

THE IMPACT OF SEQUENCE DIVERSITY ON THE GLYCAN SHIELD OF HIV-1 IMMUNOGENS

by

Gemma E. Seabright

Somerville College, University of Oxford

A thesis submitted for the degree of *Doctor of Philosophy*

Trinity term 2019

Supervised by

Prof. Max Crispin and Dr Mark Wormald

Declaration

The work presented in this thesis was undertaken in Prof. Max Crispin's laboratory at the Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford and the School of Biological Sciences, Faculty of Environmental and Life Sciences, University of Southampton between 2015 and 2019. I, Gemma Seabright, declare that all the work presented here is my own, with the following exceptions. A number of recombinant proteins were provided by collaborators, as detailed in Table 2.1. Ion mobility-mass spectrometry experiments were performed with the help of Prof. David Harvey, University of Oxford and Yasunori Watanabe, University of Oxford (Appendix D). The liquid chromatography-mass spectrometry data presented in Chapter 7 were acquired jointly with Joel Allen, University of Southampton (Figures 7.3 and 7.4). The yeast strains used in Chapter 8, and the accompanying Western blot data, were provided by Prof. Chris Taron's laboratory, New England Biolabs (Figure 8.3). The introduction of this thesis is based on my first-author review in the *Journal of Molecular Biology*, titled 'Protein and glycan mimicry in HIV vaccine design'. Chapter 5 is based on my first-author paper currently under review for publication in *Structure*, titled 'Networks of HIV-1 envelope glycans maintain antibody epitopes in the face of glycan additions and deletions'. Chapter 7 contains data published in the *Cell Reports* manuscript, 'Site-specific glycosylation of virion-derived HIV-1 Env is mimicked by a soluble trimeric immunogen', on which I am listed as co-author. A full list of the publications to which I have contributed throughout my DPhil is provided in Appendix G. This work has not been submitted for another degree at this, or any other university.

Word count (excluding references): 47'908

Table of contents

Abstract ..... **i**

Acknowledgements ..... **iii**

1. Introduction **1**

1.1 The Human Immunodeficiency Virus 1

1.1.1 The diversity of HIV 1

1.1.2 Virus structure and lifecycle 2

1.1.3 Transmission and pathogenesis 6

1.1.4 The immune response 8

1.2 The envelope spike 12

1.2.1 Biosynthesis 12

1.2.2 Structure and function 15

1.3 Fighting the HIV/AIDS pandemic 18

1.3.1 Current therapies 18

1.3.2 Cure research 19

1.3.3 Towards a vaccine 21

1.4 Challenges facing HIV-1 vaccine design 22

1.4.1 The antibody response 23

1.4.2 The glycan shield 24

1.4.3 Instability of the envelope spike 26

1.5 Protein mimicry in vaccine design 27

1.5.1 Native-like trimers display bnAb epitopes 28

1.6 Broadly neutralising antibodies 30

1.6.1 Structure and development of broadly neutralising antibodies 30

1.6.2 Understanding the nature of glycan-dependent bnAb epitopes 32

1.7 Glycan mimicry in vaccine design 34

1.7.1 Understanding Env glycosylation processing 35

1.7.2 Immunogen mimicry of viral glycosylation 38

1.7.3 Glycan clustering and the intrinsic mannose patch 39

1.7.4 Protein structure and the trimer associated mannose patch 43

1.7.5 Cell-specific glycosylation 44

1.7.6 Membrane tethering 45

1.7.7 Signal peptide 46

1.7.8 Occupancy 47

1.7.9 O-linked glycosylation 48

1.7.10 Considerations for immunogen manufacture 49
1.8 Beyond antigenic mimicry 49
1.8.1 Mimicry of glycan epitopes: chemical approaches 50
1.8.2 Mimicry of glycan epitopes: biological approaches 51
1.8.3 Mimicking bnAb development in natural infection 53
1.9 The impact of sequence diversity of the HIV-1 glycan shield 54
2. Materials and methods 56
2.1 DNA plasmids and mutagenesis 56
2.1.1 Plasmid constructs 56
2.1.2 Site-directed mutagenesis 56
2.2 Recombinant protein expression and purification 57
2.2.1 Recombinant protein expression in HEK cells 58
2.2.2 Ni^{2+} affinity purification 58
2.2.3 Protein A affinity purification 59
2.2.4 Broadly neutralising antibody affinity purification 59
2.2.5 Co-purification of antibody-Env complexes 60
2.3 Analysis of overall glycan profiles and compositions 60
2.3.1 In-gel release of N-linked glycans 60
2.3.2 Fluorescent labelling of N-linked glycans 60
2.3.3 HILIC-UPLC analysis 61
2.3.4 Glycosidase digestion of released glycans 61
2.3.5 Ion mobility mass spectrometry analysis of released N-linked glycans 61
2.4 Site-specific glycan analysis 62
2.4.1 Reduction, alkylation and proteolytic digestion 62
2.4.2 Liquid chromatography mass spectrometry analysis of glycopeptides 62
2.4.3 Glycan 'classification' analysis 64
2.5 Immunoassays 64
2.5.1 ELISA 64
2.5.2 Western blot 65
2.5.3 Dot blot 65
2.6 Model building 65
2.7 Yeast cultures and protein expression 66
2.7.1 Plasmid constructs and cloning 66
2.7.2 *Pichia pastoris* culture conditions and preparation of cells for electroporation 67
2.7.3 Generation of *Pichia pastoris* stable cell lines 67
2.7.4 Protein expression in *Pichia pastoris* 67
2.7.5 Protein expression in *Kluyveromyces lactis* 68

3. Method validation and the influences of trimer format and antibody affinity purification 69

3.1 Site-specific glycan quantification by LC-ESI MS 73
3.2 Method reproducibility based on three biological replicates 74
3.3 The influence of antibody affinity purification on Env glycosylation 77
3.3.1 Resilience of BG505 SOSIP.664 to glycan biasing 78
3.3.2 Influence of bnAb affinity purification on CZA97 SOSIP.664 78
3.4 The influence of stabilising mutations on Env glycosylation 84
3.5 Discussion 86

4. Conserved features of the HIV-1 glycan shield 90

4.1 Glycan analysis of a cross-clade panel of Envs 93
4.2 Conservation of oligomannose-type glycans 93
4.3 Integrity of the mannose patches 95
4.4 Features of complex-type glycosylation processing 98
4.5 Discussion 102

5. Networks of glycans maintain antibody epitopes in the face of glycan additions and deletions 104

5.1 The glycan network surrounding the 2G12 epitope 107
5.2 Overall resilience of the mannose patches to glycan addition and deletion 109
5.3 Glycan deletion increases mannose trimming throughout the trimer 112
5.4 Differential effects of glycan additions on processing 115
5.5 A network of glycans preserves the PG9 and PG16 epitopes 118
5.6 Discussion 119

6. Influence on the site-specific processing of the glycan shield by bnAb co-expression 124

6.1 Co-expression of BG505 SOSIP.664 trimers with bnAbs 127
6.2 Glycosylation is influenced by antibody co-expression 128
6.2.1 Outer domain-targeting bnAbs 129
6.2.2 The gp120-gp41 interface 131
6.2.3 CD4bs-targeting bnAbs 131
6.2.4 Apex-targeting bnAbs. 132
6.2.5 Non-Env targeting antibody 134
6.3 Discussion 135

7. Immunogen mimicry of the viral glycan shield 137

7.1 Production of virion-derived Env 138
7.2 Relative abundance of oligomannose- and complex-type glycans 139
7.3 Divergence and conservation of fine processing at various gp120 glycan sites 141

7.4 Discussion 145
8. Future directions: Enhancing the immunogenicity of the glycan shield using yeast 149
8.1 Yeast N-glycosylation 151
8.2 *In vivo* modification of Env glycans 154
8.3 *In vitro* modification of Env glycans 158
8.4 Discussion 161
9. Concluding remarks 163
List of abbreviations 165
Appendix A. Cloning and mutagenesis primers 168
Appendix B. Glycan and glycopeptide fragmentation 169
Appendix C. Glycan nomenclature 172
Appendix D. Glycan libraries 173
Appendix E. Env sequences 196
Appendix F. Glycan labelling reaction 202
Appendix G. List of publications and conferences 203
References 205

Abstract

Vaccine strategies for the human immunodeficiency virus type 1 (HIV-1) focus on the mimicry of its envelope spike (Env), due to its exposed location on the viral membrane and role in mediating infection. However, the virus has evolved to minimise the immunogenicity of Env, notably through extremely high sequence variability and the shielding of conserved protein epitopes by an extensive array of host-derived, immunologically 'self' N-linked glycans. Despite this, a subset of infected individuals eventually develop broadly neutralising antibodies that can recognise these glycans. Defining the composition of glycan epitopes on both viral Env and candidate immunogens is therefore a major research focus for HIV-1 vaccine design.

Using recombinant, soluble mimics of Env, this thesis sought to determine the impact of protein sequence on the processing of the glycan shield. This was addressed on three levels: the impact of virally-encoded sequence diversity, the impact of engineered stabilising mutations, and the impact of mutations affecting glycosylation sites. Despite the extensive diversity displayed by the Env sequence, many features of the glycan shield were remarkably conserved. Mutations affecting glycosylation sites, however, did impact the fine processing of nearby sites. Additionally, the glycan composition of Env mimics could be subtly altered using glycan-dependent broadly neutralising antibodies, either during the expression or purification stages. The latter effect appeared to be sequence-specific.

Although recombinant, soluble Env proteins are good structural and antigenic mimics of native, viral Env, the extent to which their glycan shields resemble that of infectious viruses is unknown. To that end, the glycosylation of a candidate immunogen was compared to that of sequence-matched Env derived from infectious virions. The glycan processing of key broadly neutralising antibody epitopes was largely conserved between virus and immunogen. However, there were subtle differences at some sites, with the viral material

typically displaying elevated glycan processing. Given that recombinant Env constructs are good mimics of the viral glycan shield, they are likely also inherently immunoquiescent. Accordingly, this thesis also explored a strategy to boost the immunogenicity of the glycan shield, using yeast to decorate a soluble Env immunogen with 'non-self' glycan structures.

Overall, features of the Env glycan shield are largely conserved between virus and immunogens, and across multiple genotypes, supporting it as a suitable target for HIV-1 vaccine design.

Acknowledgements

I am incredibly grateful to so many people for their help and support over the last four years. Firstly, thank you to Prof. Max Crispin for putting his faith in a non-Oxford educated student and inviting me back to his lab to complete a DPhil. His advice and guidance over the last few years has been invaluable. I have thoroughly enjoyed being a part of the lab and will miss it very much.

I must thank the European Aids Vaccine Initiative (EAVI2020) for funding this DPhil project and for providing me with a friendly, supportive and knowledgeable network of researchers. Many of the proteins analysed throughout this thesis were provided by our numerous collaborators, to whom I am also very grateful.

Thank you to Prof. Raymond Dwek for always encouraging me and for making the Oxford Glycobiology Institute such a welcoming place to work. To Dr Mark Wormald for co-supervising me, and to Prof. David Harvey for his endless patience and knowledge of ion-mobility mass spectrometry. I must also thank Dr Paul Skipp and Dr Mark Dixon for all their efforts in helping us to establish the lab in Southampton. They made the move from Oxford very easy and we could not have done it without their support.

I would never have completed (or started) this DPhil without the support of Snezana. She taught me everything I know in the lab, and most of what I know outside of it.

A huge thank you to all the members of the Crispin group, past and present: Lan, Martin, Weston, Kirstin, John, Charlie, Hannah, Lydia, and Al, who made every day in the lab so much fun. A particular mention must go to Laura Pritchard and Anna Behrens for their help and patience in teaching me glycan analysis, and to Joel Allen, who put up with me talking to myself for three and a half years and who is always so willing to help out. Yasa, you are alright, I suppose.

Lastly, thank you to everyone outside of the lab who made my DPhil experience so enjoyable. A special thank you to my girls, Emily, Laura, Lucy and Rose, for the twenty

million WhatsApp messages received over the last four years, and for their endless support and encouragement in matters big and small. Lastly, to my parents and brother, who still don't know what I do, but who are enthusiastic regardless. Thank you for not once doubting my abilities.

1. Introduction

1.1 The Human Immunodeficiency Virus

The human immunodeficiency virus (HIV) remains a major public health issue worldwide. The virus is the causative agent of acquired immunodeficiency syndrome (AIDS), in which progressive failure of the immune system leaves the patient susceptible to life-threatening opportunistic infections. Thus far, the HIV/AIDS pandemic has claimed over 35 million lives, with a further 37 million people currently living with the disease [1]. While HIV can be managed with antiretroviral drugs, there is no cure. A protective vaccine remains a primary goal in the efforts to control the pandemic.

1.1.1 The diversity of HIV

There are two distinct types of the virus, HIV-1 and HIV-2, though the latter is endemic only to West Africa and is therefore not a main focus of HIV vaccine research. Within HIV-1, there are four groups, M, N, O and P. Each group originates from a separate zoonotic transmission of simian immunodeficiency virus, derived from either chimpanzees or gorillas (SIVcpz or SIVgor) [2]. The only group to achieve pandemic spread is the 'main' or 'major' group, M, which is further categorised into nine genetically distinct clades (or subtypes; A, B, C, D, F, G, H, J and K). Protein sequence diversity within clades has been estimated at 8 – 17%, with between clade sequence diversity reaching 17 – 35% [3].

The clades differ in their geographical prevalence; clade A viruses are predominant in Eastern and central Africa and Eastern Europe, clade B is the main form found in Western

and central Europe, the USA, Australia, Northern Africa and the Middle East [4]. Clade C, which accounts for approximately half of all infections, is dominant in Southern Africa and India [4]. Additionally, there are several inter-subtype recombinant lineages, known as circulating recombinant forms (CRF), which are understood to have originated from individuals infected with viruses from two or more subtypes. The CRF, AE, is responsible for approximately 95% of infections in regions of Southeast Asia [5].

The vast majority of infections are established by only one virus [6-9]. However, over the course of infection, selective pressures imposed by therapy and/or the host immune response cause the virus to rapidly mutate [10, 11]. The viral envelope spike (Env) is particularly susceptible to mutation, with changes in its gene sequence often exceeding 10% [3]. The resulting set of viruses found within an infected individual, termed quasi-species, exhibit genetic diversity comparable to the global variation observed in the influenza virus in a given year (Figure 1.1) [3]. Thus, the breadth of protection that a successful HIV-1 vaccine must convey far exceeds that of an influenza vaccine. Given that influenza vaccines require updating annually, the exceptional genetic and antigenic diversity possessed by HIV-1 remains a significant obstacle in the development of a successful vaccine.

1.1.2 Virus structure and lifecycle

HIV-1 is a retrovirus, belonging to the genus *Lentivirus*. It is a single-stranded, positive-sense, enveloped RNA virus. Virions are roughly spherical, approximately 100 – 120 nm in diameter [12]. The virus contains two copies of its genome, enclosed with various viral proteins in a conical capsid. A matrix surrounds the capsid, ensuring the integrity of the virus, which is in turn surrounded by the envelope. The viral envelope is composed of a lipid bilayer obtained while budding from the infected host cell (Figure 1.2).

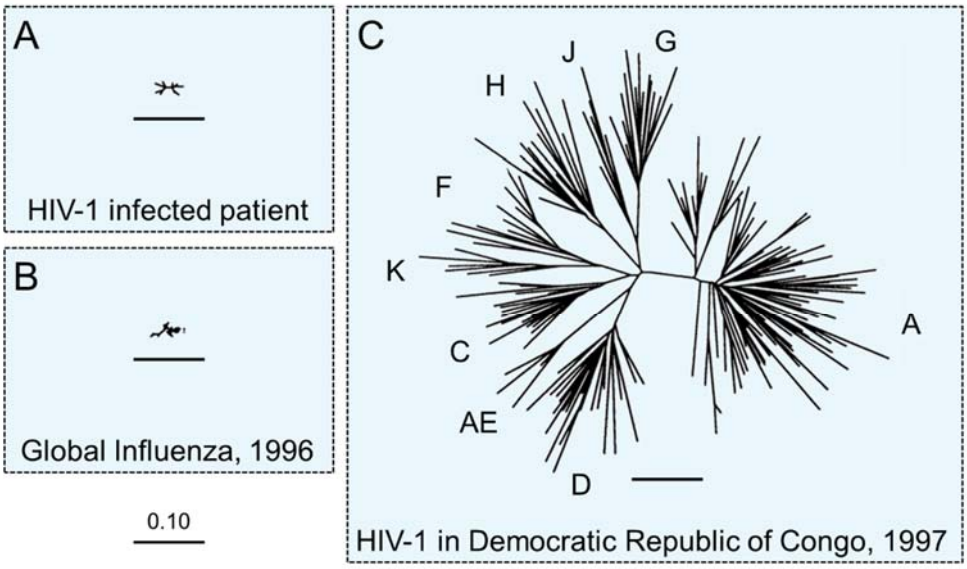


Figure 1.1. The antigenic diversity of HIV-1. Adapted from Korber *et al.*, 2001 [3]. A comparison of evolutionary distances of HIV-1 Env and influenza hemagglutinin sequences through phylogenetic analysis. A) Tree based on nine Env sequences from an individual six years after infection. B) Tree based on all hemagglutinin sequences (n = 96) in the Influenza Sequence Database, Los Alamos National Laboratory with an isolation year of 1996. C) Tree based on HIV-1 Env sequences sampled from 193 individuals in the Democratic Republic of the Congo in 1997. Subtypes are annotated. The scale bar is the same in all panels for comparison.

Embedded within the envelope are approximately 14 Env glycoproteins (Figure 1.2A) [12]. Env is a trimer of gp120-gp41 heterodimers, generated by cleavage of a gp160 pro-protein. It is extensively N-glycosylated, with approximately half the mass of gp120 comprising N-linked glycans (Section 1.2) [13]. The envelope spike defines viral tropism and mediates infection. During viral entry, gp120 attaches to the CD4 (cluster of differentiation 4) receptor, present on monocytes, macrophages, dendritic cells and T cells. Receptor binding prompts conformational changes within gp120, resulting in the exposure and formation of the CCR5 or CXCR4 co-receptor binding sites (C-C chemokine receptor type 5 or C-X-C chemokine receptor type 4). Most infections are established by CCR5-tropic viruses (R5 viruses) [9, 14, 15], thus individuals homozygous for a defect in their CCR5 gene are generally considered resistant to HIV-1 infection [16, 17]. Upon co-receptor binding,

conformational changes within gp41 result in fusion of the viral and host cell membranes, and the release of the capsid into the host cell cytoplasm (Figure 1.2B, stage 1 and 2).

Once inside the host cell, the viral genome is reverse transcribed into complementary DNA by the viral reverse transcriptase (RT) [18]. This process is extremely error-prone due to a lack of proof-reading mechanism, thus mutations are readily introduced [19]. Additionally, the viral genome is subject to mutation by host cytosine deaminases [11, 20]. A

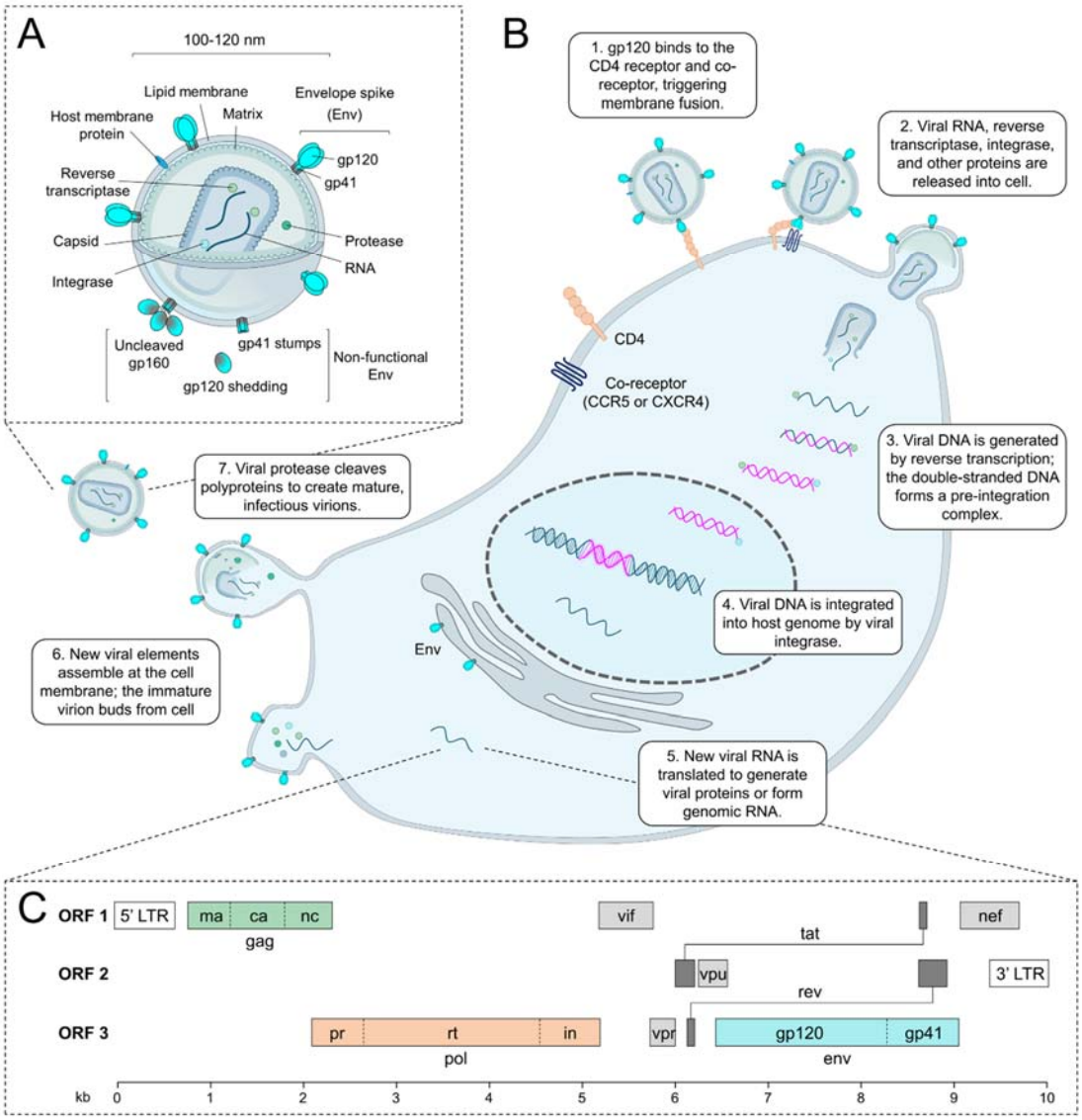


Figure 1.2. The HIV-1 structure and lifecycle. A) Structure of an HIV-1 virion. B) The viral lifecycle, see main text for details. C) The viral genome, approximately 9.8 kb, is transcribed over multiple open reading frames. MA, matrix; CA, capsid; NC, nucleocapsid; PR, protease; RT, reverse transcriptase; IN, integrase; Vif, viral infectivity factor; Nef, negative regulatory factor; Vpu, viral protein U; Vpr, viral protein R; Tat, trans-activator of transcription.

double-stranded RNA/DNA hybrid is produced initially, prior to degradation of the RNA strand by the ribonuclease H catalytic domain of RT [18]. A second DNA strand is then synthesised by RT. The double-stranded DNA remains associated with multiple viral proteins, including integrase (IN), matrix (MA), and the accessory viral protein R (Vpr), to form a 'pre-integration complex' which facilitates the transport of the viral DNA into the host cell nucleus [18]. Additionally, it is thought the capsid (CA; either in its conical form, or as individual capsid proteins) plays an integral role in nuclear import [21, 22] (Figure 1.2B, stage 3).

Upon entering the nucleus, IN integrates the viral DNA into gene-rich, transcriptionally active domains of the host's genome (Figure 1.2B, stage 4) [23, 24]. Transcription and translation of the viral genome is then carried out by the host cell (Figure 1.2B, stage 5). Initially, a small number of RNA transcripts are produced, allowing for the production of the two regulatory proteins, Tat (trans-activator of transcription) and Rev, which positively regulate transcription of the viral genome [25]. Three major genes span the majority of the viral genome: *gag*, *pol* and *env* (Figure 1.2C). All are translated as polyproteins which require cleavage by viral or host proteases in order to form functional gene products. Viral proteases cleave Gag to generate the structural proteins MA, CA and nucleocapsid (NC); cleavage of Pol generates the enzymes protease, RT and IN. The structure and synthesis of Env will be discussed in greater detail in Section 1.2.

The assembly of new viral particles occurs at the plasma membrane, most likely at lipid raft domains [26, 27]. Once the newly synthesised viral proteins and RNA genome have assembled, the virion buds off from the infected cell. The viral protein U (Vpu) aids this process by downregulating tetherin (CD317), which otherwise impedes virion release [28, 29]. Newly budded virions remain non-infectious until the viral protease cleaves Gag (Figure 1.2B, stage 6 and 7).

Alternatively, the integrated viral genome (provirus) can lie dormant within quiescent T cells for many years, replicating as part of the host cell genome. As there is little to no expression of viral proteins, latently infected cells escape immune recognition and clearance. These latent viral 'reservoirs' persist despite treatment with antiretroviral drugs.

1.1.3 Transmission and pathogenesis

The virus can be transmitted through a number of routes, most commonly across the mucosal surfaces during sexual intercourse. The genital and gastrointestinal mucosa therefore plays a crucial role in HIV-1 transmission and pathogenesis, as will be outlined in this section. Exposure to contaminated blood or blood products and mother-infant transmission (e.g. through pregnancy, delivery, or breastfeeding) are also common routes of infection. Viral load in the transmitting patient plays a major role in determining the risk of infection from one individual to another [30]. Infection is usually established by a single transmitted/founder (TF) virus, with R5 viruses preferentially transmitted across all routes [6-9, 14, 15].

During sexual transmission, the virus must penetrate the mucosal and epithelial barriers of the genital or gastrointestinal tracts in order to encounter its target cells, monocytes, macrophages, dendritic cells, and T cells. Virions may traverse the epithelium via several pathways, including endocytosis, transcytosis, or by passing through small gaps between the epithelial cells (Figure 1.3) [31, 32]. Moreover, it is possible that target cells may become directly exposed through micro-abrasions generated during sexual intercourse [31, 33]. The virus can be transmitted as both cell-free and cell-associated virus [34].

The virus first replicates locally at the infected mucosal site. Following the initial period of infection, which is often clinically silent, the virus disseminates to the lymphoid tissues, in particular the gut-associated lymphoid tissue (GALT). The classical model of virus transmission involves binding of cell-free virions to target cells. However, the virus can also spread from cell to cell. Productively infected cells, and non-infected cells harbouring intact

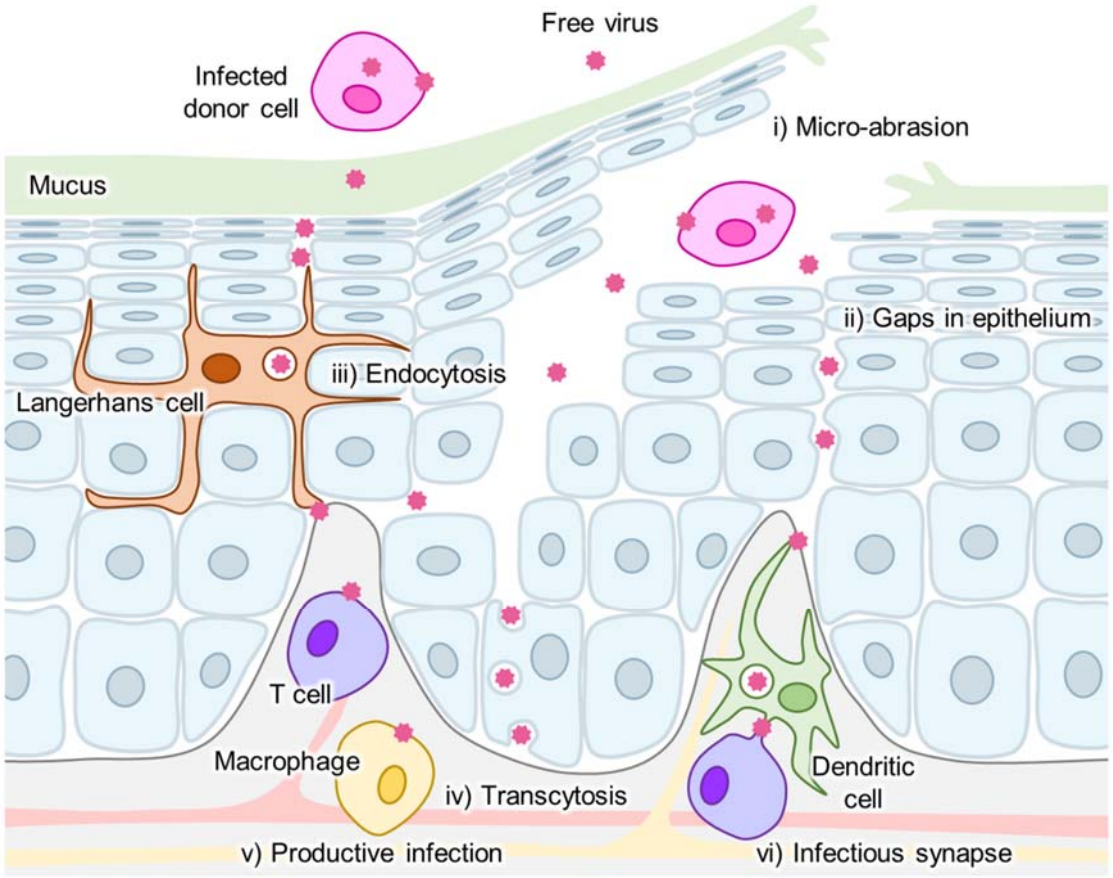


Figure 1.3. HIV-1 transmission. Virions can traverse the epithelium via several pathways. i) Shearing during sexual intercourse can generate micro-abrasions through which infected donor cells and/or free virions can migrate to contact target cells in the mucosal epithelium and stroma. ii) Virions may penetrate through small gaps between the epithelial cells. iii) Dendritic cells, for example, Langerhans cells, which extend processes towards the lumen can internalise virions via endocytosis. iv) Transcytosis of virions across the interior of epithelial cells. v) Receptor-mediated (productive) infection of CD4+ T cells and macrophages. vi) Passage of virus from stromal dendritic cells to CD4+ T cells across an infectious synapse. Figure based on Hladik and McElrath, 2008 [31].

virions, infect target cells via the formation of virological or infectious synapses, respectively (Figure 1.3). It is estimated that cell-mediated HIV-1 transmission is 100 to 1'000 times more efficient than cell-free virus spread [35]. The GALT harbours the majority of CD4+ CCR5+ T lymphocytes, thus infection at this site coincides with the loss of vast numbers of T cells [36, 37]. The depletion of CD4+ T cells due to HIV-1 infection can occur by several mechanisms, with both infected and uninfected cells susceptible [38, 39]. The vast majority of cell death is thought to occur by apoptosis. Numerous viral proteins, including Env, Tat and Nef, have been implicated in regulating cell death. For example, Env expressed on the

plasma membrane of infected cells can interact with receptors on uninfected cells to trigger membrane fusion and the formation of syncytia, triggering apoptosis [40]. Thus, viral-driven apoptosis could be considered a strategy evolved by the virus to escape immune effector cells [41].

After a few weeks of acute infection, an adaptive immune response is mounted (Section 1.1.4). Around this time the patient may experience influenza-like symptoms. Numbers of CD4+ T cells recover and viraemia decreases to reach the viral set point. However, the virus is able to repeatedly escape the host's antibody and cytotoxic T lymphocyte (CTL) responses, which are progressively weakened as CD4+ T cell numbers continue to decline. Thus, the patient becomes susceptible to opportunistic infections, marking a progression to AIDS. The rate of disease progression varies between patients, with those who maintain viraemia at < 5000 RNA copies/mL in the absence of treatment, and take more than ten years to progress to AIDS, termed long-term non-progressors. In approximately half of cases, viruses that use the CXCR4 co-receptor (X4 viruses), either exclusively or in addition to CCR5, evolve and become dominant throughout the latter stages of infection [42, 43]. The evolution of X4 variants is associated with a rapid drop in CD4+ T cell counts and accelerated disease progression (Figure 1.4) [44, 45].

1.1.4 The immune response

Both innate and adaptive immune responses are raised in response to HIV-1 infection, though they are either insufficient, or appear too late, to eliminate the virus. Furthermore, the virus has evolved to exploit the host immune response in order to enhance infection.

Upon infection, the virus must contend with the innate immune response. Dendritic cells are among the first immune cells encountered by the virus. In addition to CD4, CCR5 and CXCR4, dendritic cells express a number of pattern recognition receptors (PRRs) that are capable of interacting with the virus. Notably, C-type lectin receptors (CLRs) are able to bind to the glycans on Env. As will be discussed throughout this thesis, the Env glycan shield

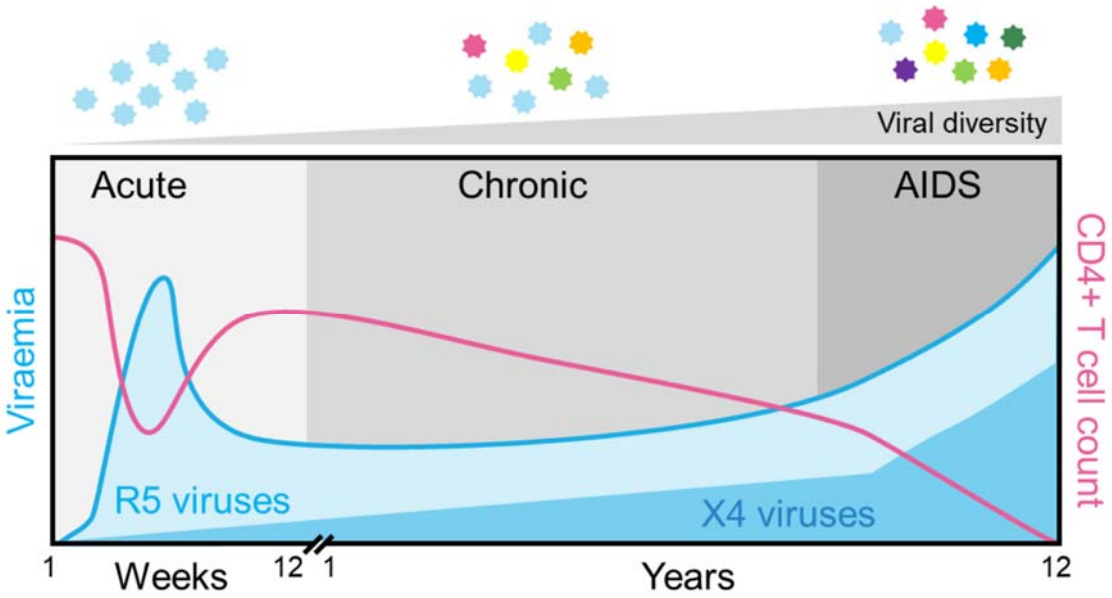


Figure 1.4. HIV-1 pathogenesis. The acute phase of infection is characterised by a dramatic increase in viraemia, corresponding with a decline in CD4+ T cell numbers. As the adaptive immune system responds, viraemia decreases and CD4+ T cell numbers recover. The virus continually evades host immune control. Later in the course of infection X4 viruses become dominant. This is associated with a rapid decrease in CD4+ T cell numbers and accelerated progression to AIDS. Figure adapted from Simon and Ho, 2003 [46].

contains a high proportion of oligomannose-type glycans, allowing for binding by mannose-specific CLRs such as Langerin and DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin). Additionally, sialic acid-binding Ig-like lectin 1 (Siglec-1) can recognise sialylated glycans on Env or on viral membrane glycolipids [47, 48]. Typically, CLRs mediate pathogen recognition and internalisation for antigen presentation [49]. However, the virus is able to subvert the cell-mediated immune response and instead actively use it to enhance infection. Binding to DC-SIGN, for example, can lead to the endocytosis and storage of intact virions, or to their retention on the cell surface [50]. The virus subsequently exploits the migratory capacity of dendritic cells to gain access to lymphoid tissues, where it can infect T cells via the formation of an infectious synapse (Figure 1.3) [51]. Siglec-1 binding has similarly been reported to enhance *trans*-infection [47, 48]. There are, however, conflicting reports on the role of Langerin in promoting viral transmission, as

Langerin-mediated capture of the virus by Langerhans cells can result in its internalisation into Birbeck granules, and subsequent degradation [52].

In contrast to the contradictory roles played by CLRs in both protecting and promoting viral transmission, soluble C-type lectins typically block infection by targeting virions for destruction by the immune system. For instance, the oligomannose-type glycans on Env are also able to interact with mannose binding lectin (MBL), which subsequently protects from infection by blocking interactions between Env and DC-SIGN, opsonising the virus, and activating the lectin pathway of the complement system [53].

The complement system, a key component of innate immunity, is a first-line of defence against HIV-1 infection. The virus can activate complement in a number of ways, including by the binding of antibodies, MBL or complement proteins to Env [54-56], resulting in complement-mediated lysis of the virus. While the complement system plays a critical role in clearing the infection, the virus is yet again able to exploit the host response to facilitate the spread of the virus within the infected patient. For instance, the deposition of complement activation products, such as C3b and C5b, on virions facilitates infection of cells expressing complement receptors CR3 or CR4, e.g. macrophages and dendritic cells [57, 58]. Furthermore, the viral envelope can contain complement regulator proteins, such as CD55 and CD59, which are derived from the host cell during budding and which act to limit complement-mediated lysis of HIV-1 [59].

An adaptive immune response is initiated in the days and weeks following infection. Mounting CD8+ T cell (CTL) responses initially follow the rise in viraemia [60]. As the CTL response peaks, the level of viraemia falls. Early CTLs are directed against Env, Gag and Nef [61], however, the virus is able to rapidly generate escape mutants. Indeed, escape from CTLs is a major force driving viral evolution, both within individuals [62] and at the population level [63, 64]. Further evidence for the role of CTLs in controlling viraemia comes from the observation that highly functional CD8+ T cell responses are correlated with slow disease

progression in non-progressors [65]. Moreover, a genome-wide association study found the most significant associations with control of viral replication were linked to human leukocyte antigen (HLA) genes [66]. Nevertheless, the virus is able to undermine total control by CTLs. In addition to escape mutations, the virus can evade CTLs via Nef-mediated downregulation of HLA class 1 expression [67, 68].

The first detectable B cell responses against HIV-1 arise a few weeks after infection [69]. Unfortunately, the initial antibody response is unable to neutralise the virus. This is likely explained by the presence of non-functional Envs embedded within the viral membrane (e.g. uncleaved gp160 or gp41 stumps; Figure 1.2). Such constructs display immunodominant epitopes that are either occluded or absent on the functional trimer, and which act to divert the antibody response away from neutralising epitopes [70-75]. That being said, antibodies are thought to play a role in reducing transmission rates between blood group discordant patients. The virus can incorporate ABO carbohydrates present on host cell glycolipids and/or membrane glycoproteins into its membrane during budding. Thus, virions derived from A- or B-positive donors may show sensitivity to anti-A- or anti-B antibodies in discordant recipients [76, 77].

Neutralising antibodies (nAbs) do not arise until several months after infection. These are typically directed against the highly variable protein surface of Env and are therefore capable of neutralising only the strain they were raised against (autologous nAbs) [78, 79]. The induction of nAbs drives the selection of viral escape mutants that, in turn, select for affinity-matured neutralising antibodies. The ensuing virus-antibody 'arms race' eventually drives the development of neutralisation breadth [80-83].

Broadly neutralising antibodies (bnAbs) develop in approximately a third of patients after several years of infection [84-86]. While bnAbs do little to alter the course of disease progression within the infected individual [87], they are able to protect non-human primates [88-99] and humanised mice [100-106] from viral challenge when passively administered.

Furthermore, they are protective at titres achievable by vaccination in other settings [89, 91]. Thus, it is widely accepted that a vaccine capable of eliciting bnAbs in the window before the establishment of latency will be able to protect from HIV-1 infection. The challenges associated with the development of a bnAb-targeting vaccine will be discussed throughout Sections 1.4 and 1.6.

1.2 The envelope spike

The envelope spike plays a crucial role in mediating infection through its interactions with the CD4 receptor and CCR5/CXCR4 co-receptors on target cells. As the sole target for bnAbs, Env is a key target for vaccine design.

1.2.1 Biosynthesis

The functional HIV-1 envelope spike comprises a trimer of non-covalently associated gp120-gp41 heterodimers. The gp120 and gp41 proteins are generated by protease cleavage of the precursor polypeptide, gp160, encoded by the *env* gene. Translation of gp160 is directed to the endoplasmic reticulum (ER) by a signal peptide, though complete release into the lumen of the ER is prevented by a stop-transfer sequence within the gp41 portion. As Env is translated, it is extensively N-glycosylated. On average, gp160 contains 30 potential N-glycosylation sites (PNGS): asparagine residues within the sequon Asn-X-Ser/Thr, where X is any amino acid except proline[*]. An overview of the mammalian N-linked glycosylation pathway is shown in Figure 1.5 and is summarised below.

The N-linked glycan precursor ($Glc_3Man_9GlcNAc_2$) is transferred from dolichol phosphate to the amide nitrogen of the PNGS asparagine by oligosaccharyltransferase (OST). The two terminal glucose residues are removed by ER glucosidases I and II, generating monoglucosylated gp160 ($Glc_1Man_9GlcNAc_2$), which forms the substrate for the folding chaperones, calnexin and calreticulin (CNX/CLN), and by extension, protein

[*]A proline in the fourth position also reduces the likelihood of glycosylation to < 1% [107]. In rare instances, asparagine residues within the Asn-X-Cys sequon can be glycosylated [108].

disulphide isomerase (PDI). Hydrolysis of the final glucose residue by glucosidase II frees gp160 from CNX/CLN, where it faces one of two fates. If correctly folded, gp160 will remain deglucosylated and continue its transit through the ER and Golgi apparatus. Alternatively, misfolded glycoproteins are reglucosylated by UDP-glucose glucosyltransferase and are able to bind to CNX/CLN for another folding cycle. Continually misfolded proteins are eventually degraded by the ER-associated protein degradation pathway. It has been reported that Env spends a considerable amount of time within the ER; slow folding events allow time for the extensive network of disulphide bonds to form and be 'shuffled' by PDI [109, 110]. The correct formation of disulphide bonds, which typically number around ten per protomer, is essential for the formation of infectious virions [111].

