylation whereas the N301 glycan site does the opposite; yet in one analysis BG505 viral Env it is the N197 glycan site that is fully processed and the N301 site which is fully high-mannose [104]. As these two sites lie proximal to one another, it could be that the fate of the processing of one site impacts the other. In other words, if the N197 glycan site is less processed in SLOSIP as it was in SOSIP, then it could enhance the enzymatic accessibility to N301, assuming that glycan processing occurs linearly. If this is the case, glycan network effects similar to those discussed in Chapter 3 could be responsible for the variability in processing changes that we observe. Whether this would impact bnAb binding or elicitation depends on where these changes are occurring relative to the bnAb and its epitope, but it is understood that some are able to tolerate promiscuity in the form of glycan micro-heterogeneity in their binding [352]. If this is the case, then addressing the glycan under-occupancy of these soluble Env constructs is of greater importance in order to prevent the elicitation of immunodominant non-nAb epitopes to glycan holes [210]; furthermore epitope areas which rely on the integrity of particular glycoforms such as the trimer apex were maintained across all SLOSIPs.

Lastly, it is important to address the weaknesses of the SLOSIP constructs in the context of vaccine design, and the ways in which they could be addressed. First and foremost is the fact that the use of viral *env* coding sequences substantially reduces the yield when producing these constructs, which makes them unsuitable as vaccine candidates in any protein-based vaccine. Further study could be undertaken to see if these effects can be abrogated whilst still enhancing glycan occupancy by only retaining native *env* sequences in and around PNGSs, but due to the sheer amount of glycans on each protomer this would likely be an ineffective approach. Another, perhaps more promising alternative would be the use of viral *env* sequences in RNA-based vaccine constructs. Due to their favourable characteristics that were

discussed in the introduction, it is possible that there would not be a need for the most optimized codon sequences in encoding Env immunogens on such platforms as they would still be able to elicit a sufficient immune response (i.e. quality over quantity). The other drawback of the SLOSIPs is that it does not completely abrogate glycan under-occupancy in all sites on the trimer. This is an issue because even if a small subset of trimers in a vaccine candidate express a glycan hole there is a danger that this response could negate the immunogenicity of the trimers which do not possess this hole [210, 353, 354]. There are many potential reasons for this, but perhaps a solution that yields the best chances of success is the use of membrane-bound constructs, which have also been shown to exhibit greater occupancy and processing to a certain extent, and have the added bonus of not possessing an exposed trimer base which can also lead to undesirable immune responses [261, 355, 356]. It is likely that many of these approaches will have to be studied in combination in order to understand what drives the occupancy and processing of these heavily glycosylated constructs, and to what end they can be exploited whilst still maintaining structural and epitopic integrity.

7 Discussion: Env glycan engineering and future perspectives

7.1 Comparing engineering approaches on Env

The purpose of this thesis was to evaluate methods by which the glycan shield of HIV-1 Env vaccine candidates could be be subjected to alterations through a range of engineering approaches. Before this thesis, the site-specific glycosylation of arguably the most widely studied vaccine candidate, BG505 SOSIP.664 had been characterized in different cell lines, as well as the strain matched viral Env soon after. These comparative studies support the idea that the HIV-1 Env glycan shield of these immunogens is comparable to that of the virus itself [261, 104]. This is crucial as the broadly neutralizing antibodies which constitute the goal of these vaccine approaches bind to or around specific N-linked glycans, highlighting the importance of the site-specific characterization of these candidates. Studies have shown that the key areas of N-linked glycans are consistent between immunogens and virus, specifically the oligomannose-type glycans that are conserved most notably at the IMP which revolves around the N332 glycan as well as the TAMP, which includes the apex glycan sites N156 and N160 [123]. Both of these N-linked glycan sites constitute key epitopes for apex and outer-domain targeting bnAbs, such as PGT145 and 2G12, respectively [69, 256].

The sites that flank these regions, such as N197, N276 or N355 show divergent glycosylation in the form of a mixture of complex-type and high-mannose glycoforms in the soluble SOSIP trimers produced in both aforementioned cell lines [261, 104]. Furthermore, areas of under-occupied glycan PNGSs underlie differences between BG505 SOSIP and viral Env. These arise most prominently in PNGSs which are located proximally to one another along the primary protein sequence, for example N133/N137, N185e/N185h, and the C-terminal sites between N611-N637 [261, 313]. Glycan under-occupancy constitutes a problem as it has significant ramifications about the type of immune response that arises from the use of these

proteins as vaccines. When used in macaque immunization studies, neutralizing antibodies were generated but they were only effective in mounting a response against an autologous or very similar virus [357]. Furthermore, studies defined non-neutralizing epitopes on BG505 SOSIP located at the N611 glycan site whose antibodies could only neutralize viruses lacking the PNGS at this site, suggesting the generation of a non-nAb response targeted at the N611 glycan hole [313, 358]. Other non-nAb epitopes that were identified localized to the base of the trimer, whose exposure arises from the solubilization of the endogenously membrane-bound Env, and glycan holes arising from under-occupancy of V1 PNGSs [358]. Base responses were found to be the most immunodominant non-nAb responses to BG505 SOSIP trimers. Lastly, areas of exposed protein surface exist on BG505 Env where PNGSs are present in other strains which draw non-nAb responses in rabbit immunization studies [359].

The findings of non-nAb responses on these vaccine candidates has called for a reevaluation in what is required for a successful Env-based HIV-1 vaccine, as the use of high fidelity mimics of the spike protein whether soluble or membrane bound have not been able to overcome these issues up until this point. Fortunately, the studies which have outlined the immunodominance of these exposed protein surfaces whether through a lack of PNGS or otherwise have yielded further observations which show that the same phenomena may be exploited to draw the initial immune response to epitopic regions [213]. This is potentially advantageous particularly in a prime-boost vaccine regimen as many germline B-cells which encode for early bnAb precursors do not arise with an immediate tolerance for the glycans in their epitopes but instead evolve such features over time [267]. In this regard, a clinical trial on humans looking into the use of an engineered Env immunogen which uses a nanoparticle to display several copies of only the conserved outer domain and is lacking in the N276 glycan site was successful in eliciting VRC01-class responses in the majority of the participants [360, 361, 170].

Whilst this constitutes a promising success in priming for germline VRC01 antibodies, this

is effectively the first hurdle in a series that lead to the induction of a mature bnAb in humans. Further studies need to consider the type of subsequent immunogens that would lead to maturity. It could be that following on from priming with these immunogens, the use of native-like trimers with or without an N276 glycan in subsequent boost vaccinations could achieve this. Depending on how strong these responses are after priming, it is possible that immunogens with non-nAb epitopes arising from glycan holes could be suitable after the initial, desired response is made; however studies will likely need to be made in order to ascertain the effects of these competing epitopes against one that was already elicited in a previous vaccination. If this is not the case, then engineering efforts will be required in order to create appropriate immunogens to follow on from priming immunogens. The most likely candidates for this would either be modified priming immunogens (eOD in this particular example), or modified native-like immunogens that are designed to mature the priming response only by abrogating distracting epitopes. This thesis explores aspects by which native-like immunogens may be subjected to modifications its glycans to achieve this whilst maintaining the overall integrity of the glycan shield using two different approaches, the modification of amino acid sequences around glycan PNGSs and co-translational engineering interventions using glycosyltransferase or bnAb co-transfection. In Chapter 3, we explored the use of glycan knock-ins and knock-outs through the insertion of PNGSs into the primary amino acid sequence in order to ascertain the ability for trimers and the glycans that occupy them to tolerate such changes.

In Chapter 4, we expanded on the use of PNGS knock ins/outs and applied this strategy as well as other amino acid modifications on a alternate, membrane-bound Env platform. We explored the use of glycosyltransferase plasmid co-transfection on soluble gp120s as well as membrane-bound VLP-presented trimers. We additionally used a combination of glycosyltransferases followed by treatment with Endo F3 to ascertain the glycan sites most susceptible to post-translational enzymatic action against those that are resistant. In Chapter 5, we built on the strategy of enzymatic co-transfection through the use of bnAb co-expression

in order to induce targeted changes in the glycan shield of SOSIP trimers. In Chapter 6, we explored the use of native *env* sequences on SOSIP trimers of different strains as a means of remediating the discrepancies that are observed between viral and immunogen glycosylation.

7.1.1 Comparing and contrasting changes to Env protein sequence

The differences in site-specific glycosylation of membrane-bound and soluble immunogens has been previously documented. In this regard, it was important to establish the suitability of these approaches and their differences in both formats. Generally, studies observing the site-specific glycosylation of membrane-bound Env constructs have shown them to be occupied by larger, branched complex-type glycans with enhanced sialylation compared to SOSIPs [350, 261, 124]. Such observations could be rationalized by the idea that membrane-bound constructs are in closer proximity to the glycan processing enzymes, thus allowing for the enhanced branching and addition of complex-type residues such as sialic acid. This is supported by data showing that some mixed sites such as N197 and N276 are more processed in strain-matched membrane-bound constructs [261].

In Chapters 3 and 4, we explored the concept of PNGS addition and deletion in two different contexts. Chapter 3 explored the ramifications of adding or removing glycan PNGSs on the site-specific glycosylation of the other glycan sites, proximal and distal, in order to ascertain how well such changes are tolerated elsewhere, as well as what the added glycan sites are comprised of. Chapter 4 combined this approach alongside methods to reduce glycan under-occupancy by modulating N-X-S/T glycosylation motifs [313] to improve two distinct epitope regions, the V2 apex and fusion peptide region.

The UPLC and site-specific glycan analysis of the BG505 SOSIP mutants in chapter 3 demonstrated that despite knocking out glycans in the IMP, these sites were somewhat resistant to changes in the overall glycosylation as they remained predominantly oligomannose-type (Figure 3.2). Despite this resilience, the site-specific glycan data of all of these mutants shows a decrease overall mannose and particularly in $Man_9GlcNAc_2$ glycoforms. This effect

was most pronounced at the glycan sites immediately adjacent to the knocked-out sites. Other effects were observed at sites not immediately adjacent to the knocked-out PNGSs, as shown by the effects on N234 in the N386, N392 and N448 knock-outs, and by the reciprocal effects in processing between the knock-out mutants of N332 and N448 on each other (Figure 3.2). The additions of the N241 glycan site did not significantly increase the abundance of oligomannose or $Man_9GlcNAc_2$ glycoforms in the IMP sites. In the +N289 and double +N241 +N289 mutant, large decreases in overall oligomannose-type glycosylation as well as $Man_9GlcNAc_2$ were observed in the BG505 SOSIP.664 format as well as the more stabilized SOSIP v5 format, albeit to a lesser extent, suggesting that the trimer was unaccommodating toward this insertion (Figure 3.3). Despite this, binding studies of the knock-in mutants to a panel of bnAbs with distinct epitopes suggested antigenic similarities with the WT (Figure 3.4B).

When these approaches were explored in Chapter 4, they yielded diverse effects. Generally speaking, the insertions of the N197 and the T49N glycan sites were well tolerated, as were the removals of the N138, N141, and N611 glycan sites (Figure 4.1C). Changes in the insertions were mostly limited to sites neighbouring the changes, for example in the +D197N + S199T JR-FL mutant (Figure 4.4F). In the membrane bound constructs, allosteric changes to distal glycan sites were more prevalent, particularly in The T49N, N611Q and double mutants, but this change was also observed in the N138A and N141A knock-out mutants (Figures 4.1C, 4.5, and 4.6). In the knock-out mutants, the N141 and N138 were substituted for alanine residues as the N135 glycan was associated with enhanced V2 nAb sensitivity. As the selectivity of CH01 to GnT1- pseudoviruses was previously shown to be increased, it is unsurprising that the VLPs which were purified using this antibody were predominantly high-mannose (Figure 4.6B, D). The fact that many distal glycan sites were affected by means of enhanced high-mannose glycans was however surprising. This could be due to a variety of factors. For one, membrane trimers show an enhanced ability to display a variety of conformations, particularly in comparison to SOSIP trimers which have enhanced

rigidity. An alternate explanation could be that in the production of VLP membrane-bound trimers, a subset of the presented trimers, and thus a subset of the analysed trimers, are not well-folded. Therefore in selecting for particular trimers using an antibody with a known specificity, we observe the subsequent allosteric effects that arise at other glycan sites when preferential high-mannose glycans at the V2 region sites are selected for.

An approach which was not explored in Chapter 3, but which has been on other studies using BG505 SOSIP.664 trimers with a degree of success is NxT sequon optimization [313]. In membrane-bound JR-FL trimers, these approaches did not have the intended effects (for example D197N vs D197N + S199T; Figure 4.4). In the D197N + S199T trimer, not only was a small proportion of glycan under-occupancy introduced at N197, but also at N295 and N463 which was fully occupied in D197N, as well as an enhancement of glycan under-occupancy at N339. Interestingly, there was an greater proportion of complex-type glycans than high-mannose at N197, suggesting that NxS to NxT changes may also impact on the level of processing (Figure 4.4A, B). Overall, whilst it is difficult to make observations between strains and immunogen platforms with a degree of certainty, from this work it appears that the additions and deletions of PNGSs as well as other modifications can be incorporated without causing averse effects in the overall glycosylation of Env proteins that are characterized by misfolded material. Most changes map to nearby glycan sites near the point of change. Careful study must be taken when imparting these changes however as they may yield unexpected effects in glycan network of Env that can result in allosteric changes to glycan sites distal from the point of mutation, particularly in membrane-bound constructs.

7.1.2 Efficacy of modulating glycan processing

Due to the propensity for alterations to the amino acid sequence to yield unexpected consequences to the site-specific glycosylation of Env constructs, methods by which the modulation of glycan processing could be exploited without affecting the amino acid sequence were explored across Chapters 4, 5, and 6. Chapters 4 and 5 focused on co-translational methods

by which to alter glycosylation using glycosyltransferase and bnAb co-expression. When assessing the ability of glycosyltransferase co-expression, we sought to define the effect on gp120s as well as VLP-expressed membrane-bound trimers.

The gp120s were co-expressed with two different galactosyltransferases and fucosyltransferases (B4GalT1/B4GalNT3, FUT6/FUT9). In the fucosyltransferase co-expressed gp120s, we observed increases in the total numbers of fucosyltated glycans, as well as increases in glycans with two or more fucose residues but we did not observe any widespread increases in the proportion of fucosylated glycans, as most changes were restricted to N156, N301, and N386 (Figure 4.10C, F). Galactosidase co-expression did not appear to impart any enzyme specific changes in the gp120s but, like FUT9 co-expression, led to increases in dominant glycan types (predominantly oligomannose) in sites previously displaying glycan under-occupancy (Figure 4.10E, F). As previously discussed, these observed changes went against current enzymatic understanding. When compared to some of the findings from Chapter 5, we observe similar changes to glycan under-occupancy that were not related to antibody co-expression (for example Figure 5.7C). Furthermore, we describe greater and at times unexpected effects in sequence changes to membrane-bound Env constructs than solbule constructs when comparing Chapters 3 and 4. Indeed, the co-expression of the same galactosyltransferases in combination appeared to exert a greater influence on VLP-expressed trimers than the gp120s (Figure 4.7). With this in mind, the changes to the glycosylation of glycosyltransferase co-expressed gp120 constructs can be explained as follows; it is possible that they were able to avoid large scale enzyme-induced changes due to their solubility and the membrane-bound nature of the co-expressed glycosyltransferases. As for the changes in glycosylation that were coupled with decreases in under-occupancy, and small scale exchanges between oligomannose and complex-type glycosylation, it is hard to understand what the underlying factors are as these effects are neither related to enzymatic activity nor do they appear to be consistent across glycan sites.

This was not the case for the B4GalT1 and ST6Gal1 co-expressed membrane bound JR-FL

trimer (Figure 4.7B, D). Indeed, in glycan sites which were complex-type, all but N611 exhibited large-scale changes in hybrid-type glycosylation compared to the parent wild-type trimer. In comparison, the NGAF-treated VLP shows the difference in accessibility of enzymes in the Env glycan shield between co-translational and post-translational application (Figure 4.7C, E). The post-translationally applied glycan trimming enzymes were only able to act upon the N88, N356 and N616 glycan sites, showing that the other sites once processed and in the final folded form of Env are occluded from the action of such enzymes. The increases in oligomannose-type glycans at sites N88, N160 and N616 additionally goes beyond explanation of the characteristics of these enzymes, and whilst it is possible that a form of quantum shift may be occurring as suggested in the discussion. It is a more feasible possibility to ascribe these changes to batch differences in glycosylation between replicates, as these changes are comparatively small to those imposed by both the B4GalT1/ST6Gal1 co-expression as well as the sites in which NGAF3 was able to cleave. This demonstrates the viability of the use of glycosyltransferase co-expression as a method to induce changes in membrane-bound Env constructs, in spite of its lack of specificity, it is able to induce significant changes whilst maintaining the integrity of the IMP and TAMP. This may only be due to their inherent steric inaccessibility, but it is nonetheless advantageous as it allows for an intervention that will not disrupt the integrity of crucial epitope regions, though consideration in the types of co-expressed glycosyltransferases must be taken with view to the desired effects.

The co-expression of bnAbs with soluble BG505 SOSIP trimers, as well as being an investigation into site-specific interventions in the glycan shield of Env, constitute a view into the effects of the mechanism between bnAb and Env binding on the glycan shield. The results of many of the Env-bnAb co-expressions can be rationalized by the location and the way in which the bnAb engages Env. For example, VRC01 was able to induce majority $Man_9GlcNAc_2$ structures at the N276 glycan site, as well as other shifts in high-mannose glycosylation in surrounding glycan sites (Figures 5.2 and 5.3). This is rationalized by the

idea that VRC01 can bind its epitope around the glycan site and occlude this and adjacent glycan sites from enzymatic action. Also previously mentioned was the dichotomy between VRC01 and VRC02 in which the latter likely provokes CD4-induced changes that lead to increased processing that we observe in both VRC02 co-expressed and purified BG505 SOSIP, with the key differences again being the steric occlusion that leads to the enrichment of high mannose at N276/N462 (Figure 5.4).

These observations were the most detailed in Chapter 5 due to their prominence and the clear relationship in their binding and the observed data. Yet other bnAbs, for example apex bnAbs, are harder to rationalize. For example, the changes in co-expression in apex-targeting bnAbs such as PGT145 and PG16 did not appear to be related to their specificities, as their epitopes did not exhibit change as they are endogenously high-mannose (Figure 5.5). Both co-expressions caused notable increases in high-mannose glycoforms at N276 and some of the adjacent glycan sites as well. Antibodies that target this area are known to possess features that allow it to penetrate the trimer apex [68]. In doing so, it is possible that there are ramifications of this binding affect the accessibility of these sites to enzymatic action, as the opening of the apex via antibody binding necessitates a degree of movement across the trimer. In comparison to glycosyltransferase co-expression, where its action is dictated only by the accessibility of its substrate, the effects of antibody co-expression are more complex. Firstly, there are direct implications enzymatic accessibility on the glycan sites that are directly part of, or immediately proximal to the bnAb epitope. Secondly, there are effects that are imparted on the structure of the Env trimer through the binding of the bnAb, and the resultant changes in enzymatic accessibility from such changes that can either restrict or enhance glycan processing. As many structural studies in binding of bnAbs to Env usually follow their elucidation, their use in co-expression can be not only advantageous in inducing specific structures as Chapter 3 demonstrated, but also in understanding what effects its binding has on the glycoprotein as a whole.

It is partly with view to the suspected translational effects observed in Chapters 4 and 5,

that Chapter 6 sought to explore the impact of translational speed on the site-specific glycosylation of Envelope constructs. The other aspect was to bridge differences in viral and immunogen glycosylation. This objective differed to those in Chapter 4, which was to use glycosyltransferases to make epitopes more preferential to bnAbs; or Chapter 5 which was a proof-of-concept in inducing targeted changes to a few glycan sites. Where Chapter 6 differs from the other two is that it did not use any external factors to reshape the site-specific glycosylation of the Envelope constructs that were investigated. As such, the implications of changing the plasmid encoding sequence to incorporate the rarer codon usage observed in viral *env* DNA sequences were inherently going to be more wide-ranging as it would affect the entire Env sequence and therefore all of the glycan sites, as opposed to those which were accessible or bound to. It is unsurprising therefore, that the changes observed between the SOSIP and SLOSIP trimers of all strains were more widespread than those observed in Chapters 4 and 5.

The changes we sought to impart in the SLOSIPs were twofold: the first being to reduce glycan under-occupancy, the second to cause a shift towards complex-type glycans in general but particularly mixed sites which are predominantly complex-type in analyses of viral env [261, 104]. With these in mind, the largest success was in reducing glycan under-occupancy. There were widespread decreases across analagous and non analagous sites in all SLOSIP trimers in glycan under-occupancy (Figures 6.1, 6.2, 6.3, 6.4), and despite some increases in other glycan sites in SLOSIP trimers (e.g. N625 in BG505 and X1193, N230 in ZM197M), overall glycan under-occupancy decreased in all SLOSIP trimers relative to SOSIP trimers (Figure 6.7). Discerning the changes in glycan processing that occurred between the SOSIP and SLOSIP trimers is more difficult, as all SLOSIP trimers displayed increases in high-mannose and complex-type glycosylation at sites with full glycan occupancy. Often, changes in one came at a compensatory cost to the other, and these changes were not analogous (for example N276 becomes more complex-type in B41/BG505 SLOSIP, more high-mannose in X1193/ZM197M; Figures 6.1, 6.2, 6.3, 6.4). With the exception of ZM197M SLOSIP,

which is the most densely glycosylated trimer in this panel, the other SLOSIP trimers observe a decrease in the dominant SOSIP high-mannose glycoforms with a shift towards more processed or intemediary high-mannose structures (Figure 6.5). Additionally, complex-type glycan structures appear to be enriched particularly in BG505 and B41 SLOSIP and between them and ZM197M there is a suggestion that they are also observing an increase, although not an established dominance, in more processed complex-type glycoforms amongst the analysed complex-type structures in these trimers (Figure 6.6).

Whilst the site-specific glycan profiles of BG505 and B41 SLOSIP do not replicate the viral profiles that have been previously established [261, 104], They do represent a shift in that direction. If we view these incomplete changes in light of the effects (enzymatic and translational) in glycosyltransferase co-expression on soluble and membrane-bound constructs as well as antibody co-expression from the other chapters, it suggests that both the rare codon usage observed in viral *env* and the enhanced enzymatic processing that arises from the use of glycosyltransferases or glycan-binding bnAbs. Both may be exploited to bring immunogens more in line with desired glycan epitopes that either recapitulate viral glycosylation, or enhance bnAb affinity.

7.2 Future perspectives

This thesis aimed to explore the ways in which the glycan shield of HIV-1 Envelope immunogens could be modified with view to improving current HIV-1 vaccine strategies. Whilst current immunogens are able to replicate key conserved areas of the glycan shield, areas in which discrepancies lie between viral and immunogen glycosylation that arise from a number of factors ranging from the expression systems, the choice of nucleotides encoding Env, as well as the choice of purification method in immunogen production. In addition, there is a variety of immunogen platforms with different engineering approaches that are undergoing trial, each with their own benefits and caveats.

This thesis demonstrated the viability of a range of engineering interventions across a range

of these formats, and that they can be applied to HIV-1 vaccine design to beneficially alter the glycan shield without compromising the integrity of conserved regions that are already in line with the virus itself. During the course of this thesis, some of these discrepancies in soluble immunogens have been shown to actively hamper bnAb development and as such, some of the focus in this area has shifted to other immunogens and strategies that have yielded early successes [353, 360, 361]. These successes in engineered subunit vaccines could indicate a turning point in the bnAb induction but such findings must be taken in context. The bnAb that is the current focus of this strategy, VRC01, is a potent and broad antibody that is able to neutralize a number of strains and yet despite this, a number of VRC01-resistant variants have already been characterized from infected patients [362, 363]. Considering this, future HIV-1 vaccine strategies will likely have to take the view of inducing bnAbs to multiple, distinct epitope regions [171]. Depending on the feasibility of each, multiple strategies could arise from either subunit-based or soluble Env based strategies. Germline-targeting strategies emulating that of eOD and VRC01 are being trialled for other epitopes such as the fusion peptide [269]. Alternatively, the use of soluble immunogens based on SIV Env trimers have been trialled as they are able to bind to early and mature apex bnAbs, and do not bind other HIV-1 bnAbs [305]. It is thought that as non-neutralizing epitopes would also not be shared across SIV and HIV-1 trimers, this could be used to prime apex bnAbs in the same way as eOD does for the CD4bs.

As a matter of priority, the concept of germline-targeting needs to be proven to the point of inducing a mature bnAb. Whilst these diverse candidates show the ability to bind and in some case stimulate early bnAb precursors, we do not yet know if this concept will work equally for all the epitopes that are being trialled. Moreover, whilst this guided immunization might work for one or many different epitopes, multiple vaccine regimens each for one epitope is not a feasible approach in combating this virus in patients, and such approaches will eventually have to be integrated. Lastly, there are factors outside of immunogen design that drive the formation of the rare mutations in bnAbs such as maintaining the encoding

B-cells in germinal centres with available epitopes [308, 364].