Typically, transit of a glycoprotein through the ER to the Golgi apparatus is accompanied by further modifications to the $Man_9GlcNAc_2$ glycan. Endoplasmic reticulum- and Golgi-resident α-mannosidases trim up to four mannose residues to generate a $Man_5GlcNAc_2$ structure. The addition of an *N*-Acetylglucosamine residue by *N*-Acetylglucosaminyltransferase I (GnT I) initiates processing to hybrid-type glycans. Alternatively, further mannose trimming by Golgi α-mannosidase II (Golgi α-Man II), followed by cell-specific processing by various glycosyltransferases, leads to the generation of a variety of complex-type glycans. The repertoire of glycan structures produced by a given cell reflects the combination of glycosyltransferase and glycosidase enzymes it expresses. Given that the human genome contains over 300 genes associated with glycan production and processing [112], glycoproteins can be exceptionally heterogeneous. The processing of Env, however, diverges from that typically observed on mammalian glycoproteins. The very high number of PNGS on gp160 result in a glycan density that is largely impenetrable to many of the host's processing enzymes. Furthermore, additional steric constraints are imposed by the correct folding of Env. Thus, many of the glycans are not trimmed much

beyond the $Man_9GlcNAc_2$ structure. The glycosylation of Env will be discussed more thoroughly in Sections 1.7 and throughout this thesis.

Trimerisation of gp160 occurs in the ER, prior to cleavage by furin protease at the C-terminus of the consensus sequence Arg-X-Lys/Arg-Arg [113]. Cleavage is essential for the structure and functionality of Env, and thus for viral infectivity [114, 115]. Furthermore, it liberates the N-terminus of gp41, which comprises the fusion peptide (Section 1.2.2). It is generally assumed that cleavage occurs in the *trans*-Golgi network, corresponding with the

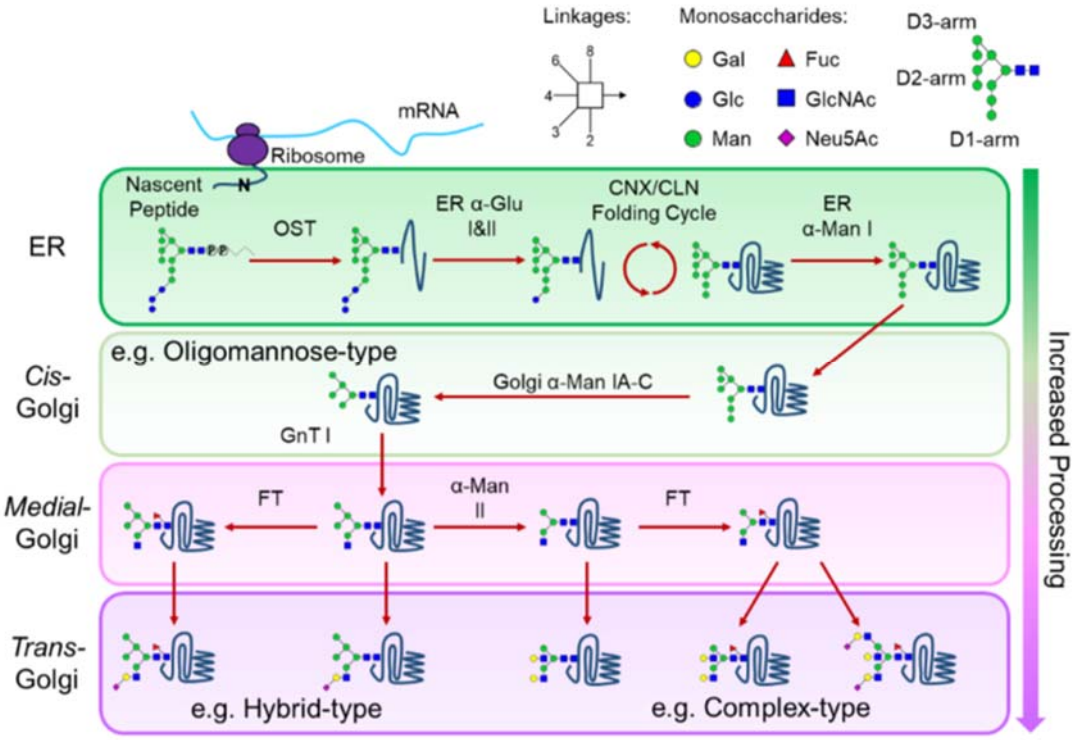


Figure 1.5. The N-glycosylation pathway. The N-linked glycan precursor, $Glc_3Man_9GlcNAc_2$, is transferred from a dolichol membrane anchor and linked to the asparagine by oligosaccharyltransferase (OST). The terminal glucose residues are removed by ER α-glucosidases I and II. Upon correct protein folding in the calnexin and calreticulin cycle, ER α-mannosidase I (ER α-Man I) removes the terminal mannose from the D2 arm of $Man_9GlcNAc_2$. Golgi α-mannosidases IA-C (Golgi α-Man IA-C) are responsible for further mannose trimming, leading to $Man_5GlcNAc_2$. Transfer of an *N*-Acetylglucosamine residue to the D3 arm, mediated by *N*-Acetylglucosaminyltransferase I (GnT I), initiates processing to hybrid-type glycans. Alternatively, further mannose trimming by Golgi α-mannosidase II (Golgi α-Man II) leads to a $GlcNAcMan_3GlcNAc_2$ structure. Modification by cell-specific glycosyltransferases, sialyltransferases and fucosyltransferases (FT) generates a variety of complex-type glycans. Glycan representation is shown according to the combined nomenclature of Harvey *et al.* and the Consortium for Functional Glycomics [116, 117]. Numbering of the D1-, D2- and D3-arms is adopted from the system of Vliegenhart *et al.* 1983 [118]. See Appendix C. Figure adapted from Watanabe *et al.*, 2019 [119].

localisation of furin [120]. However, there is evidence that furin is also able to act in the early secretory pathway [121], which is easier to reconcile with the observation that cleaved Env proteins display elevated protection from early glycan processing [122, 123].

Following cleavage, gp120 and gp41 remain associated through non-covalent interactions. The trimers then assemble at the plasma membrane, where they are incorporated into new virions. There are relatively few Envs per virion (estimated at approximately 14 for one isolate [12]), which may be partially attributed to the internalisation of Env by clathrin-mediated endocytosis, and/or weak gp41-gp120 interactions resulting in gp120 shedding [124].

1.2.2 Structure and function

The gp120 and gp41 subunits each have distinct functions. The gp120 subunit contains the CD4 receptor and co-receptor binding sites and is therefore responsible for binding to the target cell. The gp41 subunit mediates fusion of the viral and host cell membranes. The primary structure of Env is depicted in Figure 1.6A. The gp120 subunit is comprised of five constant regions (C1-5), which are reasonably conserved across different clades and isolates, interspersed with five variable regions (V1-5), which exhibit much greater sequence variation. The gp41 subunit consists of an ectodomain, transmembrane domain (TM), and a cytoplasmic tail (CT).

Efforts to solve the quaternary structure of Env were frustrated for many years due the difficulties associated with isolating Env, or designing an Env protein that was stable enough to be produced in sufficient quantities (Section 1.4.3) [125]. Efforts were further thwarted by the many variable regions and extensive glycosylation. In the interim, the majority of structural data were obtained using variations of recombinant gp120, e.g. deglycosylated truncated cores [126-129]. The development of the 'SOSIP.664' platform offered a promising solution. The SOSIP.664 constructs use an engineered disulphide bond ('SOS') to covalently link the gp120 and gp41 subunits [130], an isoleucine to proline point

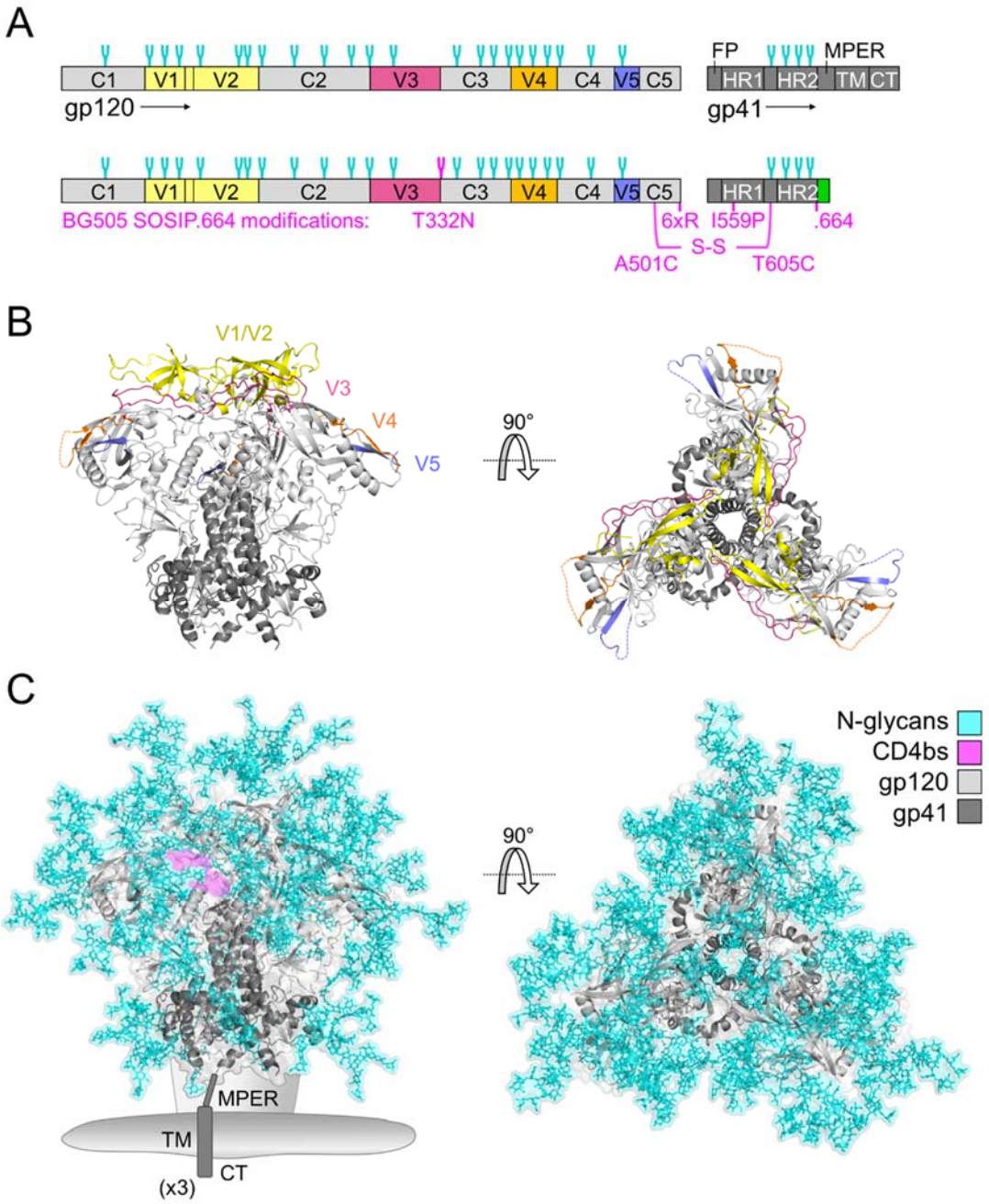


Figure 1.6. Structure of the envelope spike. A) Schematic representation of the primary structure of Env (top) and the soluble trimer mimic, BG505 SOSIP.664 (bottom). Variable regions (V1-5) are shown in colour, constant regions (C1-5) are shown in light grey, and gp41 is shown in dark grey. The envelope spike has approximately 30 potential N-glycosylation sites (cyan forks). Stabilising SOSIP.664 modifications are annotated in magenta, with optional purification tag coloured green. B) Cryo-electron microscopy structure of BG505 SOSIP.664 (based on PDB 5ACO [131]), with variable loops coloured accordingly. C) Model of a fully glycosylated envelope spike (glycans in cyan sticks). Glycans were added according to Behrens *et al.*, 2016 [132]. The membrane proximal external region (MPER), transmembrane domain (TM) and cytoplasmic tail (CT) are not present in the structure and are shown in cartoon for one protomer, along with the lipid membrane for orientation. FP, fusion peptide; HR1/2, heptad repeat 1 and 2; CD4bs, CD4 binding site.

mutation ('IP') to strengthen interactions between the three protomers, and truncation at residue 664 to ensure solubility (Figure 1.6A; Section 1.5) [133, 134]. This format was first successfully applied to a clade A strain, BG505, with the resulting 'BG505 SOSIP.664' trimers exhibiting native-like structure and antigenicity [74, 125]. The first atomic-level views of a soluble Env trimer were solved by cryo-electron microscopy (EM) and X-ray crystallography, to 5.8 Å and 4.7 Å resolution, respectively [135, 136]. Over recent years, the structure of many native-like Env trimers have been solved, including an almost full-length, membrane-bound Env [137].

The gp120 subunit is formed by two major domains (inner and outer) connected by a four-stranded, antiparallel β- 'bridging' sheet. The CD4 receptor binds to a discontinuous epitope located in a depression formed at the interface of these domains [126]. In the context of the trimer, the CD4 binding site (CD4bs) is buried in the protomer interface surrounded by a 'fence' of glycans, and is accessible to little other than CD4 [138]. Binding of CD4 induces conformational changes in Env that result in the formation of the co-receptor binding site, which includes elements of the V3 loop and bridging sheet [139]. In this context, the V3 loop has been identified as a key determinant of CCR5 or CXCR4 co-receptor tropism [140].

Co-receptor binding triggers the N-terminal fusion peptide of gp41 to insert into the host cell plasma membrane (Figure 1.7, stage 3). At this stage gp120 can be shed, in monomeric form, into the extracellular space. Subsequent rearrangement of the two heptad repeat α-helices, HR1 and HR2, to form a stable six-helix bundle, brings the virus into close proximity to the host cell and promotes membrane fusion (Figure 1.7, stage 4) [141].

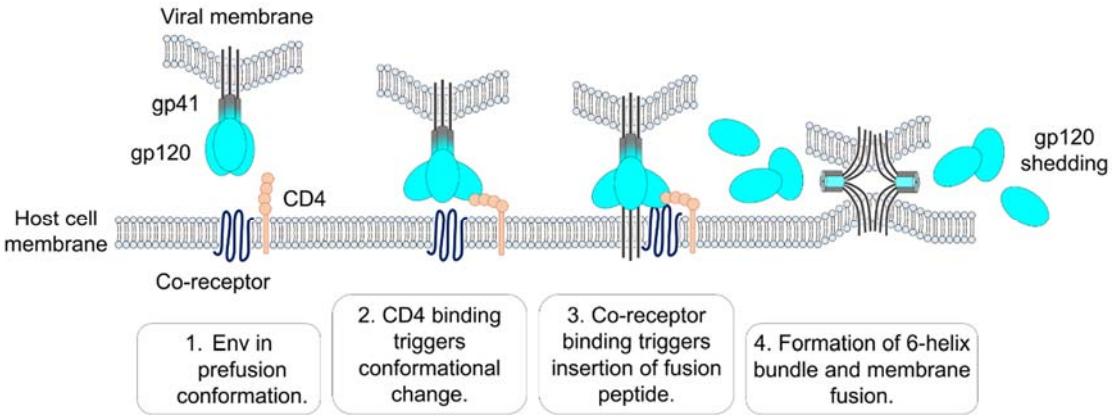


Figure 1.7. Overview of Env-mediated HIV-1 cell entry. The gp120 subunit binds the CD4 receptor. Subsequent rearrangements result in the formation of the co-receptor (CCR5 or CXCR4) binding site. Binding of the co-receptor triggers a further conformational change and the insertion of the gp41 fusion peptide into the host cell membrane. The formation of a six-helix bundle in gp41 brings the virus into close proximity to the host cell and promotes membrane fusion.

1.3 Fighting the HIV/AIDS pandemic

1.3.1 Current therapies

Infection with HIV-1 can be managed with anti-retroviral therapy (ART). There are approximately 40 US Food and Drug Administration (FDA) approved drugs available [142], categorised into six classes depending on their mechanism of action: nucleoside-analogue RT inhibitors, non-nucleoside RT inhibitors, IN inhibitors, PR inhibitors, fusion inhibitors and entry inhibitors. The large number of drugs on the market is rationalised by the speed at which HIV-1 can evolve resistance to individual drugs, again attributed to the error-prone RT and high viral replication rates [10, 11]. Indeed, drug-resistance has been documented for every class of antiretroviral currently in use [143]. In order to counter the rapid development of drug-resistance, monotherapy was replaced with highly active antiretroviral therapy (HAART; or combined antiretroviral therapy, cART), which uses a combination of at least three different anti-retrovirals, targeting two or more distinct molecular targets [144].

In some parts the world, HAART has increased life expectancies to near normal levels [145, 146]. Antiretroviral drugs can also be used as a pre-exposure prophylaxis (PrEP), to protect high risk groups from infection [147]. Despite the successes of HAART in controlling the infection, the establishment of a latent viral reservoir prevents complete eradication of the virus. Furthermore, adverse side-effects often lead to low levels of adherence, with cessation of therapy resulting in rapid viral rebound [148].

Antibody therapy presents an attractive alternative to HAART. While it is well established that the passive transfer of bnAbs is protective in animal challenge models [88-106], the use of antibodies for the treatment of an established infection has only recently become feasible with the isolation of exceptionally broad and potent antibodies. These 'next-generation' bnAbs are capable of significantly reducing the viral load in infected humanised mice and macaques [149-152], and have shown promising results in recent clinical studies [153-157]. The comparatively long half-lives (2 – 3 weeks [158]) and reduced side-effects of antibodies may alleviate some aspects of the burden of HAART treatment adherence. Furthermore, antibodies have the potential to engage the host immune system and guide host effector cells to kill HIV-1 infected cells [159, 160]. Despite the many benefits, passive transfer of bnAbs is insufficient to clear the latent viral reservoir. The development of a cure or prophylactic vaccine remains a major goal in the efforts to control the HIV-1 pandemic.

1.3.2 Cure research

There are two types of HIV-1 cure, and many conceivable strategies through which to achieve these goals. A *sterilising* cure refers to the complete elimination of replication-competent proviruses, while a *functional* cure refers to the long-term control of viral replication below the limit of detection. This allows for a functionally normal immune system and CD4+ T cell count, despite cells still harbouring proviral DNA.

'Shock and kill' strategies, for instance, aim to achieve a sterilising cure by inducing viral transcription in latently infected cells with latency-reversing agents (LRAs). The cells

would then be killed either by the virus itself, immune-mediated clearance, or another anti-HIV therapy [161-163]. Simultaneous treatment with anti-retroviral drugs and/or passively administered bnAbs would prevent the virus infecting naïve cells. Over 160 potential LRAs have been identified to date [164], though none have yet lead to a cure.

There are thought to be only two patients successfully cured of HIV-1 [165, 166]. These individuals, termed the Berlin and London patients, were both HIV-1 positive males later diagnosed with acute myeloid leukaemia or Hodgkin's lymphoma, respectively. In order to treat their life-threatening cancers, the patients received transplants of haematopoietic stem cells. In both instances, the stem cell donors were homozygous for a 32 base pair deletion mutation in the CCR5 co-receptor (CCR5Δ32/Δ32), which prevents its expression and renders the cells resistant to infection[*]. While the procedure is currently too risky and expensive to perform routinely, the manipulation of CCR5 expression on target cells by gene-editing is being investigated as a possible functional cure. Such therapies would employ technologies such as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9 to introduce targeted disruptions to the CCR5 co-receptor gene. This strategy has shown promise in pre-clinical studies [169-171].

Gene-editing technologies could also be used to incapacitate the provirus in latently infected cells. In a proof-of-concept study using humanised mice engrafted with infected human cells, CRISPR/Cas9 was used to successfully excise regions of the viral genome and decrease the level of replication-competent virus [172]. Alternatively, the transfer of genes encoding bnAbs would circumvent the problems associated with antibody half-lives by enabling long-term expression of the bnAbs [173]. For example, the delivery of bnAb genes by adeno-associated viral vectors is sufficient to both protect macaques from viral challenge [174], and to produce durable control of viremia after termination of ART in a mouse model

[*]The only other case of an HIV-1 infected patient transplanted with CCR5Δ32/Δ32 cells resulted in rebound of a population of X4 viruses [167]. No such viruses were observed in the Berlin or London patients [166, 168].

[150]. While there are still many technical barriers to overcome, including safety concerns over off-target mutagenic effects and pre-existing immunity to gene-editing components, these technologies offer significant hope for the development of an effective HIV-1 cure [175].

1.3.3 Towards a vaccine

Vaccination can be used to elicit both B and T cell responses. An advantage of T cell vaccines is that the immune response can be directed against many different viral epitopes. In contrast, B cell vaccines are only protective if the antibody response is directed to functional envelope spikes. Accordingly, T cell vaccines can be targeted to the more conserved portions of the virus, as opposed to highly variable Env. Further motivation for the development of a T cell vaccine comes from the growing body of evidence that CTL responses play a key role in controlling HIV infection (Section 1.1.4). However, a downside of T cell vaccination is that the initial wave of cell-free viruses cannot be targeted, as only processed viral antigens can be presented to T cell receptors. Thus, T cell vaccines are unlikely to be able to provide *sterilising* immunity. In contrast, a successful B cell vaccine will likely elicit antibodies capable of neutralising the virus in the window before it establishes latency. Some of the challenges facing the development of a B cell vaccine will be discussed throughout Section 1.4.

To date, six vaccine candidates have completed large scale, phase IIb or III clinical trials [176]. Of these, two consisted of recombinant gp120 proteins, aimed at generating antibody responses. While antibodies were elicited, they were unable to neutralise virus and no protection was observed [177, 178].

The Step and Phambili studies were designed to solely induce T cell responses. The vaccines, consisting of adenovirus vectors expressing Gag, Pol, and Nef proteins, were able to stimulate CD8+ T cell responses, however no protection was observed [179, 180]. Furthermore, the trials were stopped prematurely due to meeting futility boundaries.

The RV144 trial was targeted at both T and B cell responses. The RV144 vaccine consisted of a canarypox viral vector (ALVAC) expressing gp120, Gag, and Pol, followed by two booster injections of a recombinant gp120 protein. To date the RV144 trial remains the only vaccine to offer any form of protection, albeit at fairly modest levels (31.2%) [181]. Initial analyses identified IgG antibodies against the V1/V2 region of gp120 as potential correlates of protection [182], possibly conferring protection via effector functions such as antibody-dependent cellular cytotoxicity (ADCC).

Lastly, the HVTN 505 trial investigated the efficacy of a vaccine comprising a DNA prime, with plasmids encoding Env, Gag, Pol and Nef, followed by an adenovirus type 5 vector boost, expressing a Gag-Pol fusion protein and Env from various strains [183]. Unfortunately, this trial also offered no protection and was terminated early due to lack of efficacy.

1.4 Challenges facing HIV-1 vaccine design

As discussed above, the virus has evolved many features that undermine immunological control and eradication of infection, most notably, very high antigenic diversity and the establishment of a latent viral reservoir. While treatment with antiretroviral drugs can extend the life expectancy of infected individuals to near-normal [145, 146], drug-resistance has been documented for every class of antiretroviral currently in use [143], and treatment regimens are often accompanied by adverse side-effects and low levels of adherence. Furthermore, cessation of therapy results in rapid viral rebound [148]. While HIV-1 cure strategies are an important and viable field of research [184], the development of an effective prophylactic vaccine remains a primary goal in the effort to control the HIV-1 pandemic. Despite significant efforts, a protective HIV-1 vaccine has remained elusive [181]. The next section will outline some of the challenges associated with the elicitation of bnAbs in both a vaccine and infection setting.

1.4.1 The antibody response

The majority of vaccines confer protection through the elicitation of nAbs [185]. These typically work by binding an antigen on the viral surface to prevent the virus from infecting the host cell. As the only viral protein on the virion surface, and playing a key role in mediating infection, Env is the sole target for nAbs. However, it is estimated that only approximately 14 Envs are incorporated into each viral particle [12], thus functional Envs are few and far between. The wide spacing of Env is disadvantageous for the host antibody response, as B cells are more effectively activated by repetitive and organised structures [186].

The virus also produces an abundance of non-functional envelope spikes in the form of uncleaved gp160, monomeric gp120, and gp41 stumps (Figure 1.2A). These can arise either through the improper processing of the spike, or its later disintegration. This 'viral debris' displays immunodominant epitopes that are either occluded or absent on the functional trimer (e.g. the inner surface of gp120), and acts to divert the host antibody response [70-75]. Thus, the initial antibody response is incapable of binding the functional envelope spike and is 'non-neutralising' [69].

A further nAb evasion feature of HIV-1 is the relatively poor accessibility of its conserved epitopes. The CD4bs, for example, is a highly conserved region essential for infectivity and thus represents a potentially vulnerable site for antibody neutralisation. However, its recessed location within the trimer interface, surrounded by N-glycans (Figure 1.6C), reduces the accessibility of this valuable epitope for nAb recognition. Interestingly, llamas and cows are capable of generating nAbs against the CD4bs following immunisation with Env [187, 188]. Llamas naturally produce heavy chain only antibodies that are much smaller than conventional antibodies, while cow antibodies contain very long third heavy chain complementarity-determining regions (HCDR3), some over 70 amino acids in length. The unusual architectures of these antibodies enable them to easily access the CD4bs, supporting a model of steric blocking for conventional IgG molecules.

In contrast, highly variable regions occupy the more accessible regions of the trimer, providing yet another immunodominant diversion (Figure 1.6B). Within a few months of infection patients readily develop autologous nAbs, often directed at the V1/V2 and V3 loops [78, 79]. Antibody-mediated selection pressure, in combination with the error-prone viral RT, rapidly drives viral escape and results in extreme antigenic diversity [79]. Thus, an effective vaccine against HIV-1 must not only induce nAbs, but antibodies with sufficient breadth of activity to neutralise the majority of circulating strains (i.e. bnAbs).

The development of bnAbs is dependent on the activation of the appropriate naïve B cells by engaging their B cell receptor (BCR), i.e. the precursor bnAb, prior to affinity maturation of the B cell in the germinal centre by somatic hypermutation. The resulting bnAbs are often significantly mutated from their germline-encoded BCR (Section 1.6) [189]. However, this process is hindered by the low affinity of so-called 'germline'-bnAbs (gl-bnAbs) to the envelope spike. Consequently, many recombinant Env mimics also fail to bind gl-bnAbs [190-193], adding to the difficulties in eliciting bnAbs in a vaccine setting (Section 1.8).

1.4.2 The glycan shield

A contributing factor to the immunodominance of many of the non-neutralising and autologous neutralising epitopes, and the inability of Env to bind gl-bnAbs, is the presence of an array of N-linked glycans that mask much of the surface of the envelope spike (Figure 1.6). Each Env can have upwards of 90 PNGS, with glycans comprising approximately half the trimer's mass [13]. The extensive N-glycosylation presents additional challenges for the host antibody response to overcome. Firstly, the glycans are derived from the host's own glycosylation machinery during Env synthesis and are therefore considered immunologically 'self'. Auto-reactive B cells undergo strong negative selection during B cell development, constraining the development of potential anti-glycan antibodies [194]. Furthermore, glycoproteins tend to exist as heterogeneous populations, with a multitude of glycan

structures decorating the same protein backbone, thereby potentially reducing the antigenicity of each individual glycoform [195]. Lastly, protein-glycan interactions tend to have low binding affinities and often require multivalent interactions to overcome this [196]. The heavily glycosylated outer domain of gp120 has been dubbed the 'silent face' of HIV-1, due to the previous lack of antibodies described against this region [197].

The role of glycans in protecting HIV-1 from nAb responses has been well documented. The glycan shield constantly evolves to escape the host immune system, with the addition and deletion of glycan sites frequently used by the virus to escape nAbs [198-200]. The 'evolving glycan shield' escape response is typified by the N332 glycan, which has been reported to shift from the N334 position and back again after the appearance of N332-dependent nAbs [198]. Furthermore, TF viruses typically have fewer PNGS than chronic isolates [200-203]. While this suggests there may be a fitness advantage to having fewer PNGS, a balance must then be struck between maintaining viral fitness and protecting vulnerable epitopes from the emerging nAb response. More recently a study by Wagh *et al.* reported that the addition of PNGS to fill holes in the glycan shield *in vivo* resulted in increased resistance to autologous nAbs [204]. Indeed, many *in vitro* studies have also reported the increased susceptibility of the virus to neutralisation upon removal of PNGS [205-209], which can often be rationalised by clashes observed between nAbs and glycans in structural studies [210, 211].

Despite the abovementioned challenges, approximately a third of infected individuals develop some level of bnAbs after a few years of infection [84-86]. Many bnAbs are able to either penetrate the glycan shield to bind to the protein surface, or directly bind to Env glycans. Thus, while glycan shielding remains a potent immune evasion strategy, the discovery of numerous glycan-binding bnAbs has highlighted the glycan shield as an attractive target for vaccine design [212]. The importance of the glycan shield in HIV-1 vaccine design has been underlined by Wagh *et al.*, who observed that the development of

bnAbs in infected individuals correlated with the completeness of the glycan shield at transmission [204].

1.4.3 Instability of the envelope spike

Although bnAbs are increasingly being isolated from infected individuals, we are yet to elicit them in a vaccine setting in humans. Early vaccination strategies using recombinant, monomeric gp120 failed to confer protection [177, 178, 213], presumably due to the elicitation of antibodies directed against the aforementioned immunodominant non-neutralising epitopes (and/or the absence of gp41 and quaternary epitopes) [214]. Therefore, the focus of HIV-1 vaccine research has shifted to include the production of trimeric Env immunogens that display the majority of bnAb epitopes while minimising non-neutralising epitopes as much as possible, with the hope that these will be better able to induce a bnAb response [215]. However, the envelope spike is inherently unstable, reflecting its need to undergo substantial conformational changes during viral and host cell fusion. This has made the design of native-like immunogens particularly challenging. The history and design of native-like Env trimers has recently been reviewed in great detail by Sanders and Moore [215].

The desire to remove the transmembrane region of the protein in order to generate soluble mimics often amplified trimer instability. Initial efforts to stabilise the trimer involved the removal or inactivation of the furin cleavage site to prevent gp120-gp41 dissociation, and/or the introduction of a trimerisation domain at the C-terminus to prevent separation of the three gp120-gp41 heterodimers [216, 217]. Although these approaches usually generated trimers, there was often an abundance of monomers, dimers, and higher-molecular weight aggregates, owing to the inappropriate formation of intermolecular disulphide bonds [115, 218]. These trimer constructs often displayed aberrantly folded gp120s due to intramolecular disulphide bond scrambling and non-native like quaternary structures as judged by peptide mapping, negative-stain EM, hydrogen-deuterium exchange mass

spectrometry and glycosylation analysis (Section 1.7). Perhaps most importantly, these trimers display non-native-like antigenicity [115, 123, 218-220], and are commonly referred to as 'pseudotrimers' to reflect their various non-native properties [215]. With hindsight it was therefore unsurprising that immunisation studies with pseudotrimers, not unlike monomeric gp120, failed to elicit nAbs of sufficient breadth and potency [221-223].

1.5 Protein mimicry in vaccine design

The development of the SOSIP.664 platform transformed the field of native-like HIV-1 trimer design (Figure 1.6) [215]. These constructs retained the furin cleavage site, optimised for efficient cleavage, and relied on the introduction of a disulphide bond ('SOS') to covalently link the gp120 and gp41 subunits [130]. An additional point mutation in the gp41 subunits (I559P, 'IP') strengthened interactions between the three heterodimers by trapping the trimer in its pre-fusion conformation [133, 134], while truncation before the transmembrane region ('.664') ensured the solubility of the trimers and reduced aggregate formation [224]. The prototypic BG505 SOSIP.664 trimers also incorporated a further point mutation (T332N) to introduce a glycan site that contributes to the epitopes of many bnAbs*.

In recent years, an arsenal of native-like trimers has been produced. The 'SOSIP' format has since been applied to multiple strains, and mosaic and consensus sequences, though these often required further stabilising point mutations and disulphide bonds (either between gp120-gp41 and/or between protomers) [225-233]. Other native-like trimer formats focused on eliminating the requirement for furin cleavage (and therefore the need to co-express Env immunogens with furin encoding plasmids) in an attempt to simplify protein production strategies and DNA based vaccines. Native flexibly linked (NFL) and single-chain (SC) trimers achieved this feat by replacing the furin cleavage site with a flexible Gly-Ser linker [234, 235]. While the NFL and SC constructs both relied on the I559P point

*Throughout this thesis, all BG505 SOSIP.664 trimers (and other native-like trimers based on the BG505 strain) contain the T332N mutation, unless otherwise specified.

mutation to maintain the pre-fusion conformation, UFO (uncleaved prefusion-optimised) constructs achieved this through a computational redesign of the HR1 (heptad repeat 1) region [236].

The newly developed trimer formats have not only enabled the creation of native-like trimers from strains that had previously proved problematic [215], but have also resulted in 'hyper-stabilised' trimer constructs. Compared to the earlier versions, these proteins display less conformational flexibility and reduced exposure of non-nAb epitopes [225]. The immunogenicity of many of the above mentioned native-like trimers have been investigated in animal models [225, 228, 237-242], with the first Phase I clinical trials ongoing [243].

1.5.1 Native-like trimers display bnAb epitopes

Native-like trimers are often assessed by their ability to bind bnAbs and not non-neutralising antibodies. The development of soluble, native-like trimers has enabled the characterisation of many bnAb epitopes through various biophysical techniques, such as X-ray crystallography and cryo-EM [244]. Broadly neutralising antibodies are generally categorised by their recognition of five distinct and largely conserved epitopes on the envelope spike: the CD4bs, the membrane proximal external region (MPER), and the N-glycans located at the gp120-gp41 interface, the outer domain of gp120, and on the V1/V2 loops at the trimer apex. However, recent advances in high throughput B cell screening have led to a dramatic increase in the identification of new bnAbs, and have subsequently revealed a continuum of epitopes spanning the entire surface of the trimer [245], including the majority of N-glycans (Figure 1.8). The recent characterisation of bnAb VRC-PG05, which recognises glycans at positions N262, N295 and N448, could conceivably be the last class of bnAb to be identified as its discovery filled one of the few remaining gaps on the trimer surface [246].

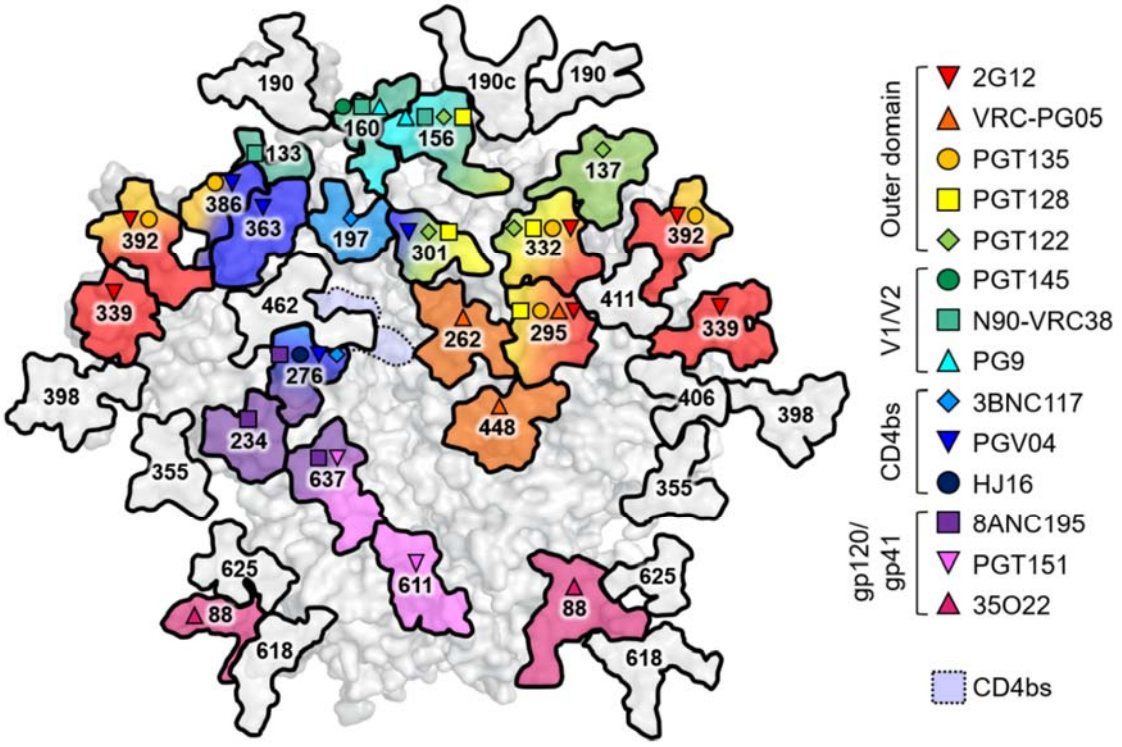


Figure 1.8. Broadly neutralising antibodies recognise glycan-containing epitopes. Model of a fully glycosylated BG505 SOSIP.664 trimer depicting the N-glycans that have been implicated in binding a variety of bnAbs (inset key). Glycans not present on the BG505 strain have been omitted, e.g. 35O22 also recognises N-glycans at positions 230 and 241. Model based on PDB 5ACO as in Figure 1.6, numbering according to the HXB2 reference sequence. CD4bs, CD4 binding site.

The extent to which individual bnAbs depend on glycans for binding and neutralisation varies. At one extreme, 2G12 recognises an epitope comprised exclusively of N-glycans [247-253]. On the other hand, bnAbs against the CD4bs, for example, have evolved to either avoid or accommodate glycans that would otherwise occlude the underlying protein epitope [254]. Most bnAb epitopes occupy a midway point, with binding dependent on both protein and glycan components.

The abundance of glycan-targeting bnAbs has earmarked N-glycans as important components of a future HIV-1 vaccine. Thus, there is a need for a detailed definition of the precise composition of the glycan shield on both viral Env and candidate immunogens. Additionally, an understanding of the principles controlling glycosylation will help guide the

design of immunogens that are able to effectively mimic viral Env glycosylation, and, hopefully, elicit a broadly neutralising anti-glycan response.

1.6 Broadly neutralising antibodies

1.6.1 Structure and development of broadly neutralising antibodies

The development of bnAbs requires repeated rounds of viral escape and antibody maturation and can therefore take several years of infection to arise [80, 83, 198, 255]. Their slow development may reflect the fact that bnAbs often depend on unusual antibody features in order to overcome the aforementioned challenges associated with the development of glycan recognition and neutralisation breadth.

The glycan-targeting antibody, 2G12, overcomes low affinity protein-glycan interactions by exhibiting a unique domain-exchanged structure. Here, two heavy chain variable regions are exchanged to create a single, domain-exchanged Fab_2 [250]. The resulting structure has an additional antigen binding site at the interface of the two arms, which allows for the binding of four Env glycans with high avidity (Figure 1.9A) [249, 250].

Many bnAbs, particularly those targeting the V1/V2 and outer domain glycans, contain very long HCDR3 sequences [256-261]. While the average length of human HCDR3 is 13 residues long [262], bnAb HCDR3s can contain upwards of 30 residues. This is particularly true of bnAbs targeting the glycans of the outer domain and V1/V2 loops, where long HCDR3s allow the antibody to penetrate the glycan shield and make contact with the underlying protein surface. The HCDR3 of PGT145, for example, contains 33 residues, and penetrates the N160 glycan triad at the apex of the trimer to contact underlying protein residues from all three protomers (Figure 1.9B) [256]. For outer domain-targeting bnAbs (e.g. PGT128 and PGT135), long HCDR3s allow penetration of the glycan shield to access the base of the V3 loop [258, 259]. In both examples, at least one glycan makes extensive interactions with the bnAb antigen binding site. A similar phenomenon has been observed for bnAbs targeting the MPER, namely 4E10, where a long (20 residue), hydrophobic

HCDR3 allows for contact with the lipid membrane as well as the gp41 protein surface [261]. Often long HCDR3s are accompanied by post-translational modifications, such as tyrosine sulphation, which have also been reported to contribute to binding and neutralisation (Figure 1.9C) [257].

Another feature present in many bnAbs is extensive somatic hypermutation [189]. Classically, antibodies accumulate such mutations in the antigen-contacting complementarity-determining regions (CDRs). However, bnAbs often require somatic mutations to the more conserved framework regions [263]. Although there is evidence that higher levels of somatic hypermutation correlate with increased breadth and potency in some bnAb lineages [264], this may be a consequence of the length of time of development rather than a necessity of bnAb activity. For instance, partially germline reverted forms of VRC01 and 10E8 bnAbs are still broad and potent neutralisers [265].

Poly- and auto-reactivity is also frequently associated with bnAbs [266-268]. These observations are somewhat unusual as both characteristics are negatively selected for during

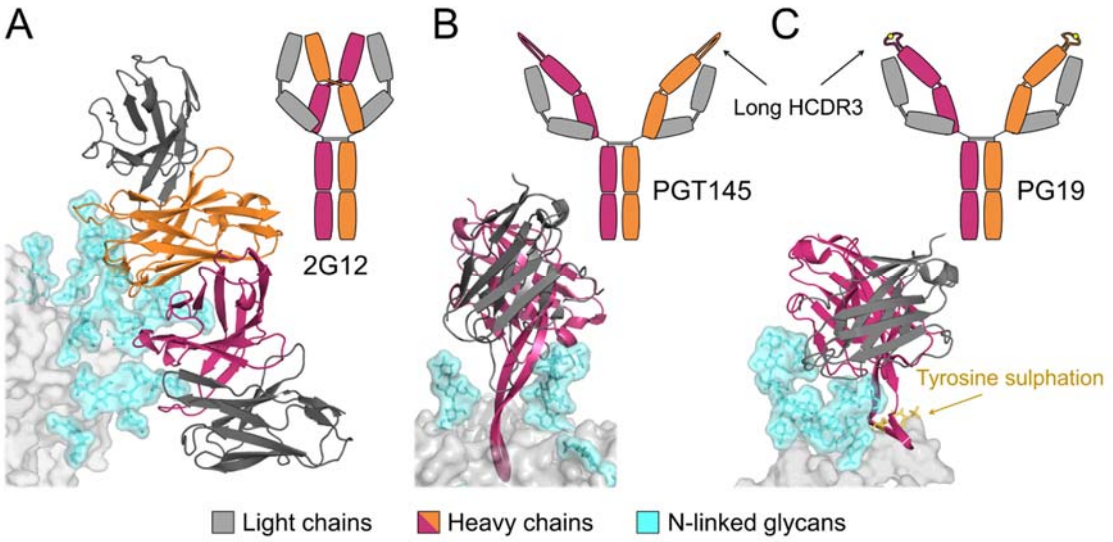


Figure 1.9. Structural features of broadly neutralising antibodies. A) Domain-exchange of the heavy chain variable regions of 2G12 creates additional antigen binding sites allowing for high avidity binding of glycans. Cryo-EM structure provided by Dr Chris Cottrell from the laboratory of Prof. Andrew Ward, Scripps Research, La Jolla, USA. B and C) Long HCDR3s on PGT145 and PG16 enable penetration of the glycan shield. The loops make contact with both glycan and protein components. Tyrosine sulphation (yellow) has been reported to contribute to bnAb binding and neutralisation (PDB 5V8L and 4DQO) [257].

B cell development [194]. It has been hypothesised that conserved HIV-1 epitopes may mimic host proteins in order to avoid host antibody responses through the down-regulation of relevant B cells by host tolerance mechanisms [266, 267, 269]. A similar host immune evasion strategy has been reported for the pathogen *Campylobacter jejuni,* whose lipooligosaccharide (LOS) mimics the gangliosides of the host's peripheral nervous system [270]. Despite these barriers, poly- and auto-reactive bnAbs are not uncommon, though their development may be dependent on prolonged exposure to the antigen, as several B cell tolerance checkpoints may need to be overcome. The elicitation of bnAbs through typical vaccination strategies may therefore prove challenging.

1.6.2 Understanding the nature of glycan-dependent bnAb epitopes

The structural characterisation of multiple bnAbs in complex with Env by both X-ray crystallography and EM has enabled a detailed description of the epitopes of many bnAbs (Table 1.1) [135, 136]. Furthermore, it has revealed many of the unusual antibody features required for broad neutralisation. Nevertheless, there remained many unknowns about the precise nature of glycan epitopes. This is, in part, due to difficulties in resolving extremely heterogeneous glycans with structural techniques that rely on the averaging of many molecules to reveal the consensus structure [244]. Glycan arrays are also a useful tool in assessing the glycan-binding properties of bnAbs, though these too have limitations (e.g. if the epitope comprises protein components). The site-specific glycan analysis of Env trimers has gone some way to bridging these knowledge gaps. While there is generally good agreement of the precise structures occupying glycan epitopes between glycan analyses and structural studies/arrays, a few discrepancies remain. For example, there are several reports on the preference of apex-targeting bnAbs PG9 and PG16 for complex- or hybrid-type glycans (specifically α2-6-linked sialylated structures) at the N156 site [271-273]. However, these structures are seldom seen at this site in glycan analyses of recombinant, soluble trimers or virally-derived Env [132, 274-276].

Table 1.1. Examples of glycan-targeting bnAbs and their glycan epitopes.

Region	bnAb (and key to Figure 1.8)	PNGS	Glycan specificity	Refs
Outer domain glycans	PGT130	N301, N332/N334	Oligomannose	[277, 278]
	2G12	N295, N332, N339, N392	Oligomannose, specifically α1-2-linked motifs on $Man_{8-9}GlcNAc_2$ structures	[247-253]
	VRC-PG05	N262, N295, N448	Oligomannose	[246]
	PGT135	N295, N332, N386, N392	Oligomannose	[259, 278, 279]
	PGT128	N156, N295, N301, N332/N334	Oligomannose	[131, 258, 277, 278]
	PGT122	N137, N156, N301, N332	Oligomannose	[135, 278, 280]
	PGT121	N137, N156, N301, N332/N334	Oligomannose or complex	[264, 277, 278, 281]
	BG18	N156, N332, N386, N392	Oligomannose, possibly complex at N156	[282]
	10-1074	N156, N301, N332	Oligomannose, possibly complex at N156	[281, 283]
V1/V2 glycans	PGT145	N160	Oligomannose	[256, 278]
	CH04	N160	Kifunensine abrogates neutralisation (can't tolerate $Man_9GlcNAc_2$)	[284]
	PGDM1400	N160	Kifunensine reduces/abrogates neutralisation (can't tolerate $Man_9GlcNAc_2$)	[285]
	VRC38.01	N133, N156, N160	*N*-Acetylglucosamine core of N133	[286]
	PG9 and PG16	N173 (N156), N160	Oligomannose at N160, hybrid or bi-antennary structures at N156/N173, specifically containing α2-6-linked terminal sialic acids	[271-273, 287-289]
	CAP256-VRC26 (lineage)	N156, N160	Oligomannose at N160, hybrid at N156, or bi-, tri-, tetra-antennary structures containing α2-6-linked terminal sialic acids	[255, 290]
	PCT64 (lineage)	N156, N160	Oligomannose, preferentially $Man_5GlcNAc_2$	[291]

CD4bs proximal	3BNC117		N197, N276	*N*-Acetylglucosamine cores	[256, 280]
	PGV04 (VRC-PG04)		N276, N301, N363, N386	Enzymatic de-glycosylation doesn't affect binding (no strong glycan-dependence)	[136, 292]
	HJ16		N276	*N*-Acetylglucosamine core	[293-295]
	IOMA		N197, N276, N363	Complex, minor contact with oligomannose glycan at N363	[283]
	179NC75		N276	Oligomannose or hybrid[a]	[296]
	VRC01		N276	*N*-Acetylglucosamine core	[254, 280]
gp120/gp41 interface	8ANC195		N234, N276, N637	Kifunensine did not affect neutralisation (can tolerate $Man_9GlcNAc_2$)	[297, 298]
	PGT151		N611, N637	Complex, specifically tri- and tetra-antennary structures	[299-301]
	VRC34.01		N88	Kifunensine, swainsonine or GnT I$^{-/-}$ cells[b] minimally affect neutralisation (not dependent on complex structures)	[302]
	ACS202		N88	Kifunensine, or GnT I$^{-/-}$ cells minimally affect neutralisation (not dependent on complex structures)	[303]
	35O22		N88, N230, N241	Oligomannose	[280, 304]

a. In Freund *et al.*, the loss of binding of 179NC75 following expression of BG505 SOSIP.664 trimers with kifunensine was used to argue a dependency on complex-type glycans. However, they also report a loss of binding to gp120 after digestion with Endoglycosidase H (which cleaves oligomannose- and hybrid-type glycans). This suggests that 179NC75 recognises $Man_{8-5}GlcNAc_2$ or hybrid-type structures. In line with this, site-specific glycan analysis of BG505 SOSIP.664 trimers reports predominantly $Man_{8-5}GlcNAc_2$ glycans at the N276 site, with a small population of hybrid-type structures [132].

b. Treatment with swainsonine (α-mannosidase II inhibitor) results in oligomannose- and hybrid-type glycosylation. Expression in GnT I$^{-/-}$ cells (deficient in *N*-Acetylglucosaminyltransferase I) results in $Man_{5-9}GlcNAc_2$ structures only.

1.7 Glycan mimicry in vaccine design

The observation that glycan-dependent bnAbs can bind to soluble mimetics of the viral spike suggests that the glycans displayed by such immunogens are largely tolerated within the bnAb glycan specificity profile. However, given the nature of bnAbs, which have evolved to tolerate microheterogeneity of glycans, cross-reactivity does not necessarily indicate that the target glycans are precisely conserved between immunogen and virus. For example, some

glycan specific bnAbs recognise the largely invariant base of the glycan [256, 280]. Furthermore, failure of an immunogen to bind a bnAb could be symptomatic of either a failure of protein mimicry, glycan mimicry, or both. For these reasons, there is significant interest in defining the glycosylation of both target viruses and candidate immunogens. Finally, detailed information about glycosylation will also help define the immunological ramifications of particular immunogen and viral glycoforms that may go beyond simply the display of particular glycan structures at a particular site. For example, are some glycoforms more inflammatory or immunogenic than others [305]?

1.7.1 Understanding Env glycosylation processing

The analysis of HIV-1 glycosylation is particularly challenging given the extensive heterogeneity displayed by glycoproteins, combined with the large number of PNGS on Env. Furthermore, isolating virally-derived Env in sufficient quantities for glycan analysis has proved difficult. Early glycan analyses were therefore generally performed on recombinant, monomeric gp120 [110, 306-311], or trimeric Env constructs derived from pseudovirions [306, 312-314], membrane-associated trimers [315, 316], and soluble trimers [122, 123, 132, 242, 274, 316-322]. These analyses soon revealed key aspects of Env glycosylation processing. Firstly, the Env glycan shield is heterogeneous. The gp120 subunit alone can contain upwards of 50 different glycan structures [132], attributed to the large number of glycan processing enzymes possessed by mammalian cells. Despite this heterogeneity, there always exists a substantial population of under-processed, oligomannose-type glycans ($Man_{5-9}GlcNAc_2$). This glycan signature arises through steric constraints within the Env glycan shield that impede the actions of some of the host glycosylation enzymes (Section 1.7.3 and 1.7.4) [310, 312, 313, 323].

Many of the above glycan analyses focused on the analysis of enzymatically released N-glycans. While this provides a useful readout of the overall glycosylation profile, no site-specific information can be gleaned. Such information, which involves the analysis of

protease-digested glycopeptides, is valuable in order to elucidate the precise composition of bnAb epitopes. The development of the SOSIP.664 platform, combined with an advancement of mass spectrometric and chromatographic technologies enabled more in-depth, site-specific glycan analyses.

In 2016, Behrens *et al.* published a quantitative, site-specific glycan analysis of the recombinant BG505 SOSIP.664 trimers [132]. This revealed that the oligomannose signature observed on Env was largely accounted for by several PNGS containing solely under-processed, oligomannose-type glycans, usually dominated by $Man_9GlcNAc_2$. The remaining sites contained either processed, complex-type glycans, or a mixed population. Mapping this information onto the structure of BG505 SOSIP.664 illustrated how large regions of oligomannose-type glycans span across the outer domain of gp120, while complex-type glycans occupied sites at the periphery of the trimer, particularly on the gp41 subunits (Figure 1.10).

The Paulson laboratory employed a complementary site-specific workflow to analyse BG505 SOSIP.664 trimers [274]. By sequential digestion of glycopeptides with Endoglycosidase H (Endo H; to cleave oligomannose-/hybrid-type glycans, leaving a GlcNAc residue) and Peptide-N-Glycosidase F (PNGase F) in the presence of O^{18}-water (to cleave the remaining complex-type glycans and convert the Asn to an O^{18}-labelled Asp), they were able to generate novel mass signatures for each category of glycan: +203 Da for oligomannose-/hybrid-type, +3 Da for complex-type glycans, and +0 for unoccupied peptides. Their data were in very close agreement to the Behrens *et al.* study. This 'classification' method is very powerful for low abundance samples (as heterogeneous glycopeptides become grouped as a single peptide), although information on the exact composition of glycan sites is not captured.

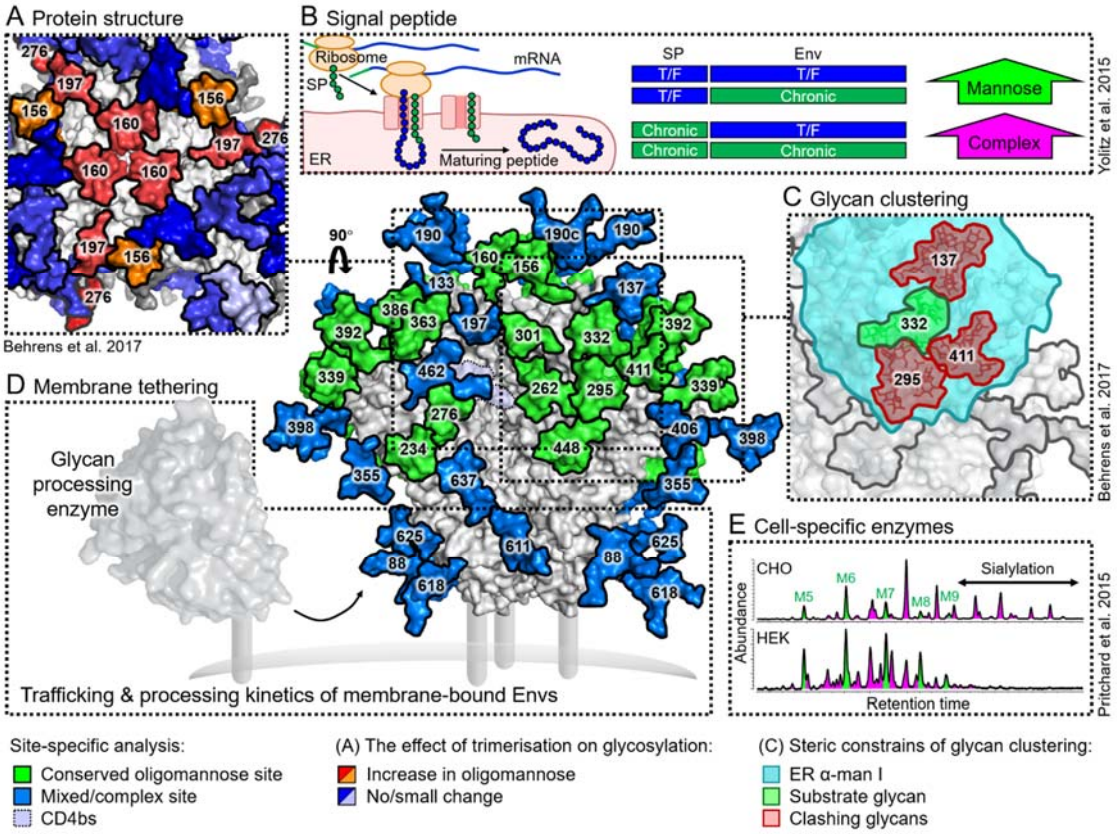


Figure 1.10. Principles controlling Env glycosylation. Site-specific glycan analysis of recombinant BG505 SOSIP.664 [122, 132, 274, 324] and virally-derived BG505 Env [275, 276] has revealed clusters of glycans displaying under-processed, oligomannose-type glycans (green). These are largely located on the outer domain of gp120 (forming the intrinsic mannose patch) and at the trimer apex and protomer interfaces (forming the trimer associated mannose patch). Model according to Figure 1.6. A) The quaternary protein structure of native-like trimers imposes steric constraints on the host's glycosylation enzymes, resulting in an increase in the amount of oligomannose-type glycans at sites near the protomer interfaces (red/orange), compared to that of monomeric or non-native trimeric Env [122]. B) Irrespective of the mature Env peptide sequence, the presence of a signal peptide (SP) from a transmitted/founder (TF) viral isolate results in an increase in oligomannose-type glycosylation, while a chronic-stage signal peptide results in increased complex-type glycosylation. The signal peptide influences Env trafficking, folding and retention through the ER [325]. C) Modelling the ER α-mannosidase I (cyan, PDB 5KIJ) on to its substrate glycan (green) reveals extensive clashes with neighbouring glycans (red) sufficient to explain the formation of the intrinsic mannose patch [323]. D) It is hypothesised that membrane-bound Env constructs display elevated glycosylation processing as they exhibit a different topology relative to the membrane-bound enzymes compared to soluble constructs, which are released into the lumen of the ER [274, 276]. Schematic of a membrane-bound glycan processing enzyme, based on the structure of a sialyltransferase. E) While the processing of many Env glycans is limited by protein-directed steric constraints, the fate of others is dependent on the glycosylation enzymes possessed by the host cell. The gp41 from BG505 SOSIP.664 expressed in Chinese hamster ovary (CHO) cells displays increased sialylation compared to the same protein expressed in human embryonic kidney (HEK) cells [123]. M5-9, $Man_{5\text{-}9}GlcNAc_2$.

In some instances, the detail obtained from intact glycopeptide analysis may explain < 100% neutralisation plateaus observed by some bnAbs [279, 326]. For instance, PGT135 is only able to neutralise around 85% of BaL pseudoviruses [279]. This bnAb recognises predominantly glycans at N332, N386, and N392 (Figure 1.8) [259]. Site-specific analysis of the N392 PNGS on recombinant $gp120_{BaL}$ (subscript denotes strain) revealed the majority of glycans at this site to be $Man_8GlcNAc_2$ structures, with a secondary population of $Man_9GlcNAc_2$ [279]. In line with this, a crystal structure of the PGT135 Fab bound to a $gp120_{JR\text{-}FL}$ core shows the N392 site to be occupied by a $Man_8GlcNAc_2$ structure [259]. Modelling an additional mannose residue to this glycan (to give $Man_9GlcNAc_2$) revealed steric clashes with the bnAb [259]. Furthermore, it has been shown that PGT135 is unable to neutralise pseudovirus displaying predominantly $Man_9GlcNAc_2$ structures [259]. Thus, $Man_9GlcNAc_2$ at the N392 site seems suboptimal for PGT135 binding, and the site-specific presence of this structure may account for the observed neutralisation plateaus.

1.7.2 Immunogen mimicry of viral glycosylation

Recent progress in the production and purification of HIV-1 virions has enabled the glycosylation analysis of virally-derived Env. The Dell laboratory presented a qualitative analysis of $gp120_{BaL}$ glycosylation derived from virions produced in a human lymphoid cell line [327]. In line with analysis of SOSIP.664 trimers, 15 of the 24 PNGS contained solely oligomannose-type structures, although the relative abundances of each glycoform were not determined. The remaining nine contained either complex-type structures or mixed populations [327].

In 2018, Struwe *et al.* performed site-specific glycan analysis on $gp120_{BG505}$ isolated from virions expressed in a similar lymphoid cell line [276]. This allowed for a comparison with BG505 SOSIP.664 trimers expressed in human embryonic kidney (HEK) cells (the key findings of which are presented in Chapter 7 of this thesis) and under equivalent conditions to current Good Manufacturing Practices (cGMP) in a Chinese hamster ovary (CHO) cell

line. As will be described in Chapter 7, the analysis revealed that sites occupied by exclusively oligomannose-type glycans on virally-derived Env were largely conserved on recombinant SOSIP.664s. This is reassuring considering a large proportion of glycan-targeting bnAbs recognise the oligomannose-type glycans at these sites (Figure 1.8). However, there were some key differences between virally-derived and recombinant Envs. Namely, virally-derived Envs displayed increased levels of glycosylation processing, both in terms of the relative amount of complex-type glycosylation, and also the structures present (i.e. complex-type glycans on virally-derived Envs were more branched). This discrepancy was generally attributed to several complex-type PNGS on virally-derived Env that contained mixed populations of glycans on the SOSIP.664 trimers. We propose this is due to the membrane-bound nature of virally-derived Envs and will discuss this further in Section 1.7.6 and Chapter 7.

The Paulson laboratory have also compared the glycosylation of virally-derived Env from three strains (JR-FL, BG505, and B41), produced in peripheral blood mononuclear cells (PBMCs), with their corresponding SOSIP.664 trimers, produced in HEK cells, using their aforementioned site-specific classification method [275]. Their results were in general agreement with Struwe *et al.*'s observations that mixed sites on SOSIP.664 trimers tended to be fully processed on the virally-derived Env. Thus, while there is some differential processing between viral and recombinant Env, many key bnAb epitopes are conserved. In terms of immunogen design, it is currently unknown whether absolute mimicry of viral glycosylation is required.

1.7.3 Glycan clustering and the intrinsic mannose patch

The abundance of oligomannose-type glycans on Env is somewhat unusual given the virus derives its glycan shield from the host cell glycosylation machinery. This typically follows a highly ordered pathway whereby oligomannose-type precursors are trimmed and rebuilt as complex- and/or hybrid-type glycans (Figure 1.5). Two hypotheses existed as to why Env

glycosylation diverges from that typically observed on mammalian glycoproteins. Either a proportion of Env glycoproteins are exiting the glycosylation pathway before encountering the later enzymes, or steric constraints exist within Env that are preventing complete processing by the earlier enzymes. Several pieces of evidence support the latter model.

Firstly, a proportion of oligomannose-type glycans are resistant to endoplasmic reticulum ER α-mannosidase I digestion *in vitro* [313]. Expressing gp120 with kifunensine (an α-mannosidase I inhibitor) generates a glycoprotein bearing almost exclusively $Man_9GlcNAc_2$ structures, replicating the immature glycoprotein found in the early ER. Kinetic analysis of the hydrolysis of $Man_9GlcNAc_2$ to $Man_8GlcNAc_2$ revealed approximately half of the $Man_9GlcNAc_2$ to be rapidly trimmed to $Man_8GlcNAc_2$, with the remaining $Man_9GlcNAc_2$ processed at a much slower rate. However, there remained a proportion of $Man_9GlcNAc_2$ (~ 30%) that could not be hydrolysed, even after exhaustive digestion. The resistance of these glycans to digestion *in vitro* supports the model of steric hindrance, likely caused by the high density of glycans on gp120 preventing α-mannosidases accessing their substrate glycans.

Each gp120 subunit contains between 18 and 33 (median 25) PNGS [3]. To elucidate whether glycan density is the limiting factor preventing complete processing of the glycan shield, the number of PNGS on the outer domain of gp120 were correlated with the abundance of oligomannose-type glycans on sequences isolated from an individual over the course of infection, and on a cross-clade panel of 29 strains [311]. In both instances, a strong correlation was observed. This correlation was not seen when comparing the abundance of oligomannose-type glycans with the total number of PNGS on gp120 for the cross-clade panel (though it was for the infected patient sequences) suggesting that the incomplete processing observed on gp120 is driven by local glycan density, rather than overall glycan number. Similarly, Stewart-Jones *et al.* observed the number of glycan processing steps to be significantly lower for 'crowded' glycans (those with more than 15 PNGS within a 50 Å

radius) than for 'dispersed' glycans (those with fewer than 15 PNGS within a 50 Å radius) [280].

The extent to which individual glycans contribute to the steric hindrance of glycosylation enzymes was investigated by the systematic site-directed mutagenesis of all the PNGS on $gp120_{BaL}$, by mutating each Asn within the glycosylation sequon to Ala [310]. Although the removal of individual sites did not have a severe effect on the abundance of oligomannose-type glycans, several site deletions resulted in a larger than expected loss of $Man_9GlcNAc_2$, often accompanied by a compensatory increase in the lower oligomannose species, $Man_{5\text{-}8}GlcNAc_2$. The glycan sites accounting for the largest decrease in $Man_9GlcNAc_2$ generally mapped to the outer domain of gp120, further supporting a model whereby localised glycan clustering is sterically hindering early glycosylation processing. Thus, the loss of a PNGS within such clusters would increase the accessibility of neighbouring glycans so that multiple nearby glycans exhibit increased processing. In further support of this hypothesis, it has recently been confirmed by site-specific glycan analysis that the disruption in glycan processing upon the loss of a PNGS is largely limited to those sites adjacent to the missing glycans [328].

Thus, while the glycan structures present on most glycoproteins are determined by the host cell, HIV-1 is able to partly modulate its glycosylation processing through the number and position of PNGS, as determined by the viral sequence. The number of PNGS on Env are conserved both across clades and longitudinally throughout infection [311, 312]. The resulting high glycan density on Env (and gp120 in particular) limits the actions of α-mannosidases and results in a large population of oligomannose-type glycans. This principle is illustrated in a model of the ER α-mannosidase I enzyme binding a substrate $Man_9GlcNAc_2$ glycan on the outer domain of gp120, which reveals extensive clashes with the surrounding glycans [323] (Figure 1.10C). A similar phenomenon has been observed on the glycoprotein complex of Lassa virus, which also exhibits high localised glycan density resulting in under

processed oligomannose-type clusters [329]. The oligomannose-type glycans of gp120 are therefore an inherent feature of the glycoprotein, known as the 'intrinsic mannose patch' (IMP), and to some extent all Env constructs will display this glycan signature (Figure 1.11). Encouragingly, many of the aforementioned glycan-targeting bnAbs recognise the oligomannose-type glycans of the IMP. The overall resilience of the oligomannose population to sequence variation, and its presence on both immunogens and viral Env, supports this feature as a conserved target for vaccine design.

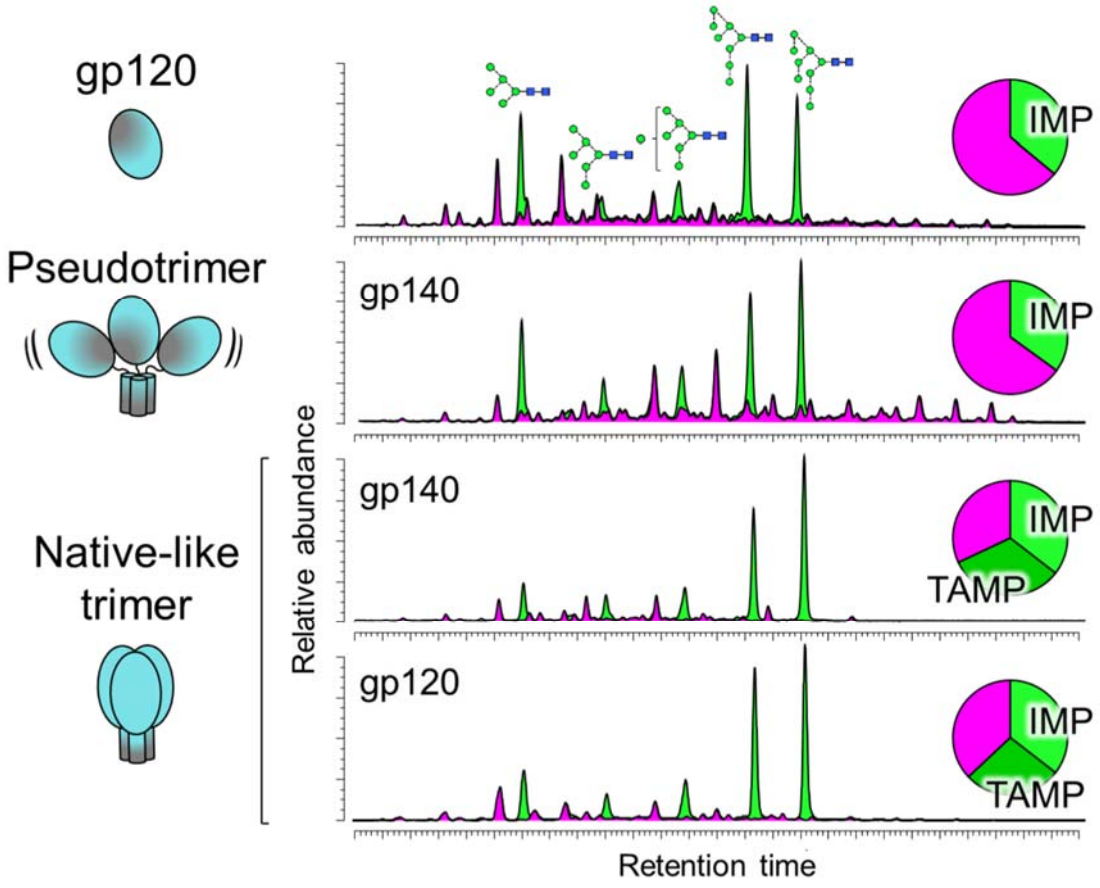


Figure 1.11. Glycan clustering and protein structure limit glycosylation processing. Quantitative glycan analysis of monomeric gp120, pseudotrimers, and native-like trimers. For comparison, data for both gp120 and gp140 (gp120 + truncated gp41) from native-like trimers have been included. Each construct is based on the BG505 sequence containing the T332N mutation. Glycans were enzymatically released, fluorescently labelled, and analysed by hydrophilic interaction liquid chromatography-ultra performance liquid chromatography. Oligomannose-type glycans were quantified by their susceptibility to digestion with Endo H. The chromatograms reveal a population of oligomannose-type glycans (green) intrinsic to all Env constructs, termed the intrinsic mannose patch (IMP). Only native-like trimers display the additional trimer-associated mannose patch (TAMP) signature, attributed to additional steric protection from processing.

1.7.4 Protein structure and the trimer associated mannose patch

As the focus of immunogen design shifted from monomeric gp120 toward trimeric Env, so did the requirement to categorise glycans from native-like immunogens. It had been observed that trimeric Envs often displayed even higher levels of under-processed oligomannose-type glycosylation [122, 123, 312, 313]. The discrepancy in the abundance of oligomannose-type structures between monomeric and trimeric Env suggests that these structures are derived from an additional mechanism to the one driving the formation of the IMP. It was hypothesised that the 'trimer associated mannose patch' (TAMP) arose due to glycan-glycan and glycan-protein interactions at the protomer interface further sterically restricting processing enzymes [312, 313, 330].

Quantitative glycan analysis soon revealed that the TAMP glycan signature was only present on native-like trimers, with non-native pseudotrimers often displaying glycosylation patterns similar to that of monomeric gp120 (Figure 1.11) [122, 123, 218]. Negative-stain EM revealed that non-native pseudotrimers often form open, irregular structures, in stark contrast to the well-ordered, propeller shaped format of native-like trimers [123, 218, 323]. This structural difference is sufficient to explain the discrepancy in glycosylation, as open, irregular structures lack the additional steric protection of native-like trimers and allow for more complete glycan processing. Thus, although most Env proteins have a population of oligomannose-type glycans atypical of mammalian glycosylation, only native-like trimers display the additional TAMP signature. As such, glycosylation profiling is becoming a widely-adopted tool that can be used to readily distinguish between native-like and misfolded trimers and guide the selection of immunogens [331].

Site-specific glycan analysis was applied to sequence-matched $gp120_{BG505}$, uncleaved pseudotrimers, and native-like, BG505 SOSIP.664 trimers. The results revealed for the first time the individual PNGS that contribute to the TAMP. As hypothesised, sites located at the protomer interfaces on the native-like BG505 SOSIP.664 trimers displayed restricted glycan

processing compared to both the pseudotrimers and monomeric gp120 (Figure 1.10A). This was most significant at sites N156, N160, N197 and N276 which had > 40 percentage point increase (pp; the arithmetic difference between two percentages) in oligomannose-type glycans on BG505 SOSIP.664 compared to gp120 [122].

Crucially, in support of the native-like configuration of SOSIP.664 trimers, the TAMP signature is present on both virally-derived Envs and native-like trimers (Chapter 7), though it is somewhat smaller on viral Env [276]. This is likely due to 'TAMP sites' N197 and N276 displaying predominantly complex-type glycosylation on viral Env [275, 276]. The N156 and N160 'TAMP sites' are, however, conserved oligomannose-type sites in most instances [275, 276]. The N160 site has been reported as a complex-type site in JR-FL virions, highlighting potential strain-specific differences [275].

1.7.5 Cell-specific glycosylation

While the abundant population of oligomannose-type glycans observed on Env are attributed to steric constraints imposed by the protein itself, the more accessible regions of the glycoprotein are subject to processing by the host cell's glycosylation pathway. Such processing is cell type-specific, dependent on the repertoire of glycosidases and glycosyltransferases expressed by the host cell [332].

A comparative analysis of BG505 SOSIP.664 derived from CHO cells and HEK cells illustrates the nature of protein- and cell-directed Env glycosylation [123]. Both CHO and HEK-derived material display similar levels of protein-directed glycosylation: oligomannose-type glycans account for 55 and 56% of the total glycan pool, respectively, with individual oligomannose species exhibiting similar distributions [123]. In contrast, the remaining complex-type glycans present considerable cell-specific differences. This is particularly evident on the gp41 subunit, with CHO-derived material containing substantially more sialylated structures than HEK-derived BG505 SOSIP.664 (Figure 1.10E) [123]. The abundance of sialic acid is known to impact the antigenicity and immunogenicity of Env.

Kong *et al.* observed that gp120 expressed in an insect cell line modified to impart asialylated mammalian-type glycans was significantly more immunogenic than sialylated gp120 expressed in HEK cells [305].

A comparison of virally-derived gp120 produced in PBMCs and gp120 from pseudovirus produced in HEK cells revealed similar cell-specific differences [333]. Namely, sialylated glycans from material produced in PBMCs displayed primarily α2-6-linkages, whereas HEK-derived material displayed only α2-3-linked structures [333]. It can be noted that α2-6-linked sialylated glycans have previously been associated with anti-inflammatory effects, thus there may be potential immunological consequences of these structures [334].

Like oligomannose-type glycans, complex-type glycans have emerged as important in regard to the formation of bnAb epitopes. Some bnAbs, such as PGT121 display promiscuous recognition of both oligomannose- and complex-type glycans in glycan arrays [281]. Antibodies such as PGT151 are dependent on tri- and tetra-antennary structures at the gp120-gp41 interface [299], while α2-6-linked sialylated hybrid- and complex-type glycans have been implicated in the development and binding of apex-targeting bnAbs such as PG16 and the CAP256-VRC26 lineage [271, 273, 290]. Interactions between the sialic acid residues of Env and Siglecs have also been observed to play an important role in the infection of macrophages, which express low levels of cell-surface CD4 [48]. Thus, cell-directed glycosylation can play a role in both viral infectivity and bnAb binding, and understanding the factors influencing complex-type glycosylation may have implications in guiding immunogen design, for instance when choosing expression cell lines.

1.7.6 Membrane tethering

The majority of glycan analyses have been performed on soluble Env constructs due to the difficulties associated with expressing and purifying full-length, membrane-bound Env. Until recently, the impact of membrane tethering on Env glycosylation had remained largely unaddressed. Analyses of virally-derived Envs, while performed on membrane-bound

material, generally did not control for expression cell-line, known to influence glycosylation processing [275, 276]. However, recent studies by Rantalainen *et al.* and Cao *et al.* directly compare the glycosylation of membrane-bound Envs with their corresponding soluble SOSIP.664 trimers expressed in the same cell lines. The differences observed between membrane-bound Env and SOSIP.664 are reminiscent of the differences observed between viral Env and SOSIP.664, in that sites containing mixed populations of glycans on the soluble SOSIP.664 trimer typically displayed only fully processed, complex-type glycan structures on the full-length construct [335]. It was hypothesised that the more complete processing observed on full-length constructs is due to membrane-bound Env being kept in closer proximity to the membrane-bound glycosylation enzymes throughout the ER and Golgi (Figure 1.10D), as opposed to soluble Env which is released into the lumen [275, 276]. In line with this hypothesis, membrane-bound CD59 also displays increased processing compared to its soluble counterpart [336].

Alternatively, it has been hypothesised that the closer proximity of the membrane to full-length constructs, particularly to gp41, would pose additional steric constraints on Env glycosylation processing [323]. In support of this hypothesis, Panico *et al.* observed exclusively oligomannose-type glycans at the membrane-proximal N88 site on virion-derived $gp120_{BaL}$ [327]. A caveat of all the above studies is that full-length constructs do not contain the SOSIP.664 stabilising mutations, which may influence glycosylation via a separate mechanism, for example, the speed at which Env transits through the ER.

1.7.7 Signal peptide

Like many membrane proteins, immature Env contains a signal peptide at the N-terminus responsible for directing the nascent peptide to the ER, which is subsequently cleaved off prior to transport of the maturing peptide through the ER and Golgi (Figure 1.10B). The signal peptide strongly influences the processing of Env as it transits through the ER, impacting factors such as trafficking to the ER, the rate of signal peptide cleavage, and the

retention time of Env within the ER [109, 337-341]. Natural variation exists within the signal peptides of Env. For example, TF viral isolates often over represent His at position 12 of the signal peptide, a signature not usually observed in chronic phase viral isolates [341]. Given the influence of the signal peptide on the molecular biology of Env, the laboratory of Fauci sought to determine the impact of the natural variation observed between the signal peptides isolated from TF viruses and from a chronically infected patient on the glycosylation, structure and antigenicity of Env. By creating four constructs containing either the gp120 sequence isolated from a TF or chronic virus, in combination with either their natural signal peptide, or that of the other virus (Figure 1.10B), Yolitz *et al.* assessed glycosylation and antigenicity by binding to various lectins and bnAbs [325]. They found that despite gp120 sequences encoding for the same mature protein, the presence of a signal peptide from the TF viral isolate resulted in an increase in oligomannose-type glycosylation (as judged by binding to the *Narcissus pseudonarcissus* lectin and bnAb 2G12, specific to α-linked mannose residues), while the presence of a signal peptide from chronic-stage virus increased the amount of complex-type glycosylation (as judged by binding to *Ricinus communis* lectin which preferentially binds structures terminating in galactose).

The results of Yolitz *et al.* are in line with the glycan analyses of Go *et al.*, who observed higher levels of oligomannose-type glycosylation on TF viruses than those isolated from chronically infected patients [318]. However, a similar comparison of SOSIP.664 trimers isolated from early and late time points by the Paulson laboratory showed very similar glycan processing [335]. Nevertheless, the findings of Yolitz *et al.*, implicate the signal peptide, a domain that does not appear in the mature protein, as a regulator of glycosylation processing.

1.7.8 Occupancy

Potential N-glycosylation site occupancy has recently emerged as an important aspect of Env glycosylation. The Paulson laboratory employed site-specific glycan classification analysis to

BG505 SOSIP.664 trimers to assess PNGS occupancy [274]. They found that overall occupancy was very high: all but four of the 28 PNGS (V1/V2 and gp41 sites N190, N197, N618, and N625) were > 90% occupied and none were < 50% occupied [274]. Extension of their method to native-like trimers derived from different strains (JR-FL, B41, CRF02_AG_250, 327c, and PC64) confirmed that the V1/V2 and gp41 sites were most susceptible to under-occupancy [275, 335]. Reports concerning occupancy of gp41 sites, particularly N625, seem to vary, perhaps due to methodological or sample variation. For example, Guttman *et al.* have reported occupancy < 20% at this site in BG505 and KNH1144 SOSIP.664 trimers [342] which is also compatible with the known epitope of the 35O22 bnAb [210].

Interestingly, virally-derived and membrane-bound Env trimers are generally more occupied than their SOSIP.664 counterparts [275, 276, 335]. Although the mechanism through which PNGS under-occupancy arises is not yet understood, it was hypothesised that the codon-optimisation process, used to increase the yield of recombinant SOSIP.664 trimers, may increase the rate of protein translation and folding and reduce the chance of an N-glycan being attached [276]. This may be particularly true of glycans in the V1/V2 region where several PNGS are in close proximity.

1.7.9 O-linked glycosylation

The vast majority of HIV-1 glycan analyses have focused on the numerous N-linked glycans on Env. There are, however, a handful of studies reporting O-glycosylation at the T499 site in various recombinant Env proteins [309, 315, 343]. Behrens *et al.* analysed the O-glycosylation present at this site on monomeric gp120, pseudotrimers and SOSIP.664 trimers [122]. They reported that the O-glycan occupancy of the T499 site was highest for monomeric gp120, at ~ 7%, compared to ~ 1% for the pseudotrimers. The presence of O-linked glycans in the BG505 SOSIP.664 trimers was negligible (~ 0.2%). The T499 site is located close to the C-terminus of gp120, thus the presence of gp41 may impose steric

constraints on the relevant glycosyltransferases. Furthermore, structures of the native-like BG505 SOSIP.664 soluble trimer and the near full-length JR-FL membrane-bound trimer show that the T499 site is not surface exposed when gp41 is present [135-137]. Therefore, an O-glycan could not be accommodated within the gp120-gp41 interface in the context of the native trimer. In line with this, $gp120_{BaL}$ derived from infectious virions lacked any detectable O-glycosylation [344].

1.7.10 Considerations for immunogen manufacture

All of the above factors influence the glycosylation processing of Env to varying degrees. Although the extent to which Env immunogens must mimic the native viral envelope spike is not yet known, there are nevertheless important considerations for immunogen design and manufacture. In particular, expression cell-type and PNGS under-occupancy may yet emerge as crucial aspects in the cGMP production of Env immunogens, as under-occupancy may introduce unfavourable, immunodominant autologous nAb epitopes. This is particularly relevant in light of Wagh *et al.*, observing increased neutralisation breadth in individuals infected with TF viruses with more intact glycan shields [204].

1.8 Beyond antigenic mimicry

Native-like trimers are excellent structural mimics of viral Env, display similar glycosylation profiles, and are capable of binding bnAbs. Yet immunogenicity studies with native-like trimers have only routinely elicited autologous Tier-2 nAbs and heterologous Tier-1 nAbs [228, 237-242]. Weakly neutralising heterologous Tier-2 antibodies have also been reported in rabbits [225]. With the exception of cows (whose antibodies naturally contain very long HCDR3), the potent bnAbs required for a protective vaccine have not yet been generated via immunisation with native-like trimers [188]. The elicitation of bnAbs against the envelope spike will inevitably be challenging. Broadly neutralising epitopes are broad due to their conserved nature, though conservation generally correlates with poor immunogenicity as

otherwise the epitope would have been selected against by immune pressure. In this section, some of the strategies aimed at increasing the immunogenicity of this target are addressed.

1.8.1 Mimicry of glycan epitopes: chemical approaches

The first of many glycan-binding bnAbs to be discovered was 2G12. To date, this the only bnAb identified to exclusively bind glycans. The 2G12 antibody recognises clusters of α1-2-linked mannose motifs present on the $Man_{8-9}GlcNAc_2$ structures of the IMP (Figure 1.8) [247-253]. Glycan based vaccine strategies have thus far centred on the design of immunogens that mimic the 2G12 epitope. Several groups have created immunogens displaying multivalent, chemically synthesised oligomannose-type glycans attached to various scaffolds, including: carbohydrate, cholic acid, cyclic peptide, dendrimer, DNA, and gold (reviewed by Wang, 2013 [345]). Although many of these constructs are able to bind 2G12, and could elicit antibodies with specificity to oligomannose-type glycans, none have induced antibodies capable of neutralising HIV-1. This may be due to the inability of these antibodies to recognise clusters of oligomannose-type glycans as they are presented on Env, perhaps due to their non-domain-exchanged structures. Domain-exchange (at least for 2G12) is a requirement for HIV-1 neutralisation [346].

The mimicry of other glycan-targeting bnAb epitopes has been explored with the design of peptide-based immunogens. Glycopeptides based on the V1/V2 region, containing the N156 and N160 PNGS, are capable of binding both apex targeting bnAbs and their unmutated common ancestors [272, 347, 348]. Similarly, glycopeptides based on the V3 region, containing the N332 PNGS, bound outer domain targeting bnAbs such as PGT128 [349]. In immunogenicity studies, V3 glycopeptides elicited antibodies capable of binding gp120 but unable to neutralise virus [349, 350]. The failure of antigenic mimics of bnAb epitopes to induce bnAbs further highlights the complicated relationship between the antigenicity and immunogenicity of these epitopes.

1.8.2 Mimicry of glycan epitopes: biological approaches

A potential failure of the above strategies may be that, while the oligomanose-type glycans of the HIV-1 glycan are not typically observed on mammalian glycoproteins, they are still fundamentally 'self' structures. As noted above, this poses a challenge for the development of glycan-targeting bnAbs, as cross-reactive B cells are likely to be deemed autoreactive and will be negatively selected for. The fact that bnAbs generally do not arise until after several years of infection could support the notion that they are operating at the edge of immunological tolerance, or may simply reflect the long maturation process. However, tolerance to self-structures can be broken. For example, in some cases, infection with *C. jejuni* elicits an antibody response against its bacterial LOS that is able to bind nearly identical structures on the gangliosides of the peripheral nervous system (Figure 1.12A) [270]. The resulting autoimmune disease, Guillain-Barré syndrome, provides proof-of-concept that immunological tolerance to 'self' glycans can be broken by exposure to micro-organisms bearing similar structures [351].