The glycan shield is intrinsic to bnAb development as it is to Env, and as such, design approaches need to assess the site-specific glycosylation of immunogens with its objective in mind. Specifically, immunogen design strategies need to ascertain how best to bridge the gaps in glycan under-occupancy and differences in glycosylation between immunogens and viral Env, and the importance of each. This thesis highlighted several strategies to achieve this, but the appropriateness of them depends on the goal in mind. For example, the use of rarer codons encoding soluble immunogens was shown to remediate some of these differences but as a consequence, the overall yield of these trimers was substantially reduced compared to immunogens with optimized sequences, which in turn makes them unsuitable for production of a protein-based vaccine. This could be addressed through harnessing the innovations from the recent coronavirus pandemic by applying them to nucleotide-based RNA vaccines, which would leverage the production and post-translational modifications onto the vaccine's recipient. However, these approaches would require a means by which to characterize the site-specific glycosylation of these vaccines to assess the feasibility of such approaches. Antibody co-expression on the other hand could be an effective means of imparting specific changes to immunogens as well as producing therapeutic bnAbs and immunogens at once, provided they could be efficiently separated. They would be unsuitable for a nucleotide based approach however as the production of both in a vaccinated individual would merely block the engineered epitope. With regards to Env, a combination of approaches including some described in this thesis will hopefully prove useful in addressing these differences and abrogating unwanted effects arising from aberrant glycosylation in future HIV-1 vaccine design.

A Analysed Env sequences

BG505 SOSIP analysed in Chapters 3 and 5

```
>BG505 SOSIP.664
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD
```

BG505 SOSIP mutants analysed in Chapter 3

```
>BG505 SOSIP.664 N295A
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQIACTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD
```

```
>BG505 SOSIP.664 N332T
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCTVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

>BG505 SOSIP.664 N339A
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWAETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD
```

```
>BG505 SOSIP.664 N386A
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCATSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

>BG505 SOSIP.664 N392A
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFASTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD
```

```
>BG505 SOSIP.664 N411A
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSASTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

>BG505 SOSIP.664 N448A
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSAITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD
```

```
>BG505 SOSIP.664 +N241
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSNSTVQCTHGIKPVNSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

>BG505 SOSIP.664 +N289
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNNKS
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD
```

```
>BG505 SOSIP.664 +N241 +N289
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHN
VWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSNSTVQCTHGIKPVNSTQLLLNGSLAEEEVMIRSENITNNNKS
ILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD

>BG505 SOSIP v5.2 +N289
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETKKHN
VWATCACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTSVQINCTRPNNNTRKSIRIGPGQWFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPECQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD
```

```
>BG505 SOSIP v5.2 +N241 +N289
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAENLWVTVYYGVPVWKDAETTLFCASDAKAYETKKHN
VWATCACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNIT
DDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFE
PIPIHYCAPAGFAILKCKDKKFNGTGPCPNVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKN
ILVQFNTSVQINCTRPNNNTRKSIRIGPGQWFYATGDIIGDIRQAHCNVSKATWNETLGKVVKQLRKHFGNN
TIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGSNSTGSNDSITLPCRIKQIINMWQR
IGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRC
KRRVVGRRRRRRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPECQQHLLKLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDNMTWLQWDKEISNYTQIIY
GLLEESQNQQEKNEQDLLALD
```

JR-FL and mutants analysed in Chapter 4

```
>JR-FL SOS E186K + N189A (Parent)
MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATH
ACVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSE
GTMERGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCDTSVITQACPKISFEPIPIHYC
APAGFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKE
SVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHS
SGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIR
GQIRCSSNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRA
VGIGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVE
RYLGDQQLLGIWGCSGKLICCTAVPWNASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKN
EQELLELDKWASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV
```

```
>JR-FL SOS gp120
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARVEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHN
VWATHACVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNT
TNDSEGTMERGEIKNCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEPI
PIHYCAPAGFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTII
VQLKESVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTI
VFNHSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMY
APPIRGQIRCSSNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVVQ
REK

>JR-FL SOS E168K + N189A + S158T
MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATH
ACVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSE
GTMERGEIKNCTFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCDTSVITQACPKISFEPIPIHYC
APAGFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKE
SVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHS
SGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIR
GQIRCSSNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRA
VGIGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVE
RYLGDQQLLGIWGCSGKLICCTAVPWNASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKN
EQELLELDKWASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV

>JR-FL SOS E168K + N189A + S364T
MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHA
CVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGT
MERGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCDTSVITQACPKISFEPIPIHYCAPA
GFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKESVEI
NCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHTSGGDP
EIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS
SNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVERYLGDQQL
LGIWGCSGKLICCTAVPWNASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELDK
WASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV
```

```
>JR-FL SOS E168K + N189A + D197N
MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHA
CVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGT
MERGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCNTSVITQACPKISFEPIPIHYCAPA
GFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKESVEI
NCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHSSGGDP
EIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS
SNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVERYLGDQQL
LGIWGCSGKLICCTAVPWNASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELDK
WASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV

>JR-FL SOS E168K + N189A + D197N + S199T
MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHA
CVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGT
MERGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCNTTVITQACPKISFEPIPIHYCAPA
GFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKESVEI
NCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHSGGDPE
IVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCSS
NITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRAVGIGAVFL
GFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVERYLGDQQLL
GIWGCSGKLICCTAVPWNASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELDKW
ASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV
```

```
>JR-FL SOS E168K + N189A + T49N
MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWATHA
CVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGT
MERGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCDTSVITQACPKISFEPIPIHYCAPA
GFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKESVEI
NCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHSSGGDP
EIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS
SNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVERYLGDQQL
LGIWGCSGKLICCTAVPWNASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELDK
WASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV

>JR-FL SOS E168K + N189A + N611Q
MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHA
CVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGT
MERGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCDTSVITQACPKISFEPIPIHYCAPA
GFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKESVEI
NCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHSSGGDP
EIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS
SNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVERYLGDQQL
LGIWGCSGKLICCTAVPWQASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELDK
WASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV
```

>JR-FL SOS E168K + N189A + T49N + N611Q

MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEANTTLFCASDAKAYDTEVHNVWATHA
CVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGT
MERGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCDTSVITQACPKISFEPIPIHYCAPA
GFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKESVEI
NCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHSSGGDP
EIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS
SNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVERYLGDQQL
LGIWGCSGKLICCTAVPWQASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELDK
WASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV

>JR-FL SOS E168K + N189A + N138A + N141A

MRVKEKYQHLWRWGWKWGTMLLGILMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHA
CVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATATTADSEGT
MERGEIKNCSFNITTSIRDKVQKEYALFYKLDVVPIDNNATSYRLISCNTTVITQACPKISFEPIPIHYCAPA
GFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNAKTIIVQLKESVEI
NCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVIKLREQFENKTIVFNHSSGGDP
EIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTITLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS
SNITGLLLTRDGGINENGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRRVVQREKRAVGIGAVF
LGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNNLLRAIEAQQRMLQLTVWGIKQLQARVLAVERYLGDQQL
LGIWGCSGKLICCTAVPWNASWSNKSLDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLELDK
WASLWNWFDITKWLWYIKIFIMIVGGLVGLRLVFTVLSIVNRV

SOSIP/SLOSIPs analysed in Chapter 6

```
>BG505 SOSIP/SLOSIP v4.1
AENLWVTVYYGVPVWKDAETTLFCASDAKAYETKKHNVWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQ
MHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFNMTTELRDKKQKVYSLFYRLDVVQI
NENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTH
GIKPVVSTQLLLNGSLAEEEVMIRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQWFYATGDII
GDIRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWI
SNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRP
GGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGAVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSG
IVQQQSNLLRAPEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLS
EIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

>B41 SOSIP/SLOSIP v4.1
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAAKKWVTVYYGVPVWKEATTTLFCASDAKAYDTKVHNV
WATHACVPTDPNPQEIVLGNVTENFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCNNVNTNNTNN
STNATISDWEKMETGEMKNCSFNVTTSIRDKIKKEYALFYKLDVVPLENKNNINNTNITNYRLINCNTSVITQ
ACPKVSFEPIPIHYCAPAGFAILKCNSKTFNGSGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEIVIRSEN
ITDNAKTIIVQLNEAVEINCTRPNNNTRKSIHIGPGQWFYATGDIIGNIRQAHCNISKARWNETLGQIVAKLE
EQFPNKTIIFNHSSGGDPEIVTHSFNCGGEFFYCNTTPLFNSTWNNTRTDDYPTGGEQNITLQCRIKQIINMW
QGVGKAMYAPPIRGQIRCSSNITGLLLTRDGGRDQNGTETFRPGGGNMRDNWRSELYKYKVVKIEPLGIAPTA
CKRRVVQRRRRRRAVGLGAFILGFLGAAGSTMGAASMALTVQARLLLSGIVQQQNNLLRAPEAQQHMLQLTVW
GIKQLQARVLAVERYLRDQQLLGIWGCSGKIICCTNVPWNDSWSNKTINEIWDNMTWMQWEKEIDNYTQHIYT
LLEVSQIQQEKNEQELLELD
```

```
>X1193 SOSIP/SLOSIP v4.1
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAGDLWVTVYYGVPVWEDADTTLFCASDAKAYSTESHNV
WATHACVPTDPNPQEIPLKNVTENFNMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLICTNVTSNSTNS
TNGVTNNSTVDYREQLKNCSFNITTEIRDKQRKEYALFYRLDIVPINDNEKNDTYRLINCNVSTIKQACPKVT
FDPIPIHYCAPAGFAILKCRDKKFNGTGPCKNVSTVQCTHGIKPVISTQLLLNGSLAEGDIMIRSENITDNAK
TIIVQLKTAVNITCTRPSNNTRKSIRFGPGQAFYATDEIIGDIRQAHCNISKTEWEDMKRNVSDKLKALFNNK
TIIFKSSSGGDLEITTHSFNCRGEFFYCNTSGLFNTSGLFNNNSNDSSGNITLPCKIKQIVRMWQRVGQAMYA
PPIAGNITCRSRITGLLLVRDGGKSNETNGTETFRPAGGDMRDNWRSELYKYKVVKIKPLGVAPTRCRRRVVG
RSSRAVGMGAVLLGFLGAAGSTMGAASITLTVQVRQLLSGIVQQQSNLLRAPEAQQHLLQLTVWGIKQLQARV
LAVERYLQDQQLLGIWGCSGRLICVTNVPWNASWSNKSYSEIWDNLTWVEWEREISNYTQHIYNLLQESQNQQ
EKNEQDLLALD

>ZM197M SOSIP/SLOSIP v4.1
MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARMEQLWVTVYYGVPVWKEAKATLFCASDAKAYEKEVHNV
WATHACVPTDPNPQEIPLGNVTENFNMWKNDMADQMHEDIISLWDQSLKPCVKLTPLCVTLNCSDATSNTTKN
ATNTNTTSTDNRNATSNDTEMKGEIKDCTFNITTEVRDRKTKQRALFYKLDVVPLEEEKNSSSKNSSYKEYRL
ISCNTSTITQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCHNVSTVQCTHGIKPVVSTQLLLNGSLA
EEEIIIRSENLTDNTKTIIVHLNESVEIECVRPNNNTRKSVRIGPGQTFFATGEIIGDIRQAHCDLSKSNWTT
TLKRIEKKLKEHFNNATIKFESSAGGDLEITTHSFNCRGEFFYCNTSGLFNSSLLNDTDGTSNSTSNATITLP
CRIKQIINMWQEVGRAMYASPIAGIITCKSNITGLLLTRDGGNKSAGIETFRPGGGNMKDNWRSELYKYKVVE
IKPLGIAPTSCKRRVVERRRRRRAAGIGAVILGFLGAAGSTMGAASVMLTVQARQLLSGIVQQQSNLLRAPEA
QQHMLQLTVWGIKQLQTRVLAIERYLKDQQLLGLWGCSGKLICCTAVPWNTSWSNKSKDEIWDNMTWMQWDRE
IDNYTQVIYQLLEVSQNQQEKNENDLLALD
```

B Automating site-specific glycan analysis

Script for analysing raw site-specific glycan data

Raw data that is analysed using the ProteinMetrics analysis suite is outputted in a .csv file that contains tabular data on True-positive glycopeptides and their respective intensities. This data is unsorted however and requires further processing in order to obtain data on the intensities of specific intensities of each glycan class. Throughout this thesis, most of the data is classified as oligomannose, hybrid, or complex-type.

Further classification of these types is also derived depending on the type of oligomannose glycan ($Man_{3-9}GlcNAc_2$), Hybrid (Hybrid or FHybrid depending on the presence or absence of fucose), and Complex (HexNAc(3)(x) - HexNAc(6+)(x) or HexNAc(3)(F)(x) - HexNAc(6+)(F)(x) depending on the presence or absence of fucose) as well as the site specific quantity of fucose, N-Glycolylneuraminic acid (NeuGc) or N-acetyl neuraminic acid (NeuAc). Most glycan data in this thesis is from the Env glycoprotein, which contains several sites, some of which may contain hundreds of individual glycans (Fig B.1 shows a screenshot of sample raw .csv data outputted post Byologic analysis).

Such large datasets create difficulties when analysing data manually not only due to time constraints when categorising glycan species but also due to potential human error that may be introduced which necessitates a solution by which these datasets may be analysed rapidly and accurately. The following script (ASAP: Automated Site-Analysis in Python) was therefore written in Python to overcome these challenges. The requirements for running this script are an installation of Python (version 3.8 or above), the pandas and numpy packages, a glycan library (contained within a .csv file entitled Glycan_s.csv), a .csv file containing a list of the N-linked glycan sites obtained from hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html (Fig B.2) and the Byologic outputted .csv file publicly available on github.com/AlessioD77/ASAP_08-06-22.

Row#	Start	End	Var. Pos.	Sequence	Glycans	Score	Validate	XIC area	Mod.	MS1	Delta	Delta Mod.
2	55	60	58	I.HLEnVT.E	HexNAc(2)Hex(5)	55.8	True-positive	2.60E+05	NGlycan/1216.4229		55.8	55.8
3	55	60	58	I.HLEnVT.E	HexNAc(3)Hex(3)Fuc(1)	53.45 - 84.13	True-positive	3.08E+06	NGlycan/1241.4545		53.4 - 84.1	53.4 - 84.1
4	55	60	58	I.HLEnVT.E	HexNAc(4)Hex(3)	38.40 - 43.94	True-positive	5.53E+06	NGlycan/1298.4760		38.4 - 43.9	38.4 - 43.9
5	55	60	58	I.HLEnVT.E	HexNAc(4)Hex(3)Fuc(1)	58.27	True-positive	5.67E+06	NGlycan/1444.5339		58.3	58.3
6	55	60	58	I.HLEnVT.E	HexNAc(4)Hex(4)	30.00 - 57.58	True-positive	2.49E+06	NGlycan/1460.5288		30.0 - 57.6	30.0 - 57.6
7	55	60	58	I.HLEnVT.E	HexNAc(4)Hex(4)Fuc(1)	30	True-positive	3.86E+06	NGlycan/1606.5867		30	30
8	55	60	58	I.HLEnVT.E	HexNAc(4)Hex(5)Fuc(1)	61.27	True-positive	6.73E+06	NGlycan/1768.6395		61.3	61.3
9	55	60	58	I.HLEnVT.E	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	146.95	True-positive	5.49E+06	NGlycan/2059.7349		137.9	137.9
10	55	60	58	I.HLEnVT.E	HexNAc(5)Hex(3)Fuc(1)	30.00 - 52.92	True-positive	9.00E+06	NGlycan/1647.6132		30.0 - 52.9	30.0 - 52.9
11	55	60	58	I.HLEnVT.E	HexNAc(5)Hex(4)Fuc(1)	30.00 - 48.27	True-positive	1.63E+07	NGlycan/1809.6661		30.0 - 48.3	30.0 - 48.3
12	55	60	58	I.HLEnVT.E	HexNAc(5)Hex(4)Fuc(2)	30	True-positive	6.31E+06	NGlycan/1955.7240		30	30
13	55	60	58	I.HLEnVT.E	HexNAc(5)Hex(4)NeuAc(1)	30	True-positive	5.64E+05	NGlycan/1954.7036		30	30
14	56	60	58	H.LEnVT.E	HexNAc(2)Hex(4)	29.97 - 30.00	True-positive	6.70E+07	NGlycan/1054.3700		0.0 - 8.0	0.0 - 8.0
15	56	60	58	H.LEnVT.E	HexNAc(5)Hex(4)Fuc(1)	57.07	True-positive	1.94E+07	NGlycan/1809.6661		23.6	23.6
16	99	105	103	T.LQCTnVT.N	HexNAc(2)Hex(5)	174.63	True-positive	5.17E+06	NGlycan/1216.4229		76.8	76.8
17	99	105	103	T.LQCTnVT.N	HexNAc(4)Hex(3)Fuc(1)	41.14	True-positive	2.21E+06	NGlycan/1444.5339		36.8	36.8
18	107	109	107	N.nIT.D	HexNAc(3)Hex(4)	49.22	True-positive	5.31E+07	NGlycan/1257.4494		29.2	29.2
19	107	109	107	N.nIT.D	HexNAc(5)Hex(3)	30	True-positive	3.68E+05	NGlycan/1501.5553		30	30
20	107	109	107	N.nIT.D	HexNAc(6)Hex(3)Fuc(3)	87.43	True-positive	1.68E+05	NGlycan/2142.8084		66.3	66.3
21	107	109	107	N.nIT.D	HexNAc(6)Hex(9)	30.75	True-positive	1.11E+07	NGlycan/2676.9516		27.5	27.5
22	114	120	118	R.GELKnCS.F	HexNAc(2)Hex(9)	29.96 - 30.00	True-positive	3.24E+07	NGlycan/1864.6342		4.6 - 30.0	4.6 - 30.0
23	121	127	122,123	S.FnmTTEL.R	HexNAc(2)Hex(5)	30	True-positive	7.00E+04	NGlycan/1216.4229; Oxidation/15.9949		0.6	0.6
24	141	153		R.LDVVQINENQGNR.S	HexNAc(1)	22.48 - 833.61	True-positive	3.61E+10			20.5 - 766.0	20.5 - 766.0
25	141	153	152	R.LDVVQINENQGnR.S	HexNAc(1)Fuc(1)	29.97 - 677.18	True-positive	7.32E+07	NGlycan/203.0794		0.0 - 677.2	0.0 - 677.2
26	141	153	152	R.LDVVQINENQGnR.S	HexNAc(2)Hex(4)	30.00 - 628.22	True-positive	3.03E+07	NGlycan/349.1373		30.0 - 628.2	30.0 - 628.2
27	141	153	152	R.LDVVQINENQGnR.S	HexNAc(2)Hex(4)	30	True-positive	2.78E+06	NGlycan/1054.3700		30	30
28	141	153	152	R.LDVVQINENQGnR.S	HexNAc(2)Hex(5)	550.08	True-positive	1.27E+07	NGlycan/1216.4229		550.1	550.1
29	141	153	152	R.LDVVQINENQGnR.S	HexNAc(3)Hex(3)	105.95	True-positive	5.14E+06	NGlycan/1095.3966		106	106
30	141	153	152	R.LDVVQINENQGnR.S	HexNAc(3)Hex(3)Fuc(1)	29.99 - 427.72	True-positive	1.47E+07	NGlycan/1241.4545		30.0 - 427.7	30.0 - 427.7
31	141	153	152	R.LDVVQINENQGnR.S	HexNAc(4)Hex(3)Fuc(1)	422.56	True-positive	4.54E+07	NGlycan/1444.5339		422.6	422.6
32	141	153	152	R.LDVVQINENQGnR.S	HexNAc(4)Hex(4)Fuc(1)	261.14	True-positive	1.59E+07	NGlycan/1606.5867		261.1	261.1
33	141	153	152	R.LDVVQINENQGnR.S	HexNAc(4)Hex(5)Fuc(1)	254.49	True-positive	1.89E+07	NGlycan/1768.6395		254.5	254.5
34	141	153	152	R.LDVVQINENQGnR.S	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	120.76 - 218.13	True-positive	3.02E+07	NGlycan/2059.7349		120.8 - 218.1	120.8 - 218.1
35	141	153	152	R.LDVVQINENQGnR.S	HexNAc(5)Hex(3)Fuc(1)	131.03 - 305.79	True-positive	4.95E+07	NGlycan/1647.6132		131.0 - 305.8	131.0 - 305.8
36	141	153	152	R.LDVVQINENQGnR.S	HexNAc(5)Hex(4)Fuc(1)	29.99 - 214.15	True-positive	1.56E+07	NGlycan/1809.6661		17.8 - 214.2	17.8 - 214.2
37	141	153	152	R.LDVVQINENQGnR.S	HexNAc(5)Hex(4)Fuc(2)	30	True-positive	1.63E+06	NGlycan/1955.7240		30	30
38	141	153	152	R.LDVVQINENQGnR.S	HexNAc(5)Hex(5)Fuc(1)	135.84 - 157.20	True-positive	7.16E+06	NGlycan/1971.7189		135.8 - 157.2	135.8 - 157.2

Figure B.1: **Screenshot of raw glycopeptide data outputted from Byologic** The columns highlighted in yellow show the columns which are used in the script for processing by the script (Start AA, End AA, Var. Pos. Protein, Sequence, Glycans, Score, Validate, XIC area summed).