It was hypothesised that 2G12 may have originally evolved to recognise high-mannose structures on pathogens other than HIV-1 [351]. Scanlan explored this further by assessing the reactivity of 2G12 to various yeast species [268]. Yeast typically display high-mannose structures that comprise a α1-6-linked mannose backbone from which branch repetitive α1-2-linked mannose motifs (i.e. a key feature of the 2G12 epitope), sometimes capped with α1-3-linked mannose residues. Of the yeast species tested, 2G12 was able to bind to two, binding *Candida albicans* with similar affinities to that of gp120 [268, 352]. This interaction was glycan dependent as binding was inhibited by D-fructose, a known ligand for the 2G12 antigen binding site. Additionally, a non-glycan-dependent control bnAb was unable to bind any of the yeasts. Thus, the 2G12 antibody may equally well be described as an anti-*Candida* antibody with cross-reactivity to HIV-1 [268].

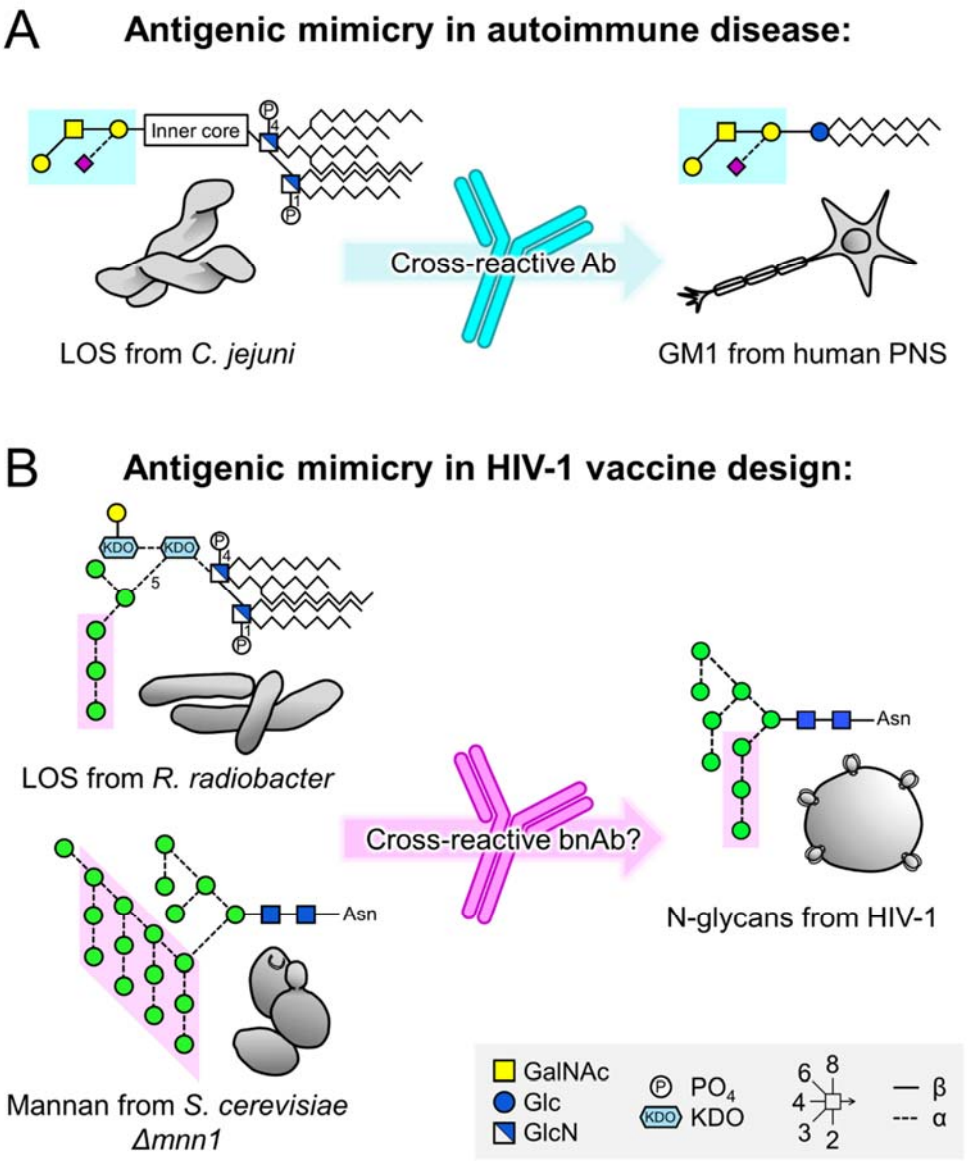


Figure 1.12. Antigenic mimicry in autoimmune disease and HIV-1 vaccine design. A) Antigenic mimicry of lipooligosaccharides (LOS) from *Campylobacter jejuni* causes Guillain-Barré syndrome, an autoimmune response against the GM1 ganglioside in the peripheral nervous system (PNS). B) LOS from *Rhizobium radiobacter* Rv3 and *Saccharomyces cerevisiae* deficient in the *mnn1* gene display glycan structures terminating in α1-2-linked mannose residues, mimicking the 2G12 epitope on HIV-1, highlighted in magenta. GalNAc, *N*-Acetylgalactosamine; Glc, gluscose; GlcN, glucosamine; KDO, 2-keto-3-deoxy-D-manno-octulosonic acid. Figure adapted from Scanlan *et al.*, 2007 [351].

Scanlan went on to formally elucidate the relationship between the antigenicity and immunogenicity of yeast glycosylation [353]. *Saccharomyces cerevisiae* deficient in *mnn1*, the gene responsible for the α1-3-linked mannose caps, displays mannans terminating in the α1-2-

linked mannose motif (Figure 1.12). Immunisation of rabbits with *S. cerevisiae* Δ*mnn1* elicited antibodies with similar glycan specificities to 2G12. Furthermore, the sera displayed weakly neutralising activity against HIV-1 [353]. The Doms laboratory also investigated using yeast as an HIV-1 immunogen. By knocking-out three genes they produced *S. cerevisiae* displaying predominantly $Man_8GlcNAc_2$ glycan structures. Although these yeasts now displayed a 'self' glycan structure, immunisation of rabbits elicited oligomannose-specific antibodies, capable of binding monomeric gp120, but unable to neutralise the virus [354]. As mentioned previously, an explanation for this apparent contradictory result may be that while the elicited antibodies were capable of recognising isolated oligomannose-type glycans, they were not able to recognise densely packed clusters of oligomannose-type glycans as they are presented on Env. Thus, a logical next step may be to present immunogenic yeast glycan structures in the context of Env, for instance, by expressing Env in yeast (Chapter 8). The resulting immunogens should display clusters of 'non-self' glycans capable of breaking immunological tolerance to the clusters of oligomannose-type glycans present on Env.

A similar phenomenon has been observed by the Pantophlet laboratory. As per *S. cerevisiae* Δ*mnn1* mannans, the LOS of *Rhizobium radiobacter* Rv3 bacteria contain a glycan motif analogous to the 2G12 epitope on HIV-1 (Figure 1.12B) [355, 356]. Immunisation of mice with *R. radiobacter* Rv3 also generated antibodies capable of weakly cross-reacting with HIV-1 [355, 357], representing another candidate micro-organism to investigate for a potential HIV-1 glycan-based vaccine.

1.8.3 Mimicking bnAb development in natural infection

The development of bnAbs in a subset of infected individuals typically only occurs after a few years of infection. Neutralising antibodies, however, arise early in infection, and exert a considerable selection pressure on the virus. Repeated rounds of viral escape and antibody evolution ultimately drive the development of nAb breadth [80, 198, 255]. The discovery that native-like trimers can induce autologous Tier-2 nAbs could therefore be an important

first step towards development of bnAbs. Accordingly, one line of research is to attempt to mimic the development of bnAbs in natural infection by immunising with a longitudinal sequence of trimers based on the Envs from an infected individual who went on to generate bnAbs. An in-depth understanding of viral and antibody co-evolution throughout the course of natural infection will no doubt prove invaluable in aiding the design of such immunisation regimens [80, 81, 291, 335, 358].

An alternative, albeit closely linked, approach is to design immunogens specifically targeted to initiate bnAb development. As noted previously, the development of bnAbs requires the activation of B cell lineages expressing gl-bnAbs, which are typically very poor at binding Env trimers [190-193]. This approach therefore involves the modification of immunogens in order to better engage gl-bnAbs [359], usually by the deletion of PNGS and variable loops to remove the steric occlusions which prevent gl-bnAb binding [360-362]. The laboratories of Schief and Nussenzweig have reported that BG505 SOSIP trimers containing, among other mutations, N133 and N137 PNGS deletions, are able to bind a germline-reverted version of the outer domain glycan-targeting bnAb, PGT121 [363]. Immunisation of mice engineered to express the predicted germline of PGT121 with this trimer, prior to boosting with a sequence of immunogens containing decreasing modifications, was able to elicit heterologous Tier-2 neutralising responses [364]. This provides proof of concept that immunisation with specifically designed immunogens can initiate bnAb development. There is, however, an apparent conflict between the need to delete PNGS sites in order to initiate bnAb development, and the recent observation that bnAbs were more likely to develop in individuals infected with isolates containing more intact glycan shields [204].

1.9 The impact of sequence diversity of the HIV-1 glycan shield

The envelope spike of HIV-1 is the sole target for bnAbs produced during infection, and it is widely acknowledged that a recombinant Env mimic will form part of a successful HIV-1

vaccine [365]. Accordingly, the development of stable, recombinant Env mimics has been a major research focus [215]. The elicitation of bnAbs will likely be challenging, due to the extreme sequence diversity of Env and the presence of an immunologically 'self' glycan shield. However, the discovery that many of the most potent bnAbs incorporate glycans into their epitopes highlights the glycan shield as a potential vaccine target. As such, there is a considerable requirement to characterise the glycan epitopes on both native, viral Env, and candidate immunogens.

To this end, Chapter 3 of this thesis builds on work by Behrens *et al.*, validating the use of a semi-quantitative mass spectrometry workflow to analyse the site-specific N-glycosylation of Env [132, 366]. Furthermore, it assesses whether bnAb affinity purification and the incorporation of stabilising mutations can bias the glycosylation of recombinant Env mimics.

Chapter 4 explores the variation in the glycosylation of candidate immunogens derived from several genotypes, in an attempt to establish the conservation of features across multiple clades of HIV-1.

Chapter 5 probes the contribution of individual glycan sites to the formation of bnAb epitopes, through the analysis of glycan site knockout and knock-in mutants.

Chapter 6 revisits the impact of bnAbs on the glycosylation of Env, to determine if co-expression of recombinant Env with a glycan targeting bnAb can alter the glycosylation profile.

The extent to which a candidate Env immunogen mimics the glycosylation of sequence-matched Env derived from infectious virions is defined in Chapter 7.

Lastly, the use of the yeast glycosylation pathway as a tool to alter the immunogenicity of the Env glycan shield is explored in Chapter 8.

2. Materials and methods

2.1 DNA plasmids and mutagenesis

2.1.1 Plasmid constructs

The pPPI4 expression plasmid encoding His-tagged BG505 SOSIP.664 and the pcDNA3.1 expression plasmid encoding Furin were provided by Prof. Ian Wilson (Scripps Research, La Jolla, USA). The 2G12, PG9, PGT121 and PGT135 light and heavy chain antibody expression plasmids were provided by Prof. Dennis Burton, the PGT145 and PGT151 plasmids by Dr Torben Schiffner, and the VRC01 and VRC02 plasmids from Prof. Rich Wyatt (all Scripps Research, La Jolla, USA). The PG16 plasmids were provided by Dr Katie Doores (King's College London, London, UK). The trastuzumab plasmids were provided by Prof. Mark Howarth (University of Oxford, Oxford UK). All DNA preps were performed using the Miniprep or Plasmid Maxi Kits (Qiagen) according to the manufacturer's instructions.

2.1.2 Site-directed mutagenesis

The pPPI4-BG505 SOSIP.664 construct was used as template DNA for the site-directed mutagenesis of N-linked glycan sites. Mutants were created by changing the asparagine within the N-glycosylation sequence (Asn-X-Ser/Thr, where X is any amino acid except proline) to alanine, using the QuikChange Lightening Mutagenesis kit (Agilent) according to the manufacturer's instructions. The primers used are listed in Appendix A. Mutations were verified by DNA sequencing.

2.2 Recombinant protein expression and purification

A number of glycoproteins were analysed throughout this thesis, many of which were kindly supplied by other laboratories. The construct, expression cell line, purification method, and laboratory of origin (where applicable) is summarised in Table 2.1.

Table 2.1. Summary of analysed glycoproteins.

Construct	Cell	Purification	Lab	Chapter
B12 antibody	CHO	Protein A	a	3
BG505 SOSIP.664 (x3)	HEK	Ni^{2+}, PGT151	-	3, 5, 7
CZA97 SOSIP.664 (x3)	CHO	2G12, SEC	b	3
CZA97 SOSIP.664 (x2)	CHO	PGT145	b	3, 4
CZA97 SOSIP.664 (x3)	CHO	PGT151	b	3
BG505 SOSIP.664	HEK	PGT145	c	3, 5
BG505 SOSIP.v4	HEK	PGT145	b	3
BG505 SOSIP.v5	HEK	PGT145	c	3, 5
BG505 SOSIP.v6	HEK	PGT145	c	3
B41 SOSIP.v4	HEK	PGT145	b	4
AMC008 SOSIP.v6	HEK	PGT145	c	4
AMC011 SOSIP.v5	HEK	PGT145	c	4
AMC016 SOSIP.v4	HEK	PGT145	c	4
BG505 NFL TD+	HEK	GNL, SEC	d	4
ZM197M NFL TD+	HEK	GNL, SEC	d	4
001428 NFL TD+	HEK	GNL, SEC	d	4
16055 NFL TD+	HEK	GNL, SEC	d	4
JR-FL NFL TD+	HEK	GNL, SEC	d	4
BG505 SOSIP.664	HEK	Ni^{2+}, PGT145	-	5
BG505 SOSIP.664 ΔN295	HEK	Ni^{2+}, PGT145	-	5
BG505 SOSIP.664 ΔN339	HEK	Ni^{2+}, PGT145	-	5
BG505 SOSIP.664 ΔN386	HEK	Ni^{2+}, PGT145	-	5
BG505 SOSIP.664 ΔN392	HEK	Ni^{2+}, PGT145	-	5
BG505 SOSIP.664 ΔN411	HEK	Ni^{2+}, PGT145	-	5
BG505 SOSIP.664 ΔN448	HEK	Ni^{2+}, PGT145	-	5
BG505 SOSIP.664 ΔN156	HEK	Ni^{2+}, PGT151	-	5
BG505 SOSIP.664 ΔN197	HEK	Ni^{2+}, PGT151	-	5
BG505 SOSIP.664 ΔN234	HEK	Ni^{2+}, PGT151	-	5
BG505 SOSIP.664 ΔN332	HEK	PGT145	c	5
BG505 SOSIP.664 +N241	HEK	PGT145	c	5
BG505 SOSIP.664 +N289	HEK	PGT145	c	5
BG505 SOSIP.664 +N241 +N289	HEK	PGT145	c	5
BG505 SOSIP.v5 +N289	HEK	PGT145	c	5
BG505 SOSIP.v5 +N241 +N289	HEK	PGT145	c	5
BG505 SOSIP.664 +2G12	HEK	Protein A	-	6
BG505 SOSIP.664 +PGT121	HEK	Protein A	-	6
BG505 SOSIP.664 +PGT135	HEK	Protein A	-	6
BG505 SOSIP.664 +VRC01	HEK	Protein A	-	6
BG505 SOSIP.664 +VRC02	HEK	Protein A	-	6
BG505 SOSIP.664 +PG16	HEK	Protein A	-	6
BG505 SOSIP.664 +PGT145	HEK	Protein A	-	6

BG505 SOSIP.664 +PGT151	HEK	Protein A	-	6
BG505 SOSIP.664	HEK	Ni^{2+}	-	6
BG505 SOSIP.664 + Trastuzumab	HEK	Ni^{2+}	-	6
BG505.T332N-LAI gp120 and gp41	A66-R5	HPLC	e	7
BG505 NFL	*K. lactis*	Ni^{2+}	f	8

CHO, Chinese hamster ovary; HEK, human embryonic kidney; *K. lactis*, *Kluyveromyces lactis*; SEC, size exclusion chromatography; GNL, *Galanthus nivalis* lectin-agarose affinity chromatography; HPLC, high pressure liquid chromatography.

a. Dr Lars Hangartner, Scripps Research, La Jolla, USA
b. Prof. John Moore, Weil Cornell Medical College, New York, USA
c. Prof. Rogier Sanders/Prof. Marit van Gils, Amsterdam University Medical Centre, Amsterdam, the Netherlands
d. Prof. Rich Wyatt, Scripps Research, La Jolla, USA
e. Dr Jeffrey Lifson, Frederick National Laboratory for Cancer Research, Frederick, USA
f. Stable transformed cell line provided by Dr Mehul Ganatra/Dr Chris Taron, Protein Expression & Modification Research division, New England Biolabs, USA.

2.2.1 Recombinant protein expression in HEK cells

Proteins were expressed transiently in HEK (293F) suspension cell cultures in FreeStyle Expression Medium (ThermoFisher Scientific). For transfection, 200 µL of FreeStyle Max transfection reagent (ThermoFisher Scientific) and 200 µg DNA were incubated in 10 mL of serum-free OptiPro medium (ThermoFisher Scientific) for 15 min. This was then added to a 200 mL culture at a concentration of 1×10^6 cells/mL. The cells were maintained at 37 °C, 8% CO_2 with shaking at 125 rpm. Cultures were incubated for 5 – 7 days prior to harvesting. Supernatant was clarified by centrifugation and sterile filtered through a 0.22 µm membrane (Merck Millipore). For BG505 SOSIP.664 expression, a 4:1 ratio of Env:furin DNA was used. For antibodies, a 2:1 ratio of light:heavy chain plasmids was used.

2.2.2 Ni^{2+} affinity purification

His-tagged Env proteins were purified using a 5 mL HisTrap column (GE Healthcare). Column equilibration and washing was performed using a 50 mM $NaPO_4$, 0.3 M NaCl, 10 mM imidazole buffer. Column loading was performed at a flow rate of 1 mL/min. Env proteins were eluted in 50 mM $NaPO_4$, 0.3 M NaCl, 150 mM imidazole. Eluted proteins were immediately buffer exchanged and concentrated using 50 kDa cut-off Vivaspin columns (Sartorius).

2.2.3 Protein A affinity purification

Antibodies used for the generation of affinity columns and immunoassays were purified using a 5 mL HiTrap Protein A affinity column (GE Healthcare). Column equilibration and washing was performed using a 20 mM $NaPO_4$, pH 7.5 buffer. Column loading was performed at a flow rate of 1 mL/min. The antibody was eluted using a 0.1 M citric acid buffer, pH 3. The pH was immediately neutralised by the addition of 1 M Tris-HCl buffer, pH 9.0, and immediately buffer-exchanged, as described previously.

2.2.4 Broadly neutralising antibody affinity purification

Following Ni^{2+} purification, recombinant Env constructs were purified by antibody affinity chromatography. Affinity columns were made by coupling protein A-purified antibodies to cyanogen bromide CNBr-Activated Sepharose 4B resin (GE Healthcare). Briefly, resin was washed with 1 mM HCl, followed by coupling buffer: 100 mM $NaHCO_3$, 0.5 M NaCl, pH8.0. The antibody was bound to the resin overnight at 4 °C in coupling buffer, at a ratio of 2.4 mg antibody per 0.5 g resin (dry weight). Unbound antibody was removed by washing the resin with coupling buffer, and the remaining active groups were blocked with 200 mM glycine, pH 8, overnight at 4 °C. Resin was packed in a Glass Econo-Column (2.5 x 10 cm; Bio-Rad).

For purification, a 20 mM Tris-HCl, 500 mM NaCl, pH 8.0 buffer was used for column equilibration and washing. Column loading was performed at a flow rate of 1 mL/min. Env proteins were eluted with 3 M $MgCl_2$, pH 7.2. Eluted proteins were immediately buffer exchanged into 10 mM Tris-HCl, 75 mM NaCl, pH 8.0 and concentrated using 50 kDa cut-off Vivaspin columns. Columns were washed extensively with three alternating cycles of 100 mM NaAc, 0.5 M NaCl, pH4.5 and 100 mM Tris-HCl, 0.5 M NaCl, pH 8.0, between purifications of different trimer constructs.

2.2.5 Co-purification of antibody-Env complexes

Recombinant Envs co-expressed with bnAbs were co-purified by Protein A affinity purification, as described in Section 2.2.3. Env trimers were separated from the antibody by resolution on a reducing SDS-PAGE gel, as described below.

2.3 Analysis of overall glycan profiles and compositions

2.3.1 In-gel release of N-linked glycans

N-linked glycans were enzymatically released from Env glycoproteins via in-gel digestion. Purified proteins were resolved by pre-cast NuPAGE 4-12% Bis-Tris polyacrylamide gel (in the presence of dithiothreitol [DTT] for reducing gels; Invitrogen) and stained with Coomassie Blue. Following destaining, the bands containing Env were excised and washed five times, alternating between acetonitrile and water. The gel bands were then digested with PNGase F (New England Biolabs) overnight at 37 °C. Released N-linked glycans were eluted with water, prior to drying in a SpeedVac.

2.3.2 Fluorescent labelling of N-linked glycans

Aliquots of released glycans were fluorescently labelled with procainamide or 2-aminobenzoic acid (2-AA) in the following solutions: 110 mg/mL procainamide hydrochloride (Abcam), 60 mg/mL sodium cyanoborohydride, dissolved in 30% dimethyl sulfoxide, 70% acetic acid; or 30 mg/mL 2-AA, 45 mg/mL sodium cyanoborohydride dissolved in a solution of 4% (w/v) sodium acetate trihydrate, 2% (w/v) boric acid in methanol. An 80 – 100 μL aliquot of labelling mixture was added to the glycan sample (dissolved in 30 μL water), vortexed and incubated at 65 °C for 4 hours or 80 °C for 1 hour, for procainamide or 2-AA, respectively. Excess label and PNGase F were removed using Spe-ed Amide 2 cartridges (Applied Separations).

2.3.3 HILIC-UPLC analysis

Glycans were analysed on a Waters Acquity H-Class UPLC instrument with a Glycan BEH Amide column (2.1 mm x 100 mm, 1.7 μm; Waters) and the following gradient: time (t) = 0: 22% A, 78% B (flow rate = 0.5 mL/min); t = 38.5: 44.1% A, 55.9% B (0.5 mL/min); t = 39.5: 100% A, 0% B (0.25 mL/min); t = 44.5: 100% A, 0% B (0.25 mL/min); t = 46.5: 22% A, 78% B (0.5 mL/min), where solvent A was 50 mM ammonium formate, pH 4.4 (Ludger) and B was acetonitrile. Fluorescence was measured at an excitation wavelength of 310 nm and an emission wavelength of 370 nm for procainamide, or 250 nm and 428 nm for 2-AA. Data were processed using Empower 3 software (Waters).

2.3.4 Glycosidase digestion of released glycans

The relative abundances of N-glycan structures were determined by digesting 15 μL of fluorescently labelled glycans with the following glycosidases overnight at 37 °C: Endo H, α2-3, 6, 8 neuraminidase, α2-3 neuraminidase, β1-4 galactosidase, β-*N*-Acetylglucosaminidase (all New England Biolabs), and α-L-fucosidase (Sigma Aldrich). Glycans were extracted using a PVDF protein-binding membrane (Merck Millipore). Digested glycans were analysed by HILIC-UPLC, as before. Glycans were quantified by integrating the peaks before and after glycosidase digestion, following normalisation of the chromatograms.

2.3.5 Ion mobility mass spectrometry analysis of released N-linked glycans

To generate glycan libraries in order to aid in the subsequent assignment of glycopeptides, a separate, unlabelled aliquot of PNGase F-released glycans were subjected to ion mobility electrospray ionisation mass spectrometry (IM-ESI MS). The glycan sample was cleaned with a Nafion membrane for approximately 2 hours, and then dissolved in a solution of methanol-water (1:1 v/v), containing a trace amount of ammonium phosphate to promote phosphate adduct formation. Glycans were analysed by nano-electrospray ionisation mass spectrometry

with direct infusion using a Synapt G2Si instrument (Waters) with the following settings: capillary voltage, 0.8 – 1.0 kV; sample cone, 100 V; extraction cone, 25 V; cone gas, 40 L/h; source temperature, 150 °C; trap collision voltage, 4 – 160 V; transfer collision voltage, 4 V; trap direct current bias, 35 – 65 V; IMS wave velocity, 450 m/s; IMS wave height, 40 V; trap gas flow, 2 mL/min; IMS gas flow, 80 mL/min. Data were acquired and processed with MassLynx v4.1 and Driftscope version 2.8 software (Waters).

2.4 Site-specific glycan analysis

2.4.1 Reduction, alkylation and proteolytic digestion

Approximately 50 – 100 μg of Env proteins were denatured, reduced, and alkylated by sequential 1 hour incubations at room temperature in the following solutions: 50 mM Tris-HCl, pH 8.0 buffer containing 6 M urea and 5 mM DTT, followed by the addition of 20 mM iodacetamide (IAA) in the dark, and then additional DTT (20 mM), to eliminate residual IAA. The alkylated trimers were buffer exchanged into 50 mM Tris-HCl, pH 8.0 using Vivaspin 3 kDa cut-off columns (GE Healthcare) and aliquots were digested with either trypsin or chymotrypsin (Mass Spectrometry Grade, Promega) at a ratio of 1:30 (w/w) overnight at 37 ºC. Reaction mixtures were dried and peptides and glycopeptides were extracted using C18 ZipTip (Merck Millipore), following the manufacturer's protocol.

2.4.2 Liquid chromatography mass spectrometry analysis of glycopeptides

Eluted glycopeptides were dried again and re-suspended in 0.1% formic acid prior to mass spectrometric analysis. Aliquots of intact glycopeptides were analysed by nano liquid chromatography-electrospray ionisation mass spectrometry (LC-ESI MS) using 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 × 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 °C. HCD collision energy was set to 50%, appropriate for the fragmentation of glycopeptide ions. Data analysis and glycopeptide identification were performed using Byonic™ (Version 2.7) and Byologic® software (Version 2.3; Protein Metrics Inc.), as follows. First, the mass spectrometry data and protein sequence (Appendix E) were processed with Byonic™ allowing for a precursor mass tolerance of 4 ppm and a fragment mass tolerance of 10 ppm. Cleavage sites were specified as the C-terminus of Arg and Lys residues for tryptic digests, and the C-terminus of Tyr, Trp, and Phe residues for chymotryptic digests. Semi-specific digestion and up to three missed cleavages were tolerated. A glycan library (generated by IM-ESI MS, Section 2.3.5; Appendix D) and the peptide modifications detailed in Table 2.2 were used to search the data. The glycopeptide fragmentation data were then evaluated manually for each identified glycopeptide using Byologic®. The peptide was validated as true-positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the identified glycan structure (Appendix B). The intensities of the extracted ion chromatograms (XIC) over all charge states for each true-positive glycopeptide (with the same amino acid sequence) were calculated as a percentage, to determine the relative amounts of each glycoform at each site.

Table 2.2. Peptide modifications for site-specific glycan analysis.

Modification	MW (Da)	Frequency[a]
Carbamidomethylation of Cys	+57.021464	Fixed
Oxidation of Met	+15.994915	Rare
Glu→pryo-Glu at N-terminus Glu	-18.010565	Rare
Gln→pryo-Glu at N-terminus Gln	-17.026549	Rare
N-glycan at N-glycan sequon	(e.g. Appendix D)	Common

a. A maximum of two common and one rare modification were permitted per peptide.

2.4.3 Glycan 'classification' analysis

In order to determine the class of glycan occupying particular sites, glycopeptides were digested with Endo H, which cleaves oligomannose- and hybrid-type glycans, leaving a single GlcNAc residue at the corresponding site. The reaction mixture was then dried completely and resuspended in a mixture containing 50 mM ammonium bicarbonate and PNGase F using only O^{18}-labelled water (Sigma Aldrich) throughout. This second reaction cleaves the remaining complex-type glycans, leaving the GlcNAc residues intact. The use of H_2O^{18} in this reaction enables complex glycan sites to be differentiated from unoccupied glycan sites as hydrolysis of the glycosidic bond by PNGase F leaves an O^{18} isotope on the resulting aspartic acid residue. The resultant peptides were purified by C18 ZipTip, as outlined above, and subjected to LC-ESI MS in a similar manner to before, but using a lower HCD energy of 27% as glycan fragmentation was not required. Data analysis was performed as above, searching for the modifications detailed in Table 2.3.

Table 2.3. Peptide modifications for glycan classification analysis.

Modification	MW (Da)	Frequency[a]
Carbamidomethylation of Cys	+57.021464	Fixed
Oxidation of Met	+15.994915	Rare
Glu→pryo-Glu at N-terminus Glu	-18.010565	Rare
Gln→pryo-Glu at N-terminus Gln	-17.026549	Rare
Deamidation of Asn (O^{18})	+2.988261	Common
GlcNAc at N-glycan sequon	+203.079373	Common

a. A maximum of four common and one rare modification were permitted per peptide.

2.5 Immunoassays

2.5.1 ELISA

High binding 96 well assay plates (Corning) were incubated with BG505 SOSIP.664 proteins (10 μg/mL in PBS) overnight at 4 °C. Plates were washed with a solution of PBS containing 0.5% Tween 20 (v/v) and blocked for 1 hour at room temperature with 5% milk in PBS + 0.5% Tween. After another wash step, the primary antibody was incubated (1:2 dilution series

with a starting concentration of 20 μg/mL) in PBS for 1 hour at room temperature. Plates were washed and an anti-human IgG conjugated to Horseradish Peroxidase (Abcam) secondary antibody was added at a 1:2'000 dilution in PBS. Plates were washed and TMB substrate solution (Thermo Fisher Scientific) was added. The reaction was stopped with sulphuric acid after 5 min and the OD 450 nm was measured.

2.5.2 Western blot

Proteins were resolved by SDS-PAGE gel and transferred onto a PVDF or nitrocellulose membrane using the Trans-Blot® Turbo™ Transfer system (BioRad). Membranes were blocked overnight at 4 °C or for 1 hour at room temperature in 5% milk in PBS + 0.5% Tween. Membranes were washed in PBS + 0.5% Tween and the primary antibody (1:1'000 – 3'000) was incubated in PBS + 0.5% Tween for 1 hour at room temperature. Membranes were washed again and incubated with the secondary antibody (goat anti-human IgG-HRP, 1:2'000; Abcam) in PBS + 0. 5% Tween for 1 hour at room temperature. Proteins of interest were visualised with Pierce ECL Western Blotting Substrate (Thermo Scientific) and images were captured with the Fuji Intelligent Dark Box ii and Image Reader 1000 Pro software.

2.5.3 Dot blot

The protocol follows that of Western blots, except that 2 μL of the solution containing protein (e.g. cell culture supernatant) were transferred onto the PVDF membrane by pipette. Membranes were allowed to dry for a few seconds and another 2 μL was added.

2.6 Model building

The model of a fully glycosylated BG505 SOSIP.664 trimer (and other native-like trimers based on the BG505 strain) used throughout this thesis was described in Behrens *et al.*, 2016 [132]. Models for all other native-like trimers were constructed as follows. Where possible, a suitable structure was obtained from the Protein Data Bank; if none were available, a homology model was generated using SWISS-MODEL (Table 2.4) [367-369]. The

predominant glycoform at each site (according to site-specific glycan data) was modelled on to the Asn residue using Coot [370, 371]. Where no site-specific glycan data were available, a $Man_5GlcNAc_2$ structure was added. Glycan structures were obtained from the following structures (PDB ID): $Man_{9\text{-}5}GlcNAc_2$, 2WAH; bi-antennary glycan, 1L6X; tri-antennary glycan, 1P8J; tetra-antennary glycan, 53TZ.

Table 2.4. Structure or reference structure used for model building.

Trimer	Structure	SWISS-MODEL reference structure
001428	-	5UM8
16055	5UM8	-
AMC008	-	6B0N
AMC011	-	6B0N
AMC016	-	6B0N
B41	6CH9	-
CZA97	-	5T3X (gp120), 5U1F (gp41)
JR-FL	-	5FUU (gp120), 5FYK (gp41)
ZM197M	-	5UM8 (gp120), 5FYJ (gp41)

PDB ID is given. Where two reference structures were used to generate one trimer, the structure used to generate each subunit is indicated in parentheses.

2.7 Yeast cultures and protein expression

2.7.1 Plasmid constructs and cloning

The pPICLIT vector (with Zeocin resistance) was provided by the laboratory of Prof. Nicole Zitzmann, University of Oxford, UK. The pJAN-S1 vector (with Nourseothricin resistance) was provided by Dr Karen Polizzi, Imperial College London, UK. Both vectors integrate into the yeast genome to generate stable transformants. For cloning in genes of interest, the pPICLIT vector was digested using the Xho1 and Xba1/Acc1 restriction sites, the pJAN-S1 vector was digested only with Xho1. The genes of interest were amplified using the primers given in Appendix A. The Xho1 site was located within the α-factor signal sequence, thus primers were designed to rebuild this sequence. Constructs were designed so that the N-terminus of the protein was flush against the Kex2 cleavage site, located downstream of the α-factor signal sequence. Amplified genes were cloned into vector using the Gibson Assembly kit (New England Biolabs), according to the manufacturer's instructions.

Constructs were verified by sequencing. Approximately 5 – 10 μg of the vector was linearised by digestion with Pme1, and resuspended in 5 – 10 μL of water, for electroporation into yeast cells.

2.7.2 *Pichia pastoris* culture conditions and preparation of cells for electroporation

A 5 mL starter culture of *Pichia pastoris* was grown overnight at 30 °C in YPD (yeast extract peptone dextrose) media with shaking at 250 rpm. Approximately 1 mL was used to inoculate 500 mL of fresh medium, which was grown overnight to an OD_{600} = 1.3 – 1.5. Cells were pelleted by centrifugation (1,500 × g for 5 minutes at 4 °C), and resuspended in 500 mL ice-cold, sterile water. The centrifugation was repeated three times, and cells resuspended as follows: 250 mL ice-cold water, 20 mL ice-cold 1 M sorbitol, 1 mL ice-cold sorbitol. The cells were kept in sorbitol, on ice prior to electroporation.

2.7.3 Generation of *Pichia pastoris* stable cell lines

An 80 μL aliquot of prepared cells was mixed with 5 – 10 μg linearised DNA and transferred to an ice-cold 0.2 cm electroporation cuvette (BioRad). Cuvettes were incubated on ice for 5 min, prior to electroporation at 2 kV, or following the manufacturer's settings. Immediately following electroporation, 1 mL ice-cold sorbitol was added to the cells, which were then transferred to a sterile 15 mL tube and incubated at 30 °C without shaking for 1 to 2 hours. The cells were then plated on YPD plates containing 1M sorbitol, 100 μg/ml Zeocin™, and incubated for 3 – 10 days at 30 °C until colonies form.

Successful transformation was additionally confirmed by PCR. Template genomic DNA was extracted from yeast colonies using the Wizard® Genomic DNA Purification Kit (Promega), following the manufacturer's instructions for yeast.

2.7.4 Protein expression in *Pichia pastoris*

A 25 mL culture of BMGY (Buffered glycerol-complex medium) medium was inoculated with transformed *P. pastoris* cells, and grown overnight to an OD_{600} = 2 – 6. This culture was

used to inoculate 1 L of BMGY media, again until an OD_{600} = 2 – 6 was reached. The cells were pelleted by centrifugation, as above, and resupended to OD_{600} = 1 in BMMY (Buffered methanol-complex medium) media, to induce expression. 100% methanol was added to a final concentration of 0.5% every 24 hours.

2.7.5 Protein expression in *Kluyveromyces lactis*

Transformed *Kluyveromyces lactis* cells were cultured in YPGal medium (yeast extract peptone galactose) at 30 °C with shaking at 250 rpm. No induction is required. Also note that the *K. lactis* transformants were plated on YCB (Yeast Carbon Base) agar plus 5 mM acetamide plates.

3. Method validation and the influences of trimer format and antibody affinity purification

The envelope spike is the sole target for bnAbs produced during infection. These bnAbs have evolved to overcome the virus' numerous defences, including a dense shield of host-derived N-linked glycans [199, 372]. While originally thought to protect the underlying protein surface from immune recognition, the discovery that many bnAbs can recognise glycan epitopes has exposed the glycan shield itself as a potential target for vaccine design [351]. Defining the composition and relative abundances of the glycans at each site of the HIV-1 Env glycan shield has therefore been a longstanding goal in HIV-1 vaccine research. Such information is invaluable in defining the precise nature of bnAb epitopes and subsequently guiding the design of immunogens.

The site-specific glycan analysis of Env is particularly challenging given the extensive heterogeneity typically displayed by glycoproteins, combined with the large number of PNGS on Env. Early glycan analyses were further limited by the difficulties associated with producing native-like trimers and/or isolating sufficient virion-derived Env. Thus, initial studies were often restricted to the analysis of released N-linked glycans, and/or performed on material unrepresentative of native, virion-associated Env [110, 123, 306-309, 311, 312, 314, 317-320]. The development of the SOSIP.664 platform combined with advances in liquid chromatography and mass spectrometry has transformed the field.

In 2016, Behrens *et al.* published a semi-quantitative, site-specific glycan analysis of the BG505 SOSIP.664 trimers [132]. The results revealed the extremes of glycan processing at individual sites on the BG505 SOSIP.664 trimers. On the one hand, large regions of densely clustered glycan sites comprised solely under-processed, oligomannose-type glycans. On the other, sites located towards the peripheries of the trimer displayed heterogeneous populations of complex-type glycans.

The methodology employed in Behrens *et al.* involved the proteolytic digestion of Env, followed by in-line liquid chromatography-electrospray ionisation mass spectrometry (LC-ESI MS) [132]. In order to aid in the assignment of glycopeptides and lower the false discovery rate, a protein-specific glycan library was generated by the ion mobility-electrospray ionisation mass spectrometry (IM-ESI MS) analysis of enzymatically released N-linked glycans. The workflow developed in Behrens *et al.* and used throughout this thesis is summarised in Figure 3.1.

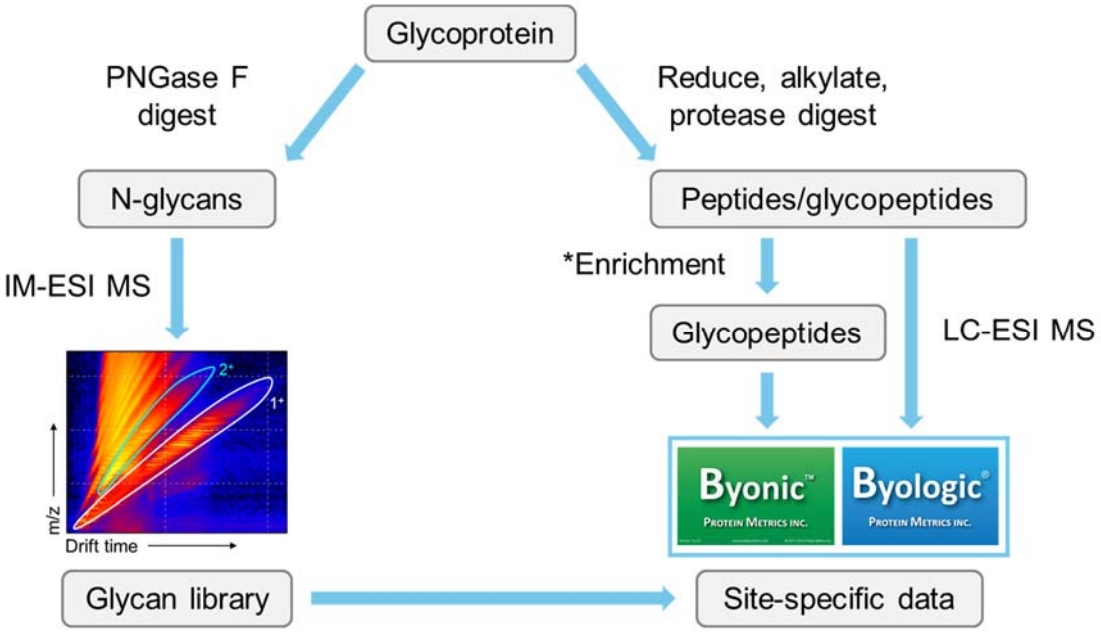


Figure 3.1. Summary of LC-ESI MS site-specific analysis workflow. A glycan library is generated by IM-ESI MS analysis of enzymatically released N-linked glycans. A separate aliquot of glycoproteins is used for LC-ESI MS analysis. Briefly, glycoproteins are denatured, reduced and alkylated prior to proteolytic digestion. In Behrens *et al.* a glycopeptide enrichment step was used, indicated with an asterisk [132]. In this thesis, both peptides and glycopeptides were extracted by C18 ZipTip. Data analysis is performed using Byonic™ and Byologic® software.

The above LC-ESI MS methodology employs higher-energy collisional dissociation (HCD) fragmentation in order to assign glycopeptides. The ions produced during HCD fragmentation of glycopeptides are predominantly B- and Y-type oxonium ions generated by protonation of glycosidic bonds (nomenclature according to Domon and Costello, 1988), and b- and y-type ions generated by fragmentation of the peptide backbone (Appendix B) [373, 374]. Byonic™ and Byologic® software is used to extract the fragmentation data and corresponding extracted ion chromatograms (XIC; ion intensities recorded as a function of retention time). Glycopeptides are manually verified by the presence of the correct b, y, and oxonium ions, and the relative abundances of individual glycan structures at a given site are then calculated from the chromatographic areas of each true-positive glycopeptide over multiple charge states.

Site-specific LC-ESI MS analysis is considered a semi-quantitative method, as it can only provide a readout of the relative abundances of a glycan structure at a given site, due to the differential ionisation efficiency of different peptides. Furthermore, ionisation biases may exist for different glycan structures within a glycopeptide. While this is unlikely to be a major source of bias, as ionisation usually occurs by protonation of the peptide backbone, negatively charged glycans, for example sialic acid, may be underestimated in positive ion mode or overestimated in negative ion mode [375, 376].

In order to validate the quantification of glycans by LC-ESI MS, Behrens *et al.* performed a parallel glycan analysis of BG505 SOSIP.664 exploiting another ionisation mode. In the second approach, tryptic glycopeptides were fractionated by reverse-phase high-performance liquid chromatography (RP-HPLC) and analysed by matrix-assisted laser desorption ionisation (MALDI) time-of-flight (TOF) mass spectrometry [132]. In contrast to electrospray ionisation, MALDI generates singly-charged ions irrespective of glycan

structure*, thus simplifying the quantification process [377]. Encouragingly, the two methodologies provided concordant results, supporting the use of the in line LC-ESI MS methodology for the analysis of Env.

While glycan quantification by mass spectrometry is subject to possible bias introduced by the ionisation properties of a glycopeptide, an alternative method is to analyse labelled glycans by hydrophilic interaction chromatography-ultra performance liquid chromatography (HILIC-UPLC). For example, fluorescently labelling the reducing end of the glycan (following enzymatic release with PNGase F), results in a 1:1 glycan-to-label stoichiometry (Appendix F). Fluorescence detection is therefore independent of the glycan structure and provides a method for the quantification of the N-linked glycans present on the glycoprotein that is devoid of some of the potential limitations of mass spectrometry.

This chapter will further validate the glycan analysis methods employed in both Behrens *et al.* and throughout this thesis [132]. Analysis of a glycoprotein containing only one PNGS by both HILIC-UPLC and LC-ESI MS confirms the accuracy of site-specific glycan quantification, while analysis of biological replicates establishes the variation inherent to both methodologies. The results are relevant for the data presented throughout this thesis, particularly in Chapter 5 and 6, as it gives confidence that observed differences in the relative abundance of glycans is not due to experimental variation.

A number of native-like Env trimers have been analysed throughout this thesis, many of which were produced by different collaborators (Table 2.1). However, the expression cell line and purification method used for the production of Env trimers can vary. Furthermore, many different native-like trimer constructs have been developed over recent years (Section 1.5), each combining a variety of different stabilising mutations. Thus, this chapter also sought to establish the extent to which stabilising mutations can impact the processing of

*Permethylation is usually required in order to detect negatively charged glycans (e.g. sialic acid) in positive ion mode.

the Env glycan shield, and whether purification can bias the population towards certain glycoforms.

3.1 Site-specific glycan quantification by LC-ESI MS

In order to assess the accuracy of LC-ESI MS quantification, a glycoprotein containing only one glycan site was analysed by site-specific glycan analysis and HILIC-UPLC. This enables a direct comparison between the glycan abundances reported by the quantitative HILIC-UPLC and semi-quantitative LC-ESI MS methodologies.

The bnAb, b12, was used for the comparative analysis as it contains only one PNGS, located at position N297 of the heavy chain. Enzymatically released glycans were fluorescently labelled prior to analysis by HILIC-UPLC, and glycan structures were assigned based on their susceptibility to glycosidase digests. The chromatogram was consistent with previously reported IgG1 glycosylation profiles, predominated by fucosylated bi-antennary structures (Figure 3.2A) [378]. The relative abundance of each glycan was quantified by integrating the area underneath the peak (Figure 3.2B).

A separate aliquot of the b12 glycoprotein was subject to analysis by LC-ESI MS. Although IgG is composed of two heavy and two light chains, and therefore contains two PNGS, the amino acid sequences of each of the two pairs of chains are identical; thus, the Byonic™ and Byologic® software processes it as a single PNGS. A comparison of the two glycan quantifications revealed the two methods to be in very close agreement (Figure 3.2B), indicating that the LC-ESI MS workflow is able to accurately quantify the glycan structures present on a given PNGS. Furthermore, the LC-ESI MS method was able to detect trace glycan structures, with a relative abundance of < 0.05%, which could not be assigned by HILIC-UPLC. Therefore, the LC-ESI MS workflow represents a sensitive, semi-quantitative method for the site-specific profiling of glycoproteins, though a limitation of this comparison is the low abundance of sialylated glycans present on b12.

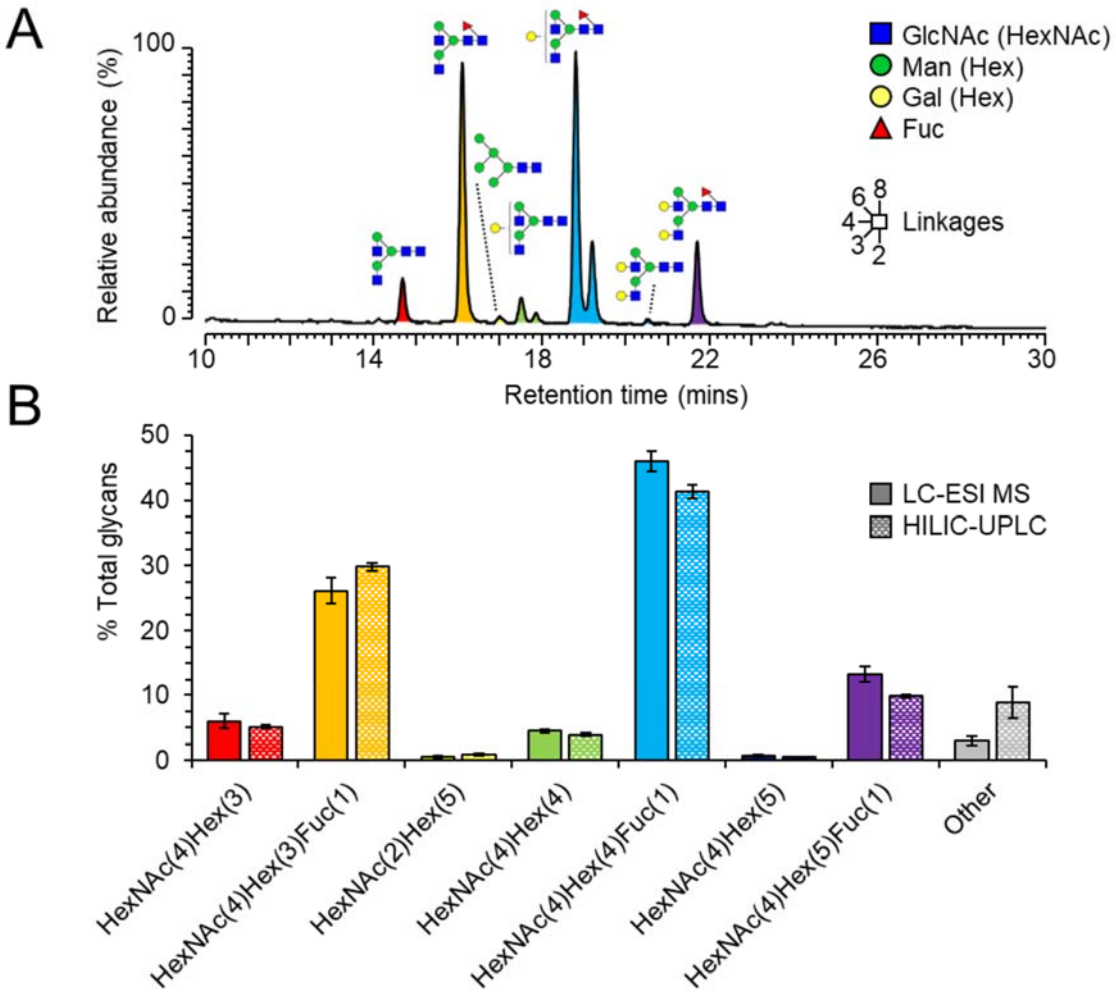


Figure 3.2. Comparison of LC-ESI MS and HILIC-UPLC glycan quantification. A) N-linked glycans from the IgG b12 expressed in CHO cells were enzymatically released and fluorescently labelled prior to HILIC-UPLC analysis. The relative abundance of glycan structures was determined by integrating the area under each peak. B) Comparison of glycan quantification by LC-ESI MS (plain) and HILIC-UPLC (hatched) methodologies. The error bars represent ±SEM for three analytical replicates. The 'other' category contains low-abundance glycans that were detected by LC-ESI MS but which could not be reliably assigned by HILIC-UPLC.

3.2 Method reproducibility based on three biological replicates

In order to determine the variation inherent to both glycan analysis methodologies, three biological replicates of BG505 SOSIP.664 trimers were analysed by LC-ESI MS and HILIC-UPLC. Pritchard *et al.* have previously reported on the reproducibility of HILIC-UPLC analysis of Env proteins expressed in parallel [310], however no such experiment has been

performed on proteins expressed several weeks apart. The reproducibility of the LC-ESI MS workflow has not been formally addressed.

Three BG505 SOSIP.664 proteins were expressed and purified by PGT151 affinity chromatography (Table 2.1). An aliquot of each protein was then subjected to HILIC-UPLC analysis. The overall glycosylation profiles of the three proteins were highly similar (Figure 3.3A). Quantification of the relative abundances of oligomannose-type glycans confirmed the minimal variation observed between the three batches (Figure 3.3B). With the exception of $Man_5GlcNAc_2$ which had an apparent outlier, the coefficient of variation was < 3% for all oligomannose-type glycans (Table 3.1). These results confirm that oligomannose-type glycan quantification by HILIC-UPLC analysis of fluorescently labelled N-linked glycans is reproducible, even when the proteins are not expressed from the same passage of cells.

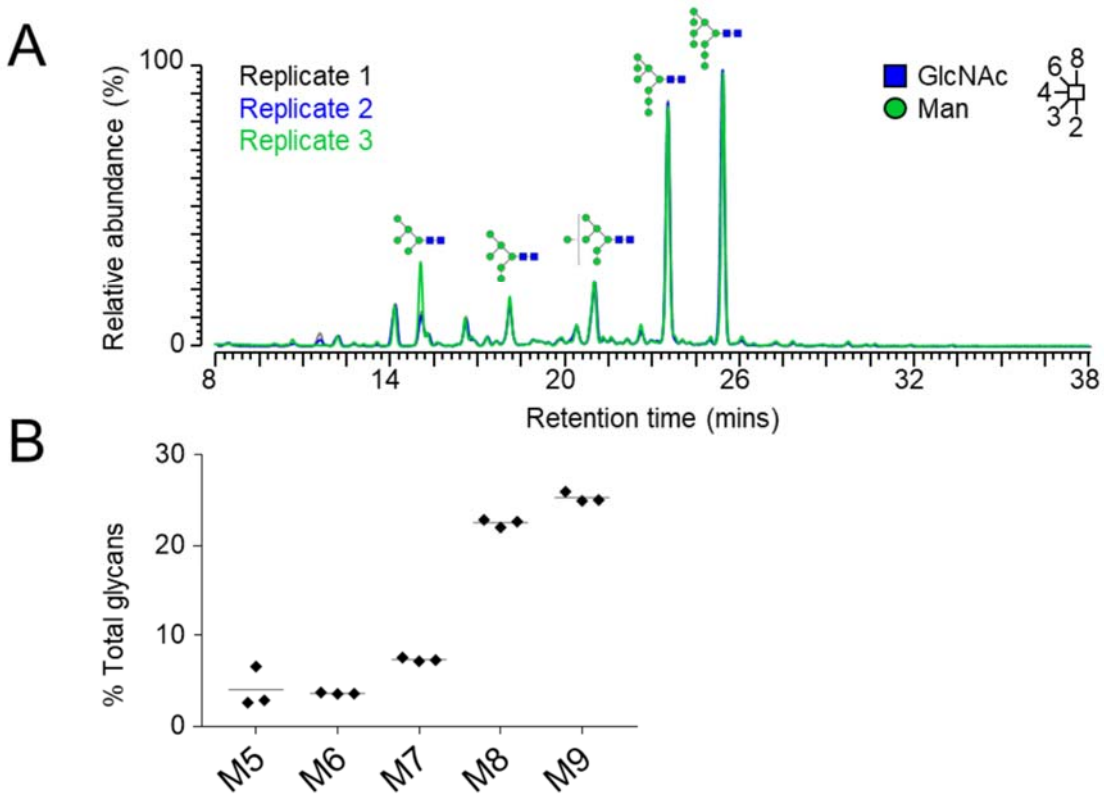


Figure 3.3. Reproducibility of HILIC-UPLC glycosylation analysis. BG505 SOSIP.664 was expressed in HEK cells and purified by PGT151 affinity chromatography, three times. A) Fluorescently labelled N-linked glycans were analysed by HILIC-UPLC. B) The relative abundances of oligomannose-type glycans were calculated by integrating the peaks before and after digestion with Endo H. M5-9, $Man_{5\text{-}9}GlcNAc_2$.

Table 3.1. Reproducibility of oligomannose quantification by HILIC-UPLC[a].

	Man5	Man6	Man7	Man8	Man9	TOTAL
Mean	4.0	3.6	7.3	22.5	25.3	63.5
SD	2.2	0.1	0.2	0.4	0.6	0.9
Coefficient of variation[b]	> 10	2.4	3.0	1.9	2.2	1.5

a. Related to Figure 3.3B.
b. The coefficient of variation is the standard deviation divided by the mean, expressed as a percentage. It is comparable to the percentage change calculated for the PNGS mutants in Chapter 5.

In order to assess the reproducibility of the LC-ESI MS workflow, the same three proteins were analysed by site-specific analysis (Figure 3.4). Glycan structures are grouped as described in Appendix C. Briefly, the classification is based on the number of mannose residues (M) within the oligomannose-type glycan series, or for complex-type glycans, by their number of antenna (A) and presence of a core fucose (F). Overall, the LC-ESI MS workflow was highly reproducible (Figure 3.4). Sites dominated by oligomannose-type glycans (e.g. N234, N262, N295, N332, N339, N363 and N386) displayed the least variation, consistent with the reproducibility of oligomannose quantification by HILIC-UPLC (Figure 3.3). The relative abundances of complex-type glycans varied slightly over the three replicates (Figure 3.4), however this effect may be exacerbated by the grouping together of multiple glycan structures.

Taken together, both the HILIC-UPLC and LC-ESI MS methodologies are very reproducible. The result is of importance for comparing the glycosylation of Env proteins, as it gives confidence that any differences observed are genuine, and not attributed to experimental error.

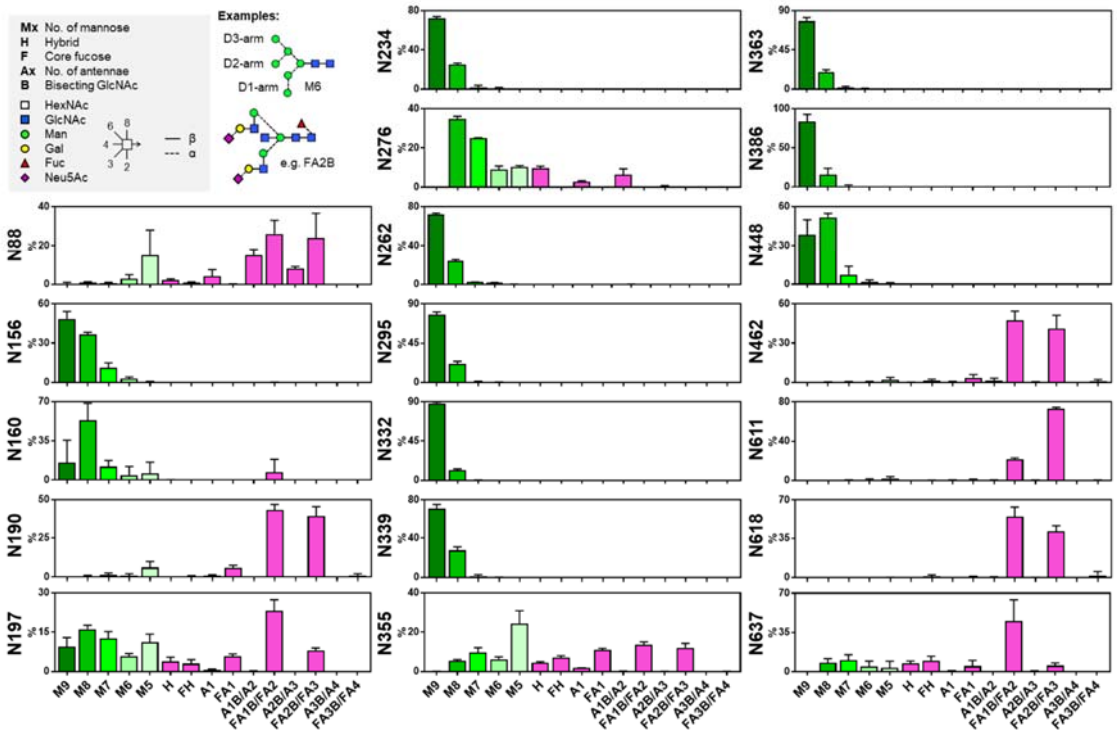


Figure 3.4. Reproducibility of site-specific glycosylation analysis. Three BG505 SOSIP.664 proteins were reduced, alkylated and digested with trypsin or chymotrypsin, prior to analysis by LC-ESI MS. Bar graphs depict the mean relative abundance of glycan structures at 19 of the 28 PNGS for the three biological replicates ±SEM. Oligomannose-type glycans are shown in green. Hybrid- and complex-type glycans are shown in magenta and are grouped according to the inset key (also see Appendix C).

3.3 The influence of antibody affinity purification on Env glycosylation

The standard protocol for the production of Env trimers involves an antibody affinity purification step. Often, bnAbs specific to cleaved, trimeric Env (e.g. PGT145 or PGT151) are used in order to select for trimers in their native-like conformation [278, 300]. The bnAb, 2G12, is also frequently used, along with a size exclusion chromatography (SEC) step to isolate the trimeric fraction. This method was used in the cGMP production of BG505 SOSIP.664 trimers [324]. Occasionally, negative selection with a non-neutralising antibody is employed to remove contaminating, non-native-like Env [227, 234]. Given the glycan specificities of the above bnAbs, it is reasonable to hypothesise that antibody affinity purification may bias the glycan profile of Env towards that of the bnAb epitope. For

example, does the use of a PGT151 affinity step increase the abundance of tri- and tetra-antennary structures at the N611 and N637 sites [300]?

3.3.1 Resilience of BG505 SOSIP.664 to glycan biasing

Pritchard *et al.* have previously analysed the overall glycosylation profiles of BG505 SOSIP.664 trimers purified using either a PGT145, PGT151 or 2G12 affinity step [123]. All three trimers had extremely similar overall glycosylation profiles and abundances of oligomannose-type glycans (Figure 3.6) [123]. The Paulson laboratory have since repeated this analysis using the site-specific glycan classification method [274]. They too reported that all three antibody affinity purified trimers had highly similar glycosylation patterns, though a handful of sites were considered significantly different from Ni^{2+} purified reference material [274]. Thus, the glycosylation of BG505 SOSIP.664 trimers appears to be independent of the bnAb affinity purification method used.

3.3.2 Influence of bnAb affinity purification on CZA97 SOSIP.664

In contrast to BG505 SOSIP.664 trimers, CZA97 SOSIP.664 (clade C) trimers display structural and antigenic differences when purified by different bnAbs (Figure 3.5). For example, while purification of the trimers with 2G12, PGT145 or PGT151 all yield 100% native-like trimers, the propensity of the trimers to adopt open or closed conformations varies (Figure 3.5C; personal communication with Dr Rajesh Ringe and Prof. John Moore, Weil Cornell Medical College, New York, USA). Furthermore, each of the trimers displays differential binding to bnAbs that cannot be explained by the different protein conformations (Figure 3.5B). Lastly, the trimers display different molecular weights when analysed by SDS-PAGE (Figure 3.5D).

To establish whether the differences in antigenicity could be attributed to glycan biasing during the bnAb purification step, glycans were enzymatically released and fluorescently labelled for analysis by HILIC-UPLC. The chromatograms revealed differences in the abundance of numerous glycans between the three populations of trimers (Figure

3.6A). To further probe these differences, oligomannose-type glycans were quantified by digestion with Endo H, again revealing the variability in the abundance of each glycan structure (Figure 3.6B). For comparison, oligomannose-type glycan quantifications for BG505 SOSIP.664 trimers expressed in CHO cells, and purified using the same three bnAbs were included (Figure 3.6C; data from Pritchard *et al.* [123]), highlighting the contrast between the susceptibility of the glycosylation of the two genotypes to bnAb affinity purification.

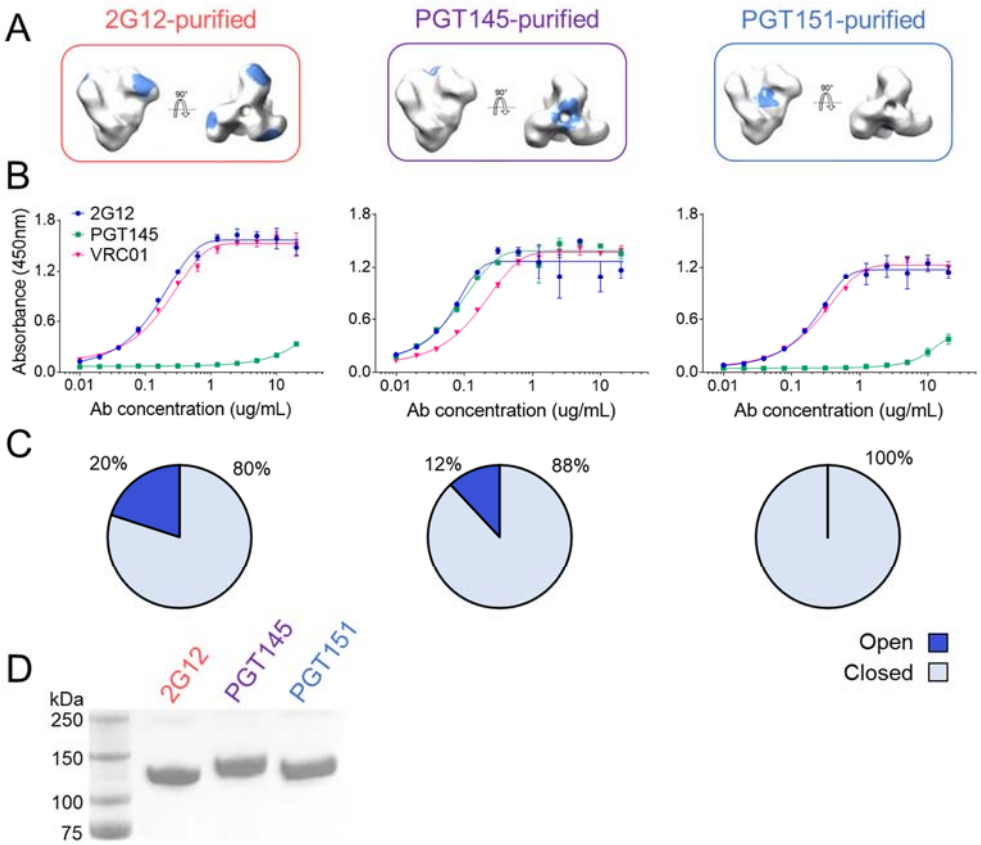


Figure 3.5. Structure and antigenicity of CZA97 SOSIP.664 trimers. CZA97 SOSIP.664 trimers were expressed in CHO cells and purified by either 2G12 (red), PGT145 (purple) or PGT151 (blue) affinity chromatography. A) The bnAbs recognise glycan epitopes at distinct regions of the trimer, adapted from Pritchard *et al.,* 2015 [123]. B) The CZA97 SOSIP.664 trimers display different binding to 2G12, PGT145 and VRC01, as judged by ELISA. No binding of PGT151 was observed to any of the trimers, suggesting that immobilisation of the trimers on the ELISA plates interrupts this epitope. C) Relative abundances of 'open' and 'closed' conformations of CZA97 SOSIP.664 trimers, as judged by negative stain EM. Conformational data provided by Dr Rajesh Ringe and Prof. John Moore, Weil Cornell Medical College, New York, USA. D) SDS-PAGE analysis of CZA97 SOSIP.664 trimers.

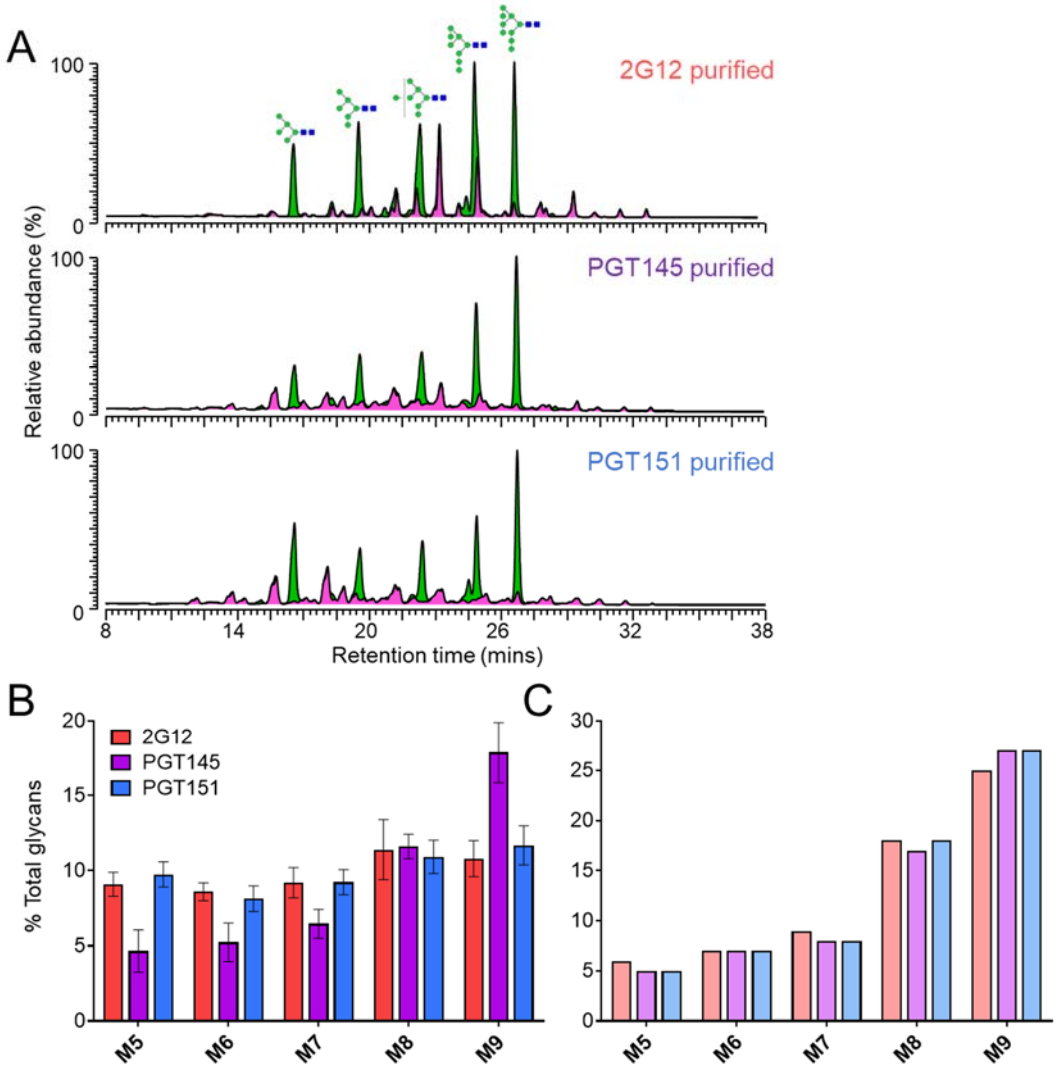


Figure 3.6. Glycosylation profiles of CZA97 SOSIP.664 trimers. A) CZA97 SOSIP.664 trimers were expressed in CHO cells and purified by either 2G12 (top/red), PGT145 (middle/purple) or PGT151 (bottom/blue) affinity chromatography. Glycans were enzymatically released and fluorescently labelled prior to analysis by HILIC-UPLC. Glycans were digested with Endo H in order to quantify the relative abundances of oligomannose- (green) and complex-type (magenta) glycans. B) Relative abundance of oligomannose-type glycans; M5-9, $Man_{5-9}GlcNAc_2$. Bars represent the mean ±SEM for three biological replicates. C) Relative abundance of oligomannose-type glycans on BG505 SOSIP.664 trimers expressed in CHO cells and purified by 2G12, PGT145 or PGT151. BG505 SOSIP.664 data reproduced from Pritchard *et al.*, 2015 [123].

To determine the sites contributing to the differences observed in the overall glycosylation profiles, the CZA97 SOSIP.664 trimers were analysed by site-specific glycan analysis. Proteins were reduced, alkylated, and digested with trypsin or chymotrypsin, prior to analysis by LC-ESI MS. Glycan compositions could be obtained for 22 of the 26 sites.

Generally, the overall trends were similar between the three populations of timers. Specifically, the distributions of oligomannose-type, hybrid-type and fucosylated glycans were comparable, irrespective of bnAb affinity purification (Figure 3.7A).

There were, however, differences in the fine processing of the glycans at most sites, which occasionally mapped to the corresponding bnAb epitope. For example, the N160 site was significantly enriched for $Man_6GlcNAc_2$ glycans on the PGT145 purified material (Figure 3.7B). This is consistent with the resolution of a $Man_6GlcNAc_2$ structure at this site in a cryo-EM structure of the PGT145 antibody in complex BG505 SOSIP.664 trimers [256]. Similarly, the N611 site, which contributes to the PGT151 glycan epitope [299, 300], contained more fucosylated tri-antennary structures on the PGT151 purified material compared to the 2G12 purified material*. In contrast, some of the observed differences did not fit a model of biasing towards the glycan epitope. For example, the N461 site also contained more fucosylated tri- and tetra-antennary glycans in the PGT151 purified material. While this fits with the glycan preference for PGT151 [299, 300], there is no evidence to suggest that PGT151 interacts with the glycan at this site.

Additionally, there was an apparent discrepancy in the abundance of $Man_9GlcNAc_2$ between the HILIC-UPLC analysis and the LC-ESI MS analysis. While PGT145 purified material contained the most $Man_9GlcNAc_2$ overall (Figure 3.6), this structure was not the most abundant at any site on the PGT145 purified trimers, exemplified by the N332 site (Figure 3.7B). This could be attributed to varying glycan site occupancy across the different populations of trimers, which was not addressed in this analysis, or possibly due to incomplete coverage.

Taken together, the results confirm that bnAb affinity purification can influence the resulting glycosylation of Env, but that these differences do not exclusively correspond to

*Note that the composition of the N611 site could only be determined for one of three biological replicates for the PGT151 and 2G12 purified material, and could not be determined for PGT145 purified material.

target epitopes. This may be symptomatic of bnAbs preferentially binding trimers adopting particular conformations. Furthermore, such effects are likely to be specific to each genotype, given that the glycosylation of BG505 SOSIP.664 trimers is independent of purification method [123, 274].

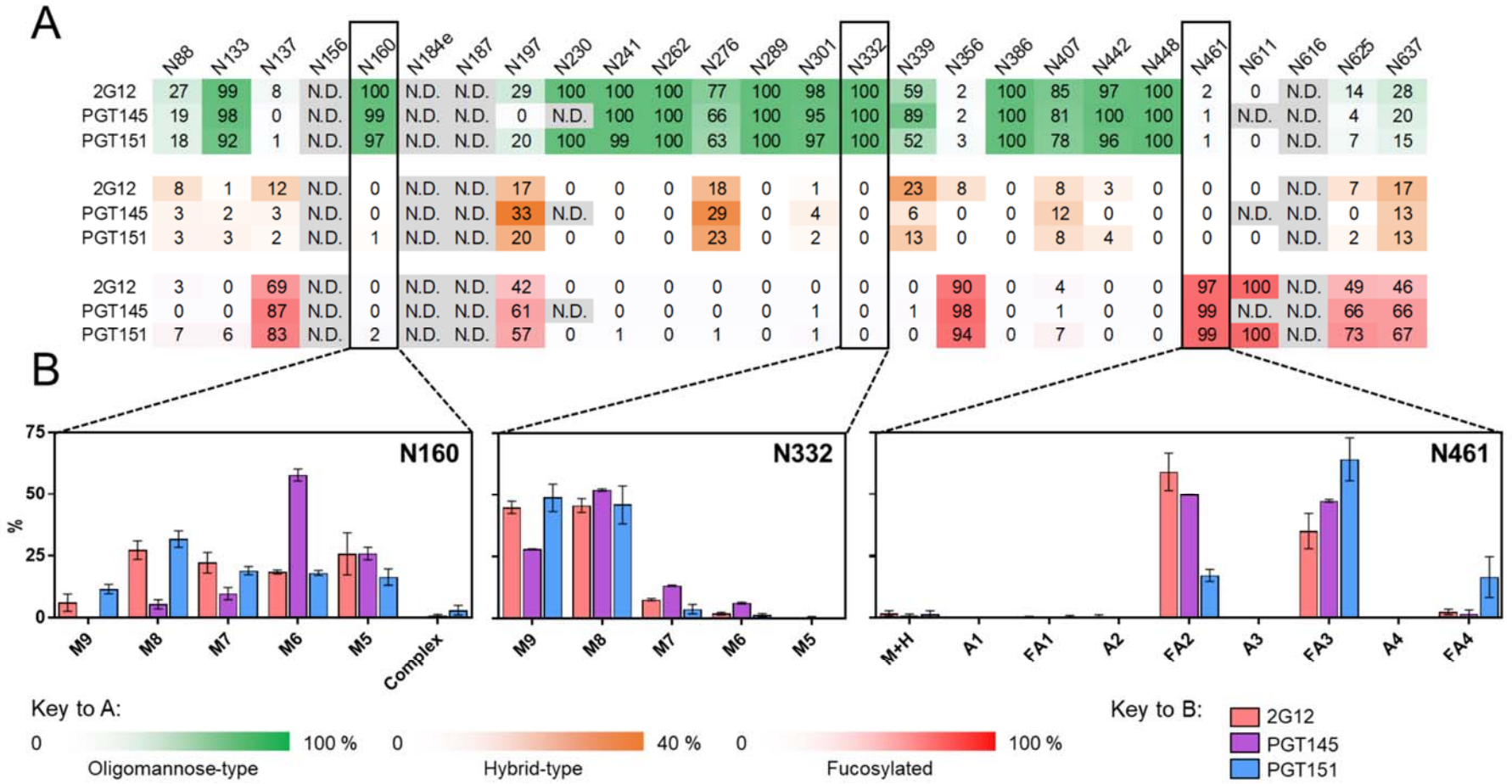


Figure 3.7. Site-specific glycosylation of CZA97 SOSIP.664 trimers. Trimers were reduced, alkylated and protease digested, prior to analysis by LC-ESI MS. A) The relative abundance of oligomannose-type, hybrid-type and fucosylated glycans at each site on CZA97 SOSIP.664 trimers purified by either 2G12, PGT145 or PGT151 affinity chromatography. B) The relative abundance of glycan structures at the N160, N332 and N461 sites. For clarity, glycan classes have been grouped as follows: Complex, hybrid- and complex-type glycans; M+H, oligomannose- and hybrid-type glycans.

3.4 The influence of stabilising mutations on Env glycosylation

Env trimers are inherently flexible and are able to transition between open and closed conformations ('breathing') [379]. A consequence of this flexibility is that non-nAb epitopes, particularly those associated with the V3 loop, can become exposed when the trimer is in the open conformation. Indeed, the V3 region is highly immunogenic when rabbits are immunised with BG505 SOSIP.664 trimers [237]. It is possible that exposure of the V3 (and associated non-nAb epitopes) may limit the presentation of bnAb epitopes associated with the closed conformation. Furthermore, the immunodominance of such epitopes may compromise the induction of bnAbs. Accordingly, it was hypothesised that a trimer stabilised in the closed conformation, for example with additional disulphide bonds or chemical cross-linking, would focus the antibody response on the desired nAb epitopes, with minimal induction of non-nAbs [228, 380].

Increasingly high-resolution structures of BG505 SOSIP.664 and other early native-like trimers have facilitated the identification of additional stabilising mutations (Appendix E) [225, 226, 381, 382]. These include additional point mutations, and the introduction of intra-subunit (i.e. within gp120), inter-subunit (i.e. gp120 and gp41), and inter-protomer disulphide bonds (Figure 3.8). The resulting next-generation, hyper-stabilised trimers display both reduced conformational flexibility and increased immunogenicity [225]. Crucially, the additional mutations have enabled the production of native-like trimers from multiple strains that had previously proved problematic [215]. This is significant as immunisation with trimers derived from one strain is insufficient to generate the required neutralisation breadth.

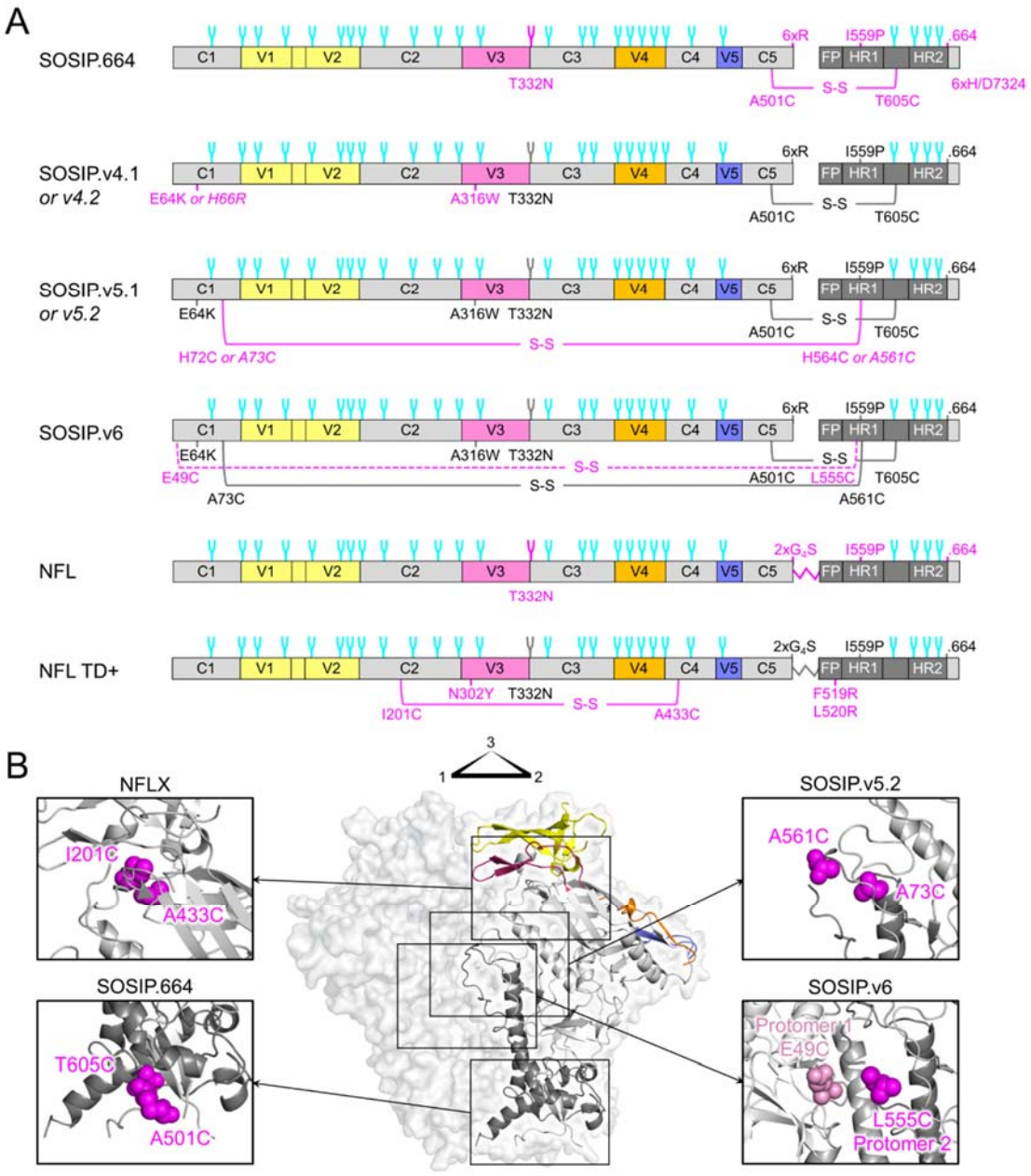


Figure 3.8. Design and structure of hyper-stabilised native-like trimers. A) Schematic of the BG505 SOSIP.664, SOSIP.v4, SOSIP.v5, SOSIP.v6, NFL and NFL TD+ constructs. New introduced modifications are annotated in magenta and subsequently greyed out. Dashed line indicates inter-protomer disulphide bond. The above trimer formats have been applied to a variety of HIV-1 strains, though the T332N mutation is specific to the BG505 genotype, which naturally lacks this key glycan epitope. Missing conserved glycan sites have been introduced into other strains. For a full details on the native-like trimer mutations used throughout this thesis, see Appendix E. B) Structure of the BG505 SOSIP.664 trimer (PDB 5CEZ) displaying the locations of introduced disulphide bonds (magenta spheres). Protomers are numbered one to three, with protomer two highlighted in cartoon.

Given the changed conformational dynamics of hyper-stabilised constructs, it is possible that the additional stabilising mutations may impact the glycosylation processing of the trimer. Torrents de la Peña *et al.* have previously performed on HILIC-UPLC analysis of BG505 SOSIP.664, SOSIP.v4.1, SOSIP.v5.1, and SOSIP.v6 trimers [225]. Overall the glycosylation profiles were extremely similar, though there was a slight trend toward increased oligomannose content on the more stabilised constructs (SOSIP.664, 64%; SOSIP.v4.1, 64%; SOSIP.v5.1, 68%; and SOSIP.v6, 69%) [225]. To determine the site-specific basis of these differences, the BG505 SOSIP.664, SOSIP.v4.1, SOSIP.v5.1, and SOSIP.v6 proteins were expressed in HEK cells and purified by PGT145 affinity chromatography, as detailed in Table 2.1. The trimers were reduced, alkylated and digested with trypsin or chymotrypsin, prior to analysis by LC-ESI MS.

As expected, the site-specific compositions of the four trimers were highly similar (Figure 3.9A). The processing of oligomannose-type sites was conserved across all trimers. Similarly, the processing of complex-type sites, was generally unaltered. There was, however, a tendency for sites containing mixed populations of oligomannose- and complex-type glycans, to contain more oligomannose-type glycans on the hyper-stabilised constructs (Figure 3.9B). The effect was subtle, though is consistent with the reduced conformational flexibility displayed by these constructs further restricting the actions of glycosylation processing enzymes.

3.5 Discussion

Defining the composition of the Env glycan shield has been a longstanding goal of HIV-1 research. Due to advancements in chromatographic and mass spectrometric techniques, detailed glycan analyses of Env are becoming more routine [331]. The LC-ESI MS workflow developed by Behrens *et al.* and used throughout this thesis, represents a highly sensitive and reproducible, semi-quantitative method for the site-specific glycan analysis of Env [132].