Pos	Pos after aligned to Hxb2	Hxb2 Env coords
58	58	88
103	103	133
107	112	142
118	126	156
122	130	160
152	160	185
155	163	185
167	175	197
204	212	234
232	240	262
246	254	276
265	273	295
271	279	301
301	311	332
308	318	339
324	334	355
332	342	363
355	365	386
361	371	392
367	377	398
374	385	406
379	390	411
416	427	448
430	441	462
581	592	611
588	599	618
595	606	625
607	618	637

Figure B.2: **Screenshot of N-linked glycan site list used in script** .csv file generated from inputting BG505 SOSIP.664 protein sequence in https://www.hiv.lanl.gov/content/sequence/GLYCOSITE/glycosite.html

ASAP: Automated Site Analysis in Python

```
import pandas as pd
import numpy as np

raw_df_filename = input('Enter the raw data filename: ') + '.csv'
final_filename = raw_df_filename[:-4] + '.xlsx'
pngs_filename = input('Enter the PNGS data filename: ') + '.csv'
try:
    raw_df = pd.read_csv(raw_df_filename)
    pngs = pd.read_csv(pngs_filename)
except:
    print('File cannot be opened: ', raw_df_filename)
    exit()

classification_df=pd.read_csv("ASAP/Glycan_s.csv")


filter_df = raw_df[['Start\r\nAA', 'End\r\nAA', 'Var. Pos.\r\nProtein', '
    Sequence', 'Glycans', 'Validate', 'Score', 'XIC area\r\nsummed']]

#make all sequence names upper case
filter_df['Sequence'] = filter_df['Sequence'].str.upper()

for i,rows in filter_df.iterrows():
    #print(rows['Glycans'])
#here is where the main work starts to find the right classfication
    class_name=classification_df[classification_df['Glycans'] == rows['
    Glycans']]
    #print(class_name['Species'])
#then its just a matter of assigning the values
    try:
```

```
29          filter_df.at[i,'Classification']=class_name.iloc[0]['Species']
30      except:
31          filter_df.at[i,'Classification']='Unoccupied'
32
33          #Code for assigning PNGS, remove below up to line 83 to not do
    PNGSs.
34  png_dict = dict(pngs[['Pos', 'Hxb2 Env coords']].values)
35
36  values = []
37  for i, row in filter_df.iterrows():
38      print(row['Var. Pos.\r\nProtein'])
39      sp = str(row['Var. Pos.\r\nProtein']).split(',')
40      print(sp)
41      if sp[0] == '' or sp[0] == 'nan':
42          val = np.NaN
43      elif len(sp)>1:
44          num1, num2 = sp
45          print(num1, num2)
46          try:
47              val = png_dict[int(num1)]
48          except KeyError:
49              try:
50                  val = png_dict[int(num2)]
51              except KeyError:
52                  val = np.NaN
53
54      else:
55
56          num = int(float(sp[0])) #this seems to crash the code with some
    CSVs for some reason. 03-22-22(ValueError: invalid literal for int()
    with base 10: '122.0')
57          #If the above line gives a ValueError then change to any integer,
    otherwise num = int(sp[0])
```

```
58          #UPDATE 03-23-22: num = int(float(sp[0])) seems to work.
59
60          print(num)
61          try:
62              val = png_dict[int(num)]
63          except KeyError or ValueError:
64              val = np.NaN
65      values.append(val)
66
67  filter_df['Hxb2 Env coords'] = values
68
69  filter_df = filter_df[['Start\r\nAA', 'End\r\nAA', 'Var. Pos.\r\nProtein',
        'Hxb2 Env coords', 'Sequence', 'Glycans', 'Validate', 'Score', 'XIC
       area\r\nsummed']]
70
71  unique_seq = filter_df['Sequence'].unique()
72
73  columns = filter_df.columns
74
75  writer = pd.ExcelWriter("ASAP/part1.xlsx", engine = 'xlsxwriter')
76
77  for col in unique_seq:
78      found_df = filter_df[filter_df['Sequence'] == col] #Matching based on
       column name
79      found_df[columns].to_excel(writer, sheet_name=col[:30], index=False)
80  writer.save()
81  writer.close()
82  print('Completed part 1')
83
84  part1 = pd.ExcelFile('ASAP/part1.xlsx')
85  sheet_names = part1.sheet_names
86  writer = pd.ExcelWriter("ASAP/part2.xlsx", engine = 'xlsxwriter')
87
```

```
for sheet_name in sheet_names:
    print(sheet_name)
    df = pd.read_excel("ASAP/part1.xlsx", sheet_name = sheet_name)
    print(df)
    total = df['XIC area_x000D_\nsummed'].sum() #Getting total XIC area
   summed for each site
    print(total)

    for i, rows in df.iterrows():
        df.at[i,'Percentage'] = ((rows['XIC area_x000D_\nsummed'] / total
   )*100)
    df.to_excel(writer, sheet_name = sheet_name, index = False)

writer.save()
writer.close()
print('Completed part 2')

part2 = pd.ExcelFile("ASAP/part2.xlsx")

sheet_names = part2.sheet_names
writer = pd.ExcelWriter("ASAP/part3.xlsx", engine = 'xlsxwriter')


for sheet_name in sheet_names:
    print(sheet_name)
    df = pd.read_excel("ASAP/part2.xlsx", sheet_name = sheet_name)

    for i, rows in df.iterrows():
        print(rows['Glycans'])
        #here is where the main work starts to find the right
   classfication
        class_name=classification_df[classification_df['Glycans'] == rows[
   'Glycans']]
```

```
        print(class_name['Species'])
        #then its just a matter of assigning the values
        try:
          df.at[i,'Classification']=class_name.iloc[0]['Species']
        except:
          df.at[i,'Classification']='Unoccupied'
          # continue
    df.to_excel(writer,sheet_name=sheet_name,index=False)
writer.save()
writer.close()
print("Completed part 3")

raw_df = pd.ExcelFile("ASAP/part3.xlsx")
to_calc = ['M9Glc','M9','M8','M7','M6','M5','M4','M3','FM','HYBRID','
   FHYBRID','HexNAc(3)(x)',
'HexNAc(3)(F)(x)','HexNAc(4)(x)','HexNAc(4)(F)(x)','HexNAc(5)(x)','HexNAc
   (5)(F)(x)',
'HexNAc(6+)(x)','HexNAc(6+)(F)(x)','Unoccupied','Core']
sheet_names=raw_df.sheet_names
writer = pd.ExcelWriter("ASAP/part4.xlsx", engine = 'xlsxwriter')

for sheet_name in sheet_names:
    print(sheet_name)
    df=pd.read_excel("ASAP/part3.xlsx",sheet_name=sheet_name)

    for index,calc in enumerate(to_calc):
        #i try to find if those names above(like M8) appear in the table
   and in which rows
        matched_records=df[df['Classification'] == calc]
        #if i do find 1 or more rows with it then i add them and assign it
    else its 0
        if(not matched_records.empty):
          value=matched_records['Percentage'].sum()
```

```
146             df.at[index,'Glycan class']=calc
147             df.at[index,'Class percentage']=value
148         else:
149             df.at[index,'Glycan class']=calc
150             df.at[index,'Class percentage']=0.00
151     df.to_excel(writer,sheet_name=sheet_name,index=False)
152
153 writer.save()
154 writer.close()
155 print("Completed part 4")
156
157 raw_df=pd.ExcelFile("ASAP/part4.xlsx")
158 #name of the things to calculate
159 to_calc=['Oligomannose','Hybrid','Complex','Unoccupied','Core','Fucose','
    NeuAc/NeuGc']
160 sheet_names=raw_df.sheet_names
161 writer = pd.ExcelWriter(final_filename, engine = 'xlsxwriter')
162
163 for sheet_name in sheet_names:
164     print(sheet_name)
165     df=pd.read_excel("ASAP/part4.xlsx",sheet_name=sheet_name)
166     #from oligomannose to core the calculation is done based on the row
    index
167     #as i know from which row i need to get the values
168     oli=df.iloc[1:9]['Class percentage'].sum()
169     df.at[0,'Oligomannose']=oli
170     hybrid=df.iloc[9:11]['Class percentage'].sum()
171     #print(hybrid)
172     df.at[0,'Hybrid']=hybrid
173     comple=df.iloc[11:19]['Class percentage'].sum()
174     df.at[0,'Complex']=comple
175     unocc=df.iloc[19]['Class percentage']
176     df.at[0,'Unoccupied']=unocc
```

```
177     core=df.iloc[20]['Class percentage']
178     df.at[0,'Core']=core
179 
180     #for Fucose i just try to find "F" in the string and if it exist i add
        it to total
181     total=0
182     for i,rows in df.iterrows():
183         if(str(rows['Glycan class']).find("F") != -1):
184           total+=rows['Class percentage']
185     df.at[0,'Fucose']=total
186 
187     #for Neu i do the same but with those keywords and an OR operator
188     total=0
189     for i,rows in df.iterrows():
190         if((str(rows['Glycans']).find("NeuAc") != -1) or (str(rows['
    Glycans']).find("NeuGc") != -1)):
191           print(i)
192           total+=rows['Percentage']
193     df.at[0,'NeuAc/NeuGc']=total
194     df.to_excel(writer,sheet_name=sheet_name,index=False)
195 
196 writer.save()
197 writer.close()
198 print("Completed Final report: ", final_filename)
```

The output of this script is an excel .xlsx file that contains processed data with all of the aforementioned glycan categories quantified. it is split into several tabs for each glycan site that are named by their sequence (Fig B.3).

Start AA	End AA	Var. Pos.	Protein	Sequence	Glycans	Validate	Score	x000D_sil	Percentage	Classification	Glycan class	Class percentage	Oligomannose	Hybrid	Complex	Unoccupied	Core	Fucose	NeuAc/NeuGc
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	123.31 - 76	6.23E+08	0.065986996	Unoccupied	M9Glc	M9	0	7.758450143	0.635191032	89.60577716	0.625234728	1.375346932	17.97689125	36.11595375
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.99 - 72(	8.4E+09	0.889712301	Core	M9	M9	0							
67	82		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	30.00 - 48;	2.05E+08	0.021713217	Core	M8	M8	0.040990317							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	511.51 - 7;	2.15E+09	0.227723982	Core	M7	M7	0.015569965							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.98 - 51(	1.08E+08	0.011439158	Core	M6	M6	0.127101757							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.99 - 71(	3.52E+08	0.037283182	Core	M5	M5	5.401824686							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.99 - 64(	1.77E+09	0.187475092	Core	M4	M4	1.009399789							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.98 - 79;	1.09E+10	1.154507629	M3	M3	M3	1.154507629							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.99 - 76;	9.53E+09	1.009399789	M4	FM	FM	0.009056							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	147.57	85500000	0.009056	FM	HYBRID	HYBRID	0.546537556							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.99 - 82;	5.1E+10	5.401824686	M5	FHYBRID	FHYBRID	0.088653476							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	599.99 - 6;	1.2E+09	0.127101757	M6	HexNAc(3)(x)	HexNAc(3)(x)	8.867254099							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	30.00 - 52;	1.47E+08	0.015569965	M7	HexNAc(3)(F)(x)	HexNAc(3)(F)(x)	2.438976805							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	30.00 - 65(	3.87E+08	0.040990317	M8	HexNAc(4)(x)	HexNAc(4)(x)	61.07811396							
67	82		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	138.11 - 7(	4.48E+10	4.745132273	HexNAc(3)(x)	HexNAc(4)(F)(x)	HexNAc(4)(F)(x)	14.40768325							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	114.38 - 5(	1.72E+09	0.182179185	HexNAc(3)(F)(x)	HexNAc(5)(x)	HexNAc(5)(x)	1.772942413							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	103.22 - 4(	3.85E+08	0.04077848	HexNAc(3)(F)(x)	HexNAc(5)(F)(x)	HexNAc(5)(F)(x)	1.017693179							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	183.48 - 8(	1.72E+10	1.821791855	HexNAc(3)(x)	HexNAc(6+)(x)	HexNAc(6+)(x)	0.008284916							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	131.03 - 5;	6.66E+08	0.070541475	HexNAc(3)(F)(x)	HexNAc(6+)(F)(x)	HexNAc(6+)(F)(x)	0.014828538							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.99 - 55(	7.56E+08	0.080074107	HexNAc(3)(F)(x)	Unoccupied	Unoccupied	0.625234728							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	124.14 - 6(	1.95E+10	2.065403556	HexNAc(3)(F)(x)	Core	Core	1.375346932							
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	223.41 - 7(	2.09E+10	2.21368894	HexNAc(3)(x)										
67	82		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	459.40 - 5(	8.18E+08	0.086641031	HexNAc(3)(x)										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	30.00 - 65;	1.23E+09	0.130279301	HYBRID										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	72.29 - 49;	8.37E+08	0.088653476	FHYBRID										
67	82		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	546.63 - 6;	1.9E+09	0.20124449	HYBRID										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	542.33 - 7(	1.68E+09	0.17794246	HYBRID										
67	82		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	679.68	3.5E+08	0.037071346	HYBRID										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	30.00 - 72(	1.22E+11	12.92201199	HexNAc(4)(x)										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	30.00 - 65(	1.45E+10	1.535812901	HexNAc(4)(F)(x)										
67	82		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	329.30 - 7(	9.68E+10	10.25287509	HexNAc(4)(x)										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.99 - 61(	1.51E+10	1.599363779	HexNAc(4)(F)(x)										
67	82		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	168.71 - 4(	1.27E+10	1.345160265	HexNAc(4)(F)(x)										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	30.00 - 44(	3.89E+10	4.120215299	HexNAc(4)(F)(x)										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	29.99 - 72(	9.3E+10	9.850386191	HexNAc(4)(x)										
67	87		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	510.64 - 5(	3.79E+09	0.401429717	HexNAc(4)(x)										
67	82		H.ACVPTDPNPQEVV	H.ACVPTDPNPQEVV	True-posi	187.77 - 6(	1.04E+11	11.01584583	HexNAc(4)(x)										

Figure B.3: **Screenshot of raw glycopeptide data outputted from Byologic** The columns highlighted in yellow show the columns which are used in the script for processing by the script (Start AA, End AA, Var. Pos. Protein, Sequence, Glycans, Score, Validate, XIC area summed).

C Glycan library used for analysis

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
1	0	0	0	0	0	0	Core
1	0	1	0	0	0	0	Core
2	0	0	0	0	0	0	Core
2	0	1	0	0	0	0	Core
2	1	0	0	0	0	0	Core
2	1	1	0	0	0	0	Core
2	2	0	0	0	0	0	Core
2	2	1	0	0	0	0	Core
2	3	0	0	0	0	0	M3
2	4	0	0	0	0	0	M4
2	5	0	0	0	0	0	M5
2	5	0	0	0	0	1	M5
2	6	0	0	0	0	0	M6
2	6	0	0	0	0	1	M6
2	7	0	0	0	0	0	M7
2	7	0	0	0	0	1	M7
2	8	0	0	0	0	0	M8
2	8	0	0	0	0	1	M8
2	9	0	0	0	0	0	M9
2	9	0	0	0	0	1	M9
2	10	0	0	0	0	0	M9GLC
2	10	0	0	0	1	0	M9GLC
2	10	0	0	0	2	0	M9GLC

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
2	10	0	0	0	3	0	M9GLC
2	11	0	0	0	0	0	M9GLC
2	12	0	0	0	0	0	M9GLC
2	3	1	0	0	0	0	FM
2	3	2	0	0	0	0	FM
2	4	1	0	0	0	0	FM
2	5	1	0	0	0	0	FM
2	5	1	0	0	1	0	FM
2	5	1	0	0	2	0	FM
2	5	1	0	0	3	0	FM
3	5	0	0	0	0	0	HYBRID
3	5	0	1	0	0	0	HYBRID
3	5	0	0	1	0	0	HYBRID
3	6	0	0	0	0	0	HYBRID
3	6	0	1	0	0	0	HYBRID
3	6	0	1	0	1	0	HYBRID
3	6	0	1	0	2	0	HYBRID
3	6	0	1	0	3	0	HYBRID
3	6	0	0	1	0	0	HYBRID
3	6	0	0	0	1	0	HYBRID
3	6	0	0	0	2	0	HYBRID
3	6	0	0	0	3	0	HYBRID
3	5	1	0	0	0	0	FHYBRID
3	5	1	0	0	1	0	FHYBRID
3	5	1	0	0	2	0	FHYBRID
3	5	1	0	0	3	0	FHYBRID