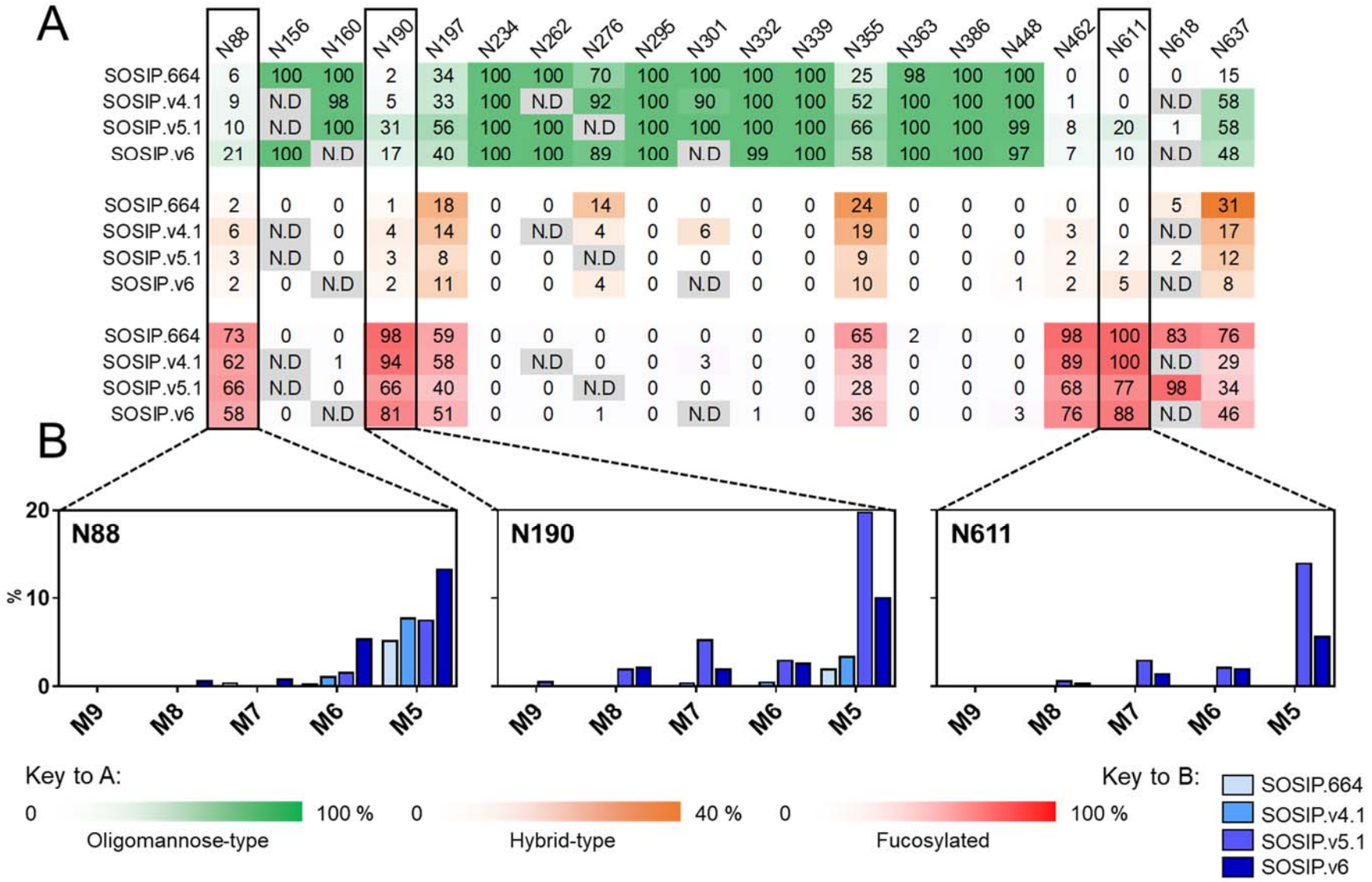

A
N88 N156 N160 N190 N197 N234 N262 N276 N295 N301 N332 N339 N355 N363 N386 N448 N462 N611 N618 N637
SOSIP.664 6 100 100 2 34 100 100 70 100 100 100 100 25 98 100 100 0 0 0 15
SOSIP.v4.1 9 N.D 98 5 33 100 N.D 92 100 90 100 100 52 100 100 100 1 0 N.D 58
SOSIP.v5.1 10 N.D 100 31 56 100 100 N.D 100 100 100 100 66 100 100 99 8 20 1 58
SOSIP.v6 21 100 N.D 17 40 100 100 89 100 N.D 99 100 58 100 100 97 7 10 N.D 48
SOSIP.664 2 0 0 1 18 0 0 14 0 0 0 0 24 0 0 0 0 0 5 31
SOSIP.v4.1 6 N.D 0 4 14 0 N.D 4 0 6 0 0 19 0 0 0 3 0 N.D 17
SOSIP.v5.1 3 N.D 0 3 8 0 0 N.D 0 0 0 0 9 0 0 0 2 2 2 12
SOSIP.v6 2 0 N.D 2 11 0 0 4 0 N.D 0 0 10 0 0 1 2 5 N.D 8
SOSIP.664 73 0 0 98 59 0 0 0 0 0 0 0 65 2 0 0 98 100 83 76
SOSIP.v4.1 62 N.D 1 94 58 0 N.D 0 0 3 0 0 38 0 0 0 89 100 N.D 29
SOSIP.v5.1 66 N.D 0 66 40 0 0 N.D 0 0 0 0 28 0 0 0 68 77 98 34
SOSIP.v6 58 0 N.D 81 51 0 0 1 0 N.D 1 0 36 0 0 3 76 88 N.D 46
B
20
%
10
0
N88
N190
N611
M9 M8 M7 M6 M5
Key to A:
0 100 % Oligomannose-type
0 40 % Hybrid-type
0 100 % Fucosylated
Key to B:
SOSIP.664
SOSIP.v4.1
SOSIP.v5.1
SOSIP.v6

Figure 3.9. Site-specific glycosylation of hyper-stabilised trimers. See overleaf. Trimers were reduced, alkylated and protease digested, prior to analysis by LC-ESI MS. A) The relative abundance of oligomannose-type, hybrid-type and fucosylated glycans at each site on BG505 SOSIP.664, SOSIP.v4.1, SOSIP.v5.1, and SOSIP.v6 trimers. B) The relative abundance of oligomannose-type glycans at the N88, N190 and N611 sites.

Over recent years, a number of hyper-stabilised trimer constructs have been developed [215]. This has enabled both the production of native-like trimers from multiple genotypes, and the reduced exposure of non-neutralising antibody epitopes [225, 228]. Here, the impact of hyper-stabilisation on the glycosylation of BG505 trimers was addressed. The fine processing of the trimers, and subsequently many glycan epitopes, was largely conserved across all constructs. However, there was some evidence to suggest that the hyper-stabilisation of the trimers can restrict some aspects of glycosylation processing, namely sites displaying mixed populations of glycans. In this context, it is pertinent that the glycosylation processing of these sites on native, virally-derived Env is generally greater than observed on the corresponding SOSIP.664 trimers (Chapter 7) [275, 276]. Thus, there may be a balance to be struck between generating stable immunogens, with reduced non-nAb epitope exposure, and those displaying glycosylation akin to the virus.

Just as there are multiple trimer formats, there are also many ways in which to purify native-like Env trimers. The most widely adopted purification methods utilise glycan-dependent bnAbs. While the glycosylation of BG505 SOSIP.664 trimers is largely independent of the bnAb used for purification [123, 274], this was not the case for CZA97 SOSIP.664 trimers, which displayed altered glycosylation, structure and antigenicity. It remains to be determined if the different CZA97 SOSIP.664 glycoforms display altered immunogenicity. In some instances, the glycosylation was biased toward that of the preferred glycan epitope of the bnAb, thus bnAb affinity purification may prove a useful tool for the production of immunogens designed to elicit specific glycan responses. However, there were several less predictable 'off-target' effects, and it is also unclear why the phenomenon appears

to be genotype-specific. The continued characterisation of Env immunogens will increase our understanding of the factors shaping glycosylation [331].

4. Conserved features of the HIV-1 glycan shield

The successful development of an HIV-1 vaccine will depend on the breadth and potency of protection it can convey. One of the challenges facing the development of an effective HIV-1 vaccine is the extremely high antigenic diversity of its Env glycoprotein, thus, suitable and conserved vaccine targets are limited. Strain-specific, autologous nAbs are routinely elicited in both natural infection and immunogenicity studies, typically directed against the highly variable protein portions of Env [78, 383]. In contrast, the bnAbs required for a protective vaccine often recognise glycan-containing epitopes (Figure 1.8, Table 1.1) [273, 278, 287-289, 297, 299, 300, 304]. A protective HIV-1 vaccine will therefore likely include a glycan component.

The envelope spike is extensively N-glycosylated by the infected host cell. Depending on strain, the gp120 subunit can contain between 18 and 33 PNGS [3], while the gp41 subunit generally contains a further four. However, the Env glycan shield is known to evolve over the course of infection, changing both the position and number of its PNGS in response to antibody-mediated selection pressure [82, 198, 199, 204]. For example, TF viruses tend to have fewer PNGS than chronic isolates; while this suggests a possible fitness advantage to having fewer PNGS, a balance must then be reached in order to protect vulnerable protein epitopes from emerging nAb responses [200-203]. In spite of the shifting nature of the Env glycan shield, a large number of PNGS are highly conserved (Figure 4.1A) [304, 384], perhaps

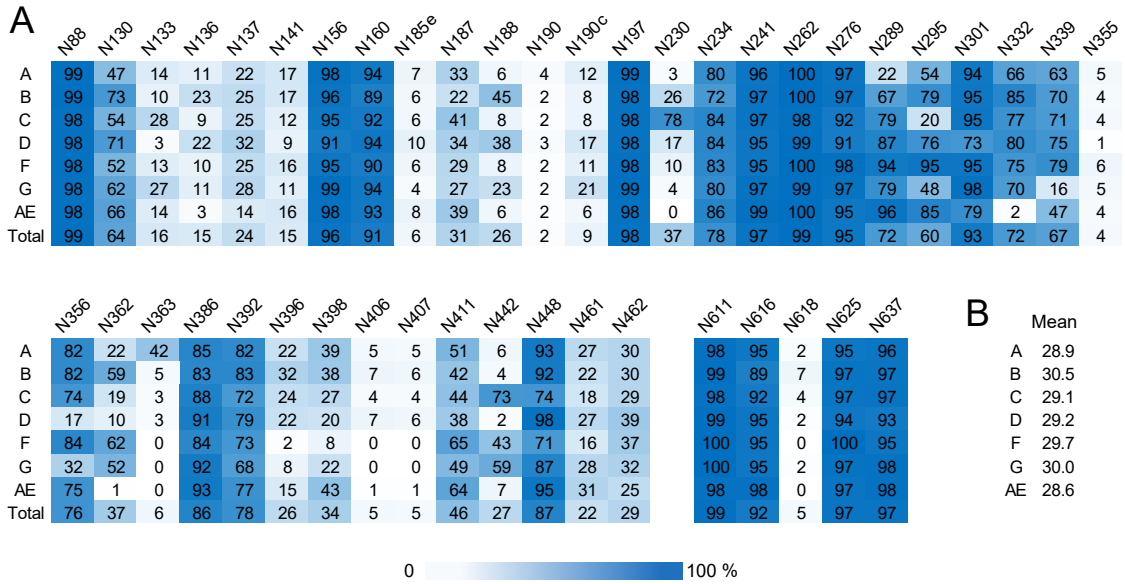


Figure 4.1 Conservation of PNGS across clades. A) Percentage conservation of individual PNGS. B) Mean number of total PNGS. The analysis was performed using sequence data obtained from the 2017 filtered web alignment from the Los Alamos HIV sequence database (https://www.hiv.lanl.gov). Clade A, n = 338; clade B, n = 2351; clade C, n = 1524; clade D, n = 127; clade F, n = 63; clade G, n = 114, clade AE, n = 535. Only the second position Asn was included for NN[S/T][S/T] motifs. Sites are only shown if present in one of the following representative strains: BG505 (clade A), AMC011 (clade B), CZA97 (clade C). The low conservation of the N332 site in clade AE is compensated by the above average use of the N334 site (92%, clade AE versus 21%, total; not shown).

reflecting their roles in protein folding and virus infectivity [385, 386], and/or lack of negative selection by the immune system. Furthermore, the total number of PNGS has remained fairly constant both across years of viral evolution [384], and multiple geographical clades (Figure 4.1B).

The large number of PNGS on Env creates a glycan density that is largely impervious to some of the host's glycan processing enzymes [310-313]. Consequently, glycan analyses of recombinant, monomeric gp120 [110, 306-311], pseudovirions [306, 312-314], membrane-associated trimers [275, 315, 316, 335], recombinant, soluble trimers [122, 123, 132, 242, 274, 316-322, 335], and virion-derived Env [275, 276, 327], have all reported an abundance of under-processed, oligomannose-type glycans. This 'mannose patch' signature is conserved both across the course of infection [311, 318, 335] and over multiple viral clades [311, 312], and the isolation of numerous bnAbs that recognise the oligomannose-type glycans of Env

confirms the antigenic importance of the mannose patches (Figure 1.8, Table 1.1). Large populations of oligomannose-type glycans are rarely observed on secreted mammalian glycoproteins, thus the divergence of Env glycosylation offers an attractive, conserved 'non-self' target for the host immune system to exploit [330, 351].

In addition to oligomannose-type glycans, complex-type glycans are emerging as similarly important targets for bnAbs. For example, α2-6-linked sialylated hybrid-type glycans have been implicated in the formation of the epitopes for apex-targeting bnAbs, such as PG9, PG16 and the CAP256-VRC26 lineage [271-273, 290], and tri- and tetra-antennary structures are required for PGT151 binding at the gp120-gp41 interface [299]. In contrast to the protein-directed processing of oligomannose-type glycans, the processing of complex-type glycans is subject to the repertoire of glycosidases and glycosyltransferases possessed by the infected host cell [333]. For instance, Env expressed in PBMCs contains primarily α2-6-linked sialylated glycans, as opposed to the α2-3-linked structures observed on HEK-expressed material [333]. In addition to influencing the antigenicity of Env, the presence or absence of specific glycan structures may also have important immunological consequences. In this context, asialylated Env has been reported to be more immunogenic than sialylated Env [305].

Although glycan targeting bnAbs can evolve after several years of infection, their elicitation through immunisation remains challenging. Firstly, the antibody response must overcome comparatively low protein-glycan binding affinities. Glycan-reactive B cells must then avoid negative selection, as the Env glycan shield is derived from the host cell's glycosylation machinery and may be considered immunologically 'self' [194]. Furthermore, given the diversity of glycosylation processing enzymes possessed by mammalian cells, glycoproteins can exist as very heterogeneous populations, consequently potentially reducing the antigenicity of each individual glycoform [195]. The presentation of inappropriate glycoforms to the immune system will therefore likely be disadvantageous to eliciting broad

anti-glycan antibody responses, and understanding the consensus of Env glycosylation processing may prove a useful tool in the design of appropriate immunogens.

With the development of numerous native-like trimers, spanning several clades, comes the ability to probe the conservation of key features of the glycan shield in greater detail. In this chapter, the glycosylation profiles of ten native-like trimers from three clades were assessed. The results confirm the cross-clade conservation of the mannose patch signature, and reveal new insights into the distribution of complex-type glycosylation processing.

4.1 Glycan analysis of a cross-clade panel of Envs

To determine the conservation of the processing of the Env glycan shield, a panel of ten native-like trimers were expressed and purified, as detailed in Table 2.1. The panel contained native-like trimers based on genotypes from clade A (BG505), B (B41, AMC008, AMC011, AMC016, and JR-FL), and C (001428, 16055, CZA97, and ZM197M). The amino acid sequences for the native-like trimers are located in Appendix E.

4.2 Conservation of oligomannose-type glycans

To assess the overall glycosylation profiles of the native-like trimers, the N-linked glycans were enzymatically released by in-gel digestion with PNGase F and fluorescently labelled prior to analysis by HILIC-UPLC (Figure 4.2). Labelled glycans were subsequently digested with Endo H and re-analysed. Integration of the chromatograms before and after digestion enabled quantification of the abundance of oligomannose-type glycans.

A substantial population of oligomannose-type glycans was observed across the panel of trimers, though the exact abundance varied (~ 40 – 60%). The predominant glycan on all ten constructs was $Man_9GlcNAc_2$ (Figure 4.2). The glycosylation profiles of all ten trimers were reminiscent of the BG505 SOSIP.664 profile and are consistent with the native-like architecture adopted by the trimers [122, 123, 331].

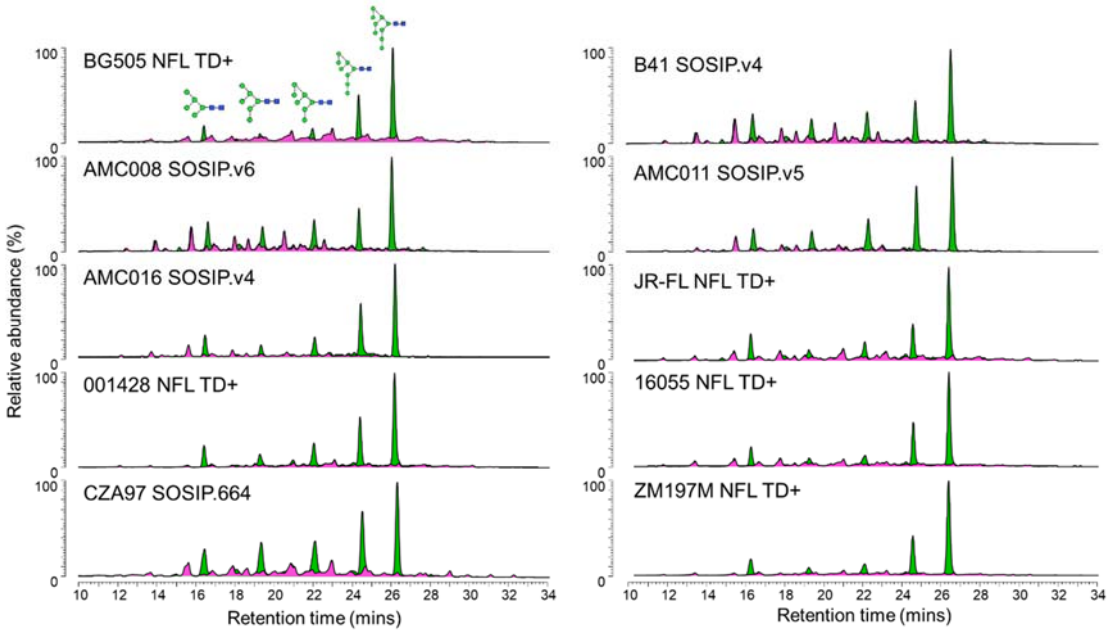


Figure 4.2. Glycosylation profiles of native-like Env trimers. N-linked glycans were enzymatically released and fluorescently labelled prior to resolution by HILIC-UPLC. Oligomannose-type glycans were quantified by integrating the peaks before (green) and after (magenta) digestion with Endo H. The $Man_{5\text{-}9}GlcNAc_2$ series is annotated according to the key in Appendix C.

Previous glycan analysis of recombinant, monomeric gp120 observed a correlation between the abundance of oligomannose-type glycans and the number of PNGS on the outer domain [311], in line with the role that glycan density plays in limiting glycosylation processing. To investigate the cause of differences observed in the abundance of oligomannose-type glycans on native-like trimers, a similar analysis was performed (Figure 4.3). In contrast to recombinant, monomeric gp120, no significant correlation was observed between the number of PNGS located on the outer domain (residues 252 – 482) and the total abundance of oligomannose-type glycans on the native-like trimers (Figure 4.3A). However, when the selection was expanded to include PNGS in the C2 region (i.e. N230, N234 and N241), a significant correlation was observed for both the total abundance of oligomannose-type glycans (Figure 4.3B, $p < 0.05$) and $Man_9GlcNAc_2$ (Figure 4.3E, $p < 0.01$). In line with monomeric gp120, no correlation was observed when compared to the total number of PNGS on the trimer (Figure 4.3C and F).

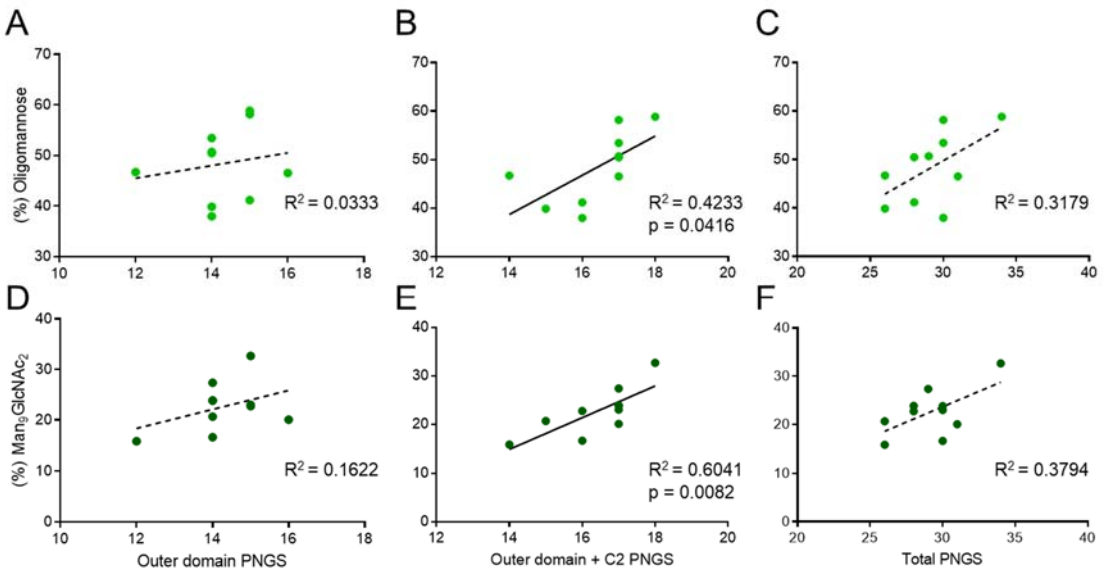


Figure 4.3. Correlation between the number of PNGS and abundance of oligomannose-type glycans. The abundance of oligomannose-type glycans (A-C) and $Man_9GlcNAc_2$ glycans (D-F) were plotted against the number of PNGS on the outer domain (residues 252 – 482; A and D), the outer domain and C2 region (residues 200 – 482; B and E), or the total number of PNGS (C and F). Correlations were performed using linear regression analysis in GraphPad Prism (version 7.05).

4.3 Integrity of the mannose patches

Previous site-specific glycan analyses of BG505 SOSIP.664 trimers have reported the Env oligomannose signature to be largely accounted for by several PNGS comprising solely oligomannose-type glycans [122, 132, 274]. In the context of the trimer, these sites are primarily located on the densely glycosylated outer domain of gp120, forming the IMP, and at the protomer interfaces and trimer apex, forming the TAMP. To determine the site-specific compositions of glycans across the panel of native-like trimers, the proteins were subject to site-specific glycan analysis. Proteins were reduced, alkylated and digested with trypsin or chymotrypsin, prior to LC-ESI MS analysis. Glycan compositions could be obtained for the majority of the sites on each protein, however where this was not readily achievable (e.g. when two PNGS co-occupied a peptide), the oligomannose- versus complex-type glycan content was determined by glycan classification analysis (by sequential digestion of the glycopeptide with Endo H and PNGase F; Section 2.4.3). Fully glycosylated models of the trimers were then generated based on published crystal or cryo-EM structures, or by

homology modelling where structural information was not available. Mapping the glycan composition data onto the models of the trimers revealed the conservation in the distribution of oligomannose-type glycans across the panel of trimers (Figure 4.4).

In line with the previously reported analyses of BG505 SOSIP.664 trimers, the distribution of sites containing predominantly oligomannose-type glycans was non-random. Sites contributing to the IMP on BG505 SOSIP.664 trimers (i.e. N262, N295, N301, N332, N363, N386, N392, and N448) were conserved oligomannose-type sites across all ten trimers (Figure 4.4 and 4.5). Additionally, the N230, N241, N289 and N442 sites, where present, were occupied by predominantly oligomannose-type glycans. The models confirmed that these glycans form large clusters of under-processed glycans, consistent with the role of glycan-glycan interactions limiting access to the host's mannosidases [310, 323]. The oligomannose-type glycans on the outer domain of gp120 form the epitopes for numerous bnAbs, and their prevalence across all ten trimers analysed support this target for vaccine design [212, 330, 351].

Additionally, the highly conserved N156 and N160 'TAMP sites' formed a ring of under-processed oligomannose-type glycans at the apex of each trimer (Figure 4.4 and 4.5). The restricted processing at these sites can be attributed to steric constraints imposed by protein and glycan elements from the neighbouring protomers [122]. The conserved processing of these sites is also significant given that numerous bnAbs target this region (Table 1.1). However, the observation that these sites are occupied by exclusively oligomannose-type glycans is inconsistent with several studies detailing the glycan preference of such bnAbs as sialylated hybrid-type glycans [271-273, 290]. The discrepancy between structural studies and glycopeptide analysis may reflect the tolerance of bnAbs to recognise multiple glycan structures [387].

Where present, the glycan at the N234 site was also comprised of predominantly oligomannose-type glycans. This glycan is located very nearby the highly conserved, CD4bs-

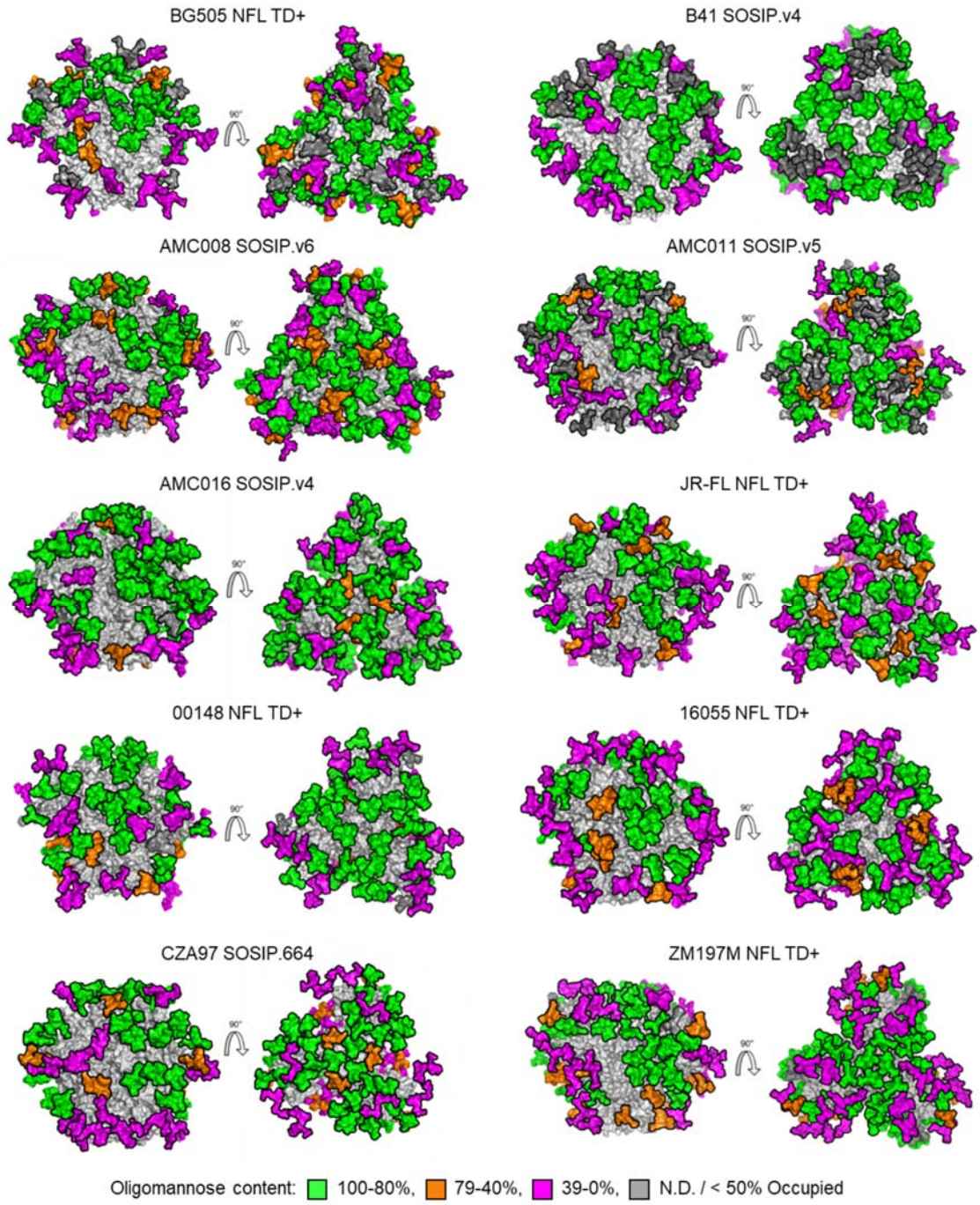


Figure 4.4. Distribution of oligomannose-type glycans across a panel of native-like trimers. Trimers were reduced, alkylated and digested with trypsin or chymotrypsin prior to analysis by LC-ESI MS. Fully glycosylated models were constructed according to Section 2.6, and the glycans were coloured according to their oligomannose content: > 80%, green; 40 – 79%, orange; 0 – 39%, magenta. Where the composition of sites could not be determined (e.g. when two PNGS co-occupied a peptide), the oligomannose content was determined by glycan classification analysis (by sequential digestion of the glycopeptide with Endo H and PNGase F; Section 2.4.3). Sites that still could not be determined, or were found to be < 50% occupied are coloured grey.

proximal N276 site. Across the panel of native-like trimers, the processing of the N276 site varied substantially (Figure 4.5), ranging from 0 – 100% oligomannose-type glycans. The oligomannose content of the N276 site was significantly lower on strains lacking the N234 site ($p < 0.05$), suggesting a reciprocal protection from enzymatic processing of these two sites (Chapter 5). The presence of the N234 site likely contributes to the correlation observed between PNGS on the outer domain and C2 region and the abundance of under-processed glycans by HILIC-UPLC analysis (Figure 4.3).

4.4 Features of complex-type glycosylation processing

The remaining sites were mostly occupied by complex-type glycans, or mixed populations of complex- and oligomannose-type. These sites largely correspond to the N- and C-termini of the trimers, including the N88 site, the V1/V2 sites, and the four sites on gp41. Additionally, the site located at N460 – 463 (depending on strain) constituted a conserved complex-type glycan located very near to the CD4bs (Figure 4.4).

The abundance of hybrid-type glycans, the preferred epitope for some bnAbs [271-273, 290], was low across the PNGS on all trimers, though appeared slightly enriched at the N276 and N637 sites. In contrast, fucosylated glycans were abundant on gp41, and at N355 and N460 – 463. Unfortunately, the close proximity of V1/V2 loop glycan sites to one another often resulted in doubly-occupied peptides, thus the glycan compositions of many of these sites could not be readily determined. Those that could, also displayed high levels of fucosylation. Sialylation was also abundant, though to a lesser extent, at the gp41, N355 and N460 – 463 sites. The CZA97 SOSIP.664 trimers appeared to be the most sialylated of the panel, though this was likely a reflection of the expression cell line used. In contrast to the other trimers, CZA97 SOSIP.664 was expressed in CHO cells. Although the protein-directed processing of oligomannose-type glycans is conserved between CHO and HEK cells, the former have previously been reported to display increased sialylation [123].

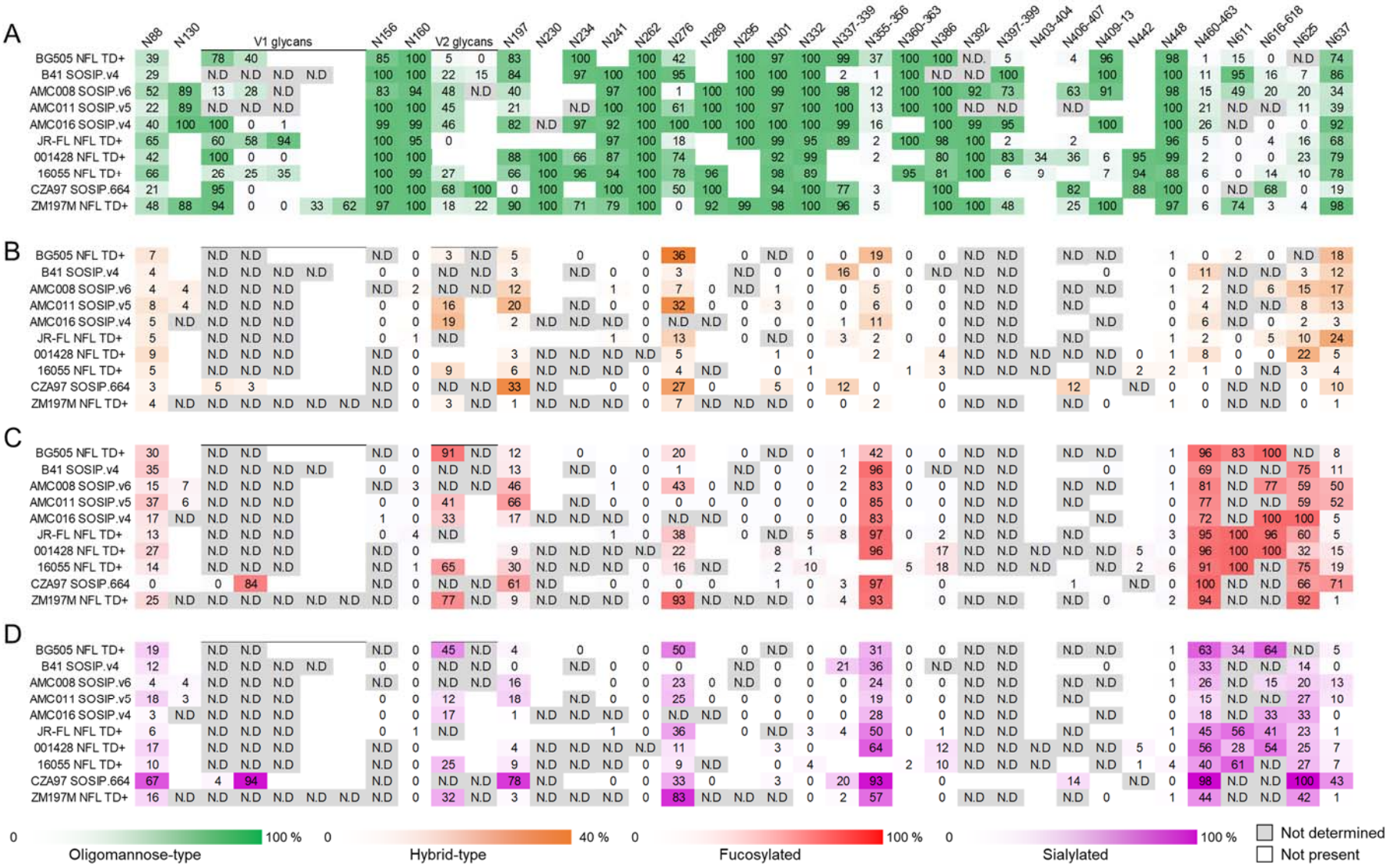

A
B
C
D
V1 glycans
V2 glycans
N88 N130 N156 N160 N197 N230 N234 N241 N262 N276 N289 N295 N301 N332 N337-339 N355-356 N360-363 N386 N392 N397-399 N403-404 N406-407 N409-13 N442 N448 N460-463 N611 N616-618 N625 N637
BG505 NFL TD+
B41 SOSIP.v4
AMC008 SOSIP.v6
AMC011 SOSIP.v5
AMC016 SOSIP.v4
JR-FL NFL TD+
001428 NFL TD+
16055 NFL TD+
CZA97 SOSIP.664
ZM197M NFL TD+
0 100 % Oligomannose-type
0 40 % Hybrid-type
0 100 % Fucosylated
0 100 % Sialylated
Not determined
Not present

Figure 4.5. Site-specific glycosylation processing of a panel of native-like trimers. See overleaf. A) Relative abundance of oligomannose-type glycans at each PNGS. Where oligomannose content could not be determined by compositional analysis, glycan classification analysis was included, with values adjusted to account for < 100% occupancy. B) Relative abundance of hybrid-type glycans at each PNGS. C) Relative abundance of fucosylated glycans. D) Relative abundance of sialylated glycans. Sites that could not be determined are coloured in grey, cells are left blank where the glycan is not present in the strain.

The restricted processing of oligomannose-type glycans is attributed to steric constraints limiting the access of α-mannosidases to their substrate glycans. If a glycan passes this initial hurdle, it could conceivably be accessible to most subsequent glycosylation processing enzymes. To test this hypothesis, the data were reanalysed (for conserved, non-oligomannose sites only), expressing the abundance of complex-type glycosylation features as a percentage of the complex-type glycans at a given site. For example, a glycan site that comprises 60% oligomannose-type glycans, 20% fucosylated complex-type glycans and 20% afucoslylated complex-type glycans, is reported as 50% fucosylated. While comparing the absolute abundances of specific glycan types is useful for defining the compositions of glycan epitopes, this analysis provides further insight into the dynamics of complex-type glycan processing.

The analysis revealed the processing of hybrid-type glycans was not uniform across complex-type and mixed sites, as determined by a one-way ANOVA (Figure 4.6A; $p < 0.0001$). Post-hoc testing confirmed the N276 and N637 sites to contain significantly more hybrid-type glycans (Figure 4.6B). This suggests that there is a population of glycans at each of these sites that is susceptible to processing by α-mannosidase I, but resistant to processing by α-mannosidase II enzymes. Such resistance could again be conferred by steric constraints imposed by the close proximity of these glycans to one another, and to the N234 site (where present), or to protein or glycan elements from the neighbouring protomer.

Similarly, the distribution of fucose was not evenly spread (Figure 4.6A; $p < 0.0001$). As before, the N276 and N637 sites were significantly less fucosylated, as was the N88 site

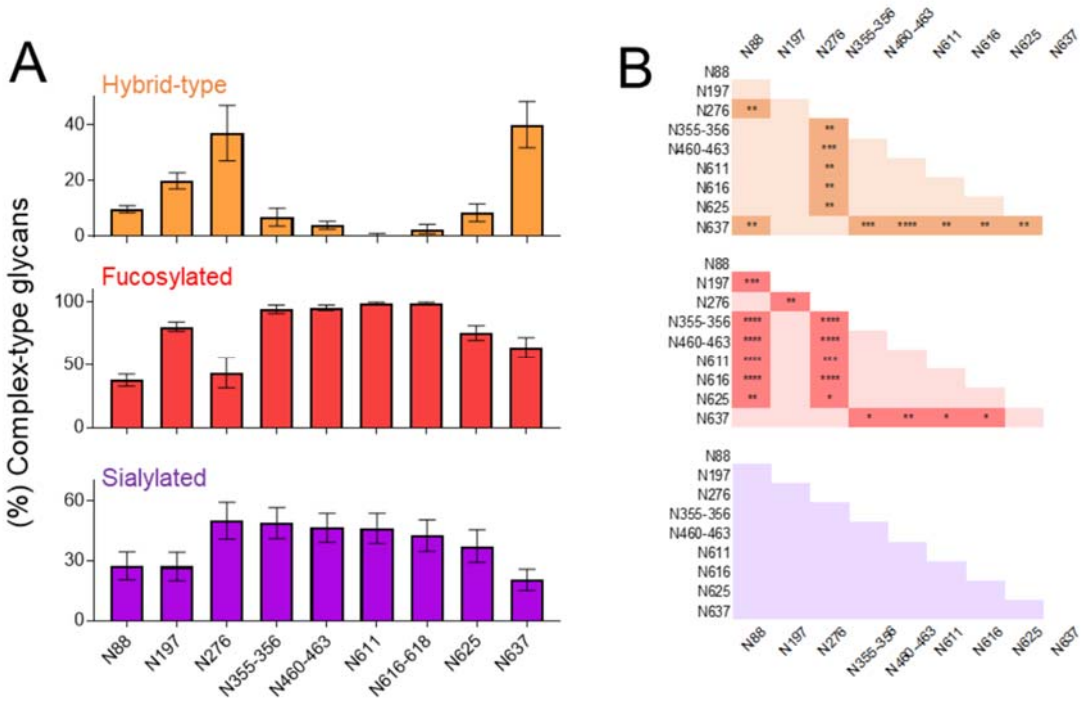


Figure 4.6. Features of complex-type glycan processing. A) The abundance of hybrid-type glycans, fucosylated glycans and sialylated glycans, expressed as a percentage of the complex-type glycans at each site. For example, a glycan site that comprises 60% oligomannose-type glycans, 20% fucosylated complex-type glycans and 20% afucoslylated complex-type glycans, is reported as 50% fucosylated. Bars represent mean ±SEM. B) Results of post-hoc testing using Tukey's multiple comparison test: ****, $p < 0.0001$; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$. Analysis was performed in GraphPad Prism (version 7.05).

(Figure 4.6B). The N88 site is located proximal to the gp41 subunit, and it is possible that the presence of gp41 restricts access of fucosyltransferases to base of N88. Consistent with this view, the N88 site on recombinant, monomeric gp120 contains almost exclusively fucosylated glycans [122].

The distribution of sialic acids, on the other hand, was more uniform across the complex-type sites (Figure 4.6A). While a one-way ANOVA returned a significant result ($p < 0.05$), post-hoc analysis did not identify any significant differences between the sites (Figure 4.6B). As sialic acids are terminal glycan modifications, they may be less prone to steric restrictions by neighbouring glycan elements. Taken together, the results suggest that the glycosylation processing of Env can be restricted at many levels, not just by limiting α-mannosidase trimming.

4.5 Discussion

Native-like Env trimers display extremes of glycan processing, which are conserved across multiple clades and strains. On one hand, very high glycan density results in large regions of under-processed oligomannose-type glycans. On the other hand, less densely packed glycans are susceptible to cell-directed glycosylation processing. Despite the potential for vast glycan heterogeneity at processed sites, these PNGS often displayed similar features across a panel of ten native-like trimers. Thus, despite diversity at the level of the amino acid sequence, the glycosylation processing of the Env is largely conserved. The relative homogeneity of the glycan shield supports it as a target for HIV-1 vaccine design [212, 330, 351], though it will also be important to determine the glycosylation processing of virally-derived Env from multiple strains.

The extent to which certain glycan features help or hinder the development of bnAbs is not yet known. For immunogen design this must be considered at two levels. Firstly, what is the total number of PNGS required by an immunogen to elicit bnAb responses? While there is evidence to suggest that more complete glycan shields on TF viruses correlate with the development of bnAbs in natural infection [204], immunogens containing fewer glycan sites display increased binding to gl-bnAbs, and are thus able to initiate the development of bnAbs [360, 363]. Secondly, are specific glycan epitopes required? In this context, sialic acid at the trimer apex has been implemented in binding gl-bnAbs and driving the development of neutralisation breadth of antibodies targeting this region [290]. However, glycan density is known to impact the fine processing of Env glycans [310], thus the two are interlinked. For example, the processing of the N276 site appears related to the presence or absence of the N234 glycan. The fine processing of glycan epitopes through manipulation of glycosylation density will be investigated more thoroughly in the next chapter.

While targeting conserved epitopes may result in increased neutralisation breadth, it is important to consider that conserved features are inherently immunoquiescent. Thus, it

may still emerge that glycan-based HIV-1 immunogens require additional strategies in order to boost their immunogenicity. This will be explored further in Chapter 8.