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
3	6	1	0	0	0	0	FHYBRID
3	6	1	1	0	0	0	FHYBRID
3	6	1	0	1	0	0	FHYBRID
3	6	1	0	0	1	0	FHYBRID
3	6	1	0	0	2	0	FHYBRID
3	6	1	0	0	3	0	FHYBRID
3	3	0	0	0	0	0	HexNAc(3)(x)
3	4	0	0	0	0	0	HexNAc(3)(x)
3	4	0	1	0	0	0	HexNAc(3)(x)
3	4	0	0	1	0	0	HexNAc(3)(x)
3	4	0	0	0	1	0	HexNAc(3)(x)
3	4	0	0	0	2	0	HexNAc(3)(x)
3	4	0	0	0	3	0	HexNAc(3)(x)
3	3	1	0	0	0	0	HexNAc(3)(F)(x)
3	3	1	0	0	1	0	HexNAc(3)(F)(x)
3	3	1	0	0	2	0	HexNAc(3)(F)(x)
3	3	1	0	0	3	0	HexNAc(3)(F)(x)
3	3	2	0	0	0	0	HexNAc(3)(F)(x)
3	3	2	0	0	1	0	HexNAc(3)(F)(x)
3	3	2	0	0	2	0	HexNAc(3)(F)(x)
3	3	2	0	0	3	0	HexNAc(3)(F)(x)
3	4	1	0	0	0	0	HexNAc(3)(F)(x)
3	4	1	1	0	0	0	HexNAc(3)(F)(x)
3	4	1	1	0	1	0	HexNAc(3)(F)(x)
3	4	1	1	0	2	0	HexNAc(3)(F)(x)
3	4	1	1	0	3	0	HexNAc(3)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
3	4	1	0	1	0	0	HexNAc(3)(F)(x)
3	4	1	0	0	3	0	HexNAc(3)(F)(x)
3	4	2	0	0	0	0	HexNAc(3)(F)(x)
3	4	2	0	0	1	0	HexNAc(3)(F)(x)
3	4	2	0	0	2	0	HexNAc(3)(F)(x)
3	4	2	0	0	3	0	HexNAc(3)(F)(x)
4	3	0	0	0	0	0	HexNAc(4)(x)
4	3	0	1	0	0	0	HexNAc(4)(x)
4	3	0	0	1	0	0	HexNAc(4)(x)
4	3	0	0	0	1	0	HexNAc(4)(x)
4	3	0	0	0	2	0	HexNAc(4)(x)
4	3	0	0	0	3	0	HexNAc(4)(x)
4	4	0	0	0	0	0	HexNAc(4)(x)
4	4	0	1	0	0	0	HexNAc(4)(x)
4	4	0	0	1	0	0	HexNAc(4)(x)
4	5	0	0	0	0	0	HexNAc(4)(x)
4	5	0	1	0	0	0	HexNAc(4)(x)
4	5	0	1	1	0	0	HexNAc(4)(x)
4	5	0	2	0	0	0	HexNAc(4)(x)
4	5	0	0	1	0	0	HexNAc(4)(x)
4	5	0	0	2	0	0	HexNAc(4)(x)
4	5	0	0	0	1	0	HexNAc(4)(x)
4	5	0	0	0	2	0	HexNAc(4)(x)
4	5	0	0	0	3	0	HexNAc(4)(x)
4	6	0	0	0	0	0	HexNAc(4)(x)
4	6	0	1	0	0	0	HexNAc(4)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
4	6	0	0	1	0	0	HexNAc(4)(x)
4	7	0	0	0	0	0	HexNAc(4)(x)
4	7	0	1	0	0	0	HexNAc(4)(x)
4	7	0	0	0	1	0	HexNAc(4)(x)
4	7	0	0	0	2	0	HexNAc(4)(x)
4	7	0	0	0	3	0	HexNAc(4)(x)
4	3	1	0	0	0	0	HexNAc(4)(F)(x)
4	3	2	0	0	0	0	HexNAc(4)(F)(x)
4	3	3	0	0	0	0	HexNAc(4)(F)(x)
4	4	1	0	0	0	0	HexNAc(4)(F)(x)
4	4	1	1	0	0	0	HexNAc(4)(F)(x)
4	4	1	0	1	0	0	HexNAc(4)(F)(x)
4	4	1	0	0	1	0	HexNAc(4)(F)(x)
4	4	1	0	0	2	0	HexNAc(4)(F)(x)
4	4	1	0	0	3	0	HexNAc(4)(F)(x)
4	4	2	0	0	0	0	HexNAc(4)(F)(x)
4	4	2	1	0	0	0	HexNAc(4)(F)(x)
4	4	2	0	1	0	0	HexNAc(4)(F)(x)
4	4	2	0	0	1	0	HexNAc(4)(F)(x)
4	4	2	0	0	2	0	HexNAc(4)(F)(x)
4	4	2	0	0	3	0	HexNAc(4)(F)(x)
4	5	1	0	0	0	0	HexNAc(4)(F)(x)
4	5	1	1	0	0	0	HexNAc(4)(F)(x)
4	5	1	1	1	0	0	HexNAc(4)(F)(x)
4	5	1	2	0	0	0	HexNAc(4)(F)(x)
4	5	1	0	1	0	0	HexNAc(4)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
4	5	1	0	2	0	0	HexNAc(4)(F)(x)
4	5	2	0	0	0	0	HexNAc(4)(F)(x)
4	5	2	1	0	0	0	HexNAc(4)(F)(x)
4	5	2	1	0	1	0	HexNAc(4)(F)(x)
4	5	2	1	0	2	0	HexNAc(4)(F)(x)
4	5	2	1	0	3	0	HexNAc(4)(F)(x)
4	5	2	0	1	0	0	HexNAc(4)(F)(x)
4	5	3	0	0	0	0	HexNAc(4)(F)(x)
4	5	3	1	0	0	0	HexNAc(4)(F)(x)
4	5	3	1	1	0	0	HexNAc(4)(F)(x)
4	5	3	2	0	0	0	HexNAc(4)(F)(x)
4	5	3	0	2	0	0	HexNAc(4)(F)(x)
4	5	3	0	0	1	0	HexNAc(4)(F)(x)
4	5	3	0	0	2	0	HexNAc(4)(F)(x)
4	5	3	0	0	3	0	HexNAc(4)(F)(x)
4	5	4	0	0	0	0	HexNAc(4)(F)(x)
4	6	1	0	0	0	0	HexNAc(4)(F)(x)
4	6	1	1	0	0	0	HexNAc(4)(F)(x)
4	6	1	0	1	0	0	HexNAc(4)(F)(x)
4	6	1	0	0	1	0	HexNAc(4)(F)(x)
4	6	1	0	0	2	0	HexNAc(4)(F)(x)
4	6	1	0	0	3	0	HexNAc(4)(F)(x)
4	6	2	0	0	0	0	HexNAc(4)(F)(x)
4	6	3	0	0	0	0	HexNAc(4)(F)(x)
4	7	1	0	0	0	0	HexNAc(4)(F)(x)
4	7	2	0	0	0	0	HexNAc(4)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
5	3	0	0	0	0	0	HexNAc(5)(x)
5	3	0	0	0	1	0	HexNAc(5)(x)
5	3	0	0	0	2	0	HexNAc(5)(x)
5	3	0	0	0	3	0	HexNAc(5)(x)
5	4	0	0	0	0	0	HexNAc(5)(x)
5	4	0	1	0	0	0	HexNAc(5)(x)
5	4	0	1	1	0	0	HexNAc(5)(x)
5	4	0	2	0	0	0	HexNAc(5)(x)
5	4	0	0	1	0	0	HexNAc(5)(x)
5	4	0	0	2	0	0	HexNAc(5)(x)
5	4	0	0	0	1	0	HexNAc(5)(x)
5	4	0	0	0	2	0	HexNAc(5)(x)
5	4	0	0	0	3	0	HexNAc(5)(x)
5	5	0	0	0	0	0	HexNAc(5)(x)
5	5	0	1	0	0	0	HexNAc(5)(x)
5	5	0	1	1	0	0	HexNAc(5)(x)
5	5	0	1	0	1	0	HexNAc(5)(x)
5	5	0	1	0	2	0	HexNAc(5)(x)
5	5	0	1	0	3	0	HexNAc(5)(x)
5	5	0	2	0	0	0	HexNAc(5)(x)
5	5	0	0	1	0	0	HexNAc(5)(x)
5	5	0	0	2	0	0	HexNAc(5)(x)
5	6	0	0	0	0	0	HexNAc(5)(x)
5	6	0	1	0	0	0	HexNAc(5)(x)
5	6	0	1	1	0	0	HexNAc(5)(x)
5	6	0	1	2	0	0	HexNAc(5)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
5	6	0	2	0	0	0	HexNAc(5)(x)
5	6	0	2	1	0	0	HexNAc(5)(x)
5	6	0	3	0	0	0	HexNAc(5)(x)
5	6	0	3	0	1	0	HexNAc(5)(x)
5	6	0	3	0	2	0	HexNAc(5)(x)
5	6	0	3	0	3	0	HexNAc(5)(x)
5	6	0	0	1	0	0	HexNAc(5)(x)
5	6	0	0	2	0	0	HexNAc(5)(x)
5	6	0	0	3	0	0	HexNAc(5)(x)
5	7	0	0	0	0	0	HexNAc(5)(x)
5	8	0	0	0	0	0	HexNAc(5)(x)
5	3	1	0	0	0	0	HexNAc(5)(F)(x)
5	3	1	1	0	0	0	HexNAc(5)(F)(x)
5	3	1	0	1	0	0	HexNAc(5)(F)(x)
5	3	1	0	0	1	0	HexNAc(5)(F)(x)
5	3	1	0	0	2	0	HexNAc(5)(F)(x)
5	3	1	0	0	3	0	HexNAc(5)(F)(x)
5	4	1	0	0	0	0	HexNAc(5)(F)(x)
5	4	1	1	0	0	0	HexNAc(5)(F)(x)
5	4	1	1	1	0	0	HexNAc(5)(F)(x)
5	4	1	1	0	1	0	HexNAc(5)(F)(x)
5	4	1	1	0	2	0	HexNAc(5)(F)(x)
5	4	1	1	0	3	0	HexNAc(5)(F)(x)
5	4	1	2	0	0	0	HexNAc(5)(F)(x)
5	4	1	0	1	0	0	HexNAc(5)(F)(x)
5	4	1	0	2	0	0	HexNAc(5)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
5	4	1	0	0	1	0	HexNAc(5)(F)(x)
5	4	1	0	0	2	0	HexNAc(5)(F)(x)
5	4	1	0	0	3	0	HexNAc(5)(F)(x)
5	4	2	0	0	0	0	HexNAc(5)(F)(x)
5	4	2	1	0	0	0	HexNAc(5)(F)(x)
5	4	2	0	1	0	0	HexNAc(5)(F)(x)
5	4	2	0	0	1	0	HexNAc(5)(F)(x)
5	4	2	0	0	2	0	HexNAc(5)(F)(x)
5	4	2	0	0	3	0	HexNAc(5)(F)(x)
5	5	1	0	0	0	0	HexNAc(5)(F)(x)
5	5	1	1	0	0	0	HexNAc(5)(F)(x)
5	5	1	1	1	0	0	HexNAc(5)(F)(x)
5	5	1	2	0	0	0	HexNAc(5)(F)(x)
5	5	1	0	1	0	0	HexNAc(5)(F)(x)
5	5	1	0	2	0	0	HexNAc(5)(F)(x)
5	5	2	0	0	0	0	HexNAc(5)(F)(x)
5	5	2	0	0	1	0	HexNAc(5)(F)(x)
5	5	2	0	0	2	0	HexNAc(5)(F)(x)
5	5	2	0	0	3	0	HexNAc(5)(F)(x)
5	5	3	0	0	0	0	HexNAc(5)(F)(x)
5	6	1	0	0	0	0	HexNAc(5)(F)(x)
5	6	1	1	0	0	0	HexNAc(5)(F)(x)
5	6	1	1	1	0	0	HexNAc(5)(F)(x)
5	6	1	1	2	0	0	HexNAc(5)(F)(x)
5	6	1	2	0	0	0	HexNAc(5)(F)(x)
5	6	1	2	1	0	0	HexNAc(5)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
5	6	1	3	0	0	0	HexNAc(5)(F)(x)
5	6	1	0	1	0	0	HexNAc(5)(F)(x)
5	6	1	0	2	0	0	HexNAc(5)(F)(x)
5	6	1	0	3	0	0	HexNAc(5)(F)(x)
5	6	2	0	0	0	0	HexNAc(5)(F)(x)
5	6	2	1	0	0	0	HexNAc(5)(F)(x)
5	6	2	0	1	0	0	HexNAc(5)(F)(x)
5	6	3	0	0	0	0	HexNAc(5)(F)(x)
5	6	3	1	0	0	0	HexNAc(5)(F)(x)
5	6	3	0	1	0	0	HexNAc(5)(F)(x)
5	7	1	0	0	0	0	HexNAc(5)(F)(x)
5	7	1	1	0	0	0	HexNAc(5)(F)(x)
5	7	1	1	1	0	0	HexNAc(5)(F)(x)
5	7	1	1	0	1	0	HexNAc(5)(F)(x)
5	7	1	1	0	2	0	HexNAc(5)(F)(x)
5	7	1	1	0	3	0	HexNAc(5)(F)(x)
5	7	1	2	0	0	0	HexNAc(5)(F)(x)
5	7	1	0	2	0	0	HexNAc(5)(F)(x)
5	8	1	0	0	0	0	HexNAc(5)(F)(x)
5	8	4	0	0	0	0	HexNAc(5)(F)(x)
5	9	1	0	0	0	0	HexNAc(5)(F)(x)
6	3	0	0	0	0	0	HexNAc(6+)(x)
6	4	0	0	0	0	0	HexNAc(6+)(x)
6	4	0	1	0	0	0	HexNAc(6+)(x)
6	4	0	0	0	1	0	HexNAc(6+)(x)
6	4	0	0	0	2	0	HexNAc(6+)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
6	4	0	0	0	3	0	HexNAc(6+)(x)
6	5	0	0	0	0	0	HexNAc(6+)(x)
6	6	0	0	0	0	0	HexNAc(6+)(x)
6	6	0	1	0	0	0	HexNAc(6+)(x)
6	6	0	1	1	0	0	HexNAc(6+)(x)
6	6	0	2	0	0	0	HexNAc(6+)(x)
6	6	0	0	2	0	0	HexNAc(6+)(x)
6	7	0	0	0	0	0	HexNAc(6+)(x)
6	7	0	1	0	0	0	HexNAc(6+)(x)
6	7	0	1	1	0	0	HexNAc(6+)(x)
6	7	0	2	0	0	0	HexNAc(6+)(x)
6	7	0	2	1	0	0	HexNAc(6+)(x)
6	7	0	2	2	0	0	HexNAc(6+)(x)
6	7	0	3	0	0	0	HexNAc(6+)(x)
6	7	0	3	1	0	0	HexNAc(6+)(x)
6	7	0	4	0	0	0	HexNAc(6+)(x)
6	7	0	0	1	0	0	HexNAc(6+)(x)
6	7	0	0	2	0	0	HexNAc(6+)(x)
6	8	0	1	0	0	0	HexNAc(6+)(x)
6	9	0	0	0	0	0	HexNAc(6+)(x)
7	3	0	0	0	0	0	HexNAc(6+)(x)
7	4	0	0	0	0	0	HexNAc(6+)(x)
7	6	0	0	0	0	0	HexNAc(6+)(x)
7	7	0	0	0	0	0	HexNAc(6+)(x)
7	8	0	0	0	0	0	HexNAc(6+)(x)
7	8	0	1	0	0	0	HexNAc(6+)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
7	8	0	0	1	0	0	HexNAc(6+)(x)
7	8	0	0	0	1	0	HexNAc(6+)(x)
7	8	0	0	0	2	0	HexNAc(6+)(x)
7	8	0	0	0	3	0	HexNAc(6+)(x)
8	3	0	0	0	0	0	HexNAc(6+)(x)
8	4	0	0	0	0	0	HexNAc(6+)(x)
8	5	0	0	0	0	0	HexNAc(6+)(x)
8	5	0	0	0	1	0	HexNAc(6+)(x)
8	5	0	0	0	2	0	HexNAc(6+)(x)
8	5	0	0	0	3	0	HexNAc(6+)(x)
8	6	0	0	0	0	0	HexNAc(6+)(x)
8	7	0	0	0	0	0	HexNAc(6+)(x)
8	8	0	0	0	0	0	HexNAc(6+)(x)
8	9	0	0	0	0	0	HexNAc(6+)(x)
9	10	0	0	0	0	0	HexNAc(6+)(x)
9	3	0	0	0	0	0	HexNAc(6+)(x)
9	4	0	0	0	0	0	HexNAc(6+)(x)
9	6	0	0	0	0	0	HexNAc(6+)(x)
10	7	0	0	0	0	0	HexNAc(6+)(x)
11	11	0	1	0	0	0	HexNAc(6+)(x)
11	11	0	0	1	0	0	HexNAc(6+)(x)
6	3	1	0	0	0	0	HexNAc(6+)(F)(x)
6	3	1	1	0	0	0	HexNAc(6+)(F)(x)
6	3	1	1	1	0	0	HexNAc(6+)(F)(x)
6	3	1	1	0	1	0	HexNAc(6+)(F)(x)
6	3	1	1	0	2	0	HexNAc(6+)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
6	3	1	1	0	3	0	HexNAc(6+)(F)(x)
6	3	1	2	0	0	0	HexNAc(6+)(F)(x)
6	3	1	0	1	0	0	HexNAc(6+)(F)(x)
6	3	1	0	2	0	0	HexNAc(6+)(F)(x)
6	3	1	0	0	1	0	HexNAc(6+)(F)(x)
6	3	1	0	0	2	0	HexNAc(6+)(F)(x)
6	3	1	0	0	3	0	HexNAc(6+)(F)(x)
6	3	2	0	0	0	0	HexNAc(6+)(F)(x)
6	3	2	1	0	0	0	HexNAc(6+)(F)(x)
6	3	2	0	1	0	0	HexNAc(6+)(F)(x)
6	3	2	0	0	1	0	HexNAc(6+)(F)(x)
6	3	2	0	0	2	0	HexNAc(6+)(F)(x)
6	3	2	0	0	3	0	HexNAc(6+)(F)(x)
6	3	3	0	0	0	0	HexNAc(6+)(F)(x)
6	3	3	0	0	1	0	HexNAc(6+)(F)(x)
6	3	3	0	0	2	0	HexNAc(6+)(F)(x)
6	3	3	0	0	3	0	HexNAc(6+)(F)(x)
6	4	1	0	0	0	0	HexNAc(6+)(F)(x)
6	4	1	0	0	1	0	HexNAc(6+)(F)(x)
6	4	1	0	0	2	0	HexNAc(6+)(F)(x)
6	4	1	0	0	3	0	HexNAc(6+)(F)(x)
6	4	2	0	0	0	0	HexNAc(6+)(F)(x)
6	4	2	0	0	1	0	HexNAc(6+)(F)(x)
6	4	2	0	0	2	0	HexNAc(6+)(F)(x)
6	4	2	0	0	3	0	HexNAc(6+)(F)(x)
6	5	1	0	0	0	0	HexNAc(6+)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
6	5	1	1	0	0	0	HexNAc(6+)(F)(x)
6	5	1	1	1	0	0	HexNAc(6+)(F)(x)
6	5	1	1	2	0	0	HexNAc(6+)(F)(x)
6	5	1	1	0	1	0	HexNAc(6+)(F)(x)
6	5	1	1	0	2	0	HexNAc(6+)(F)(x)
6	5	1	1	0	3	0	HexNAc(6+)(F)(x)
6	5	1	2	0	0	0	HexNAc(6+)(F)(x)
6	5	1	2	1	0	0	HexNAc(6+)(F)(x)
6	5	1	2	0	1	0	HexNAc(6+)(F)(x)
6	5	1	2	0	2	0	HexNAc(6+)(F)(x)
6	5	1	2	0	3	0	HexNAc(6+)(F)(x)
6	5	1	3	0	0	0	HexNAc(6+)(F)(x)
6	5	1	0	3	0	0	HexNAc(6+)(F)(x)
6	5	1	0	0	1	0	HexNAc(6+)(F)(x)
6	5	1	0	0	2	0	HexNAc(6+)(F)(x)
6	5	1	0	0	3	0	HexNAc(6+)(F)(x)
6	5	2	0	0	0	0	HexNAc(6+)(F)(x)
6	5	2	0	0	1	0	HexNAc(6+)(F)(x)
6	5	2	0	0	2	0	HexNAc(6+)(F)(x)
6	5	2	0	0	3	0	HexNAc(6+)(F)(x)
6	6	1	0	0	0	0	HexNAc(6+)(F)(x)
6	6	1	1	0	0	0	HexNAc(6+)(F)(x)
6	6	1	1	1	0	0	HexNAc(6+)(F)(x)
6	6	1	1	2	0	0	HexNAc(6+)(F)(x)
6	6	1	2	0	0	0	HexNAc(6+)(F)(x)
6	6	1	2	1	0	0	HexNAc(6+)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
6	6	1	3	0	0	0	HexNAc(6+)(F)(x)
6	6	1	0	2	0	0	HexNAc(6+)(F)(x)
6	6	1	0	3	0	0	HexNAc(6+)(F)(x)
6	6	2	0	0	0	0	HexNAc(6+)(F)(x)
6	6	3	0	0	0	0	HexNAc(6+)(F)(x)
6	7	1	0	0	0	0	HexNAc(6+)(F)(x)
6	7	1	1	0	0	0	HexNAc(6+)(F)(x)
6	7	1	2	0	0	0	HexNAc(6+)(F)(x)
6	7	1	3	0	0	0	HexNAc(6+)(F)(x)
6	7	1	3	1	0	0	HexNAc(6+)(F)(x)
6	7	1	4	0	0	0	HexNAc(6+)(F)(x)
6	7	1	0	1	0	0	HexNAc(6+)(F)(x)
6	7	4	0	0	0	0	HexNAc(6+)(F)(x)
6	7	4	1	0	0	0	HexNAc(6+)(F)(x)
6	7	4	0	1	0	0	HexNAc(6+)(F)(x)
6	7	4	0	0	1	0	HexNAc(6+)(F)(x)
6	7	5	0	0	0	0	HexNAc(6+)(F)(x)
6	8	1	1	0	0	0	HexNAc(6+)(F)(x)
6	9	1	1	1	0	0	HexNAc(6+)(F)(x)
6	9	1	2	0	0	0	HexNAc(6+)(F)(x)
6	10	1	1	0	0	0	HexNAc(6+)(F)(x)
6	10	1	0	1	0	0	HexNAc(6+)(F)(x)
6	11	1	0	0	0	0	HexNAc(6+)(F)(x)
7	3	1	0	0	0	0	HexNAc(6+)(F)(x)
7	3	1	0	0	1	0	HexNAc(6+)(F)(x)
7	4	1	0	0	0	0	HexNAc(6+)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
7	4	1	0	0	1	0	HexNAc(6+)(F)(x)
7	4	1	0	0	2	0	HexNAc(6+)(F)(x)
7	4	1	0	0	3	0	HexNAc(6+)(F)(x)
7	4	2	0	0	0	0	HexNAc(6+)(F)(x)
7	4	2	0	0	1	0	HexNAc(6+)(F)(x)
7	4	2	0	0	2	0	HexNAc(6+)(F)(x)
7	4	2	0	0	3	0	HexNAc(6+)(F)(x)
7	6	1	0	0	0	0	HexNAc(6+)(F)(x)
7	6	1	3	1	0	0	HexNAc(6+)(F)(x)
7	6	1	4	0	0	0	HexNAc(6+)(F)(x)
7	7	1	0	0	0	0	HexNAc(6+)(F)(x)
7	7	1	1	2	0	0	HexNAc(6+)(F)(x)
7	7	1	2	0	0	0	HexNAc(6+)(F)(x)
7	7	1	2	1	0	0	HexNAc(6+)(F)(x)
7	7	1	2	0	1	0	HexNAc(6+)(F)(x)
7	7	1	2	0	2	0	HexNAc(6+)(F)(x)
7	7	1	2	0	3	0	HexNAc(6+)(F)(x)
7	7	1	3	0	0	0	HexNAc(6+)(F)(x)
7	7	1	4	0	0	0	HexNAc(6+)(F)(x)
7	8	1	0	0	0	0	HexNAc(6+)(F)(x)
7	8	1	1	0	0	0	HexNAc(6+)(F)(x)
7	8	1	1	1	0	0	HexNAc(6+)(F)(x)
7	8	1	1	2	0	0	HexNAc(6+)(F)(x)
7	8	1	2	0	0	0	HexNAc(6+)(F)(x)
7	8	1	2	1	0	0	HexNAc(6+)(F)(x)
7	8	1	3	0	0	0	HexNAc(6+)(F)(x)

HexNAc	Hex	Fuc	NeuAc	NeuGc	Sulfo	Phospho	Classification
7	8	1	4	0	0	0	HexNAc(6+)(F)(x)
7	8	1	0	2	0	0	HexNAc(6+)(F)(x)
7	8	1	0	3	0	0	HexNAc(6+)(F)(x)
8	3	1	0	0	0	0	HexNAc(6+)(F)(x)
8	5	1	0	0	0	0	HexNAc(6+)(F)(x)
8	8	1	0	0	0	0	HexNAc(6+)(F)(x)
8	9	1	0	0	0	0	HexNAc(6+)(F)(x)
8	9	1	1	2	0	0	HexNAc(6+)(F)(x)
8	9	1	2	1	0	0	HexNAc(6+)(F)(x)
8	9	1	3	0	0	0	HexNAc(6+)(F)(x)
8	9	1	4	0	0	0	HexNAc(6+)(F)(x)
9	10	1	0	0	0	0	HexNAc(6+)(F)(x)
9	10	1	4	0	0	0	HexNAc(6+)(F)(x)
9	3	1	0	0	0	0	HexNAc(6+)(F)(x)
9	4	1	0	0	0	0	HexNAc(6+)(F)(x)
9	6	1	0	0	0	0	HexNAc(6+)(F)(x)
9	9	1	0	0	0	0	HexNAc(6+)(F)(x)
10	10	1	0	0	0	0	HexNAc(6+)(F)(x)

D Publications and confrerences

Publications

- Gemma E.Seabright, Christopher A.Cottrell, Marit J.van Gils, **Alessio D'addabbo**, David J.Harvey, Anna-Janina Behrens, Joel D.Allen, Yasunori Watanabe, Nicole Scaringi Thomas, M. Polveroni, Allison Maker, Snezana Vasiljevic, Nataliade Val, Rogier W.Sanders, Andrew B. Ward, Max Crispin "Networks of HIV-1 Envelope Glycans Maintain Antibody Epitopes in the Face of Glycan Additions and Deletions". In:Structure(2020).issn:18784186.doi:10.1016/j.str.2020.04.022.

- Emma T Crooks, Francisco Almanza, **Alessio D'addabbo**, Erika Duggan, Jinsong Zhang, Kshitij Wagh, Huihui Mou, Joel D Allen, Alyssa Thomas, Keiko Osawa, Bette T Korber, Yaroslav Tsybovsky, Evan Cale, John Nolan, Max Crispin, Laurent K Verkoczy, James M Binley "Engineering well-expressed, V2-immunofocusing HIV-1 envelope glycoprotein membrane trimers for use in heterologous prime-boost vaccine regimens".
  In:PLOS Pathogens(2021).issn: 15537374.doi:10.1371/journal.ppat.1009807.

- Tommy Tong, **Alessio D'Addabbo**, Jiamin Xu, Himanshi Chawla, Albert Nguyen, Paola Ochoa, Max Crispin, James M. Binley "Impact of stabilizing mutations on the antigenic profile and glycosylation of membrane-expressed HIV-1 envelope glycoprotein".
  In: PLOS Pathogens(2023).issn:e1011452.doi:10.1371/journal.ppat.1011452

Conferences

Date	Conference attended
09/2019	Protein Metrics 2019 EU users meeting; Geneva, Switzerland
10/2019	European aids vaccine initiative (EAVI) Annnual meeting; Paris, France
02/2020	Scripps Consortium for AIDS Vaccine Development (CHAVD); La Jolla, USA
11/2020	EAVI PhD training course: HIV glycoprotein engineering for vaccines Online, hosted by the Academic Medical Centre (AMC), University of Amsterdam
02/2022	European aids vaccine initiative (EAVI) Annnual meeting; Brussels, Belgium

Bibliography

[1] Nuno R Faria et al. "HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations." In: *Science (New York, N.Y.)* 346.6205 (2014), pp. 56–61. ISSN: 1095-9203. DOI: 10.1126/science.1256739. URL: http://www.ncbi.nlm.nih.gov/pubmed/25278604.

[2] UNAIDS. "2017 GLOBAL HIV STATISTICS". In: (2018), pp. 1–6. URL: http://www.unaids.org/en/resources/fact-sheet.

[3] Elizabeth Fee and Theodore M. Brown. "Michael S. Gottlieb and the Identification of AIDS". In: *American Journal of Public Health* 96.6 (June 2006), pp. 982–983. ISSN: 0090-0036. DOI: 10.2105/AJPH.2006.088435. URL: http://ajph.aphapublications.org/doi/10.2105/AJPH.2006.088435.

[4] Michael S. Gottlieb et al. "Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men". In: *New England Journal of Medicine* 305.24 (Dec. 1981), pp. 1425–1431. ISSN: 0028-4793. DOI: 10.1056/NEJM198112103052401. URL: http://www.ncbi.nlm.nih.gov/pubmed/6272109.

[5] Michael S Gottlieb. "Pneumocystis pneumonia–Los Angeles. 1981." In: *American journal of public health* 96.6 (June 2006), pp. 980–1. ISSN: 0090-0036. URL: http://www.ncbi.nlm.nih.gov/pubmed/16714472.

[6] The Antiretroviral Therapy Cohort Antiretroviral Therapy Cohort Collaboration. "Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies." In: *Lancet (London, England)* 372.9635 (July 2008), pp. 293–9. ISSN: 1474-547X. DOI: 10.1016/S0140-6736(08)61113-7. URL: http://www.ncbi.nlm.nih.gov/pubmed/18657708.

[7] Rajesh Kumar et al. "Broadly neutralizing antibodies in HIV-1 treatment and prevention." In: *Therapeutic advances in vaccines and immunotherapy* 6.4 (Aug. 2018), pp. 61–68. ISSN: 2515-1355. DOI: 10.1177/2515135518800689. URL: http://www.ncbi.nlm.nih.gov/pubmed/30345419.

[8] Paul M Sharp and Beatrice H Hahn. "Origins of HIV and the AIDS pandemic." In: *Cold Spring Harbor perspectives in medicine* 1.1 (Sept. 2011), a006841. ISSN: 2157-1422. DOI: 10.1101/cshperspect.a006841. URL: http://www.ncbi.nlm.nih.gov/pubmed/22229120.

[9] Martine Peeters et al. "Risk to Human Health from a Plethora of Simian Immunodeficiency Viruses in Primate Bushmeat". In: *Emerging Infectious Diseases* 8.5 (May 2002), pp. 451–457. ISSN: 1080-6040. DOI: 10.3201/eid0805.010522. URL: http://www.ncbi.nlm.nih.gov/pubmed/11996677.

[10] Omobolaji T Campbell-Yesufu and Rajesh T Gandhi. "Update on human immunodeficiency virus (HIV)-2 infection." In: *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 52.6 (Mar. 2011), pp. 780–7. ISSN: 1537-6591. DOI: 10.1093/cid/ciq248. URL: http://www.ncbi.nlm.nih.gov/pubmed/21367732.

[11] R Marlink et al. "Reduced rate of disease development after HIV-2 infection as compared to HIV-1." In: *Science (New York, N.Y.)* 265.5178 (Sept. 1994), pp. 1587–90. ISSN: 0036-8075. URL: http://www.ncbi.nlm.nih.gov/pubmed/7915856.

[12] François Simon et al. "Identification of a new human immunodeficiency virus type 1 distinct from group M and group O". In: *Nature Medicine* 4.9 (Sept. 1998), pp. 1032–1037. ISSN: 1078-8956. DOI: 10.1038/2017. URL: http://www.nature.com/articles/nm0998_1032.

[13] Jean-Christophe Plantier et al. "A new human immunodeficiency virus derived from gorillas". In: *Nature Medicine* 15.8 (Aug. 2009), pp. 871–872. ISSN: 1078-8956. DOI: 10.1038/nm.2016. URL: http://www.nature.com/articles/nm.2016.

[14] Ana Vallari et al. "Confirmation of putative HIV-1 group P in Cameroon." In: *Journal of virology* 85.3 (Feb. 2011), pp. 1403–7. ISSN: 1098-5514. DOI: 10.1128/JVI.02005-10. URL: http://www.ncbi.nlm.nih.gov/pubmed/21084486.

[15] L. Buonaguro, M. L. Tornesello, and F. M. Buonaguro. "Human Immunodeficiency Virus Type 1 Subtype Distribution in the Worldwide Epidemic: Pathogenetic and Therapeutic Implications". In: *Journal of Virology* 81.19 (Oct. 2007), pp. 10209–10219. ISSN: 0022-538X. DOI: 10.1128/JVI.00872-07. URL: https://jvi.asm.org/content/81/19/10209.

[16] S Campbell and A Rein. "In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain." In: *Journal of virology* 73.3 (Mar. 1999), pp. 2270–9. ISSN: 0022-538X. URL: http://www.ncbi.nlm.nih.gov/pubmed/9971810.

[17] Stefan G Sarafianos et al. "Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition." In: *Journal of molecular biology* 385.3 (Jan. 2009), pp. 693–713. ISSN: 1089-8638. DOI: 10.1016/j.jmb.2008.10.071. URL: http://www.ncbi.nlm.nih.gov/pubmed/19022262.

[18] Cordelia K. Leonard et al. "Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells". In: *Journal of Biological Chemistry* 265.18 (June 1990), pp. 10373–82. ISSN: 00219258. URL: http://www.ncbi.nlm.nih.gov/pubmed/2355006.

[19] Ping Zhu et al. "Distribution and three-dimensional structure of AIDS virus envelope spikes". In: *Nature* 441.7095 (June 2006), pp. 847–852. ISSN: 0028-0836. DOI: 10.1038/nature04817. URL: http://www.nature.com/articles/nature04817.

[20] Md Munan Shaik et al. "Structural basis of coreceptor recognition by HIV-1 envelope spike". In: *Nature* 565.7739 (Jan. 2019), pp. 318–323. ISSN: 0028-0836. DOI: 10.1038/s41586-018-0804-9. URL: http://www.nature.com/articles/s41586-018-0804-9.

[21] Craig B Wilen, John C Tilton, and Robert W Doms. "HIV: cell binding and entry." In: *Cold Spring Harbor perspectives in medicine* 2.8 (Aug. 2012). ISSN: 2157-1422. DOI: 10.1101/cshperspect.a006866. URL: http://www.ncbi.nlm.nih.gov/pubmed/22908191.

[22] Robert Craigie and Frederic D Bushman. "HIV DNA integration." In: *Cold Spring Harbor perspectives in medicine* 2.7 (July 2012), a006890. ISSN: 2157-1422. DOI: 10.1101/cshperspect.a006890. URL: http://www.ncbi.nlm.nih.gov/pubmed/22762018.

[23] María Eugenia González. "Vpu Protein: The Viroporin Encoded by HIV-1." In: *Viruses* 7.8 (Aug. 2015), pp. 4352–68. ISSN: 1999-4915. DOI: 10.3390/v7082824. URL: http://www.ncbi.nlm.nih.gov/pubmed/26247957.

[24] Neil M. Bell and Andrew M.L. Lever. "HIV Gag polyprotein: processing and early viral particle assembly". In: *Trends in Microbiology* 21.3 (Mar. 2013), pp. 136–144. ISSN: 0966842X. DOI: 10.1016/j.tim.2012.11.006. URL: http://www.ncbi.nlm.nih.gov/pubmed/23266279.

[25] Robert F Siliciano and Warner C Greene. "HIV latency." In: *Cold Spring Harbor perspectives in medicine* 1.1 (Sept. 2011), a007096. ISSN: 2157-1422. DOI: 10.1101/cshperspect.a007096. URL: http://www.ncbi.nlm.nih.gov/pubmed/22229121.

[26] Arjun Rustagi and Michael Gale. "Innate Antiviral Immune Signaling, Viral Evasion and Modulation by HIV-1". In: *Journal of Molecular Biology* 426.6 (Mar. 2014), pp. 1161–1177. ISSN: 00222836. DOI: 10.1016/j.jmb.2013.12.003. URL: http://www.ncbi.nlm.nih.gov/pubmed/24326250.

[27] Daniel Sauter and Frank Kirchhoff. "HIV replication: a game of hide and sense." In: *Current opinion in HIV and AIDS* 11.2 (Mar. 2016), pp. 173–81. ISSN: 1746-6318. DOI: 10.1097/COH.0000000000000233. URL: http://www.ncbi.nlm.nih.gov/pubmed/26628325.

[28] Greg J. Towers and Mahdad Noursadeghi. "Interactions between HIV-1 and the Cell-Autonomous Innate Immune System". In: *Cell Host & Microbe* 16.1 (July 2014),

pp. 10–18. ISSN: 1931-3128. DOI: 10.1016/J.CHOM.2014.06.009. URL: https://www.sciencedirect.com/science/article/pii/S193131281400225X?via%3Dihub.

[29] Takamitsu Matsuzawa et al. "Immunological function of Langerhans cells in HIV infection". In: *Journal of Dermatological Science* 87.2 (Aug. 2017), pp. 159–167. URL: https://www.sciencedirect.com/science/article/pii/S0923181116310921?via%3Dihub.

[30] Yvette van Kooyk and Gabriel A Rabinovich. "Protein-glycan interactions in the control of innate and adaptive immune responses." In: *Nature immunology* 9.6 (June 2008), pp. 593–601. ISSN: 1529-2916. DOI: 10.1038/ni.f.203. URL: http://www.nature.com/articles/ni.f.203.

[31] Sergey Pustylnikov et al. "Short Communication: Inhibition of DC-SIGN-Mediated HIV-1 Infection by Complementary Actions of Dendritic Cell Receptor Antagonists and Env-Targeting Virus Inactivators." In: *AIDS research and human retroviruses* 32.1 (Jan. 2016), pp. 93–100. ISSN: 1931-8405. DOI: 10.1089/aid.2015.0184. URL: http://www.ncbi.nlm.nih.gov/pubmed/26383762.

[32] Lot de Witte et al. "Langerin is a natural barrier to HIV-1 transmission by Langerhans cells". In: *Nature Medicine* 13.3 (Mar. 2007), pp. 367–371. ISSN: 1078-8956. DOI: 10.1038/nm1541. URL: http://www.ncbi.nlm.nih.gov/pubmed/17334373.

[33] Alexandra A Lambert et al. "The C-type lectin surface receptor DCIR acts as a new attachment factor for HIV-1 in dendritic cells and contributes to trans- and cis-infection pathways." In: *Blood* 112.4 (Aug. 2008), pp. 1299–307. ISSN: 1528-0020. DOI: 10.1182/blood-2008-01-136473. URL: http://www.ncbi.nlm.nih.gov/pubmed/18541725.

[34] Li Wu and Vineet N KewalRamani. "Dendritic-cell interactions with HIV: infection and viral dissemination." In: *Nature reviews. Immunology* 6.11 (Nov. 2006), pp. 859–68. ISSN: 1474-1733. DOI: 10.1038/nri1960. URL: http://www.ncbi.nlm.nih.gov/pubmed/17063186.

[35] Nilu Goonetilleke et al. "The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection". In: *The Journal of Experimental Medicine* (2009).

[36] Viviana Simon and David D. Ho. "HIV-1 dynamics in vivo: implications for therapy". In: *Nature Reviews Microbiology* 1.3 (Dec. 2003), pp. 181–190. ISSN: 1740-1526. DOI: 10.1038/nrmicro772. URL: http://www.nature.com/articles/nrmicro772.

[37] Liu et al. "Selection on the human immunodeficiency virus type 1 proteome following primary infection". In: *Journal of virology* (2006). ISSN: 0022-538X. DOI: 10.1128/JVI.00575-06.

[38] Persephone Borrow et al. "Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus". In: *Nature Medicine* (1997). ISSN: 10788956. DOI: 10.1038/nm0297-205.

[39] Todd M Allen et al. "Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution." In: *Journal of virology* 79.21 (Nov. 2005), pp. 13239–49. ISSN: 0022-538X. DOI: 10.1128/JVI.79.21.13239-13249.2005. URL: http://www.ncbi.nlm.nih.gov/pubmed/16227247.

[40] Joseph Kim et al. "CD8+ Cytotoxic T Lymphocyte Responses and Viral Epitope Escape in Acute HIV-1 Infection". In: *Viral Immunology* 31.7 (Sept. 2018), pp. 525–536. DOI: 10.1089/vim.2018.0040. URL: https://www.liebertpub.com/doi/10.1089/vim.2018.0040.

[41] Bruce D. Walker et al. "HIV-specific cytotoxic T lymphocytes in seropositive individuals". In: *Nature* 328.6128 (July 1987), pp. 345–348. ISSN: 0028-0836. DOI: 10.1038/328345a0. URL: http://www.nature.com/articles/328345a0.

[42] X N Xu et al. "A novel approach to antigen-specific deletion of CTL with minimal cellular activation using alpha3 domain mutants of MHC class I/peptide complex." In: *Immunity* 14.5 (May 2001), pp. 591–602. ISSN: 1074-7613. DOI: 10.1016/S1074-7613(01)00133-9. URL: http://www.ncbi.nlm.nih.gov/pubmed/11371361.

[43] G. D. Tomaras et al. "Initial B-Cell Responses to Transmitted Human Immunodeficiency Virus Type 1: Virion-Binding Immunoglobulin M (IgM) and IgG Antibodies Followed by Plasma Anti-gp41 Antibodies with Ineffective Control of Initial Viremia". In: *Journal of Virology* 82.24 (Dec. 2008), pp. 12449–12463. ISSN: 0022-538X. DOI: 10.1128/JVI.01708-08. URL: http://www.ncbi.nlm.nih.gov/pubmed/18842730.

[44] Brandon F. Keele et al. "Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection". In: *Proceedings of the National Academy of Sciences of the United States of America* 105.21 (May 2008), pp. 7552–7557. ISSN: 1091-6490. DOI: 10.1073/PNAS.0802203105. URL: https://pubmed.ncbi.nlm.nih.gov/18490657/.

[45] E. S. Gray et al. "Neutralizing Antibody Responses in Acute Human Immunodeficiency Virus Type 1 Subtype C Infection". In: *Journal of Virology* 81.12 (June 2007), pp. 6187–6196. ISSN: 0022-538X. DOI: 10.1128/JVI.00239-07. URL: http://www.ncbi.nlm.nih.gov/pubmed/17409164.

[46] Bing Li et al. "Evidence for potent autologous neutralizing antibody titers and compact envelopes in early infection with subtype C human immunodeficiency virus type

1." In: *Journal of virology* 80.11 (June 2006), pp. 5211–8. ISSN: 0022-538X. DOI: 10.1128/JVI.00201-06. URL: http://www.ncbi.nlm.nih.gov/pubmed/16699001.

[47] Manish Sagar et al. "Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity." In: *Journal of virology* 80.19 (Oct. 2006), pp. 9586–98. ISSN: 0022-538X. DOI: 10.1128/JVI.00141-06. URL: http://www.ncbi.nlm.nih.gov/pubmed/16973562.

[48] P. L. Moore et al. "The C3-V4 Region Is a Major Target of Autologous Neutralizing Antibodies in Human Immunodeficiency Virus Type 1 Subtype C Infection". In: *Journal of Virology* 82.4 (Feb. 2008), pp. 1860–1869. ISSN: 0022-538X. DOI: 10.1128/JVI.02187-07. URL: http://www.ncbi.nlm.nih.gov/pubmed/18057243.

[49] Penny L. Moore et al. "Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape". In: *Nature Medicine* 18.11 (Nov. 2012), pp. 1688–1692. ISSN: 10788956. DOI: 10.1038/nm.2985. URL: http://www.ncbi.nlm.nih.gov/pubmed/23086475.

[50] David C. Montefiori et al. "Neutralization tiers of HIV-1". In: *Current Opinion in HIV and AIDS* 13.2 (Mar. 2018), pp. 128–136. ISSN: 1746-630X. DOI: 10.1097/COH.0000000000000442. URL: http://www.ncbi.nlm.nih.gov/pubmed/29266013.

[51] A. Karpas et al. "Lytic infection by British AIDS virus and development of rapid cell test for antiviral antibodies". In: *Lancet* 2.8457 (Sept. 1985), pp. 695–696. ISSN: 01406736. DOI: 10.1016/s0140-6736(85)92934-4. URL: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(85)92934-4/abstract.

[52] John R. Mascola et al. "Two Antigenically Distinct Subtypes of Human Immunodeficiency Virus Type 1: Viral Genotype Predicts Neutralization Serotype". In: *The Journal of Infectious Diseases* 169.1 (Jan. 1994), pp. 48–54. ISSN: 0022-1899. DOI: 10.1093/INFDIS/169.1.48. URL: https://academic.oup.com/jid/article/169/1/48/896258.

[53] Alice K. Pilgrim et al. "Neutralizing Antibody Responses to Human Immunodeficiency Virus Type 1 in Primary Infection and Long-Term-Nonprogressive Infection". In: *The Journal of Infectious Diseases* 176.4 (Oct. 1997), pp. 924–932. ISSN: 0022-1899. DOI: 10.1086/516508. URL: https://academic.oup.com/jid/article/176/4/924/883950.

[54] Thomas Muster et al. "A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1". In: *Journal of Virology* 67.11 (Nov. 1993), pp. 6642–6647. ISSN: 0022-538X. DOI: 10.1128/JVI.67.11.6642-6647.1993. URL: https://journals.asm.org/doi/10.1128/jvi.67.11.6642-6647.1993.

[55] A Trkola et al. "Cross-clade neutralization of primary isolates of human immunodeficiency virus type 1 by human monoclonal antibodies and tetrameric CD4-IgG". In: *Journal of virology* 69.11 (Nov. 1995), pp. 6609–6617. ISSN: 0022-538X. DOI: 10.1128/JVI.69.11.6609-6617.1995. URL: https://pubmed.ncbi.nlm.nih.gov/7474069/.

[56] Marit J. van Gils and Rogier W. Sanders. "Broadly neutralizing antibodies against HIV-1: Templates for a vaccine". In: *Virology* 435.1 (Jan. 2013), pp. 46–56. ISSN: 00426822. DOI: 10.1016/j.virol.2012.10.004. URL: http://www.ncbi.nlm.nih.gov/pubmed/23217615.

[57] Devin Sok and Dennis R. D.R. Burton. *Recent progress in broadly neutralizing antibodies to HIV*. Nov. 2018. DOI: 10.1038/s41590-018-0235-7.

[58] J. T. Safrit et al. "hu-PBL-SCID mice can be protected from HIV-1 infection by passive transfer of monoclonal antibody to the principal neutralizing determinant of envelope gp120". In: *AIDS (London, England)* 7.1 (1993), pp. 15–21. ISSN: 0269-9370. DOI: 10.1097/00002030-199301000-00002. URL: https://pubmed.ncbi.nlm.nih.gov/7680205/.

[59] Yukari Okamoto et al. "In SCID-hu Mice, Passive Transfer of a Humanized Antibody Prevents Infection and Atrophic Change of Medulla in Human Thymic Implant due to Intravenous Inoculation of Primary HIV-1 Isolate". In: *The Journal of Immunology* 160.1 (Jan. 1998), pp. 69–76. ISSN: 0022-1767. DOI: 10.4049/JIMMUNOL.160.1.69. URL: https://journals.aai.org/jimmunol/article/160/1/69/74473/In-SCID-hu-Mice-Passive-Transfer-of-a-Humanized.

[60] Timothy W. Baba et al. "Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian–human immunodeficiency virus infection". In: *Nature Medicine 2000 6:2* 6.2 (Feb. 2000), pp. 200–206. ISSN: 1546-170X. DOI: 10.1038/72309. URL: https://www.nature.com/articles/nm0200_200.

[61] Ann J. Hessell et al. "Broadly Neutralizing Human Anti-HIV Antibody 2G12 Is Effective in Protection against Mucosal SHIV Challenge Even at Low Serum Neutralizing Titers". In: *PLOS Pathogens* 5.5 (May 2009), e1000433. ISSN: 1553-7374. DOI: 10.1371/JOURNAL.PPAT.1000433. URL: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1000433.

[62] Ruth M. Ruprecht et al. "Antibody protection: passive immunization of neonates against oral AIDS virus challenge". In: *Vaccine* 21.24 (July 2003), pp. 3370–3373. ISSN: 0264-410X. DOI: 10.1016/S0264-410X(03)00335-9.

[63] Lotta Von Boehmer et al. "Sequencing and cloning of antigen-specific antibodies from mouse memory B cells". In: *Nature Protocols 2016 11:10* 11.10 (Sept. 2016), pp. 1908–1923. ISSN: 1750-2799. DOI: 10.1038/nprot.2016.102. URL: https://www.nature.com/articles/nprot.2016.102.

[64] Philipp Schommers et al. "Restriction of HIV-1 Escape by a Highly Broad and Potent Neutralizing Antibody". In: *Cell* 180.3 (Feb. 2020), pp. 471–489. ISSN: 0092-8674. DOI: 10.1016/J.CELL.2020.01.010.

[65] Kenneth Smith et al. "Rapid generation of fully human monoclonal antibodies specific to a vaccinating antigen". In: *Nature Protocols 2009 4:3* 4.3 (Feb. 2009), pp. 372–384. ISSN: 1750-2799. DOI: 10.1038/nprot.2009.3. URL: https://www.nature.com/articles/nprot.2009.3.

[66] Peter D. Kwong et al. *Structure of an HIV gp 120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody*. 1998. DOI: 10.1038/31405.

[67] Dmitry Lyumkis et al. "Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer." In: *Science (New York, N.Y.)* (2013). ISSN: 1095-9203. DOI: 10.1126/science.1245627.

[68] Jeong Hyun Lee et al. "A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic β-Hairpin Structure". In: *Immunity* 46.4 (Apr. 2017), pp. 690–702. ISSN: 10974180. DOI: 10.1016/j.immuni.2017.03.017. URL: http://www.ncbi.nlm.nih.gov/pubmed/28423342.

[69] Anna-Janina Janina Behrens et al. "Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein". In: *Cell Reports* 14.11 (Mar. 2016), pp. 2695–2706. ISSN: 22111247. DOI: 10.1016/j.celrep.2016.02.058. URL: http://www.ncbi.nlm.nih.gov/pubmed/26972002.

[70] Anna-Janina Behrens and Max Crispin. "Structural principles controlling HIV envelope glycosylation". In: *Current Opinion in Structural Biology* 44 (June 2017), pp. 125–133. URL: http://www.ncbi.nlm.nih.gov/pubmed/28363124.

[71] Jeong Hyun Lee et al. "Model Building and Refinement of a Natively Glycosylated HIV-1 Env Protein by High-Resolution Cryoelectron Microscopy." In: *Structure (London, England : 1993)* 23.10 (Oct. 2015), pp. 1943–51. ISSN: 1878-4186. DOI: 10.1016/j.str.2015.07.020.

[72] D. G. Myszka et al. "Energetics of the HIV gp120-CD4 binding reaction". In: *Proceedings of the National Academy of Sciences* (2002). ISSN: 0027-8424. DOI: 10.1073/pnas.97.16.9026.

[73] Chih Chin Huang et al. "Structural biology: Structure of a V3-containing HIV-1 gp120 core". In: *Science* (2005). ISSN: 00368075. DOI: 10.1126/science.1118398.

[74] Robert Blumenthal, Stewart Durell, and Mathias Viard. "HIV entry and envelope glycoprotein-mediated fusion." In: *The Journal of biological chemistry* 287.49 (Nov.

2012), pp. 40841–9. ISSN: 1083-351X. DOI: 10.1074/jbc.R112.406272. URL: http://www.ncbi.nlm.nih.gov/pubmed/23043104.

[75] Mary Ann Checkley, Benjamin G Luttge, and Eric O Freed. "HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation." In: *Journal of molecular biology* 410.4 (July 2011), pp. 582–608. ISSN: 1089-8638. DOI: 10.1016/j.jmb.2011.04.042. URL: http://www.ncbi.nlm.nih.gov/pubmed/21762802.

[76] R W Doms, P L Earl, and B Moss. "The assembly of the HIV-1 env glycoprotein into dimers and tetramers." In: *Advances in experimental medicine and biology* 300 (1991), pp. 203–19. ISSN: 0065-2598. URL: http://www.ncbi.nlm.nih.gov/pubmed/1781345.

[77] Jeffrey D Esko Hudson H Freeze Pamela Stanley Carolyn R Bertozzi Gerald W Hart Ajit Varki Richard D Cummings et al. *Essentials of Glycobiology, 2nd edition*. 2009. ISBN: 9780879697709. DOI: 10.1016/S0962-8924(00)01855-9.

[78] Milan Raska et al. "Glycosylation patterns of HIV-1 gp120 depend on the type of expressing cells and affect antibody recognition." In: *The Journal of biological chemistry* 285.27 (July 2010), pp. 20860–9. ISSN: 1083-351X. DOI: 10.1074/jbc.M109.085472. URL: http://www.ncbi.nlm.nih.gov/pubmed/20439465.

[79] Adnan Halim et al. "Site-specific characterization of threonine, serine, and tyrosine glycosylations of amyloid precursor protein/amyloid beta-peptides in human cerebrospinal fluid." In: *Proceedings of the National Academy of Sciences of the United States of America* 108.29 (July 2011), pp. 11848–53. ISSN: 1091-6490. DOI: 10.1073/pnas.1102664108. URL: http://www.ncbi.nlm.nih.gov/pubmed/21712440.

[80] Yasunori Watanabe et al. "Exploitation of glycosylation in enveloped virus pathobiology". In: 1863.10 (2019), pp. 1480–1497. URL: https://www.sciencedirect.com/science/article/pii/S0304416519301333?via%3Dihub.

[81] Ikuko Nishikawa et al. "Computational Prediction of O-linked Glycosylation Sites that Preferentially Map on Intrinsically Disordered Regions of Extracellular Proteins". In: *International Journal of Molecular Sciences* 11.12 (Dec. 2010), pp. 4991–5008. ISSN: 1422-0067. DOI: 10.3390/ijms11124991. URL: http://www.mdpi.com/1422-0067/11/12/4991.

[82] Catharina Steentoft et al. "Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology". In: *The EMBO Journal* 32.10 (Apr. 2013), pp. 1478–1488. ISSN: 0261-4189. DOI: 10.1038/emboj.2013.79. URL: http://emboj.embopress.org/cgi/doi/10.1038/emboj.2013.79.

[83] Etienne Decroly et al. “The convertases furin and PC1 can both cleave the human immunodeficiency virus (HIV)-1 envelope glycoprotein gp160 into gp120 (HIV-I SU) and gp41 (HIV-I TM)”. In: *Journal of Biological Chemistry* (1994). ISSN: 00219258.

[84] Jessica Shapiro et al. “Localization of Endogenous Furin in Cultured Cell Lines”. In: *Journal of Histochemistry & Cytochemistry* 45.1 (Jan. 1997), pp. 3–12. ISSN: 0022-1554. DOI: 10.1177/002215549704500102. URL: http://www.ncbi.nlm.nih.gov/pubmed/9010463.

[85] S Takahashi et al. “Localization of furin to the trans-Golgi network and recycling from the cell surface involves Ser and Tyr residues within the cytoplasmic domain.” In: *The Journal of biological chemistry* 270.47 (Nov. 1995), pp. 28397–401. ISSN: 0021-9258. DOI: 10.1074/JBC.270.47.28397. URL: http://www.ncbi.nlm.nih.gov/pubmed/7499343.

[86] Eric O. Freed. “HIV-1 assembly, release and maturation”. In: *Nature Reviews Microbiology* 13.8 (Aug. 2015), pp. 484–496. ISSN: 1740-1526. DOI: 10.1038/nrmicro3490. URL: http://www.nature.com/articles/nrmicro3490.

[87] Marc C Johnson. “Mechanisms for Env glycoprotein acquisition by retroviruses.” In: *AIDS research and human retroviruses* 27.3 (Mar. 2011), pp. 239–47. ISSN: 1931-8405. DOI: 10.1089/AID.2010.0350. URL: http://www.ncbi.nlm.nih.gov/pubmed/21247353.

[88] J. Zavada. “The Pseudotypic Paradox”. In: *Journal of General Virology* 63.1 (Nov. 1982), pp. 15–24. ISSN: 0022-1317. DOI: 10.1099/0022-1317-63-1-15. URL: http://jgv.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-63-1-15.

[89] Philip R Tedbury and Eric O Freed. “The role of matrix in HIV-1 envelope glycoprotein incorporation.” In: *Trends in microbiology* 22.7 (July 2014), pp. 372–8. ISSN: 1878-4380. DOI: 10.1016/j.tim.2014.04.012. URL: http://www.ncbi.nlm.nih.gov/pubmed/24933691.

[90] Junghwa Kirschman et al. “HIV-1 Envelope Glycoprotein Trafficking through the Endosomal Recycling Compartment is Required for Particle Incorporation.” In: *Journal of virology* (2017). ISSN: 1098-5514. DOI: 10.1128/JVI.01893-17.

[91] Tsutomu Murakami. “Retroviral Env Glycoprotein Trafficking and Incorporation into Virions”. In: *Molecular Biology International* 2012 (2012), p. 682850. ISSN: 2090-2190. DOI: 10.1155/2012/682850. URL: http://www.ncbi.nlm.nih.gov/pubmed/22811910.

[92] M Huang et al. "p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease." In: *Journal of virology* (1995). ISSN: 0022-538X.

[93] H. G. Gottlinger et al. "Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release." In: *Proceedings of the National Academy of Sciences* (2006). ISSN: 0027-8424. DOI: 10.1073/pnas.88.8.3195.

[94] D. G. Demirov et al. "Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function". In: *Proceedings of the National Academy of Sciences* (2002). ISSN: 0027-8424. DOI: 10.1073/pnas.032511899.

[95] Jennifer E. Garrus et al. "Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding". In: *Cell* (2001). ISSN: 00928674. DOI: 10.1016/S0092-8674(01)00506-2.

[96] J. Martin-Serrano, T. Zang, and P. D. Bieniasz. "HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress". In: *Nature Medicine* (2001). ISSN: 10788956. DOI: 10.1038/nm1201-1313.

[97] Jörg Votteler and Wesley I. Sundquist. *Virus budding and the ESCRT pathway*. 2013. DOI: 10.1016/j.chom.2013.08.012.

[98] Paola Sette et al. "Ubiquitin conjugation to Gag is essential for ESCRT-mediated HIV-1 budding". In: *Retrovirology* 10.1 (July 2013), p. 79. ISSN: 1742-4690. DOI: 10.1186/1742-4690-10-79. URL: http://www.ncbi.nlm.nih.gov/pubmed/23895345.

[99] David J. Harvey et al. *Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds*. 2009. DOI: 10.1002/pmic.200900096.

[100] AAFKE LAND, DUCO ZONNEVELD, and INEKE BRAAKMAN. "Folding of HIV-1 Envelope glycoprotein involves extensive isomerization of disulfide bonds and conformation dependent leader peptide cleavage". In: *The FASEB Journal* 17.9 (June 2003), pp. 1058–1067. ISSN: 0892-6638. DOI: 10.1096/fj.02-0811com. URL: http://www.ncbi.nlm.nih.gov/pubmed/12773488.

[101] Yan Li et al. "The HIV-1 Env Protein Signal Sequence Retards Its Cleavage and Down-regulates the Glycoprotein Folding". In: *Virology* 272.2 (July 2000), pp. 417–428. ISSN: 00426822. DOI: 10.1006/viro.2000.0357. URL: http://www.ncbi.nlm.nih.gov/pubmed/10873786.

[102] Camille Bonomelli et al. "The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade". In: *PLoS ONE* 6.8 (2011), e23521.

ISSN: 19326203. DOI: 10.1371/journal.pone.0023521. URL: http://www.ncbi.nlm.nih.gov/pubmed/21858152.

[103] Jurgen H. Gross. *Mass spectrometry : a textbook*. Third Edit. Springer US, 2017. ISBN: 9783319543987.

[104] W.B. Weston B. Struwe et al. "Site-Specific Glycosylation of Virion-Derived HIV-1 Env Is Mimicked by a Soluble Trimeric Immunogen". In: *Cell Reports* 24.8 (Aug. 2018), pp. 1958–1966. ISSN: 22111247. DOI: 10.1016/j.celrep.2018.07.080.

[105] Aaron M Robitaille et al. "Real-Time Search enables a new gold standard for TMT quantitation accuracy on the Orbitrap Eclipse Tribrid mass spectrometer". In: ().

[106] Yehia Mechref. "Use of CID/ETD mass spectrometry to analyze glycopeptides". In: *Current Protocols in Protein Science* (2012). ISSN: 19343663. DOI: 10.1002/0471140864.ps1211s68.

[107] Laura K Pritchard et al. "Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers." In: *Cell reports* 11.10 (June 2015), pp. 1604–13. ISSN: 2211-1247. DOI: 10.1016/j.celrep.2015.05.017. URL: http://www.ncbi.nlm.nih.gov/pubmed/26051934.

[108] Antu K. Dey et al. "cGMP production and analysis of BG505 SOSIP.664, an extensively glycosylated, trimeric HIV-1 envelope glycoprotein vaccine candidate". In: *Biotechnology and Bioengineering* (2018). ISSN: 10970290. DOI: 10.1002/bit.26498.

[109] R.W. Rogier W. Sanders et al. "A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies". In: *PLoS Pathogens* 9.9 (Sept. 2013). Ed. by Alexandra Trkola, e1003618. ISSN: 15537366. DOI: 10.1371/journal.ppat.1003618. URL: https://dx.plos.org/10.1371/journal.ppat.1003618.

[110] Brian Clas, Norbert Schuelke, and Aditi Master. "A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41". In: *J. Virol.* (2000). DOI: 10.1128/JVI.74.2.627-643.2000.Updated.