5. Networks of glycans maintain antibody epitopes in the face of glycan additions and deletions

The HIV-1 envelope spike is the sole target for bnAbs produced during infection and is consequently a main target for vaccine design. Accordingly, vaccine strategies typically involve producing recombinant mimics of the native, virion-associated Env trimer that present multiple bnAb epitopes, and/or immunogens specifically designed to target the germline-encoded bnAb precursors (gl-bnAbs) [215, 359].

The majority of bnAbs have evolved to recognise epitopes that are either entirely or partially composed of N-linked glycans (Figure 1.8) [273, 278, 287-289, 297, 299, 300, 304]. These bnAbs typically recognise the glycans at four distinct regions of Env: the gp120/gp41 protomer interface (e.g. PGT151), surrounding the CD4 binding site (e.g. HJ16), the V1/V2 loops at the trimer apex (e.g. PG9 and PG16), and the oligomannose-type glycans centred around the highly conserved N332 site on the outer domain of gp120 (e.g. PGT135 and 2G12) [244]. Given the number of bnAbs targeting the N332 glycan, it has previously been termed the 'supersite of immune vulnerability' [259].

The prototypic recombinant Env mimics are the BG505 SOSIP.664 trimers, based on the subtype A TF virus sequence, BG505. These trimers display native-like structure and antigenicity [74, 125], and are lead candidates in ongoing human immunogenicity studies [243, 324]. The BG505 TF virus naturally lacks the conserved N332 glycan site, thus this glycan was included (T332N) in the BG505 SOSIP.664 trimers in order to introduce the

'supersite' epitope [74]. However, the BG505 sequence also lacks glycans at the 241 and 289 positions, despite their presence in 97% and 72% of HIV-1 isolates, respectively (Figure 4.1). The presence or absence of holes within the glycan shield has recently received a lot of attention because of the putative role of holes in initiating neutralising antibody (nAb) responses.

Immunisation of rabbits with BG505 SOSIP.664 trimers elicits autologous nAb responses centred on glycan holes at positions 241 and 289 [240, 383]. Filling this hole (i.e. by introducing a glycan site) blocks antibody neutralisation [383]. Similar results have been observed with nAbs targeting holes at the 130, 197 and 465 positions in immunogenicity studies with native-like trimers from different isolates [240, 388-390]. Such autologous nAb responses can be readily redirected *in vivo* by closing the glycan holes, and opening new ones elsewhere on the trimer [391]. This phenomenon is echoed in natural infection, as the glycan shield 'shifts' to escape arising nAbs [198-200, 204]. The N332 glycan, for example, has been observed to shift from the N334 position and back again after the appearance of nAbs [198]. While it is accepted that glycan holes offer an immunodominant distraction capable of eliciting autologous nAbs, the extent to which holes hinder the development of bnAbs remains largely unknown. For example, there is evidence to suggest that more complete glycan shields in transmitted/founder viruses correlate with the development of greater neutralisation breadth in infected individuals [204]. Future immunisation strategies may therefore include immunogens with closed glycan holes, in order to redirect the nAb response away from the immunodominant protein surface toward more broadly neutralising glycan-based epitopes [383, 391].

The elicitation of a bnAb response requires the activation of bnAb precursor B cells. Effective immunogens must therefore be capable of engaging the B cell receptor (i.e. the gl-bnAb), prior to affinity maturation of the bnAb in the germinal centres. However, this process is hampered by the low affinity of gl-bnAbs to Env, often due to their inability to

accommodate conserved N-linked glycans [190-193, 392]. Thus an alternative, albeit closely linked, approach to eliciting bnAbs, is to prime with glycan-depleted immunogens, capable of engaging gl-bnAbs, and subsequently boost with their 'filled-in' derivatives in order to drive the development of neutralisation breadth [359-363].

Glycan density, however, impacts glycosylation processing, which can in turn influence epitope presentation. The unusually high density of N-linked glycans on gp120 limit the extent to which individual sites can be processed by the host's α-mannosidases [323]. Thus, gp120 displays a significant population of under-processed oligomannose-type glycans, termed the intrinsic mannose patch (IMP) [309, 310, 312, 313]. Analysis of recombinant, monomeric gp120 revealed that the removal of individual glycan sites from within the IMP often results in larger-than-expected decreases in the abundance of oligomannose-type glycans, as sites surrounding the deletion become more susceptible to glycan processing [310]. In Env trimers displaying native-like conformations, additional steric hindrances imposed by glycan and protein elements from neighbouring protomers give rise to a further trimer associated mannose patch (TAMP) [122, 123, 274]. Analysis of glycan-depleted, trimeric immunogens also revealed increased glycan processing at sites proximal to the glycan deletions [274, 328]. Furthermore, a number of studies have reported correlations between glycan density and the abundance of under-processed oligomannose-type glycans [280, 311]. Thus, while oligomannose-type glycans are a conserved feature of the Env glycan shield, and a key bnAb target, in some circumstances they can become susceptible to enzymatic processing.

Given the propensity for glycan density to influence the processing of glycans, this chapter sought to determine the impact of individual glycan site additions and deletions on bnAb epitopes. Site-specific glycan analysis of BG505 SOSIP.664 trimers reveals that glycan site addition and deletion not only influences the fine processing of glycans proximal to the mutated glycan site, but also elsewhere on the trimer. The tolerance of bnAbs targeting two

distinct glycan epitopes to this altered processing was investigated further, revealing the differing dependencies of mannose patch-targeting and apex-targeting bnAbs on the surrounding network of N-linked glycans.

5.1 The glycan network surrounding the 2G12 epitope

The first glycan-targeting bnAb to be identified was 2G12. This antibody adopts an unusual structure in order to overcome low protein-glycan affinities, whereby the two Fabs adopt a domain-exchanged dimer conformation (Fab_2) via exchange of their VH domains [250, 346, 393]. This unique architecture creates two primary binding sites at the VH/VL interfaces and two secondary binding sites on either side of the VH/VH interface, allowing for high avidity binding of four N-linked glycans. The structure of the 2G12 Fab_2 bound to BG505 SOSIP.664 was solved to 3.8 Å by cryo-EM (structure provided by Dr Chris Cottrell from the laboratory of Prof. Andrew Ward, Scripps Research, La Jolla, USA; Figure 5.1A). The 2G12 Fab_2 binds to a 1,879 $Å^2$ epitope composed entirely of N-linked glycans [247, 248, 251, 253]. The primary binding sites make contact with the terminal α1-2-linked mannose residues of the D1-arms of the oligomannose-type glycans at positions N392 and N295 (Figure 5.1A-C). The secondary binding sites make contact with the D2- and D3-arms of the oligomannose-type glycans at positions N332 and N339 (Figure 5.1A-C). Although not directly in contact with the 2G12 Fab_2, the glycans at N363 and N386 may also play a role in 2G12 binding by providing support to the N392 glycan via glycan/glycan interactions (Figure 5.1A) [394]. Previous studies have reported a potential contact between the 2G12 Fab_2 and glycans at positions N137 and N411 [249, 395], however, no coordinated density was observed for these glycans in this reconstruction, indicating that they do not play a direct structural role in 2G12 binding.

To further gauge the contribution of individual glycans to the 2G12 epitope, the binding to glycan knockout mutants was assessed by ELISA. As expected, the N295A and N392A deletions resulted in the largest decrease in antibody binding (Figure 5.1D). Though

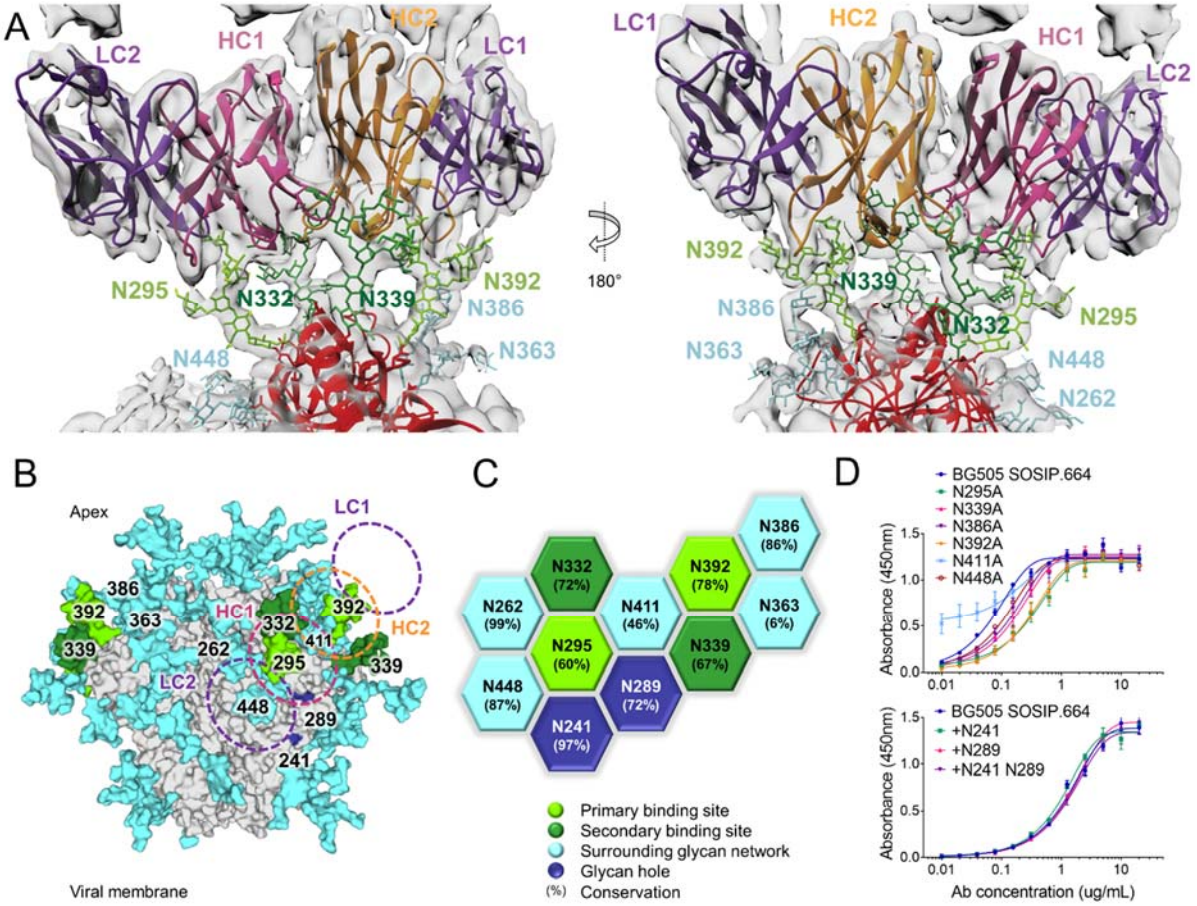


Figure 5.1. The 2G12 glycan epitope. A) The cryo-EM structure of the 2G12 Fab_2 in complex with BG505 SOSIP.664 was resolved to 3.8 Å by Dr Chris Cottrell from the laboratory of Prof. Andrew Ward, Scripps Research, La Jolla, USA. Electron density is shown in grey. No density was observed for the glycan at N411. Glycans bound by the 2G12 primary binding site are coloured light green, glycans bound by the secondary binding site are coloured dark green, the surrounding network of glycans are shown in cyan. B) Model of a fully glycosylated BG505 SOSIP.664 trimer based on PDB 5ACO with glycans added according to Behrens *et al.*, 2016 [132]. The glycan holes at the 241 and 289 positions are highlighted in dark blue. The footprint of 2G12 Fab_2 is shown for orientation. C) The network of glycans surrounding the 2G12 epitope, and their conservation (Figure 4.1). D) The binding of 2G12 to glycan knockouts and knock-ins was assessed by ELISA, mean ±SEM.

not directly contributing to the 2G12 epitope, the N448A and N386A knockouts diminished binding to a similar extent as a secondary binding site deletion. In contrast, the N411A mutation increased 2G12 binding (Figure 5.1D). Additionally, the impact of filling the nearby 241/289 glycan hole on 2G12 binding was determined. Neither the +N241 or +N289 glycan

knock-ins, nor the +N241 N289 double knock-in had a substantial impact on 2G12 binding (Figure 5.1D).

5.2 Overall resilience of the mannose patches to glycan addition and deletion

To assess whether the observed differences in 2G12 binding could be attributed to differences in glycan processing, the glycosylation profiles of the glycan site mutants were compared to that of BG505 SOSIP.664 trimers (Figure 5.2). As the removal of glycan sites can conceivably negatively impact trimer integrity, a quaternary structure-dependent antibody (PGT145), targeting a region distal to the site of mutation, was used for purification. The N-glycans from the target glycoproteins were enzymatically released, fluorescently labelled, and analysed by HILIC-UPLC (Figure 5.2A and B). Quantification of oligomannose-type species was performed by the integration of chromatograms before and after digestion with Endo H, as previously described.

As per the BG505 SOSIP.664 trimers, the chromatograms of all the glycan mutants were dominated by oligomannose-type glycans, though the distribution of individual oligomannose-type glycans, particularly $Man_9GlcNAc_2$ and $Man_8GlcNAc_2$ varied slightly (Figure 5.2). Previous site-specific glycan analyses of BG505 SOSIP.664 trimers have reported the N295, N332, N339, N386, N392, and N448 sites to be occupied by oligomannose-type glycans, predominantly $Man_9GlcNAc_2$ [132, 274]. Accordingly, deletion of each of these sites resulted in a decrease in the abundance of both total oligomannose-type glycans and $Man_9GlcNAc_2$ structures (Figure 5.2D). The quantification of the oligomannose-type glycans from three biological replicates of BG505 SOSIP.664 in Chapter 3 gives confidence that the observed differences were not artefactual (coefficient of variation of 2.2% for $Man_9GlcNAc_2$ and 1.5% for total oligomannose-type glycans; Figure 3.3). The largest changes were observed for the N448A and N386A glycan-knockouts, which resulted

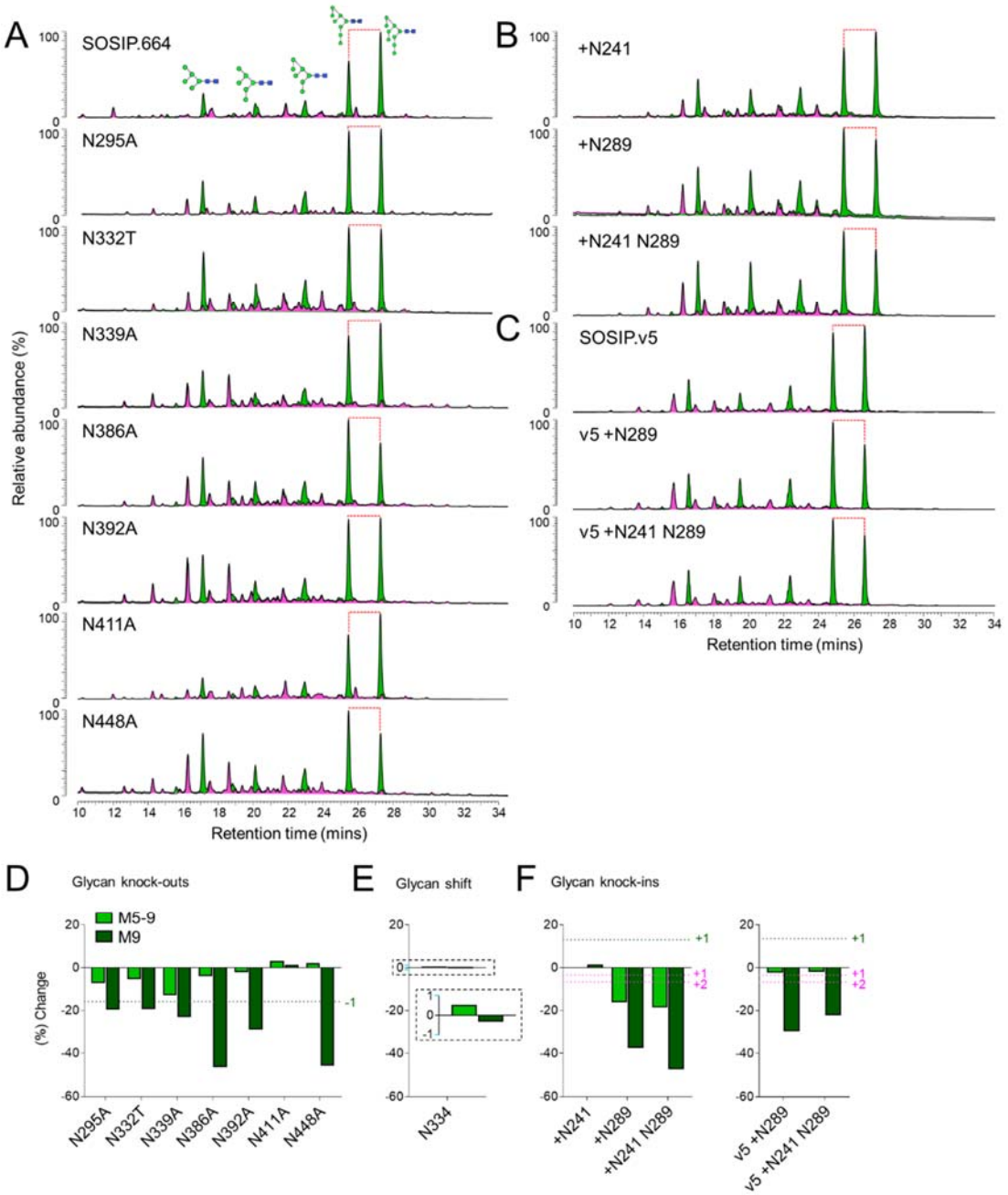


Figure 5.2. Impact of glycan site additions and deletions on overall glycosylation profiles. HILIC-UPLC profile of fluorescently labelled N-linked glycans released from A) BG505 SOSIP.664 glycan site knockouts and B) glycan site knock-ins, and C) BG505 SOSIP.v5 glycan site knock-ins. Oligomannose-type glycans (green) were quantified by integration of peaks before and after digestion with Endo H (magenta). D-F) The impact of glycan site deletion, shift, or addition on the abundance of $Man_9GlcNAc_2$ (M9) and total oligomannose-type glycans (M5-9). Values represent the percentage change in abundance, relative to BG505 SOSIP.664: percentage change = [(% glycan mutant - % BG505 SOSIP.664)/% BG505 SOSIP.664] × 100. The dashed green lines represent the change in $Man_9GlcNAc_2$ expected upon either the deletion (D) or addition (F) of a glycan site comprising solely $Man_9GlcNAc_2$ structures. The dashed magenta line (F) represents the decrease in the abundance of $Man_9GlcNAc_2$ expected upon the addition of one or two sites comprising solely complex-type glycans. The data for panel E was obtained from Behrens, 2017 [366]; the commonality in methodologies used justifies this comparison.

in a 46% and 47% decrease in the abundance of $Man_9GlcNAc_2$, respectively. Many of the observed decreases were somewhat larger than the decrease predicted upon the loss of a glycan site comprising solely $Man_9GlcNAc_2$ structure (16%; Figure 5.2D, dashed line). Thus, glycan site deletion on BG505 SOSIP.664 trimers can result in widespread increased glycosylation processing, as the surrounding glycans become more accessible to the host's processing enzymes [310]. In this context, it has previously been shown that the N332 to N334 glycan 'shift' escape mutation, which maintains local glycan density, has negligible impact on glycosylation processing (Figure 5.2E) [366]. However, the effect was not universally observed, for example, the N411A glycan site knockout had minimal impact on overall glycosylation processing despite its location at the centre of the IMP (Figure 5.1 and 5.2D).

Given that glycan site deletion generally resulted in increased glycosylation processing, and a glycan shift mutation did not impact glycosylation processing, it was hypothesised that glycan site addition may restrict processing. Analysis of the +N241 glycan knock-in, however, revealed only a minimal increase in the relative abundance of $Man_9GlcNAc_2$ structures (2%; Figure 5.2F). In contrast, the +N289 and +N241 N289 knock-ins resulted in a decrease in both the abundance of oligomannose-type glycans, and $Man_9GlcNAc_2$ structures. This result may be expected if the knocked-in sites were composed of predominantly complex-type glycans. However, the observed decrease in the abundance of $Man_9GlcNAc_2$ (38% and 48%, respectively) far exceeds the predicted decreases upon the addition of one or two sites containing only complex-type glycans (3% and 7%, respectively; Figure 5.2F, dashed line). Thus, both glycan site deletions and additions including the N289 site appear to be increasing the glycosylation processing of the trimer.

A possible explanation for the unexpected increase in processing is that the BG505 SOSIP.664 trimer may be unable to accommodate the N289 glycan, and that its addition may be inducing conformational changes which in turn influence glycosylation processing.

To address this, the analysis of the +N289 and +N241 N289 knock-ins was repeated on hyper-stabilised BG505 SOSIP.v5 trimers (Figure 3.8), which incorporate further stabilising mutations, including an additional inter-subunit disulphide bond, and display reduced conformational flexibility [225]. While the decrease in the abundance of $Man_9GlcNAc_2$ on the SOSIP.v5 background was not as severe as on the SOSIP.664 trimers, it still exceeded the decrease predicted if the knocked-in sites were composed of only complex-type glycans (Figure 5.2F, dashed line).

5.3 Glycan deletion increases mannose trimming throughout the trimer

To elucidate the impact of glycan site mutations at the site-specific level, site-specific LC-ESI MS analysis was performed on the glycan knockout and knock-in constructs. To aid the assignment of glycopeptides, a glycan library was generated by ion-mobility mass spectrometry analysis of an aliquot of unlabelled N-glycans from the BG505 SOSIP.664 protein (Appendix D). The reproducibility of the LC-ESI MS analysis of three biological replicates of BG505 SOSIP.664 in Chapter 3 confirms that observed differences were not due to experimental variation (Figure 3.4).

The deletion of glycan sites from the network of glycans surrounding the 2G12 epitope generally resulted in increased glycosylation processing at the immediately adjacent sites. This was most significant for the N332T, N411A and N448A glycan deletions, surrounding the N295 glycan site. The loss of each of these glycans resulted in a 61, 80 and 78 percentage point (pp; the arithmetic difference between two percentages) decrease in $Man_9GlcNAc_2$ at the N295 site, respectively, generally accompanied by a compensatory increase in $Man_{5\text{-}8}GlcNAc_2$ structures (Figure 5.3). The effect was reciprocal, though not equal, with the N295A glycan knockout increasing processing at the N332 and N448 sites (18 and 37 pp decrease in $Man_9GlcNAc_2$, respectively; Figure

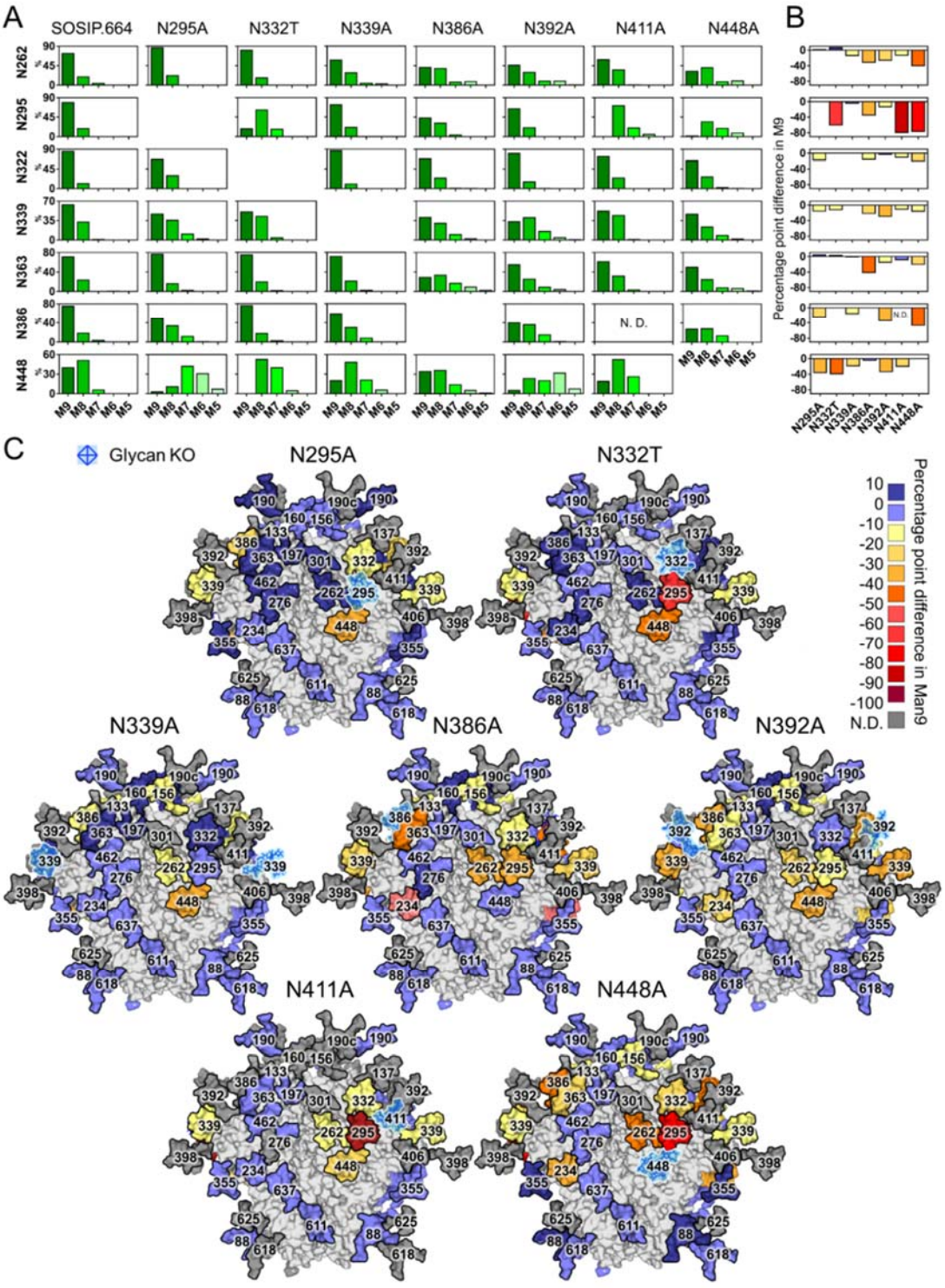


Figure 5.3. Glycan deletion increases mannose trimming throughout the trimer. A) Relative quantification of IMP sites from BG505 SOSIP.664 and the N295A, N332T, N339A, N386A, N392A, N411A and N448A glycan knockouts. M9, $Man_9GlcNAc_2$ (dark green) to M5, $Man_5GlcNAc_2$ (pale green). B) The percentage point difference in the abundance of $Man_9GlcNAc_2$ at IMP sites in the glycan knockouts, compared to BG505 SOSIP.664. Decreases in the abundance of $Man_9GlcNAc_2$ are coloured as per the key in panel C. C) Heat map demonstrating the percentage point difference in the abundance of $Man_9GlcNAc_2$ at each site in the glycan knockouts, compared to BG505 SOSIP.664. Differences are calculated as follows: (% $Man_9GlcNAc_2$ in knockout − % $Man_9GlcNAc_2$ in BG505 SOSIP.664). KO, knockout; N.D., not determined.

5.3). Similarly, the N386A mutation resulted in a 42 pp decrease in $Man_9GlcNAc_2$ at the adjacent N363 site, and the N392A knockout increased processing at the surrounding N339 and N386 sites by 30 and 35 pp, respectively (Figure 5.3).

Increased glycosylation processing, however, was not entirely limited to sites adjacent to the glycan knockout. In many instances, glycan site deletion resulted in increased processing emanating from the point of knockout. For example, the N332T knockout also increased processing at the N448 and N339 sites (40 and 13 pp decrease in $Man_9GlcNAc_2$, respectively; Figure 5.3). As before, the effect appeared reciprocal, as the N448A knockout increased processing at the N332 site (21 pp decrease), located beyond the N295 glycan (Figure 5.3). Additionally, the N234 glycan, which forms a small cluster with the N276 glycan distal glycan to the IMP, demonstrated increased processing upon the deletion of the N386, N392, or N448 sites (Figure 5.3).

In Chapter 4, the absence of the N234 site was typically associated with increased processing of the N276 site. Here, this relationship was explored further by the analysis of the N234A glycan knockout (Figure 5.4). The deletion of the N234 glycan site resulted in significantly increased processing at the N276 site, causing the complete conversion of oligomannose-type glycans to complex-type. The dominant peak shifted from a $Man_8GlcNAc_2$ to a bi-antennary structure. Despite its increased processing, the N276 site remained predominantly afucosylated (76%), consistent with observations in Chapter 4. The association between the two glycans can be rationalised by their very close proximity. In BG505 SOSIP.664 trimers the Asn Cα positions are separated by only 10.6 Å (PDB 5ACO). Furthermore, Stewart-Jones *et al.* have observed how these two glycans can interact via their oligomannose branches in X1193.c1 SOSIP.665 (clade G) trimers [280].

Taken together, the results reveal the varying impact of glycan site knockouts on the processing of glycans across the glycan network, which may relate to underlying structural interactions between the glycans [396].

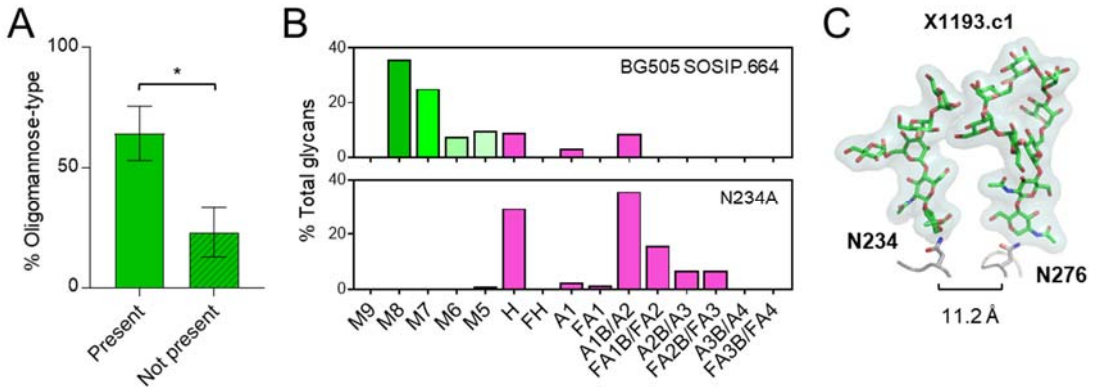


Figure 5.4. Processing of the N276 site. A) The abundance of oligomannose-type glycans at the N276 site is higher in strains where the N234 site is present ($p = 0.0422$), related to Chapter 4. Data were compared using an unpaired t-test in GraphPad Prism (version 7.05). Mean ±SEM is plotted. B) Relative abundance of the glycans at the N276 site on BG505 SOSIP.664 trimers, and upon deletion of the N234 glycan. M9, $Man_9GlcNAc_2$ (dark green) to M5, $Man_5GlcNAc_2$ (pale green); complex-type glycans (magenta) are grouped according to their number of antenna (A1-4) and/or presence of a bisecting GlcNAc (B) and/or the presence of core fucose (F). C) Crystal structure of the X1193.c1 trimer (PDB 5FYJ) illustrating the proximity of the N234 and N276 glycans, distance between Asn Cα is annotated.

5.4 Differential effects of glycan additions on processing

In line with high glycan density limiting mannosidase trimming, glycopeptide analysis of the glycan knock-in constructs revealed that the +N241 glycan site addition restricted processing at the neighbouring N448 site, resulting in a 21 pp increase in the abundance of $Man_9GlcNAc_2$ (Figure 5.5).

In contrast, the +N289 site knock-in resulted in increased glycosylation processing, both at sites neighbouring the introduced glycan and throughout the trimer. The N262, N295 and N332 sites all displayed a decrease in the abundance of $Man_9GlcNAc_2$ (37, 20, 28 pp decrease, respectively; Figure 5.5). The N363 and N234 sites were also affected to a similar extent, despite their greater distance from the +N289 knock-in. The +N241 N289 double glycan knock-in displayed both restricted glycosylation processing at the N448 site (12 pp increase in $Man_9GlcNAc_2$), and increased processing at the N262, N295, N332 and N363 sites (Figure 5.5).

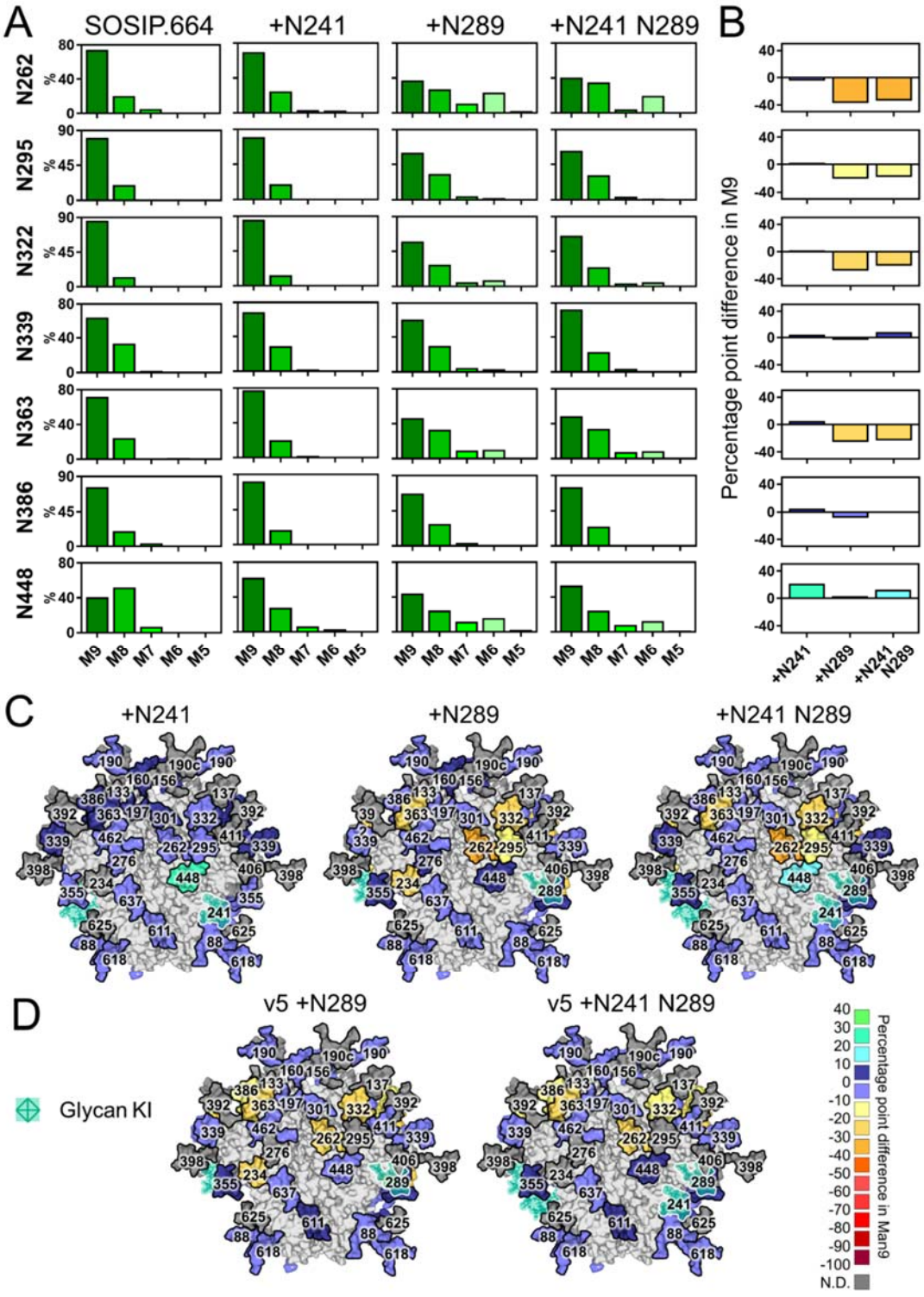


Figure 5.5. Differential effect of glycan additions on glycosylation processing. A) Relative quantification of IMP sites from BG505 SOSIP.664 and the +N241, +N289 and +N241 N289 glycan knock-ins. M9, $Man_9GlcNAc_2$ (dark green) to M5, $Man_5GlcNAc_2$ (pale green). B) The percentage point difference in the abundance of $Man_9GlcNAc_2$ in the glycan knock-ins, compared to BG505 SOSIP.664. Graphs are coloured according to the key in panel D. C) Heat map demonstrating the percentage point difference in the abundance of $Man_9GlcNAc_2$ at each site in the glycan knock-ins, compared to BG505 SOSIP.664: (% $Man_9GlcNAc_2$ in knock-in − % $Man_9GlcNAc_2$ in BG505 SOSIP.664). D) Percentage point difference in the abundance of $Man_9GlcNAc_2$ at each site in the SOSIP.v5 glycan knock-ins compared to BG505 SOSIP.v5. KI, knock-in.

The increased glycan processing associated with the +N289 glycan knock-in is surprising, given that high glycan density is generally associated with restricted glycosylation processing. It was hypothesised that the BG505 SOSIP.664 protein might be unable to accommodate a glycan at this site, and thus the addition of a glycan may be causing wider conformational changes to the protein. The impact of these glycan knock-ins was therefore investigated on the hyper-stabilised SOSIP.v5 background [225] (Figure 3.8). In line with the HILIC-UPLC analysis, glycopeptide analysis confirmed that the addition of the N289 glycan to the BG505 SOSIP.v5 background resulted in increased glycan processing, though to a slightly less extent than that observed on the SOSIP.664 background (Figure 5.5D). The SOSIP.v5 +N241 N289 double glycan knock-in also exhibited changes in processing similar to that of the SOSIP.664 background, but not as pronounced.

The decrease in the total abundance of oligomannose-type glycans observed by HILIC-UPLC analysis could be, at least partially, explained by the addition of a site(s) comprising predominantly complex-type glycans. However, the precise compositions of the N241 and N289 glycan additions could not be readily determined as they co-occupy peptides with the N234 and N295 sites, respectively. To classify the glycan-type occupying these sites, the glycopeptides were subjected to sequential digests with Endo H (which cleaves oligomannose-type glycans) and PNGase F (to cleave remaining complex-type glycans) [274]. Both the N241 and N289 sites were found to be almost exclusively occupied by oligomannose-type glycans, irrespective of SOSIP.664 or SOSIP.v5 background (Figure 5.6), confirming that the observed decreases in the abundance of oligomannose-type glycans, particularly $Man_9GlcNAc_2$ structures, were solely due to increased processing at other glycan sites on the trimer.

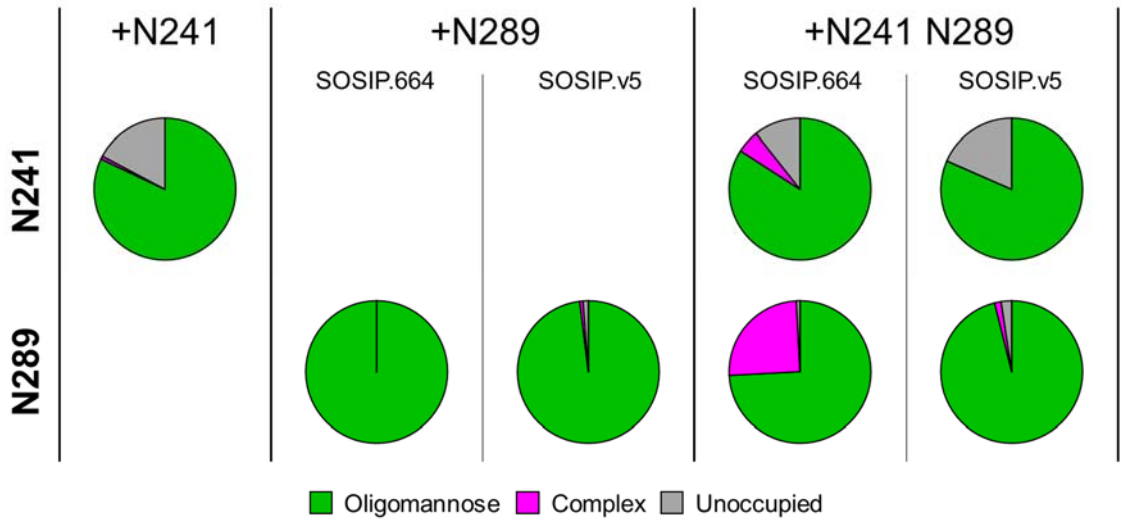


Figure 5.6. Glycan-type occupying the knocked-in N241 and N289 sites. Glycopeptides were subject to sequential digestion with Endo H and PNGase F (in the presence of O^{18}-labelled water) and analysed by LC-ESI MS. This generates novel mass signatures on the peptides to allow broad classification into oligomannose-type glycans (green), complex-type glycans (magenta), or unoccupied peptides.

5.5 A network of glycans preserves the PG9 and PG16 epitopes

A large proportion of glycan-targeting bnAbs recognise the glycans of the V1/V2 loops, located at the trimer apex [287]. This class of bnAbs are typified by the PG9 and PG16 antibodies, which contain very long HCDR3, allowing for penetration of the N160 glycan triad at the trimer apex (Figure 5.7D) [273, 288]. The N160 glycan sits in the centre of a network of glycans spanning all three protomers, including the highly conserved N156 and N197 sites (96% and 98%, respectively; Figure 5.7B).

In contrast to 2G12, which was largely tolerant of glycan site deletions, the binding of PG16 and, to a lesser extent, PG9 was significantly reduced upon the loss of the N156 glycan (Figure 5.7E). As before, this is somewhat expected as the N156 glycan directly contributes to the antibodies' epitopes [273, 288]. However, consistent with previous reports, the N197A glycan knockout also reduced antibody binding [132]. Given that the N197 glycan does not contribute to either antibody epitope it was hypothesised that the knockout may be

disrupting the glycosylation processing at the epitope; glycopeptide analysis was performed to address this.

Removal of either the N156 or N197 glycan sites resulted in increased processing at the proximal N160 site (Figure 5.7A and C). This was particularly true of the N156A glycan knockout, which resulted in the complete processing of $Man_9GlcNAc_2$ to smaller oligomannose-type structures and complex-type glycosylation, such that the dominant peak shifted from $Man_9GlcNAc_2$ to afucosylated bi-antennary structures (Figure 5.7A). The N197A glycan knockout resulted in increased oligomannose trimming at the N160 site, resulting in $Man_8GlcNAc_2$ predominating (Figure 5.7A). The N156A and N197A glycan knockouts also affected each other reciprocally, with each resulting in a slight loss of $Man_9GlcNAc_2$ at the other site (13.3 and 16.4 pp, respectively; Figure 5.7A and B). The N363 site also exhibited slightly increased oligomannose trimming upon the loss of both glycan sites, though the rest of the trimer appeared unaffected.