[111] R. W. Sanders et al. "Stabilization of the Soluble, Cleaved, Trimeric Form of the Envelope Glycoprotein Complex of Human Immunodeficiency Virus Type 1". In: *Journal of Virology* (2002). ISSN: 0022-538X. DOI: 10.1128/jvi.76.17.8875-8889.2002.

[112] James M Binley et al. "Enhancing the proteolytic maturation of human immunodeficiency virus type 1 envelope glycoproteins." In: *Journal of virology* 76.6 (Mar. 2002),

pp. 2606–16. ISSN: 0022-538X. DOI: 10.1128/JVI.76.6.2606-2616.2002. URL: http://www.ncbi.nlm.nih.gov/pubmed/11861826.

[113] Jürgen Haas, Eun-Chung C Park, and Brian Seed. "Codon usage limitation in the expression of HIV-1 envelope glycoprotein". In: *Current biology : CB* 6.3 (Mar. 1996), pp. 315–24. ISSN: 0960-9822. DOI: 10.1016/S0960-9822(02)00482-7. URL: http://www.ncbi.nlm.nih.gov/pubmed/8805248.

[114] Steven W S.W. Steven W. de Taeye et al. "Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes". In: *Cell* 163.7 (Dec. 2015). ISSN: 1097-4172. DOI: 10.1016/j.cell.2015.11.056. URL: http://www.ncbi.nlm.nih.gov/pubmed/26687358.

[115] A.T. Andrew T. McGuire et al. "Antigen modification regulates competition of broad and narrow neutralizing HIV antibodies". In: *Science* 346.6215 (Dec. 2014), pp. 1380–1383. ISSN: 10959203. DOI: 10.1126/science.1259206. URL: https://science.sciencemag.org/content/346/6215/1380.

[116] Yang Zhang et al. "Germinal center B cells govern their own fate via antibody feedback". In: *The Journal of Experimental Medicine* (2013). ISSN: 0022-1007. DOI: 10.1084/jem.20120150.

[117] Dirk Eggink et al. "HIV-1 anchor inhibitors and membrane fusion inhibitors target distinct but overlapping steps in virus entry". In: *Journal of Biological Chemistry* 294.15 (Apr. 2019), pp. 5736–5746. ISSN: 0021-9258. DOI: 10.1074/JBC.RA119.007360. URL: http://www.jbc.org/content/294/15/5736.full.

[118] Marie Pancera et al. "Structure and immune recognition of trimeric pre-fusion HIV-1 Env". In: *Nature* 514.7253 (Oct. 2014), pp. 455–461. ISSN: 0028-0836. DOI: 10.1038/nature13808. URL: http://www.nature.com/articles/nature13808.

[119] Richard Wyatt et al. "The antigenic structure of the HIV gp120 envelope glycoprotein". In: *Nature* 393.6686 (June 1998), pp. 705–711. ISSN: 0028-0836. DOI: 10.1038/31514.

[120] Katie J. Doores et al. "Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens". In: *Proceedings of the National Academy of Sciences* 107.31 (Aug. 2010), pp. 13800–5. ISSN: 0027-8424. DOI: 10.1073/pnas.1006498107. URL: http://www.ncbi.nlm.nih.gov/pubmed/20643940.

[121] Kabamba B. Alexandre et al. "Mannose-rich glycosylation patterns on HIV-1 subtype C gp120 and sensitivity to the lectins, Griffithsin, Cyanovirin-N and Scytovirin". In: *Virology* (2010). ISSN: 00426822. DOI: 10.1016/j.virol.2010.03.021.

[122] Karen P. Coss et al. "HIV-1 Glycan Density Drives the Persistence of the Mannose Patch within an Infected Individual". In: *Journal of Virology* (2016). ISSN: 0022-538X. DOI: 10.1128/jvi.01542-16.

[123] Anna-Janina Behrens et al. "Molecular Architecture of the Cleavage-Dependent Mannose Patch on a Soluble HIV-1 Envelope Glycoprotein Trimer". In: *Journal of Virology* 91.2 (Jan. 2017). Ed. by Susan R. Ross, pp. 01894–16. ISSN: 0022-538X. DOI: 10.1128/jvi.01894-16. URL: http://www.ncbi.nlm.nih.gov/pubmed/27807235.

[124] Maria Panico et al. "Mapping the complete glycoproteome of virion-derived HIV-1 gp120 provides insights into broadly neutralizing antibody binding". In: *Scientific Reports* 6.1 (Dec. 2016), p. 32956. ISSN: 20452322. DOI: 10.1038/srep32956. URL: http://www.ncbi.nlm.nih.gov/pubmed/27604319.

[125] Max Crispin and Katie J. Doores. *Targeting host-derived glycans on enveloped viruses for antibody-based vaccine design*. 2015. DOI: 10.1016/j.coviro.2015.02.002.

[126] Laura K. Pritchard et al. "Cell- and Protein-Directed Glycosylation of Native Cleaved HIV-1 Envelope". In: *Journal of Virology* 89.17 (Sept. 2015), pp. 8932–44. ISSN: 0022-538X. DOI: 10.1128/jvi.01190-15. URL: http://www.ncbi.nlm.nih.gov/pubmed/26085151.

[127] Heribert Stoiber et al. "Human Complement Proteins C3b, C4b, Factor H and Properdin React with Specific Sites in gpl20 and gp4l, the Envelope Proteins of HIV1". In: *Immunobiology* 193.1 (June 1995), pp. 98–113. ISSN: 0171-2985. DOI: 10.1016/S0171-2985(11)80158-0. URL: https://www.sciencedirect.com/science/article/pii/S0171298511801580?via%3Dihub.

[128] Suzanne Willey et al. "Extensive complement-dependent enhancement of HIV-1 by autologous non-neutralising antibodies at early stages of infection". In: *Retrovirology* 8.1 (Mar. 2011), p. 16. ISSN: 1742-4690. DOI: 10.1186/1742-4690-8-16. URL: http://www.ncbi.nlm.nih.gov/pubmed/21401915.

[129] L D Powell et al. "Natural ligands of the B cell adhesion molecule CD22 beta carry N-linked oligosaccharides with alpha-2,6-linked sialic acids that are required for recognition." In: *The Journal of biological chemistry* 268.10 (Apr. 1993), pp. 7019–27. ISSN: 0021-9258. URL: http://www.ncbi.nlm.nih.gov/pubmed/8463235.

[130] Inka Brockhausen, Harry Schachter, and Pamela Stanley. *O-GalNAc Glycans*. Cold Spring Harbor Laboratory Press, 2009. ISBN: 9780879697709. URL: http://www.ncbi.nlm.nih.gov/pubmed/20301232.

[131] H B Bernstein et al. "Human immunodeficiency virus type 1 envelope glycoprotein is modified by O-linked oligosaccharides." In: *Journal of virology* 68.1 (Jan. 1994), pp. 463–8. ISSN: 0022-538X. URL: http://www.ncbi.nlm.nih.gov/pubmed/8254757.

[132] J E Hansen et al. "Sensitivity of HIV-1 to neutralization by antibodies against O-linked carbohydrate epitopes despite deletion of O-glycosylation signals in the V3 loop." In: *Archives of virology* 141.2 (1996), pp. 291–300. ISSN: 0304-8608. URL: http://www.ncbi.nlm.nih.gov/pubmed/8634021.

[133] Eden P. Go et al. "Characterization of Host-Cell Line Specific Glycosylation Profiles of Early Transmitted/Founder HIV-1 gp120 Envelope Proteins". In: *Journal of Proteome Research* 12.3 (Mar. 2013), pp. 1223–1234. ISSN: 1535-3893. DOI: 10.1021/pr300870t. URL: http://www.ncbi.nlm.nih.gov/pubmed/23339644.

[134] Eden P. Go, David Hua, and Heather Desaire. "Glycosylation and Disulfide Bond Analysis of Transiently and Stably Expressed Clade C HIV-1 gp140 Trimers in 293T Cells Identifies Disulfide Heterogeneity Present in Both Proteins and Differences in O -Linked Glycosylation". In: *Journal of Proteome Research* 13.9 (Sept. 2014), pp. 4012–4027. ISSN: 1535-3893. DOI: 10.1021/pr5003643. URL: http://www.ncbi.nlm.nih.gov/pubmed/25026075.

[135] Eden P Go et al. "Comparative Analysis of the Glycosylation Profiles of Membrane-Anchored HIV-1 Envelope Glycoprotein Trimers and Soluble gp140." In: *Journal of virology* 89.16 (Aug. 2015), pp. 8245–57. ISSN: 1098-5514. DOI: 10.1128/JVI.00628-15. URL: http://www.ncbi.nlm.nih.gov/pubmed/26018173.

[136] Elizabeth Stansell et al. "Gp120 on HIV-1 Virions Lacks O-Linked Carbohydrate". In: *PLOS ONE* 10.4 (Apr. 2015). Ed. by William A Paxton, e0124784. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0124784. URL: https://dx.plos.org/10.1371/journal.pone.0124784.

[137] Weiming Yang et al. "Glycoform analysis of recombinant and human immunodeficiency virus envelope protein gp120 via higher energy collisional dissociation and spectral-aligning strategy." In: *Analytical chemistry* 86.14 (July 2014), pp. 6959–67. ISSN: 1520-6882. DOI: 10.1021/ac500876p. URL: http://www.ncbi.nlm.nih.gov/pubmed/24941220.

[138] Jeong Hyun Lee, Gabriel Ozorowski, and Andrew B. Ward. "Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer". In: *Science* (2016). ISSN: 10959203. DOI: 10.1126/science.aad2450.

[139] M. Zhang et al. "Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin". In: *Glycobiology* 14.12 (July 2004), pp. 1229–1246. ISSN: 1460-2423. DOI: 10.1093/glycob/cwh106. URL: http://www.ncbi.nlm.nih.gov/pubmed/15175256.

[140] Christopher N. Scanlan et al. "The Broadly Neutralizing Anti-Human Immunodeficiency Virus Type 1 Antibody 2G12 Recognizes a Cluster of 1-¿2 Mannose Residues

on the Outer Face of gp120". In: *Journal of Virology* 76.14 (July 2002), pp. 7306–21. ISSN: 0022-538X. DOI: 10.1128/jvi.76.14.7306-7321.2002. URL: http://www.ncbi.nlm.nih.gov/pubmed/12072529.

[141] Xiping Wei et al. "Antibody neutralization and escape by HIV-1". In: *Nature* 422.6929 (Mar. 2003), pp. 307–312. ISSN: 00280836. DOI: 10.1038/nature01470. URL: http://www.ncbi.nlm.nih.gov/pubmed/12646921.

[142] Leopold Kong, Robyn L. Stanfield, and Ian A. Wilson. "Molecular Recognition of HIV Glycans by Antibodies". In: *HIV glycans in infection and immunity*. New York, NY: Springer New York, 2014, pp. 117–141. DOI: 10.1007/978-1-4614-8872-9{_}5. URL: http://link.springer.com/10.1007/978-1-4614-8872-9_5.

[143] Jinghe Huang et al. "Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface". In: *Nature* 515.7525 (Nov. 2014), pp. 138–142. ISSN: 14764687. DOI: 10.1038/nature13601.

[144] Devin Sok et al. "Promiscuous glycan site recognition by antibodies to the high-mannose patch of gp120 broadens neutralization of HIV". In: *Science Translational Medicine* 6.236 (May 2014), 236ra63. ISSN: 19466242. DOI: 10.1126/scitranslmed.3008104. URL: http://www.ncbi.nlm.nih.gov/pubmed/24828077.

[145] A Trkola et al. "Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1." In: *Journal of virology* 70.2 (Feb. 1996), pp. 1100–8. ISSN: 0022-538X. URL: http://www.ncbi.nlm.nih.gov/pubmed/8551569.

[146] Daniel A. Calarese et al. "Antibody domain exchange is an immunological solution to carbohydrate cluster recognition". In: *Science* 300.5628 (2003). ISSN: 00368075. DOI: 10.1126/science.1083182.

[147] M. Huber et al. "Potent Human Immunodeficiency Virus-Neutralizing and Complement Lysis Activities of Antibodies Are Not Obligatorily Linked". In: *Journal of Virology* (2008). ISSN: 0022-538X. DOI: 10.1128/jvi.02569-07.

[148] Alexey M. Eroshkin et al. "bNAber: database of broadly neutralizing HIV antibodies". In: *Nucleic Acids Research* 42.D1 (Jan. 2014), pp. D1133–D1139. ISSN: 0305-1048. DOI: 10.1093/nar/gkt1083. URL: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1083.

[149] Amir Dashti et al. *Broadly Neutralizing Antibodies against HIV: Back to Blood*. 2019. DOI: 10.1016/j.molmed.2019.01.007.

[150] Barton F. Haynes. "New Approaches to HIV Vaccine Development". In: *Current opinion in immunology* 35 (2015), p. 39. DOI: 10.1016/J.COI.2015.05.007. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4553082/.

[151] Leopold Kong et al. "Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120". In: *Nature Structural & Molecular Biology* 20.7 (July 2013), pp. 796–803. ISSN: 1545-9993. DOI: 10.1038/nsmb.2594. URL: http://www.nature.com/articles/nsmb.2594.

[152] Max Crispin, Andrew B. Ward, and Ian A. Wilson. "Structure and Immune Recognition of the HIV Glycan Shield". In: *Annual Review of Biophysics* 47.1 (May 2018), pp. 060414–034156. ISSN: 1936-122X. DOI: 10.1146/annurev-biophys-060414-034156.

[153] Marie Pancera et al. "Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16". In: *Nature Structural and Molecular Biology* 20.7 (July 2013), pp. 804–13. ISSN: 15459993. DOI: 10.1038/nsmb.2600. URL: http://www.ncbi.nlm.nih.gov/pubmed/23708607.

[154] Vidya S. Shivatare et al. "Unprecedented Role of Hybrid ¡i¿N-¡/i¿ Glycans as Ligands for HIV-1 Broadly Neutralizing Antibodies". In: *Journal of the American Chemical Society* 140.15 (Apr. 2018), pp. 5202–5210. ISSN: 0002-7863. DOI: 10.1021/jacs.8b00896. URL: http://pubs.acs.org/doi/10.1021/jacs.8b00896.

[155] Liwei Cao et al. "Global site-specific N-glycosylation analysis of HIV envelope glycoprotein". eng. In: *Nature Communications* 8 (Mar. 2017), p. 14954. ISSN: 2041-1723. DOI: 10.1038/ncomms14954.

[156] J. A. Horwitz et al. "HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice". In: *Proceedings of the National Academy of Sciences* 110.41 (Oct. 2013), pp. 16538–16543. ISSN: 0027-8424. DOI: 10.1073/pnas.1315295110. URL: http://www.ncbi.nlm.nih.gov/pubmed/24043801.

[157] Dan H. Barouch et al. "Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys". In: *Nature* 503.7475 (Nov. 2013), pp. 224–228. ISSN: 0028-0836. DOI: 10.1038/nature12744. URL: http://www.nature.com/articles/nature12744.

[158] Masashi Shingai et al. "Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia". In: *Nature* 503.7475 (Nov. 2013), pp. 277–280. ISSN: 0028-0836. DOI: 10.1038/nature12746. URL: http://www.nature.com/articles/nature12746.

[159] Marina Caskey et al. "Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117". In: *Nature* 522.7557 (June 2015), pp. 487–491. ISSN: 0028-0836. DOI: 10.1038/nature14411. URL: http://www.nature.com/articles/nature14411.

[160] Punnee Pitisuttithum et al. "Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand". In: *Journal of Infectious Diseases* (2006). ISSN: 00221899. DOI: 10.1086/508748.

[161] N. M. Flynn et al. "Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection". In: *Journal of Infectious Diseases* (2005). ISSN: 00221899. DOI: 10.1086/428404.

[162] Susan P. Buchbinder et al. "Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial". In: *The Lancet* (2008). ISSN: 01406736. DOI: 10.1016/S0140-6736(08)61591-3.

[163] Glenda E. Gray et al. "Safety and efficacy of the HVTN 503/Phambili Study of a clade-B-based HIV-1 vaccine in South Africa: A double-blind, randomised, placebo-controlled test-of-concept phase 2b study". In: *The Lancet Infectious Diseases* (2011). ISSN: 14733099. DOI: 10.1016/S1473-3099(11)70098-6.

[164] Supachai Rerks-Ngarm et al. "Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand". In: *N. Engl. J. Med.* 361.23 (Dec. 2009), pp. 2209–2220. ISSN: 0028-4793. DOI: 10.1056/nejmoa0908492.

[165] Barton F. Haynes et al. "Immune-correlates analysis of an HIV-1 vaccine efficacy trial". In: *N. Engl. J. Med.* 366.14 (Apr. 2012), pp. 1275–1286. ISSN: 0028-4793. DOI: 10.1056/nejmoa1113425.

[166] Glenda E. Gray et al. "Vaccine efficacy of ALVAC-HIV and bivalent subtype C gp120-MF59 in adults". In: *N. Engl. J. Med.* 384.12 (Mar. 2021), pp. 1089–1100. ISSN: 0028-4793. DOI: 10.1056/nejmoa2031499.

[167] James B. Whitney et al. "Prevention of SIVmac251 reservoir seeding in rhesus monkeys by early antiretroviral therapy". In: *Nat. Commun.* 9.1 (Dec. 2018). ISSN: 20411723. DOI: 10.1038/s41467-018-07881-9.

[168] *Johnson & Johnson and Global Partners Announce Results from Phase 2b Imbokodo HIV Vaccine Clinical Trial in Young Women in Sub-Saharan Africa — Johnson & Johnson*. URL: https://www.jnj.com/johnson-johnson-and-global-partners-announce-results-from-phase-2b-imbokodo-hiv-vaccine-clinical-trial-in-young-women-in-sub-saharan-africa.

[169] *Evaluating the Safety and Immunogenicity of HIV-1 BG505 SOSIP.664 gp140 With TLR Agonist and/or Alum Adjuvants in Healthy, HIV-uninfected Adults - Full Text View - ClinicalTrials.gov*. URL: https://clinicaltrials.gov/ct2/show/study/NCT04177355.

[170] *A Study to Assess the Efficacy of a Heterologous Prime/Boost Vaccine Regimen of Ad26.Mos4.HIV and Aluminum Phosphate-Adjuvanted Clade C gp140 in Preventing Human Immunodeficiency Virus (HIV) -1 Infection in Women in Sub-Saharan Africa - No Study Results Posted - ClinicalTrials.gov*. URL: https://clinicaltrials.gov/ct2/show/results/NCT03060629.

[171] Barton F. Haynes et al. "Strategies for HIV-1 vaccines that induce broadly neutralizing antibodies". In: *Nature Reviews Immunology 2022* (Aug. 2022), pp. 1–17. ISSN: 1474-1741. DOI: 10.1038/s41577-022-00753-w. URL: https://www.nature.com/articles/s41577-022-00753-w.

[172] Peter B. Gilbert et al. "Correlation between immunologic responses to a recombinant glycoprotein 120 vaccine and incidence of HIV-1 infection in a phase 3 HIV-1 preventive vaccine trial". In: *Journal of Infectious Diseases* (2005). ISSN: 00221899. DOI: 10.1086/428405.

[173] Peter Gilbert et al. "Magnitude and breadth of a nonprotective neutralizing antibody response in an efficacy trial of a candidate HIV-1 gp120 vaccine". In: *Journal of Infectious Diseases* (2010). ISSN: 00221899. DOI: 10.1086/654816.

[174] David C. Montefiori et al. "Magnitude and breadth of the neutralizing antibody response in the RV144 and Vax003 HIV-1 vaccine efficacy trials". In: *Journal of Infectious Diseases* (2012). ISSN: 00221899. DOI: 10.1093/infdis/jis367.

[175] Sarah Sterrett et al. "Low multiplicity of HIV-1 infection and no vaccine enhancement in VAX003 injection drug users". In: *Open Forum Infectious Diseases* (2014). ISSN: 23288957. DOI: 10.1093/ofid/ofu056.

[176] Glenda E. Gray et al. "Recombinant adenovirus type 5 HIV gag/pol/nef vaccine in South Africa: Unblinded, long-term follow-up of the phase 2b HVTN 503/Phambili study". In: *The Lancet Infectious Diseases* (2014). ISSN: 14744457. DOI: 10.1016/S1473-3099(14)70020-9.

[177] T. Hertz et al. "A study of vaccine-induced immune pressure on breakthrough infections in the Phambili phase 2b HIV-1 vaccine efficacy trial". In: *Vaccine* (2016). ISSN: 18732518. DOI: 10.1016/j.vaccine.2016.09.054.

[178] Ellen M. Leitman et al. "Lower Viral Loads and Slower CD4+ T-Cell Count Decline in MRKAd5 HIV-1 Vaccinees Expressing Disease-Susceptible HLA-B58:02". In: *Journal of Infectious Diseases* (2016). ISSN: 15376613. DOI: 10.1093/infdis/jiw093.

[179] Zoe Moodie et al. "Continued follow-up of Phambili phase 2b randomized HIV-1 vaccine trial participants supports increased HIV-1 acquisition among vaccinated men". In: *PLoS ONE* (2015). ISSN: 19326203. DOI: 10.1371/journal.pone.0137666.

[180] Merlin L. Robb et al. "Risk behaviour and time as covariates for efficacy of the HIV vaccine regimen ALVAC-HIV (vCP1521) and AIDSVAX B/E: A post-hoc analysis of the Thai phase 3 efficacy trial RV 144". In: *The Lancet Infectious Diseases* (2012). ISSN: 14733099. DOI: 10.1016/S1473-3099(12)70088-9.

[181] Siriwat Akapirat et al. "Characterization of HIV-1 gp120 antibody specificities induced in anogenital secretions of RV144 vaccine recipients after late boost immunizations". In: *PLoS ONE* (2018). ISSN: 19326203. DOI: 10.1371/journal.pone.0196397.

[182] Raphael Gottardo et al. "Plasma IgG to Linear Epitopes in the V2 and V3 Regions of HIV-1 gp120 Correlate with a Reduced Risk of Infection in the RV144 Vaccine Efficacy Trial". In: *PLoS ONE* (2013). ISSN: 19326203. DOI: 10.1371/journal.pone.0075665.

[183] Glenda E. Gray et al. "Immune correlates of the Thai RV144 HIV vaccine regimen in South Africa". In: *Science Translational Medicine* (2019). ISSN: 19466242. DOI: 10.1126/scitranslmed.aax1880.

[184] Kalpana Dommaraju et al. "CD8 and CD4 epitope predictions in RV144: No strong evidence of a T-cell driven sieve effect in HIV-1 Breakthrough sequences from trial participants". In: *PLoS ONE* (2014). ISSN: 19326203. DOI: 10.1371/journal.pone.0111334.

[185] Andrew J. Gartland et al. "Analysis of HLA A*02 Association with Vaccine Efficacy in the RV144 HIV-1 Vaccine Trial". In: *Journal of Virology* (2014). ISSN: 0022-538X. DOI: 10.1128/jvi.01164-14.

[186] Yunda Huang et al. "Predictors of durable immune responses six months after the last vaccination in preventive HIV vaccine trials". In: *Vaccine* (2017). ISSN: 18732518. DOI: 10.1016/j.vaccine.2016.09.053.

[187] Punnee Pitisuttithum et al. "Safety and reactogenicity of canarypox ALVAC-HIV (vCP1521) and HIV-1 gp120 AIDSVAX B/E vaccination in an efficacy trial in Thailand". In: *PLoS ONE* (2011). ISSN: 19326203. DOI: 10.1371/journal.pone.0027837.

[188] Lue Ping Zhao et al. "Landscapes of binding antibody and T-cell responses to poxprotein HIV vaccines in Thais and South Africans". In: *PLoS ONE* (2020). ISSN: 19326203. DOI: 10.1371/journal.pone.0226803.

[189] Susan Zolla-Pazner et al. "Vaccine-induced IgG antibodies to V1V2 regions of multiple HIV-1 subtypes correlate with decreased risk of HIV-1 infection". In: *PLoS ONE* (2014). ISSN: 19326203. DOI: 10.1371/journal.pone.0087572.

[190] Sushma Boppana et al. "HLA-I Associated Adaptation Dampens CD8 T-Cell Responses in HIV Ad5-Vectored Vaccine Recipients". In: *Journal of Infectious Diseases* (2019). ISSN: 15376613. DOI: 10.1093/infdis/jiz368.

[191] Allan C. DeCamp et al. "Sieve analysis of breakthrough HIV-1 sequences in HVTN 505 identifies vaccine pressure targeting the CD4 binding site of Env-gp120". In: *PLoS ONE* (2017). ISSN: 19326203. DOI: 10.1371/journal.pone.0185959.

[192] Youyi Fong et al. "Modification of the Association between T-Cell Immune Responses and Human Immunodeficiency Virus Type 1 Infection Risk by Vaccine-Induced Antibody Responses in the HVTN 505 Trial". In: *Journal of Infectious Diseases* (2018). ISSN: 15376613. DOI: 10.1093/infdis/jiy008.

[193] Scott M. Hammer et al. "Efficacy Trial of a DNA/rAd5 HIV-1 Preventive Vaccine". In: *New England Journal of Medicine* (2013). ISSN: 0028-4793. DOI: 10.1056/nejmoa1310566.

[194] Holly E. Janes et al. "Higher T-Cell responses induced by DNA/rAd5 HIV-1 preventive vaccine are associated with lower HIV-1 infection risk in an efficacy trial". In: *Journal of Infectious Diseases* (2017). ISSN: 15376613. DOI: 10.1093/infdis/jix086.

[195] Shuying S. Li et al. "Fc Gamma Receptor Polymorphisms Modulated the Vaccine Effect on HIV-1 Risk in the HVTN 505 HIV Vaccine Trial". In: *Journal of Virology* (2019). ISSN: 0022-538X. DOI: 10.1128/jvi.02041-18.

[196] Scott D. Neidich et al. "Antibody Fc effector functions and IgG3 associate with decreased HIV-1 risk". In: *Journal of Clinical Investigation* (2019). ISSN: 15588238. DOI: 10.1172/JCI126391.

[197] Wilton B. Williams et al. "Diversion of HIV-1 vaccine-induced immunity by gp41-microbiota cross-reactive antibodies". In: *Science* (2015). ISSN: 10959203. DOI: 10.1126/science.aab1253.

[198] Lawrence Corey et al. "Two Randomized Trials of Neutralizing Antibodies to Prevent HIV-1 Acquisition". In: *New England Journal of Medicine* (2021). ISSN: 0028-4793. DOI: 10.1056/nejmoa2031738.

[199] Srilatha Edupuganti et al. "Feasibility and Successful Enrollment in a Proof-of-Concept HIV Prevention Trial of VRC01, a Broadly Neutralizing HIV-1 Monoclonal Antibody". In: *Journal of Acquired Immune Deficiency Syndromes* (2021). ISSN: 10779450. DOI: 10.1097/QAI.0000000000002639.

[200] Nyaradzo M. Mgodi et al. "A Phase 2b Study to Evaluate the Safety and Efficacy of VRC01 Broadly Neutralizing Monoclonal Antibody in Reducing Acquisition of HIV-1 Infection in Women in Sub-Saharan Africa: Baseline Findings". In: *Journal of*

Acquired Immune Deficiency Syndromes (2021). ISSN: 10779450. DOI: 10.1097/QAI.0000000000002649.

[201] Fatima Laher et al. "Willingness to use HIV prevention methods among vaccine efficacy trial participants in Soweto, South Africa: discretion is important". In: *BMC Public Health* (2020). ISSN: 14712458. DOI: 10.1186/s12889-020-09785-0.

[202] Dan H. Barouch et al. "Evaluation of a mosaic HIV-1 vaccine in a multicentre, randomised, double-blind, placebo-controlled, phase 1/2a clinical trial (APPROACH) and in rhesus monkeys (NHP 13-19)". In: *The Lancet* (2018). ISSN: 1474547X. DOI: 10.1016/S0140-6736(18)31364-3.

[203] Douglas D. Richman et al. "Rapid evolution of the neutralizing antibody response to HIV type 1 infection". In: *Proceedings of the National Academy of Sciences of the United States of America* (2003). ISSN: 00278424. DOI: 10.1073/pnas.0630530100.

[204] R.W. Rogier W. Sanders et al. "HIV-1 neutralizing antibodies induced by native-like envelope trimers". In: *Science* 349.6244 (July 2015), aac4223–aac4223. ISSN: 10959203. DOI: 10.1126/science.aac4223. URL: http://www.ncbi.nlm.nih.gov/pubmed/26089353.

[205] Yu Feng et al. "Thermostability of Well-Ordered HIV Spikes Correlates with the Elicitation of Autologous Tier 2 Neutralizing Antibodies". In: *PLoS Pathogens* (2016). ISSN: 15537374. DOI: 10.1371/journal.ppat.1005767.

[206] Alba Torrents de la Peña et al. "Immunogenicity in Rabbits of HIV-1 SOSIP Trimers from Clades A, B, and C, Given Individually, Sequentially, or in Combination". In: *Journal of Virology* (2018). ISSN: 0022-538X. DOI: 10.1128/jvi.01957-17.

[207] P.J. J. Klasse et al. "Sequential and Simultaneous Immunization of Rabbits with HIV-1 Envelope Glycoprotein SOSIP.664 Trimers from Clades A, B and C". In: *PLoS Pathogens* 12.9 (Sept. 2016). ISSN: 15537374. DOI: 10.1371/journal.ppat.1005864.

[208] Cheng Cheng et al. "Immunogenicity of a Prefusion HIV-1 Envelope Trimer in Complex with a Quaternary-Structure-Specific Antibody". In: *Journal of Virology* (2016). ISSN: 0022-538X. DOI: 10.1128/jvi.02380-15.

[209] Yoann Aldon et al. "Rational Design of DNA-Expressed Stabilized Native-Like HIV-1 Envelope Trimers". In: *Cell Reports* 24.12 (Sept. 2018), pp. 3324–3338. ISSN: 22111247. DOI: 10.1016/j.celrep.2018.08.051.

[210] Bartek Nogal et al. "Mapping Polyclonal Antibody Responses in Non-human Primates Vaccinated with HIV Env Trimer Subunit Vaccines". In: *Cell Reports* 30.11 (Mar. 2020), pp. 3755–3765. ISSN: 2211-1247.

[211] Kshitij Wagh et al. "Completeness of HIV-1 Envelope Glycan Shield at Transmission Determines Neutralization Breadth". In: *Cell Reports* 25.4 (Oct. 2018), pp. 893–908. ISSN: 22111247. DOI: 10.1016/j.celrep.2018.09.087. URL: http://www.ncbi.nlm.nih.gov/pubmed/30355496.

[212] Colin Havenar-Daughton, Jeong Hyun Lee, and Shane Crotty. *Tfh cells and HIV bnAbs, an immunodominance model of the HIV neutralizing antibody generation problem*. 2017. DOI: 10.1111/imr.12512.

[213] Rajesh P. Ringe et al. "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". In: *Journal of Virology* 93.4 (Nov. 2018). Ed. by Guido Silvestri. ISSN: 0022-538X. DOI: 10.1128/jvi.01656-18. URL: http://www.ncbi.nlm.nih.gov/pubmed/30487280.

[214] P.J. J. Klasse et al. "Epitopes for neutralizing antibodies induced by HIV-1 envelope glycoprotein BG505 SOSIP trimers in rabbits and macaques". In: *PLoS Pathogens* 14.2 (Feb. 2018). ISSN: 15537374. DOI: 10.1371/journal.ppat.1006913.

[215] Keith Elkon and Paolo Casali. "Nature and functions of autoantibodies". In: *Nat. Clin. Pract. Rheumatol.* 4.9 (2008), pp. 491–498. ISSN: 17458382. DOI: 10.1038/ncprheum0895.

[216] Hugo Mouquet and Michel C. Nussenzweig. "Polyreactive antibodies in adaptive immune responses to viruses". In: *Cell Mol. Life Sci.* 69.9 (May 2012), pp. 1435–1445. ISSN: 1420682X. DOI: 10.1007/s00018-011-0872-6.

[217] Hedda Wardemann et al. "Predominant autoantibody production by early human B cell precursors". In: *Science* 301.5638 (Sept. 2003), pp. 1374–1377. ISSN: 00368075. DOI: 10.1126/science.1086907.

[218] Colin Havenar-Daughton et al. "The human naive B cell repertoire contains distinct subclasses for a germline-targeting HIV-1 vaccine immunogen". In: *Sci. Transl. Med.* 10.448 (July 2018). ISSN: 19466242. DOI: 10.1126/scitranslmed.aat0381.

[219] Jon M. Steichen et al. "A generalized HIV vaccine design strategy for priming of broadly neutralizing antibody responses". In: *Science* 366.6470 (Dec. 2019). ISSN: 10959203. DOI: 10.1126/science.aax4380.

[220] Jordan R. Willis et al. "Long antibody HCDR3s from HIV-naive donors presented on a PG9 neutralizing antibody background mediate HIV neutralization". In: *Proc. Natl Acad. Sci. USA* 113.16 (Apr. 2016), pp. 4446–4451. ISSN: 10916490. DOI: 10.1073/pnas.1518405113.

[221] Barton F. Haynes et al. "B-cell–lineage immunogen design in vaccine development with HIV-1 as a case study". In: *Nature Biotechnology 2012 30:5* 30.5 (May 2012), pp. 423–433. ISSN: 1546-1696. DOI: 10.1038/nbt.2197. URL: https://www.nature.com/articles/nbt.2197.

[222] Tysheena P. T.P. Charles et al. "The C3/465 glycan hole cluster in BG505 HIV-1 envelope is the major neutralizing target involved in preventing mucosal SHIV infection". In: 17.2 (Feb. 2021), e1009257. ISSN: 1553-7374. DOI: 10.1371/JOURNAL.PPAT.1009257. URL: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1009257.

[223] Max Medina-Ramírez et al. "Design and crystal structure of a native-like HIV-1 envelope trimer that engages multiple broadly neutralizing antibody precursors in vivo". In: 214.9 (Sept. 2017), pp. 2573–2590. ISSN: 1540-9538. URL: http://www.ncbi.nlm.nih.gov/pubmed/28847869.

[224] Whitaker N et al. "Developability Assessment of Physicochemical Properties and Stability Profiles of HIV-1 BG505 SOSIP.664 and BG505 SOSIP.v4.1-GT1.1 gp140 Envelope Glycoprotein Trimers as Candidate Vaccine Antigens". In: *Journal of pharmaceutical sciences* 108.7 (July 2019), pp. 2264–2277. ISSN: 1520-6017. DOI: 10.1016/J.XPHS.2019.01.033. URL: https://pubmed.ncbi.nlm.nih.gov/30776383/.

[225] E.T. Ema T. Crooks et al. "Glycoengineering HIV-1 Env creates 'supercharged' and 'hybrid' glycans to increase neutralizing antibody potency, breadth and saturation". In: *PLoS Pathogens* 14.5 (2018). ISSN: 15537374. DOI: 10.1371/journal.ppat.1007024.

[226] *HIV-1 isolate 9032.08_A1 from USA envelope glycoprotein (env) gene, co - Nucleotide - NCBI*. URL: https://www.ncbi.nlm.nih.gov/nuccore/EU576114.

[227] *HIV-1 isolate BG505.M27P.ENV.K1 from Kenya envelope glycoprotein (env) - Nucleotide - NCBI*. URL: https://www.ncbi.nlm.nih.gov/nuccore/MW650637.1.

[228] *HIV-1 isolate X1193_1 from Spain envelope glycoprotein (env) gene, com - Nucleotide - NCBI*. URL: https://www.ncbi.nlm.nih.gov/nuccore/EU885761.

[229] *HIV-1 isolate ZM197M clone PB7 from Zambia truncated rev protein (rev) - Nucleotide - NCBI*. URL: https://www.ncbi.nlm.nih.gov/nuccore/DQ388515.

[230] Daniel Wrapp et al. "Structure-Based Stabilization of SOSIP Env Enhances Recombinant Ectodomain Durability and Yield". In: *Journal of Virology* (Jan. 2023). ISSN: 0022-538X. DOI: 10.1128/JVI.01673-22/SUPPL{_}FILE/JVI.01673-22-S0001.PDF. URL: https://journals.asm.org/doi/10.1128/jvi.01673-22.

[231] Rogier W. Sanders and John P. Moore. "Native-like Env trimers as a platform for HIV-1 vaccine design". In: *Immunological Reviews* 275.1 (Jan. 2017), pp. 161–182. ISSN: 01052896. DOI: 10.1111/imr.12481. URL: http://doi.wiley.com/10.1111/imr.12481.

[232] Leonidas Stamatatos, Marie Pancera, and Andrew T. McGuire. "Germline-targeting immunogens". In: *Immunological Reviews* 275.1 (Jan. 2017), pp. 203–216. ISSN: 1600-065X. DOI: 10.1111/IMR.12483. URL: https://onlinelibrary.wiley.com/doi/10.1111/imr.12483.

[233] Emilia Falkowska et al. "Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers". In: *Immunity* 40.5 (May 2014), pp. 657–68. ISSN: 10974180. DOI: 10.1016/j.immuni.2014.04.009.

[234] Claudia Blattner et al. "Structural delineation of a quaternary, cleavage-dependent epitope at the gp41-gp120 interface on intact HIV-1 env trimers". In: *Immunity* 40.5 (May 2014), pp. 669–680. ISSN: 10974180. DOI: 10.1016/j.immuni.2014.04.008.

[235] Katie J. Doores and Dennis R. Burton. "Variable Loop Glycan Dependency of the Broad and Potent HIV-1-Neutralizing Antibodies PG9 and PG16". In: *Journal of Virology* 84.20 (Oct. 2010), pp. 10510–10521. ISSN: 0022-538X. DOI: 10.1128/jvi.00552-10.

[236] Jason S. McLellan et al. "Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9". In: *Nature* 480.7377 (Dec. 2011), pp. 336–343. ISSN: 00280836. DOI: 10.1038/nature10696.

[237] R. Pejchal et al. "A Potent and Broad Neutralizing Antibody Recognizes and Penetrates the HIV Glycan Shield". In: *Science* 334.6059 (Nov. 2011), pp. 1097–1103. ISSN: 0036-8075. DOI: 10.1126/science.1213256.

[238] Louise Scharf et al. "Antibody 8ANC195 Reveals a Site of Broad Vulnerability on the HIV-1 Envelope Spike". In: *Cell Reports* (2014). ISSN: 22111247. DOI: 10.1016/j.celrep.2014.04.001.

[239] Laura M. Walker et al. "Broad neutralization coverage of HIV by multiple highly potent antibodies". In: *Nature* (2011). ISSN: 00280836. DOI: 10.1038/nature10373.

[240] Dennis R. Burton et al. "Broad and potent neutralizing antibodies from an african donor reveal a new HIV-1 vaccine target". In: *Science* (2009). ISSN: 00368075. DOI: 10.1126/science.1178746.

[241] Johannes S. Gach et al. "Antibody Responses Elicited by Immunization with BG505 Trimer Immune Complexes". In: *Journal of Virology* (2019). ISSN: 0022-538X. DOI: 10.1128/jvi.01188-19.

[242] E.T. Ema T. Crooks et al. "Vaccine-Elicited Tier 2 HIV-1 Neutralizing Antibodies Bind to Quaternary Epitopes Involving Glycan-Deficient Patches Proximal to the CD4 Binding Site". In: *PLoS Pathogens* 11.5 (May 2015). ISSN: 15537374. DOI: 10.1371/journal.ppat.1004932.

[243] E.T. Ema T. Crooks et al. "Effects of partially dismantling the CD4 binding site glycan fence of HIV-1 Envelope glycoprotein trimers on neutralizing antibody induction". In: *Virology* 505 (2017), pp. 193–209. ISSN: 10960341. DOI: 10.1016/j.virol.2017.02.024.

[244] James E. J.E. Voss et al. "Elicitation of Neutralizing Antibodies Targeting the V2 Apex of the HIV Envelope Trimer in a Wild-Type Animal Model". In: *Cell Reports* 21.1 (Oct. 2017), pp. 222–235. ISSN: 22111247. DOI: 10.1016/j.celrep.2017.09.024. URL: https://www.cell.com/cell-reports/abstract/S2211-1247(17)31299-8.

[245] Laura E. L.E. McCoy et al. "Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies". In: *Cell Reports* 16.9 (Aug. 2016), pp. 2327–2338. ISSN: 22111247. DOI: 10.1016/j.celrep.2016.07.074.

[246] L. Dacheux et al. "Evolutionary Dynamics of the Glycan Shield of the Human Immunodeficiency Virus Envelope during Natural Infection and Implications for Exposure of the 2G12 Epitope". In: *Journal of Virology* (2004). ISSN: 0022-538X. DOI: 10.1128/jvi.78.22.12625-12637.2004.

[247] K. J. Doores et al. "2G12-Expressing B Cell Lines May Aid in HIV Carbohydrate Vaccine Design Strategies". In: *Journal of Virology* (2013). ISSN: 0022-538X. DOI: 10.1128/jvi.02820-12.

[248] Sam Hoot et al. "Recombinant HIV Envelope Proteins Fail to Engage Germline Versions of Anti-CD4bs bNAbs". In: *PLoS Pathogens* (2013). ISSN: 15537366. DOI: 10.1371/journal.ppat.1003106.

[249] Ben Jiang Ma et al. "Envelope deglycosylation enhances antigenicity of hiv-1 gp41 epitopes for both broad neutralizing antibodies and their unmutated ancestor antibodies". In: *PLoS Pathogens* (2011). ISSN: 15537366. DOI: 10.1371/journal.ppat.1002200.

[250] Bryan Briney et al. "Tailored Immunogens Direct Affinity Maturation toward HIV Neutralizing Antibodies". In: *Cell* 166.6 (Sept. 2016), pp. 1459–1470. ISSN: 00928674. URL: https://pubmed.ncbi.nlm.nih.gov/27610570/.

[251] Andrew T. McGuire et al. "Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies". In: *Journal of Experimental Medicine* 210.4 (Apr. 2013), pp. 655–63. ISSN: 00221007. DOI: 10.1084/jem.20122824. URL: http://www.ncbi.nlm.nih.gov/pubmed/23530120.

[252] Joseph Jardine et al. "Rational HIV immunogen design to target specific germline B cell receptors". In: 340.6133 (May 2013). ISSN: 1095-9203. DOI: 10.1126/science.1234150. URL: http://www.ncbi.nlm.nih.gov/pubmed/23539181.

[253] Jon M. J.M. Steichen et al. "HIV Vaccine Design to Target Germline Precursors of Glycan-Dependent Broadly Neutralizing Antibodies". In: *Immunity* 45.3 (Sept. 2016), pp. 483–496. ISSN: 10974180. DOI: 10.1016/j.immuni.2016.08.016.

[254] G.B.E. Guillaume B.E. B.E. Stewart-Jones et al. "Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G". In: *Cell* 165.4 (May 2016), pp. 813–826. ISSN: 10974172. DOI: 10.1016/j.cell.2016.04.010.

[255] Anna Janina Behrens et al. "Integrity of Glycosylation Processing of a Glycan-Depleted Trimeric HIV-1 Immunogen Targeting Key B-Cell Lineages". In: *Journal of Proteome Research* (2018). ISSN: 15353907. DOI: 10.1021/acs.jproteome.7b00639.

[256] Laura K. Pritchard et al. "Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies". In: *Nature Communications* 6.1 (2015), p. 7479. ISSN: 2041-1723. DOI: 10.1038/ncomms8479. URL: http://www.nature.com/articles/ncomms8479.

[257] Gemma E. Seabright et al. "Networks of HIV-1 Envelope Glycans Maintain Antibody Epitopes in the Face of Glycan Additions and Deletions". In: *Structure* (2020). ISSN: 18784186. DOI: 10.1016/j.str.2020.04.022.

[258] Thomas Lemmin et al. "Microsecond Dynamics and Network Analysis of the HIV-1 SOSIP Env Trimer Reveal Collective Behavior and Conserved Microdomains of the Glycan Shield". In: *Structure* (2017). ISSN: 18784186. DOI: 10.1016/j.str.2017.07.018.

[259] Alba Torrents de la Peña et al. "Improving the Immunogenicity of Native-like HIV-1 Envelope Trimers by Hyperstabilization". In: *Cell Reports* 20.8 (Aug. 2017), pp. 1805–1817. ISSN: 2211-1247. DOI: 10.1016/J.CELREP.2017.07.077.

[260] Christopher N. Scanlan et al. *Exploiting the defensive sugars of HIV-1 for drug and vaccine design*. 2007. DOI: 10.1038/nature05818.

[261] Liwei Cao et al. "Differential processing of HIV envelope glycans on the virus and soluble recombinant trimer". In: *Nature Communications* 9.1 (Dec. 2018), p. 3693. ISSN: 2041-1723. DOI: 10.1038/s41467-018-06121-4.

[262] R. W. Sanders et al. "The Mannose-Dependent Epitope for Neutralizing Antibody 2G12 on Human Immunodeficiency Virus Type 1 Glycoprotein gp120". In: *Journal of Virology* (2002). ISSN: 0022-538X. DOI: 10.1128/jvi.76.14.7293-7305.2002.

[263] Wenbo Wang et al. "A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization". In: 10.1 (Feb. 2013). ISSN: 1742-4690. DOI: 10.1186/1742-4690-10-14. URL: https://retrovirology.biomedcentral.com/articles/10.1186/1742-4690-10-14.

[264] Leopold Kong, Ian A Wilson, and Peter D Kwong. "Crystal structure of a fully glycosylated HIV-1 gp120 core reveals a stabilizing role for the glycan at Asn262." In: *Proteins* 83.3 (Mar. 2015), pp. 590–6. ISSN: 1097-0134. DOI: 10.1002/prot.24747.

[265] Laura K. Pritchard et al. "Glycan Microheterogeneity at the PGT135 Antibody Recognition Site on HIV-1 gp120 Reveals a Molecular Mechanism for Neutralization Resistance". In: *Journal of Virology* (2015). ISSN: 0022-538X. DOI: 10.1128/jvi.00230-15.

[266] Raiees Andrabi et al. "Glycans Function as Anchors for Antibodies and Help Drive HIV Broadly Neutralizing Antibody Development". In: *Immunity* 47.3 (Sept. 2017), pp. 524–537. ISSN: 10974180. DOI: 10.1016/j.immuni.2017.08.006.

[267] Viktoriya Dubrovskaya et al. "Targeted N-glycan deletion at the receptor-binding site retains HIV Env NFL trimer integrity and accelerates the elicited antibody response". In: *PLoS Pathogens* 13.9 (Sept. 2017).

[268] Dennis R. D.R. Burton and Lars Hangartner. "Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design". In: *Annual Review of Immunology* 34.1 (May 2016), pp. 635–659. ISSN: 0732-0582. DOI: 10.1146/annurev-immunol-041015-055515.

[269] Kai Xu et al. "Epitope-based vaccine design yields fusion peptide-directed antibodies that neutralize diverse strains of HIV-1". In: 24.6 (June 2018), pp. 857–867. ISSN: 1546-170X. URL: https://pubmed.ncbi.nlm.nih.gov/29867235/.

[270] Dubrovskaya V et al. "Vaccination with Glycan-Modified HIV NFL Envelope Trimer-Liposomes Elicits Broadly Neutralizing Antibodies to Multiple Sites of Vulnerability". In: *Immunity* 51.5 (Nov. 2019), pp. 915–929. ISSN: 1097-4180. URL: https://pubmed.ncbi.nlm.nih.gov/31732167/.

[271] Rui Kong et al. "Antibody Lineages with Vaccine-Induced Antigen-Binding Hotspots Develop Broad HIV Neutralization". In: *Cell* 178.3 (July 2019), pp. 567–584. ISSN: 1097-4172. DOI: 10.1016/J.CELL.2019.06.030. URL: https://pubmed.ncbi.nlm.nih.gov/31348886/.

[272] Chuang GY et al. "Development of a 3Mut-Apex-Stabilized Envelope Trimer That Expands HIV-1 Neutralization Breadth When Used To Boost Fusion Peptide-Directed Vaccine-Elicited Responses." In: *Journal of Virology* 94.13 (June 2020). ISSN: 0022-538X. DOI: 10.1128/JVI.00074-20. URL: https://europepmc.org/articles/PMC7307166.

[273] Laurent Verkoczy, F.W. Frederick W. Alt, and Ming Tian. "Human Ig knockin mice to study the development and regulation of HIV-1 broadly neutralizing antibodies". In: *Immunological Reviews* 275.1 (Jan. 2017), pp. 89–107.

[274] Peter D. P.D. Kwong and J.R. John R. Mascola. "HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure". In: *Immunity* 48.5 (May 2018), pp. 855–871.

[275] Max Medina-Ramírez, R.W. Rogier W. Sanders, and P.J. Per Johan Klasse. "Targeting B-cell germlines and focusing affinity maturation: The next hurdles in HIV-1-vaccine development?" In: *Expert Review of Vaccines* 13.4 (2014), pp. 449–452. DOI: 10.1586/14760584.2014.894469.

[276] Verkoczy L and L. Verkoczy. *Humanized Immunoglobulin Mice: Models for HIV Vaccine Testing and Studying the Broadly Neutralizing Antibody Problem*. 2017. URL: https://pubmed.ncbi.nlm.nih.gov/28413022/.

[277] Nicole A. N.A. Doria-Rose et al. "Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies". In: *Nature* 509.7498 (May 2014), pp. 55–62. ISSN: 0028-0836. DOI: 10.1038/nature13036. URL: http://www.ncbi.nlm.nih.gov/pubmed/24590074.

[278] Raiees Andrabi et al. "Identification of Common Features in Prototype Broadly Neutralizing Antibodies to HIV Envelope V2 Apex to Facilitate Vaccine Design". In: *Immunity* 43.5 (Nov. 2015), pp. 959–973. ISSN: 10974180. DOI: 10.1016/j.immuni.2015.10.014.

[279] Jason Gorman et al. "Structures of HIV-1 Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design". In: *Nature Structural and Molecular Biology* 23.1 (Jan. 2016), pp. 81–90.

[280] E.M. Evan M. Cale et al. "Virus-like Particles Identify an HIV V1V2 Apex-Binding Neutralizing Antibody that Lacks a Protruding Loop". In: *Immunity* 46.5 (May 2017), pp. 777–791.

[281] Mattia Bonsignori et al. "Analysis of a clonal lineage of HIV-1 envelope V2/V3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors". In: *Journal of virology* 85.19 (Oct. 2011), pp. 9998–10009.

ISSN: 1098-5514. DOI: 10.1128/JVI.05045-11. URL: https://pubmed.ncbi.nlm.nih.gov/21795340/.

[282] Kwinten Sliepen et al. "Binding of inferred germline precursors of broadly neutralizing HIV-1 antibodies to native-like envelope trimers." In: *Virology* 486 (Dec. 2015), pp. 116–20. ISSN: 1096-0341. DOI: 10.1016/j.virol.2015.08.002. URL: http://www.ncbi.nlm.nih.gov/pubmed/26433050.

[283] Takayuki Ota et al. "Anti-HIV B Cell lines as candidate vaccine biosensors". In: *Journal of immunology (Baltimore, Md. : 1950)* 189.10 (Nov. 2012), pp. 4816–4824. ISSN: 1550-6606. DOI: 10.4049/JIMMUNOL.1202165. URL: https://pubmed.ncbi.nlm.nih.gov/23066156/.

[284] J.G. Joseph G. Jardine et al. "Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen". In: *Science* 349.6244 (July 2015), pp. 156–161.

[285] Tongqing Zhou et al. "Quantification of the Impact of the HIV-1-Glycan Shield on Antibody Elicitation". In: *Cell Reports* 19.4 (Apr. 2017), pp. 719–732.

[286] C.C. Celia C. LaBranche et al. "Neutralization-guided design of HIV-1 envelope trimers with high affinity for the unmutated common ancestor of CH235 lineage CD4bs broadly neutralizing antibodies". In: *PLoS Pathogens* 15.9 (2019).

[287] Celia C.C. LaBranche et al. "HIV-1 envelope glycan modifications that permit neutralization by germline-reverted VRC01-class broadly neutralizing antibodies". In: *PLoS Pathogens* 14.11 (Nov. 2018).

[288] Elise Landais et al. "HIV Envelope Glycoform Heterogeneity and Localized Diversity Govern the Initiation and Maturation of a V2 Apex Broadly Neutralizing Antibody Lineage". In: *Immunity* 47.5 (Nov. 2017), pp. 990–1003. ISSN: 1074-7613. URL: https://www.cell.com/immunity/abstract/S1074-7613(17)30479-X.

[289] Ralph Pantophlet et al. "Bacterially derived synthetic mimetics of mammalian oligomannose prime antibody responses that neutralize HIV infectivity". In: *Nature Communications* 8.1 (Dec. 2017).

[290] Cheng Cheng et al. "Consistent elicitation of cross-clade HIV-neutralizing responses achieved in Guinea pigs after fusion peptide priming by repetitive envelope trimer boosting". In: *PLoS ONE* 14.4 (Apr. 2019).

[291] Adam S. A.S. Dingens et al. "Complete functional mapping of infection- and vaccine-elicited antibodies against the fusion peptide of HIV". In: *PLoS Pathogens* 14.7 (July 2018).

[292] Muriel Lavie, Xavier Hanoulle, and Jean Dubuisson. "Glycan Shielding and Modulation of Hepatitis C Virus Neutralizing Antibodies." In: *Frontiers in immunology* 9 (2018), p. 910. ISSN: 1664-3224. DOI: 10.3389/fimmu.2018.00910.

[293] Alemu Tekewe Mogus et al. "Virus-Like Particle Based Vaccines Elicit Neutralizing Antibodies against the HIV-1 Fusion Peptide". In: *Vaccines* 8.4 (Dec. 2020), pp. 1–15. ISSN: 2076-393X. DOI: 10.3390/VACCINES8040765. URL: https://pubmed.ncbi.nlm.nih.gov/33333740/.

[294] Mattia Bonsignori et al. "Maturation Pathway from Germline to Broad HIV-1 Neutralizer of a CD4-Mimic Antibody". In: *Cell* 165.2 (Apr. 2016), pp. 449–463. ISSN: 1097-4172. DOI: 10.1016/J.CELL.2016.02.022. URL: https://pubmed.ncbi.nlm.nih.gov/26949186/.

[295] Cynthia A. C.A. Derdeyn, P.L. Penny L. Moore, and Lynn Morris. "Development of broadly neutralizing antibodies from autologous neutralizing antibody responses in HIV infection". In: *Current Opinion in HIV and AIDS* 9.3 (2014), pp. 210–216. DOI: 10.1097/COH.0000000000000057.

[296] Elise Landais and P.L. Penny L. Moore. "Development of broadly neutralizing antibodies in HIV-1 infected elite neutralizers". In: *Retrovirology* 15.1 (Sept. 2018).

[297] P.L. Penny L. Moore et al. "Ontogeny-based immunogens for the induction of V2-directed HIV broadly neutralizing antibodies". In: *Immunological Reviews* 275.1 (Jan. 2017), pp. 217–229.

[298] Tongqing Zhou et al. "Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01". In: *Science* 329.5993 (Aug. 2010), pp. 811–817. ISSN: 00368075. DOI: 10.1126/science.1192819.

[299] Tongqing Zhou et al. "Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors". In: *Cell* 161.6 (June 2015), pp. 1280–1292.

[300] Peter D. Kwong and John R. Mascola. "Human Antibodies that Neutralize HIV-1: Identification, Structures, and B Cell Ontogenies". In: *Immunity* 37.3 (Sept. 2012), pp. 412–425.

[301] Jinghe Huang et al. "Identification of a CD4-Binding-Site Antibody to HIV that Evolved Near-Pan Neutralization Breadth". In: *Immunity* 45.5 (Nov. 2016), pp. 1108–1121.

[302] Amelia Escolano et al. "Sequential Immunization Elicits Broadly Neutralizing Anti-HIV-1 Antibodies in Ig Knockin Mice". In: *Cell* 166.6 (Sept. 2016), pp. 1445–1458. ISSN: 10974172. DOI: 10.1016/j.cell.2016.07.030.

[303] Christine A. C.A. Bricault et al. "HIV-1 Neutralizing Antibody Signatures and Application to Epitope-Targeted Vaccine Design". In: *Cell Host and Microbe* 25.1 (Jan. 2019), pp. 59–72.

[304] Raiees Andrabi, Jinal N. Bhiman, and Dennis R. D.R. Burton. "Strategies for a multistage neutralizing antibody-based HIV vaccine". In: *Current Opinion in Immunology* 53 (Aug. 2018), pp. 143–151. ISSN: 1879-0372. DOI: 10.1016/j.coi.2018.04.025. URL: http://www.ncbi.nlm.nih.gov/pubmed/29775847.

[305] Raiees Andrabi et al. "The Chimpanzee SIV Envelope Trimer: Structure and Deployment as an HIV Vaccine Template". In: *Cell Reports* 27.8 (May 2019), pp. 2426–2441. ISSN: 2211-1247. DOI: 10.1016/J.CELREP.2019.04.082.

[306] Alba Torrents de la Peña et al. "Immunogenicity in rabbits of SOSIP trimers from clades A, B and C, given individually, sequentially or in combinations". In: *Journal of Virology* (2018), pp. 01957–17. ISSN: 0022-538X. DOI: 10.1128/JVI.01957-17.

[307] Susan Zolla-Pazner et al. "Rationally Designed Vaccines Targeting the V2 Region of HIV-1 gp120 Induce a Focused, Cross-Clade-Reactive, Biologically Functional Antibody Response". In: *Journal of virology* 90.24 (Dec. 2016), pp. 10993–11006. ISSN: 1098-5514. DOI: 10.1128/JVI.01403-16. URL: https://pubmed.ncbi.nlm.nih.gov/27630234/.

[308] Kevin Wiehe et al. "Functional Relevance of Improbable Antibody Mutations for HIV Broadly Neutralizing Antibody Development". In: *Cell host & microbe* 23.6 (June 2018), pp. 759–765. ISSN: 1934-6069. DOI: 10.1016/J.CHOM.2018.04.018. URL: https://pubmed.ncbi.nlm.nih.gov/29861171/.

[309] N A Doria-Rose et al. "A short segment in the HIV-1 gp120 V1/V2 region is a major determinant of neutralization resistance to PG9-like antibodies". In: *Retrovirology 2012 9:2* 9.2 (Sept. 2012), pp. 1–1. ISSN: 1742-4690. DOI: 10.1186/1742-4690-9-S2-O29. URL: https://retrovirology.biomedcentral.com/articles/10.1186/1742-4690-9-S2-O29.

[310] M.N. Mohammed N. Amin et al. "Synthetic glycopeptides reveal the glycan specificity of HIV-neutralizing antibodies". In: 9.8 (Aug. 2013). ISSN: 1552-4469. DOI: 10.1038/nchembio.1288. URL: http://www.ncbi.nlm.nih.gov/pubmed/23831758.

[311] Rui Kong et al. "Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody". In: *Science* 352.6287 (May 2016), pp. 828–833.

[312] van Gils MJ et al. "An HIV-1 antibody from an elite neutralizer implicates the fusion peptide as a site of vulnerability". In: 2 (Nov. 2016). ISSN: 2058-5276. URL: https://pubmed.ncbi.nlm.nih.gov/27841852/.

[313] Ronald Derking et al. "Enhancing glycan occupancy of soluble HIV-1 envelope trimers to mimic the native viral spike". In: *Cell Reports* 35.1 (Apr. 2021), p. 108933. ISSN: 2211-1247. DOI: 10.1016/J.CELREP.2021.108933.

[314] Mayumi Igura and Daisuke Kohda. "Quantitative assessment of the preferences for the amino acid residues flanking archaeal N-linked glycosylation sites". In: *Glycobiology* 21.5 (May 2011), pp. 575–583. DOI: 10.1093/GLYCOB/CWQ196.

[315] Emma T. Crooks et al. "Engineering well-expressed, V2-immunofocusing HIV-1 envelope glycoprotein membrane trimers for use in heterologous prime-boost vaccine regimens". In: *PLOS Pathogens* (2021). ISSN: 15537374. DOI: 10.1371/journal.ppat.1009807.

[316] M. Gordon Joyce et al. "Soluble Prefusion Closed DS-SOSIP.664-Env Trimers of Diverse HIV-1 Strains". In: *Cell Reports* 21.10 (Dec. 2017), pp. 2992–3002. DOI: 10.1016/J.CELREP.2017.11.016.

[317] Nicholas McCaul et al. "Intramolecular quality control: HIV-1 envelope gp160 signal-peptide cleavage as a functional folding checkpoint". In: *Cell Reports* (2021). ISSN: 22111247. DOI: 10.1016/j.celrep.2021.109646.

[318] Lifei Yang et al. "Structure-guided redesign improves NFL HIV Env Trimer integrity and identifies an inter-protomer disulfide permitting post-expression cleavage". In: *Frontiers in Immunology* (2018). ISSN: 16643224. DOI: 10.3389/fimmu.2018.01631.

[319] Daniel P. Leaman et al. "Immunogenic Display of Purified Chemically Cross-Linked HIV-1 Spikes". In: *Journal of Virology* (2015). ISSN: 0022-538X. DOI: 10.1128/jvi.03738-14.

[320] Wen Han Yu et al. "Exploiting glycan topography for computational design of Env glycoprotein antigenicity". In: *PLoS Computational Biology* (2018). ISSN: 15537358. DOI: 10.1371/journal.pcbi.1006093.

[321] Samantha Townsley et al. "Conserved Role of an N-Linked Glycan on the Surface Antigen of Human Immunodeficiency Virus Type 1 Modulating Virus Sensitivity to Broadly Neutralizing Antibodies against the Receptor and Coreceptor Binding Sites". In: *Journal of Virology* (2016). ISSN: 0022-538X. DOI: 10.1128/jvi.02321-15.

[322] Fawzi Khoder-Agha et al. "Assembly of B4GALT1/ST6GAL1 heteromers in the Golgi membranes involves lateral interactions via highly charged surface domains". In: *Journal of Biological Chemistry* (2019). ISSN: 1083351X. DOI: 10.1074/jbc.RA119.009539.

[323] O'Flaherty R et al. "The sweet spot for biologics: recent advances in characterization of biotherapeutic glycoproteins". In: *Expert review of proteomics* 15.1 (Jan. 2018),

pp. 13–29. ISSN: 1744-8387. DOI: 10.1080/14789450.2018.1404907. URL: https://pubmed.ncbi.nlm.nih.gov/29130774/.

[324] Tejwani V et al. "Glycoengineering in CHO Cells: Advances in Systems Biology". In: *Biotechnology journal* 13.3 (Mar. 2018). ISSN: 1860-7314. DOI: 10.1002/BIOT.201700234. URL: https://pubmed.ncbi.nlm.nih.gov/29316325/.

[325] J. M. Binley et al. "Role of Complex Carbohydrates in Human Immunodeficiency Virus Type 1 Infection and Resistance to Antibody Neutralization". In: *Journal of Virology* (2010). ISSN: 0022-538X. DOI: 10.1128/jvi.00105-10.

[326] Joel D. Allen et al. "Subtle Influence of ACE2 Glycan Processing on SARS-CoV-2 Recognition". In: *Journal of Molecular Biology* 433.4 (Feb. 2021), p. 166762. ISSN: 0022-2836. DOI: 10.1016/J.JMB.2020.166762.

[327] Keith Wanzeck, Kelli L. Boyd, and Jonathan A. McCullers. "Glycan Shielding of the Influenza Virus Hemagglutinin Contributes to Immunopathology in Mice". In: *American Journal of Respiratory and Critical Care Medicine* 183.6 (Mar. 2011), pp. 767–773. ISSN: 1073-449X. DOI: 10.1164/rccm.201007-1184OC.

[328] Katie J. Doores. "The HIV glycan shield as a target for broadly neutralizing antibodies". In: *FEBS Journal* 282.24 (Dec. 2015), pp. 4679–4691. ISSN: 1742464X. DOI: 10.1111/febs.13530.

[329] Christopher N. Scanlan et al. "The carbohydrate epitope of the neutralizing anti-HIV-1 antibody 2G12". In: *Advances in Experimental Medicine and Biology*. Springer, Boston, MA, 2002, pp. 205–218. DOI: 10.1007/978-1-4615-0065-0{_}13. URL: http://link.springer.com/10.1007/978-1-4615-0065-0_13.

[330] Kwinten Sliepen et al. "Structure and immunogenicity of a stabilized HIV-1 envelope trimer based on a group-M consensus sequence". In: *Nature Communications* (2019). ISSN: 20411723. DOI: 10.1038/s41467-019-10262-5.

[331] Joseph G. Jardine et al. "HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen". In: 351.6280 (Mar. 2016), pp. 1458–1463. URL: https://pubmed.ncbi.nlm.nih.gov/27013733/.

[332] Byrne G et al. "CRISPR/Cas9 gene editing for the creation of an MGAT1-deficient CHO cell line to control HIV-1 vaccine glycosylation". In: *PLoS biology* 16.8 (Aug. 2018). ISSN: 1545-7885. DOI: 10.1371/JOURNAL.PBIO.2005817. URL: https://pubmed.ncbi.nlm.nih.gov/30157178/.

[333] Jean Philippe Julien et al. "Crystal structure of a soluble cleaved HIV-1 envelope trimer". In: *Science* (2013). ISSN: 10959203. DOI: 10.1126/science.1245625.

[334] Erin E.H. Tran et al. "Structural mechanism of trimeric HIV-1 envelope glycoprotein activation". In: *PLoS Pathogens* (2012). ISSN: 15537366. DOI: 10.1371/journal.ppat.1002797.

[335] Rajesh P. Ringe et al. "Improving the Expression and Purification of Soluble, Recombinant Native-Like HIV-1 Envelope Glycoprotein Trimers by Targeted Sequence Changes". In: *Journal of Virology* 91.12 (June 2017). Ed. by Guido Silvestri. ISSN: 0022-538X. DOI: 10.1128/JVI.00264-17. URL: https://journals.asm.org/doi/10.1128/JVI.00264-17.

[336] Xueling Wu et al. "Rational Design of Envelope Identifies Broadly Neutralizing Human Monoclonal Antibodies to HIV-1". In: *Science* 329.5993 (Aug. 2010), pp. 856–861. ISSN: 0036-8075. DOI: 10.1126/science.1187659. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.1187659.

[337] Devin Sok et al. "The Effects of Somatic Hypermutation on Neutralization and Binding in the PGT121 Family of Broadly Neutralizing HIV Antibodies". In: *PLoS Pathogens* (2013). ISSN: 15537366. DOI: 10.1371/journal.ppat.1003754.

[338] Hugo Mouquet et al. "Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies". In: *Proceedings of the National Academy of Sciences of the United States of America* (2012). ISSN: 00278424. DOI: 10.1073/pnas.1217207109.

[339] Daniel A. Calarese et al. "Dissection of the carbohydrate specificity of the broadly neutralizing anti-HIV-1 antibody 2G12". In: *Proceedings of the National Academy of Sciences of the United States of America* (2005). ISSN: 00278424. DOI: 10.1073/pnas.0505763102.

[340] C. D. Murin et al. "Structure of 2G12 Fab2 in Complex with Soluble and Fully Glycosylated HIV-1 Env by Negative-Stain Single-Particle Electron Microscopy". In: *Journal of Virology* (2014). ISSN: 0022-538X. DOI: 10.1128/jvi.01229-14.

[341] X. Ji et al. "Mannose-binding lectin binds to Ebola and Marburg envelope glycoproteins, resulting in blocking of virus interaction with DC-SIGN and complement-mediated virus neutralization". In: *Journal of General Virology* 86.9 (Sept. 2005), pp. 2535–2542. ISSN: 0022-1317. DOI: 10.1099/vir.0.81199-0.

[342] Leopold Kong et al. "Expression-System-Dependent Modulation of HIV-1 Envelope Glycoprotein Antigenicity and Immunogenicity". In: *Journal of Molecular Biology* (2010). ISSN: 00222836. DOI: 10.1016/j.jmb.2010.08.033.

[343] Hannah L. Turner et al. "Disassembly of HIV envelope glycoprotein trimer immunogens is driven by antibodies elicited via immunization". In: *Science Advances* 7.31 (July 2021). ISSN: 23752548. DOI: 10.1126/SCIADV.ABH2791/SUPPL{_}FILE/

SCIADV.ABH2791{_}SM.PDF. URL: https://www.science.org/doi/10.1126/sciadv.abh2791.

[344] Gemma E. Seabright et al. *Protein and Glycan Mimicry in HIV Vaccine Design*. 2019. DOI: 10.1016/j.jmb.2019.04.016.

[345] Joel D. Allen et al. "Harnessing post-translational modifications for next-generation HIV immunogens". In: *Biochemical Society Transactions* 46.3 (June 2018), pp. 691–698. ISSN: 0300-5127. DOI: 10.1042/BST20170394. URL: http://www.ncbi.nlm.nih.gov/pubmed/29784645.

[346] Masashi Shingai et al. "Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques." In: *The Journal of experimental medicine* 211.10 (Sept. 2014), pp. 2061–74. ISSN: 1540-9538. DOI: 10.1084/jem.20132494.

[347] Max Medina-Ramírez, Rogier W. Sanders, and Quentin J. Sattentau. "Stabilized HIV-1 envelope glycoprotein trimers for vaccine use". In: *Current Opinion in HIV and AIDS* 0 (2017), p. 1. ISSN: 1746-630X. DOI: 10.1097/COH.0000000000000363.

[348] Rajesh P Ringe et al. "Reducing V3 antigenicity and immunogenicity on soluble, native-like HIV-1 Env SOSIP trimers". In: *Journal of Virology* 91.15 (Aug. 2017), pp. 00677–17. ISSN: 0022-538X. DOI: 10.1128/JVI.00677-17.

[349] Pavel Pugach et al. "A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene." In: *Journal of virology* 89.6 (Mar. 2015), pp. 3380–95. ISSN: 1098-5514. DOI: 10.1128/JVI.03473-14.

[350] Alba Torrents de la Peña et al. "Similarities and differences between native HIV-1 envelope glycoprotein trimers and stabilized soluble trimer mimetics". In: *PLOS Pathogens* 15.7 (July 2019). Ed. by Alexandra Trkola, e1007920. ISSN: 1553-7374. DOI: 10.1371/journal.ppat.1007920.

[351] Talar Tokatlian et al. "Innate immune recognition of glycans targets HIV nanoparticle immunogens to germinal centers". In: *Science* (Dec. 2018), eaat9120. ISSN: 0036-8075. DOI: 10.1126/SCIENCE.AAT9120.

[352] Anna Janina Behrens, Weston B. Struwe, and Max Crispin. *Glycosylation profiling to evaluate glycoprotein immunogens against HIV-1*. Oct. 2017. DOI: 10.1080/14789450.2017.1376658. URL: https://www.tandfonline.com/doi/abs/10.1080/14789450.2017.1376658.

[353] Anna Schorcht et al. "The Glycan Hole Area of HIV-1 Envelope Trimers Contributes Prominently to the Induction of Autologous Neutralization". In: *Journal of virology*

96.1 (Jan. 2022). ISSN: 1098-5514. DOI: 10.1128/JVI.01552-21. URL: https://pubmed.ncbi.nlm.nih.gov/34669426/.

[354] Kshitij Wagh, Beatrice H. Hahn, and Bette Korber. "Hitting the sweet spot: Exploiting HIV-1 glycan shield for induction of broadly neutralizing antibodies". In: *Current Opinion in HIV and AIDS* 15.5 (Sept. 2020), pp. 267–274. ISSN: 17466318. DOI: 10.1097/COH.0000000000000639. URL: https://journals.lww.com/cohivandaids/Fulltext/2020/09000/Hitting_the_sweet_spot__exploiting_HIV_1_glycan.2.aspx.

[355] Matteo Bianchi et al. "Electron-Microscopy-Based Epitope Mapping Defines Specificities of Polyclonal Antibodies Elicited during HIV-1 BG505 Envelope Trimer Immunization". In: *Immunity* 49.2 (Aug. 2018), pp. 288–300. DOI: 10.1016/J.IMMUNI.2018.07.009.

[356] Joyce K. Hu et al. "Murine Antibody Responses to Cleaved Soluble HIV-1 Envelope Trimers Are Highly Restricted in Specificity". In: *Journal of Virology* 89.20 (Oct. 2015), pp. 10383–10398. DOI: 10.1128/JVI.01653-15.

[357] Matthias Pauthner et al. "Elicitation of Robust Tier 2 Neutralizing Antibody Responses in Nonhuman Primates by HIV Envelope Trimer Immunization Using Optimized Approaches". In: *Immunity* (2017). ISSN: 10974180. DOI: 10.1016/j.immuni.2017.05.007.

[358] Christopher A. Cottrell et al. "Mapping the immunogenic landscape of near-native HIV-1 envelope trimers in non-human primates". In: *PLoS Pathogens* (2020). ISSN: 15537374. DOI: 10.1371/journal.ppat.1008753.

[359] Yuhe R. Yang et al. "Autologous Antibody Responses to an HIV Envelope Glycan Hole Are Not Easily Broadened in Rabbits". In: *Journal of Virology* (2020). ISSN: 0022-538X. DOI: 10.1128/jvi.01861-19.

[360] David J. Leggat et al. "Vaccination induces HIV broadly neutralizing antibody precursors in humans". In: *Science* 378.6623 (Dec. 2022). ISSN: 10959203. DOI: 10.1126/SCIENCE.ADD6502/SUPPL{_}FILE/SCIENCE.ADD6502{_}DATA{_}S1{_}AND{_}S2.ZIP. URL: https://www.science.org/doi/10.1126/science.add6502.

[361] Priya Venkatesan. "Preliminary phase 1 results from an HIV vaccine candidate trial". In: *The Lancet Microbe* 2.3 (Mar. 2021), e95. ISSN: 26665247. DOI: 10.1016/s2666-5247(21)00042-2. URL: https://www.thelancet.com/journals/lanmic/article/PIIS2666-5247(21)00042-2/abstract.

[362] Rebecca M. Lynch et al. "HIV-1 fitness cost associated with escape from the VRC01 class of CD4 binding site neutralizing antibodies". In: *Journal of virology* 89.8 (Apr.

2015), pp. 4201–4213. ISSN: 1098-5514. DOI: 10.1128/JVI.03608-14. URL: https://pubmed.ncbi.nlm.nih.gov/25631091/.

[363] Yuka Otsuka et al. "Diverse pathways of escape from all well-characterized VRC01-class broadly neutralizing HIV-1 antibodies". In: *PLoS Pathogens* 14.8 (Aug. 2018). ISSN: 15537374. DOI: 10.1371/JOURNAL.PPAT.1007238. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6117093/.

[364] Kevin O. Saunders et al. "Targeted selection of HIV-specific antibody mutations by engineering B cell maturation". In: *Science* 366.6470 (Dec. 2019). ISSN: 10959203. DOI: 10.1126/science.aay7199.