5.6 Discussion

There are wide ranging influences of glycan additions and deletions on HIV-1 immune evasion, both in the context of natural infection and in immunisation regimens. At one level, the very high density of glycans on Env indicates a selective advantage in using glycans to evade elimination by the host immune system [351, 372]. However, the glycan shield doesn't simply evolve to a maximum number of glycans. Instead, the creation and filling of holes in the shield during the course of infection illustrates that active rearrangements are required for effective immune evasion [198-200, 204].

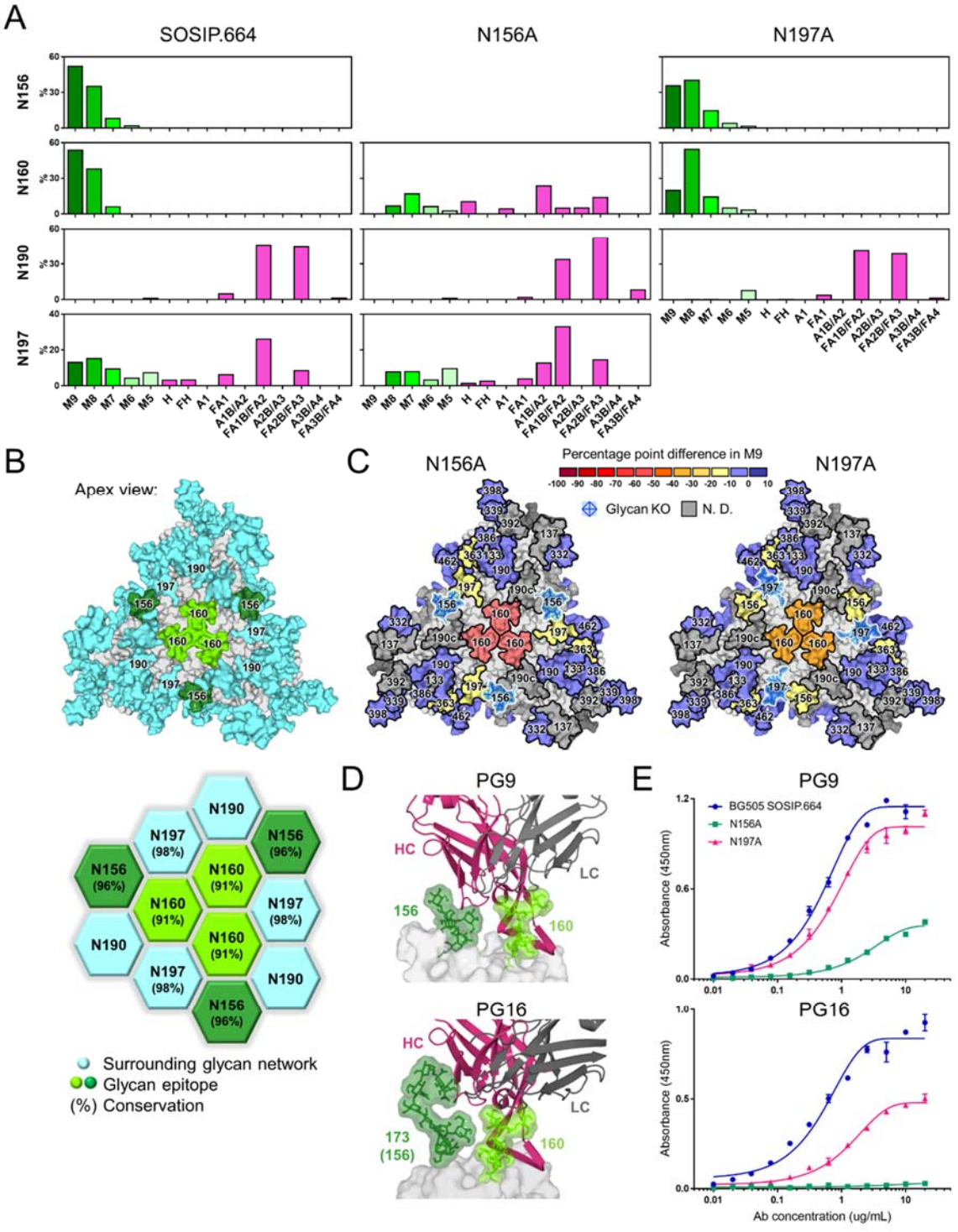


Figure 5.7. A network of glycans preserves the PG9 and PG16 epitopes. A) Relative quantification of apex glycan sites on BG505 SOSIP.664 and N156A and N197A glycan knockouts. B) Model of a fully glycosylated trimer (top; as described previously) illustrating the network of glycans at the trimer apex (bottom). (C) Heat maps displaying the percentage point difference in $Man_9GlcNAc_2$ on N156A and N197A glycan knockouts compared to BG505 SOSIP.664. D) Structure of the PG9 antibody in complex with the V1/V2 region of the CAP45 strain (PDB 3U4E) and PG16 in complex with the V1/V2 region of ZM109 (PDB 4DQO). E) ELISA data of PG9 and PG16 binding to BG505 SOSIP.664 and N156A and N197A glycan knockouts.

In the context of vaccination, the opening and closing of glycan holes in immunogens may prove a useful tool for driving the development of neutralisation breadth [359-363, 391]. The role of glycans in forming and blocking epitopes is of particular interest as bnAbs can evolve to recognise these structures [198]. In some instances, the precise processing state of the glycan target is essential for bnAb recognition and neutralisation, with glycan heterogeneity manifesting as < 100% neutralisation plateaus [259, 279, 289, 326]. The heterogeneity exhibited at a given glycan site can be influenced by the proximity of neighbouring glycans [274, 310, 328]. Using the BG505 SOSIP.664 trimers as a model system, this chapter sought to establish the impact of individual glycan site mutations on the glycan networks at two key antigenic regions of the HIV-1 glycan shield, the intrinsic mannose patch and the trimer apex.

The contribution of glycans to many bnAb epitopes has been defined by structural methods, such as X-ray crystallography and cryo-EM, complemented by glycopeptide analysis [125, 244]. For example, the 2G12 bnAb directly contacts the oligomannose-type glycans at four sites, though surrounding glycans are also thought to contribute. It has previously been shown that mutations affecting glycan sites lying outside of bnAb epitopes can disrupt binding and/or neutralisation [132, 253, 383]. In line with this, the deletion of glycan sites in the networks surrounding the 2G12, PG9 and PG16 glycan epitopes perturbed antibody binding, in some cases to the same extent as mutations directly impacting the epitope. Such observations may be partially explained by disruptions to the fine processing of the glycan epitope upon the mutation of proximal glycan sites, specifically increased trimming by mannosidases.

Despite causing significant disruption to the fine processing of the 2G12 epitope, the N411A knockout displayed increased 2G12 binding, consistent with previous reports [253]. The N411 glycan was unresolved in a cryo-EM structure provided by Dr Chris Cottrell and Prof. Andrew Ward (Scripps Research, La Jolla USA), though the asparagine residue was

observed to occupy the middle of 2G12-glycan complex. Thus, a glycan at this site could potentially result in a clash or the entropy of the glycan could be reduced upon 2G12 binding.

This study was restricted to the analysis of soluble BG505 SOSIP.664 immunogens, an important experimental model for viral glycosylation. However, soluble SOSIP.664 trimers display somewhat elevated oligomannose levels compared to virion-derived Env (discussed further in Chapter 7) [275, 276]. It could, therefore, be argued that mutations that increase the glycan processing of SOSIP.664 trimers are beneficial to generating immunogens that mimic the glycosylation of the native virus. However, oligomannose sites tend to be conserved between SOSIP.664 trimers and virion-derived Env [275, 276], thus the integrity of the mannose patch is associated with well folded trimers. Within the BG505 SOSIP.664 experimental system, substantial changes in the abundance of oligomannose-type glycans can indicate deviations away from native-like conformations [123].

In this context, it is possible that the mutation of N-linked glycan sites can induce glycoprotein misfolding or conformational changes [251, 310, 385, 386]. Here, the glycan mutants were purified using a quaternary structure-dependent antibody affinity step, PGT145 or PGT151 [278, 300], to minimise the contribution of misfolded proteins to the analysis. Accordingly, all the mutants displayed a glycosylation profile dominated by oligomannose-type glycans, a signature of native-like trimer configuration [122, 123, 323]. However, the unexpected increase in glycosylation processing observed upon the knock-in of the N289 glycan, in both the SOSIP.664 trimers and hyper-stabilised SOSIP.v5 trimers, suggests that this mutation may be causing a degree of protein instability. Notably, the +N289 knock-in was generated by mutating a proline at the 291 position. Ongoing work will explore the glycosylation of +N289 knock-ins known to display native-like structures. Nevertheless, the results highlight the importance of characterising the glycosylation of all candidate immunogens in depth [331].

Understanding the interdependence of glycans and their processing states is important in revealing how viral mutations can influence distant epitopes. Similarly, in immunogen design, the presence or absence of holes in the glycan shield could have wider antigenic and immunogenic consequences through glycan-glycan network effects. While the processing state of some glycans is key to the formation of bnAb epitopes [259, 279], the network may be dominated simply by the presence or absence of glycans independent of their processing state. Consistent with this view, both the deletion of glycan sites surrounding the 2G12 epitope, and the addition of the N289 glycan, increased the mannosidase trimming of the epitope. However, only glycan site knockouts had a measurable impact on 2G12 binding. Thus, the epitope is likely maintained by the network of glycans surrounding the epitope providing structural support to the four glycans directly involved in 2G12 binding.

The same network effects appear to apply to the epitopes of the apex binding antibodies, PG9 and PG16. Previous studies have reported the dependence of apex targeting bnAbs on sialylated glycans at the N156 site [273, 288, 290]. However, glycopeptide analyses presented both here and previously, report this site to be predominantly occupied by oligomannose-type glycans [122, 132, 274, 275]. Thus, these bnAbs are able to tolerate glycan heterogeneity. The considerable loss of binding upon glycan site deletion is therefore symptomatic of the structural role the network of surrounding glycans plays in stabilising this glycan epitope. At the apex in particular, the impact of the deletion of an individual glycan site may be amplified three-fold.

While it is not yet known the extent to which a successful vaccine candidate must display precise glycan epitopes, the results presented here shed light on the role of individual glycan sites in the fine processing of bnAb epitopes. A growing understanding of the factors shaping the glycosylation of Env will aid the continued development of HIV-1 immunogens.

6. Influence on the site-specific processing of the glycan shield by bnAb co-expression

It is widely believed that a successful HIV-1 vaccine will be capable of inducing bnAbs, similar to those developed by a subset of infected individuals after several years of infection. The Env glycoprotein is the sole target for bnAbs. Env is extensively glycosylated, and consequently most bnAbs are capable of recognising glycans to varying extents. Generally, bnAbs can be categorised by their recognition of the glycans at four distinct regions of the trimer: the outer domain of gp120, the trimer apex, the gp120-gp41 interface and proximal to the CD4bs (Figure 1.8) [273, 278, 287-289, 297, 299, 300, 304]. Defining the glycan epitopes on HIV-1 Env is a major research focus; the rationale being that in order to elicit an antibody response against a specific glycan epitope, that epitope must be presented to the immune system.

Classically, the precise glycan structures recognised by bnAbs have been defined using structural methods and glycan arrays, supplemented with glycopeptide analysis [125, 132, 244, 290]. There are, however, discrepancies between the glycans present on Env (according to glycopeptide analysis) and the preferred glycan epitopes of bnAbs (according to arrays and structural data). The most prominent example of this being the apex-targeting bnAbs, PG9 and PG16, which were reported to preferentially bind α2-6-linked sialylated hybrid-type glycans at the N156 site in structural studies with a V1/V2 scaffold construct [273, 288]. Furthermore, sialylated glycans were found to be key in the development of

another apex-targeting bnAb lineage [290]. However, these bnAbs are able to cross react with native-like Env mimics despite the absence of such structures at the N156 site, which is instead occupied by oligomannose-type glycans [132, 274]. Thus, the best glycan structure for the initiation and development of a bnAb response may not be one that is preferentially displayed on Env immunogens. Strategies to modify the glycan epitopes of Env could therefore prove a useful tool in the development of glycan-targeting Env immunogens.

The Env glycan shield can be manipulated by many means, including: expression with glycosylation enzyme inhibitors (e.g. kifunensine), expression in glycosyltransferase knockout cell lines (e.g. GnT $I^{-/-}$ cells), co-transfection with glycosyltransferase-encoding plasmids, and *in vitro* enzyme reactions [313, 387, 397]. The beneficial impacts of glycan engineering on antibody binding and neutralisation have been well documented (Table 6.1). For example, treatment with kifunensine, which prevents trimming of $Man_9GlcNAc_2$ structures, increases the sensitivity of the virus to neutralisation by PGT121, PGT125, PGT128, and 35O22 [304, 387]. Similarly, pseudovirions produced in GnT I deficient cell lines, which generate $Man_{5\text{-}9}GlcNAc_2$ glycoforms, are more susceptible to neutralisation by the CD4bs-targeting bnAb, VRC01, and the apex-targeting PGT145 [387]. In line with the aforementioned glycan preference of the apex-targeting PG9 and CAP256.09 bnAbs, their neutralisation was reportedly 30-fold more potent when pseudovirions were co-transfected with an α2-6 sialyltransferase [387].

However, many of the abovementioned glycan engineering techniques, particularly those acting early in the glycan pathway, alter the glycosylation processing of the entire glycoprotein. As a result, the same treatments that increase neutralisation sensitivity to one bnAb, can result in decreased sensitivity to another (Table 6.1). For example, kifunensine treatment decreases the sensitivity of HIV-1 to neutralisation by some apex-targeting bnAbs, and expression in GnT $I^{-/-}$ cells decreases the neutralisation sensitivity of PG9, PG16 and PGT151 [278, 289, 299, 387].

Table 6.1. Examples of the differential impact of glycan-engineering on bnAb neutralisation.

Treatment (glycoform)	Increased neutralisation	Decreased neutralisation
Expression with Kifunensine ($Man_9GlcNAc_2$)	2G12 35022 PGT121 PGT125 PGT128 VRC01	PG9 PG16 PGT151
Expression with Swainsonine ($Man_{5\text{-}9}GlcNAc_2$ and hybrid-type)	35022 PG9 PGT121 PGT145	PGT128 PGT151
Expression in GnT $I^{-/-}$ cells ($Man_{5\text{-}9}GlcNAc_2$)	35022 PGT121 PGT125 PGT128 PGT145 VRC01	PG9 PG16 PGT151

References: [289, 299, 304, 387].

Alternatively, the glycosylation processing of individual sites can be influenced by controlling the surrounding glycan density. Broadly speaking, the removal of a glycan site results in the increased processing of proximal glycans [274, 310, 328]. However, as discussed in the previous chapter, the mutation of glycan sites does not always have predictable effects. Furthermore, the mutation of a glycan site in order to alter the processing of a glycan epitope can have more severe consequences for the integrity of that epitope due to glycan network effects. The targeted site-specific manipulation of the Env glycan shield thus remains a challenge.

Using different glycan-targeting bnAbs to purify Env trimers may offer a solution, though this may only be applicable to certain trimer constructs (Section 3.3). Furthermore, given that the starting material has the same average glycosylation pattern, using bnAb affinity purification to bias the population towards certain glycoforms may result in the loss of considerable yields. With this in mind, this chapter explored the potential for using glycan-specific bnAbs to co-translationally alter the glycosylation of Env. Two main hypotheses were proposed for how bnAb co-expression might influence glycosylation processing.

Firstly, as the preferred glycan epitope is generated by the relevant glycosidases/glycosyltransferases, the bnAb will bind the epitope, 'locking' it in that form by protecting it from further processing. Alternatively, bnAb co-expression may accelerate Env folding, resulting in increased protection from early mannosidase trimming [122, 123]. The results suggest that the co-expression of BG505 SOSIP.664 trimers with glycan-targeting bnAbs can influence the glycosylation processing of the trimers, however the mechanism through which it achieves this may be slightly more complex.

6.1 Co-expression of BG505 SOSIP.664 trimers with bnAbs

To determine if the glycosylation processing of Env could be altered by co-expression with bnAbs, HEK cells were co-transfected with BG505 SOSIP.664 trimers and a panel of bnAbs targeting different regions of the trimer: the outer domain of gp120 (2G12, PGT135, PGT121), the trimer apex (PG16, PGT145), the gp120-gp41 interface (PGT151) and proximal to the CD4bs (VRC01 and VRC02). The cells were transfected with a 1:1 (w/w) ratio of BG505 SOSIP.664 (4:1, BG505 SOSIP.664:furin) and bnAb plasmids (2:1, light:heavy chain).

The trimer-bnAb complexes were co-purified via Protein A affinity chromatography. In this regard, the choice of BG505 SOSIP.664 trimers is key, as any observed differences in glycosylation are unlikely to be attributed to glycan biasing during purification (Chapter 3) [274]. The trimer-bnAb complexes were analysed by reducing SDS-PAGE gel (Figure 6.1) in order to separate the BG505 SOSIP.664 from the bnAbs, which have a very similar molecular weight. There was, however, an unexpected band at 140 kDa that likely corresponds to unreduced BG505 SOSIP.664. To minimise the possibility for contamination with unreduced antibody (150 kDa), only the gp120-containing band was excised for subsequent glycan analysis.

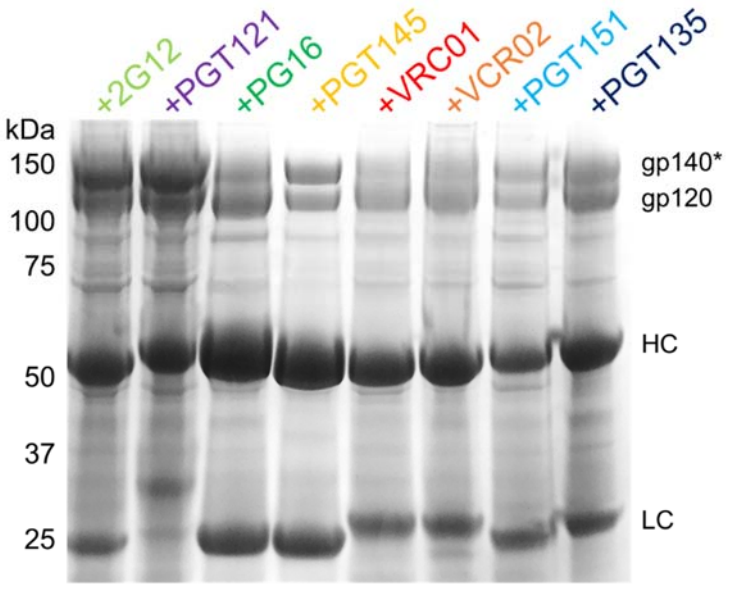


Figure 6.1. Co-expression of BG505 SOSIP.664 and bnAbs. HEK cells were co-transfected with the BG505 SOSIP.664, furin, bnAb heavy chain (HC), and bnAb light chain (LC) plasmids at a ratio of 40:10:17:33. The trimer-bnAb complexes were co-purified via Protein A chromatography and resolved by reducing SDS-PAGE. The band corresponding to gp120 was excised for HILIC-UPLC analysis. *Unreduced gp140 band possibly containing unreduced bnAb (150 kDa).

6.2 Glycosylation is influenced by antibody co-expression

Glycans were enzymatically released from the gp120-containing gel bands by digestion with PNGase F. The glycans were fluorescently labelled and analysed by HILIC-UPLC (Figure 6.2A). Integration of the peaks before and after digestion of the labelled glycans with Endo H enabled the quantification of oligomannose-type glycans (Figure 6.2B).

Overall, the chromatograms were highly similar. All the proteins contained a large proportion of oligomannose-type glycans, with a predominant peak corresponding to $Man_9GlcNAc_2$ (Figure 6.2). This is consistent with the trimers adopting native-like architectures [122, 123], and subsequently their ability to bind to bnAbs [74]. There were, however, subtle differences between the abundance of individual glycan structures. The difference was most pronounced for the $Man_9GlcNAc_2$ glycan, which varied from 17% for PGT135 co-expressed constructs, to 27% for PGT145 co-expressed constructs (Figure 6.2B).

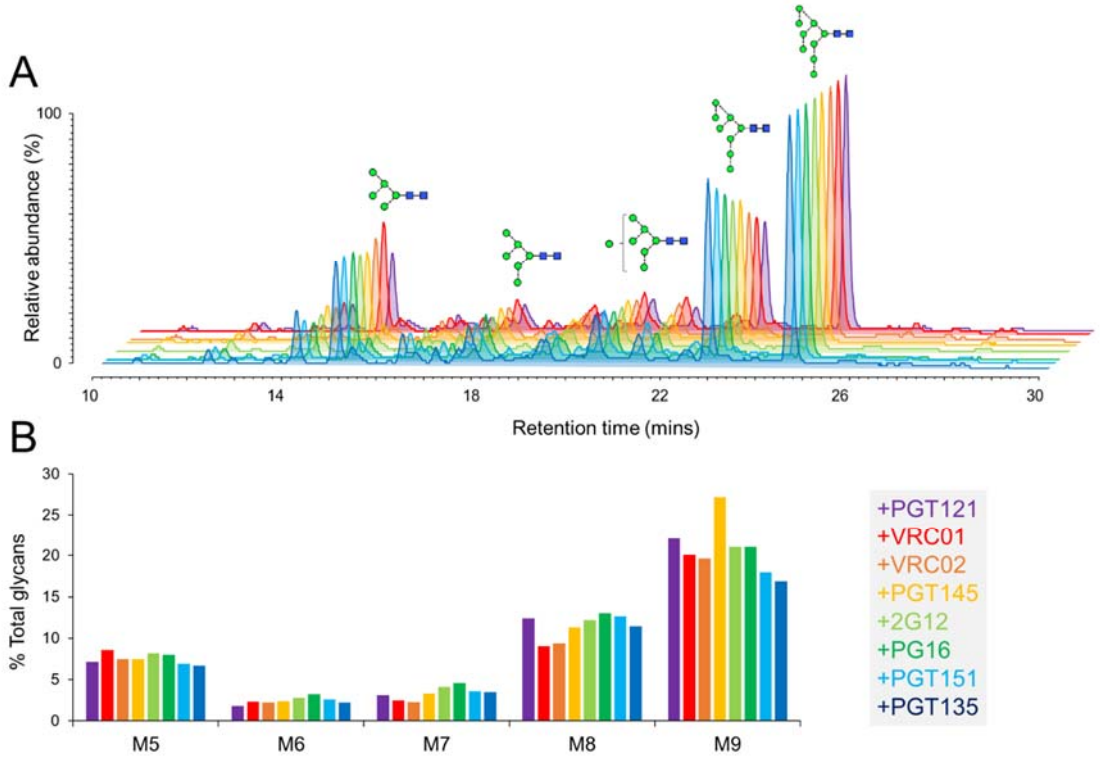


Figure 6.2. Glycosylation profiles of BG505 SOSIP.664 co-expressed with bnAbs. Bands corresponding to gp120 were excised, and N-linked glycans were enzymatically released by in-gel digestion with PNGase F. A) Glycans were fluorescently labelled and analysed by HILIC-UPLC. B) Abundance of oligomannose-type glycans was determined by integration of the chromatograms before and after digestion with Endo H. M5-9, $Man_{5-9}GlcNAc_2$.

To probe whether the subtle changes in glycosylation mapped to the bnAb epitopes, or otherwise, the proteins were subject to site-specific glycan analysis. The trimer-bnAb complexes were reduced, alkylated and digested with trypsin or chymotrypsin, prior to analysis by LC-ESI MS. There were several notable differences in the site-specific glycosylation of the trimers (Figure 6.3), which generally corresponded to the glycan sites directly contributing, or proximal, to the bnAb epitope.

6.2.1 Outer domain-targeting bnAbs

Three bnAbs targeting the outer domain of gp120 were chosen for co-expression: PGT121, PGT135 and 2G12. The PGT121 bnAb recognises the N137, N156, N301 and N332 glycans [264, 277, 278, 281]. It is the only bnAb of the panel to recognise the N301 glycan, which is typically occupied by $Man_{5-8}GlcNAc_2$ structures [132]. However, co-expression with

PGT121 appeared to restrict the processing of the N301 glycan, resulting in approximately 25% $Man_9GlcNAc_2$ structures at this site (Figure 6.4). No $Man_9GlcNAc_2$ was detected at this site on any of the other trimers. Unfortunately, the compositions of the N137 and N156 glycans, which also form a key part of the epitope could not be determined. However, it is noted that the N190 site also contained marginally more oligomannose- and hybrid-type glycans compared to the rest of the panel (Figure 6.3).

Co-expression with PGT135, which recognises the N295, N332, N386 and N392 glycans [259, 278, 279], did not appear to drastically alter the glycosylation processing at any site, aside from a small decrease in the abundance of oligomannose at the N160 site (Figure 6.3 and 6.4). The composition of the N392 glycan could not be determined, though the presence of $Man_9GlcNAc_2$ at this site has previously been reported to limit PGT135 binding [259, 279].

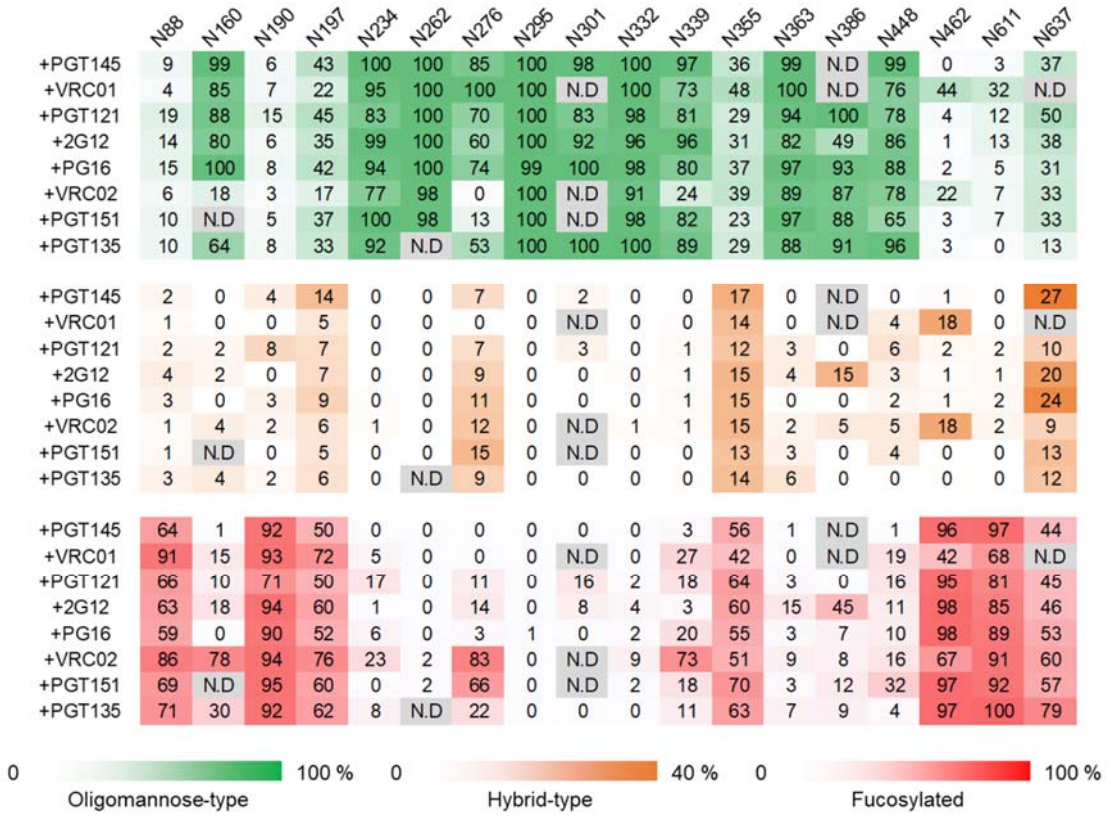


Figure 6.3. Site-specific glycosylation of BG505 SOSIP.664 trimers co-expressed with bnAbs. Trimer-bnAb complexes were reduced, alkylated and proteolytically digested prior to analysis by LC-ESI MS. The relative abundance of oligomannose-, hybrid-type, and fucosylated glycans at each site is given here.

The 2G12 epitope is comprised of the N295, N332, N339, and N392 glycans [247-253]. These sites are typically occupied exclusively by oligomannose-type glycans, and co-expression with 2G12 did not alter this (Figure 6.3). There was, however a notable decrease in the abundance of oligomannose-type glycans at the N386 site. The oligomannose content of this site is typically 100% (Figure 3.4) [132], and was > 80% for the other trimers in the panel, but was reduced to 49% when co-expressed with 2G12 (Figure 6.3). The N386 glycan forms part of the glycan network immediately surrounding the 2G12 epitope, and has been implicated in the stabilisation of the epitope (Chapter 5; Dr Chris Cottrell and Prof. Andrew Ward, Scripps Research, La Jolla, USA). The increased processing of this glycan could be attributed to changed conformational dynamics of the glycan upon antibody binding making it more accessible to glycan processing.

6.2.2 The gp120-gp41 interface

The PGT151 bnAb recognises tri- and tetra-antennary structures at the N611 and N637 sites [299-301]. However, co-expression with PGT135 did not appear to dramatically alter the processing of these sites. Structural studies suggest that the PGT151 light chain CDR loops are also in close proximity to, and possibly interact with, the N262 and N448 glycans from one protomer, and the light chain framework region with the N276 glycan from another protomer [300]. The oligomannose content of the N448 glycan, which is typically 100% in BG505 SOSIP.664 trimers (Figure 3.4) [132], was lowest for PGT151 co-expressed material (65%; Figure 6.3). Similarly, the N276 glycan contained an uncharacteristically low oligomannose content and a high abundance of fucosylated bi-antennary glycans (Figure 6.4). Again, it is unclear why the processing of sites located proximal to the bnAb epitope is increased upon co-expression.

6.2.3 CD4bs-targeting bnAbs

The VRC01 and VRC02 bnAbs, derived from somatic variants of the same IgG1 clone [398], arguably caused the largest disruptions in glycosylation processing. The VRC01 bnAb

recognises a largely protein-based epitope at the CD4bs, but has evolved to accommodate the N276 site immediately adjacent to the CD4bs [254]. The N276 site is typically occupied by a mixed population of glycans (Figure 3.4) [132]. However, when co-expressed with VRC01 this site was occupied by exclusively oligomannose-type glycans, with a large abundance of $Man_9GlcNAc_2$ glycans (65%, Figure 6.4). VRC01 was the only bnAb in the panel to have this effect. Similarly, the processing of the N462 glycan, also located very close to the CD4bs, was somewhat restricted. While this site is usually composed of predominantly complex-type glycans (Figure 3.4) [132], the oligomannose content reached 44%, mainly accounted for by $Man_5GlcNAc_2$ structures, with a further population of hybrid-type glycans (18%; Figure 6.3 and 6.4). In contrast, the processing of the N197 glycan, which also surrounds the CD4bs, was moderately increased (Figure 6.3). In a molecular dynamics simulation of the BG505 SOSIP.664 Env trimer the N197, N276, N363, and N462 glycans each had considerable overlap with VRC01 bound to the CD4bs [280].

The highly related VRC02 bnAb also restricted processing at the N462 site (40% oligomannose- and hybrid-type glycans; Figure 6.3 and 6.4). In contrast, the processing of the N276 site was dramatically increased. All oligomannose-type glycans were fully processed, with an increased abundance of fucosylated bi-, tri-, and even tetra-antennary glycans (Figure 6.3 and 6.4). The oligomannose-content of the N160 was also unusually low (18%), replaced by fucosylated bi- and tri-antennary structures (Figure 6.4). The processing of this site can provide a read out whether the trimer adopts a native-like architecture [122]. Thus, VRC02 binding may be inducing different trimer conformations. In line with this, the binding of soluble CD4 and some CD4bs-targetting antibodies can lead to the formation of open conformations of Env. In contrast, VRC01 appears to lock Env in a closed state [399].

6.2.4 Apex-targeting bnAbs.

Co-expression with the apex-targeting bnAbs, PG16 and PGT145, did not appear to have a major impact on the glycosylation processing of any site (Figure 6.3). Unfortunately, it was

not possible to determine the glycan composition of the N156 site. Structural studies with PG16 have reported a preference for α2-6-linked sialylated hybrid-type glycans at this site [273, 288]. However, given that PG16 can also bind to the oligomannose-type glycans typically observed at this site [132, 274], and which are formed earlier in the glycosylation pathway, it is unlikely that co-expression with PG16 would increase the abundance of α2-6-linked sialylated hybrid-type glycans at the N156 site. Of note is that hybrid-type glycans were somewhat enriched at the N637 site, located at the opposite end of the trimer, on both the PGT145 and PG16 co-expressed trimers (Figure 6.3).

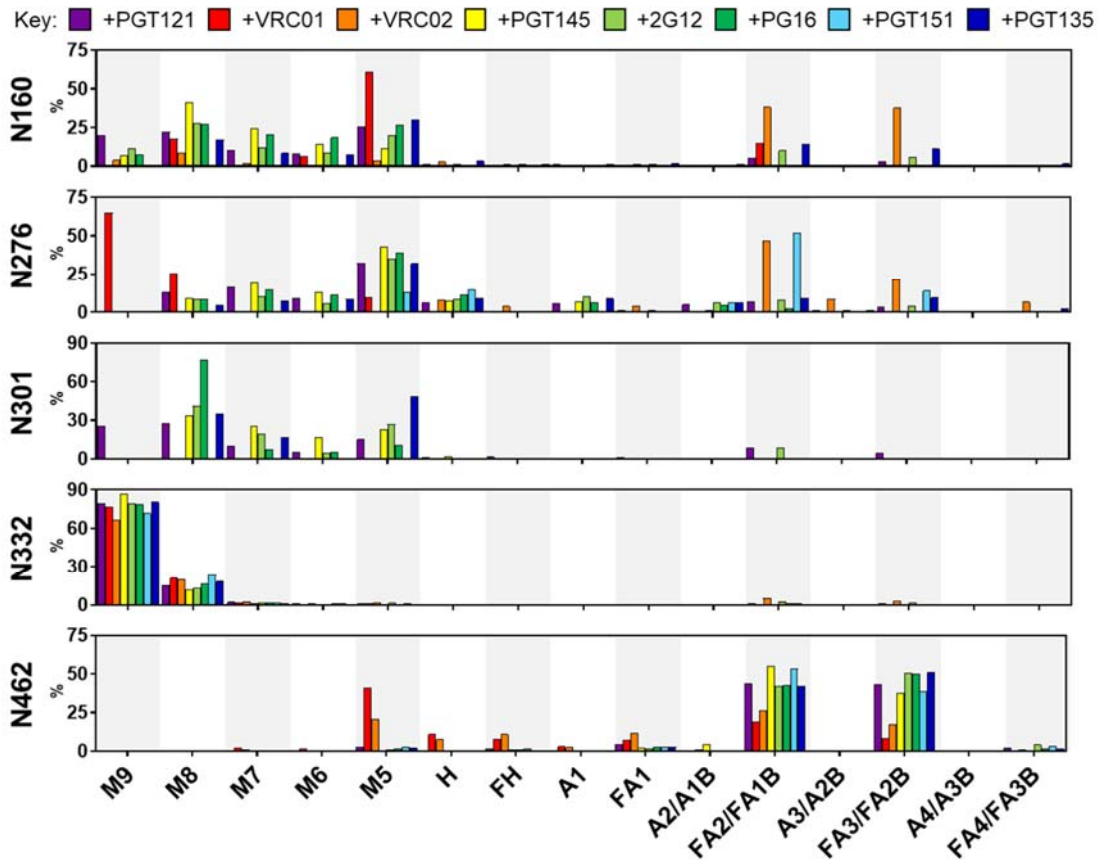


Figure 6.4. Site-specific glycosylation of BG505 SOSIP.664 co-expressed with bnAbs. The relative abundance of glycan structures at the N160, N276, N301, N332 and N462 sites on BG505 SOSIP.664 proteins co-expressed with bnAbs. Glycans are grouped according to Appendix C.

6.2.5 Non-Env targeting antibody

To test whether the observed differences in glycosylation could be directly attributed to bnAb binding, the experiment was repeated with co-expression of an antibody that is not directed against Env. HEK cells were co-transfected with BG505 SOSIP.664 and trastuzumab (Herceptin®) DNA, as described previously. Trastuzumab is an anti-breast cancer antibody directed to the human epidermal growth factor receptor 2. Additionally, BG505 SOSIP.664 was expressed alone. The trimers were purified by Ni^{2+} affinity chromatography and analysed by reducing SDS-PAGE. As before, bands corresponding to the gp120 subunit were excised for glycan analysis by HILIC-UPLC.

Again, the overall glycosylation profiles of the two BG505 SOSIP.664 trimers were extremely similar, though there were minor differences in the abundance of some of the complex-type glycans (Figure 6.5). The relative abundance of oligomannose-type glycans was not influenced by trastuzumab co-expression. The result suggests that there may be a small influence of antibody co-expression on glycosylation processing, irrespective of target protein-antibody interactions. This could perhaps be attributed to the increased burden placed on the ER and Golgi processing enzymes.

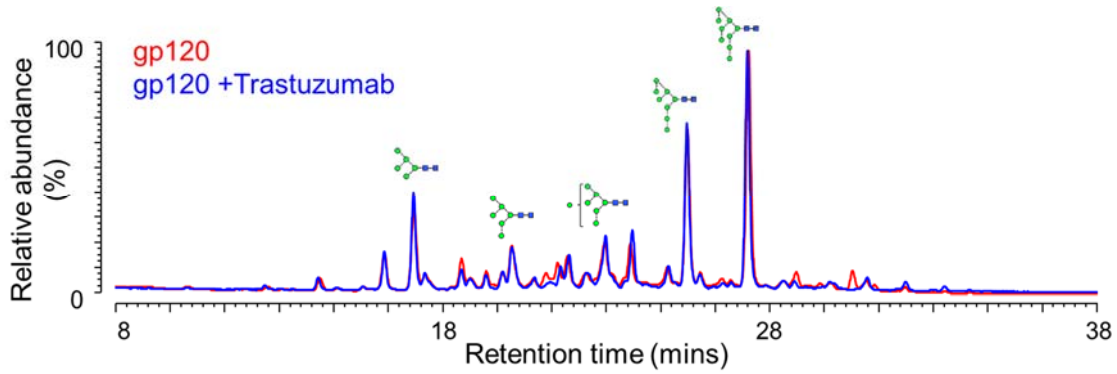


Figure 6.5. Glycosylation of BG505 SOSIP.664 co-expressed with trastuzumab. BG505 SOSIP.664 was expressed, or co-expressed with trastuzumab, in HEK cells. Trimers were purified by Ni^{2+} affinity chromatography and analysed by reducing SDS-PAGE. Glycans were enzymatically released from the gel bands corresponding to the gp120 subunit, fluorescently labelled, and analysed by HILIC-UPLC.

6.3 Discussion

The modification of the glycan shield is a powerful tool for altering the antigenicity and immunogenicity of Env trimers [305, 313, 387, 397]. To date, most of the research in this area has made use of glycosylation inhibitors or engineered cell lines; the outcome being alterations to the global glycosylation processing of the trimer that can benefit the formation of some bnAb epitopes, but adversely affect others [289, 299, 304, 387]. A strategy for the site-specific modification of glycan epitopes may therefore prove beneficial for the production of immunogens aimed at eliciting particular bnAbs [290, 387]. This chapter investigated using co-expression with glycan-dependent bnAbs to alter the glycosylation of BG505 SOSIP.664 trimers.

Co-expression with a panel of bnAbs targeting distinct regions of the trimer was found to influence the glycosylation processing of the trimers. Encouragingly, the effects largely mapped to glycan sites in the vicinity of the bnAb epitope, thus targeted alterations to bnAb epitopes may be possible. However, the use of some bnAbs resulted in considerable deviations in the processing of sites associated with the native-like architecture of BG505 SOSIP.664 trimers. Such effects could be indicative of non-native-like trimer architectures and consequently, sub-optimal immunogens [215].

It was hypothesised that co-expression with bnAbs would generally restrict glycan processing, either through the direct steric blocking of individual sites, or by increasing the protection of the trimer as a whole through accelerated folding. While this may be true for some bnAbs, others resulted in increased glycosylation at neighbouring sites. Furthermore, highly related antibodies appeared to have contradictory effects. It is unlikely that the observed differences were attributed to glycan biasing during purification [123, 274]. However, co-expression of the BG505 SOSIP.664 trimers with an anti-cancer antibody also imparted minor changes on glycosylation processing. Thus, there may be multiple mechanisms through which glycan processing may be altered. Nevertheless, this chapter

demonstrates that targeted alterations to the Env glycan shield can be achieved via co-expression with bnAbs, representing another tool for the glycan engineering of HIV-1 Env immunogens.

7. Immunogen mimicry of the viral glycan shield

Antigenic mimicry is fundamental to the majority of licensed vaccines. As the sole target for bnAbs, it is likely that a successful HIV-1 vaccine will be based on mimicry of the Env glycoprotein. Most bnAbs recognise epitopes that are composed, to varying degrees, of N-linked glycans, or have evolved to accommodate glycans adjacent to their epitopes (Table 1.1) [273, 278, 287-289, 297, 299, 300, 304]. It is therefore likely that Env-based immunogens, with glycosylation matching that of viral Env, will be required at some stage in the vaccination program to elicit bnAbs.

Native-like trimers are good structural and antigenic mimics of virion-derived Env [74, 125]. The observation that glycan-dependent bnAbs can bind to soluble Env mimics suggests that their glycosylation processing is largely in agreement with that of the virus. Indeed, glycan analyses of both virion-derived and recombinantly produced Env both report a large abundance of oligomannose-type glycans [110, 122, 123, 132, 242, 274, 306-322]. However, given that bnAbs can tolerate glycan microheterogeneity, cross-reactivity does not necessarily indicate that target glycans are precisely conserved between immunogen and virus. For example, some glycan specific bnAbs recognise the largely invariant base of the glycan [254, 286]. Defining the glycosylation on native, virion-derived Env will be valuable for guiding immunogen design.

In this chapter, the glycosylation of BG505 SOSIP.664 trimers is compared to Env derived from infectious virions of the same genotype. The results reveal that glycans relevant

to key bnAb epitopes are generally conserved between immunogen and virion, though the latter displays hotspots of elevated glycan processing.

7.1 Production of virion-derived Env

Virion-derived Env was provided by Dr Elena Chertova from the laboratory of Dr Jeffrey Lifson, Frederick National Laboratory for Cancer Research, USA, and was produced as follows. Briefly, an infectious stock of virus was derived by transfecting HEK (293T) cells with the infectious molecular clone (IMC), BG505.T332N-LAI (Figure 7.1E) [360]. The Env component of the IMC was derived from the BG505 sequence, with the N332 glycan knocked-in to mimic that of the BG505 SOSIP.664 trimer sequence [74]. An aliquot of cell-free HEK (293T) culture supernatant (infectious stock) was used to infect a CD4+ CCR5+ A66-R5 T cell line, to generate the virus required for glycan analysis. The use of the

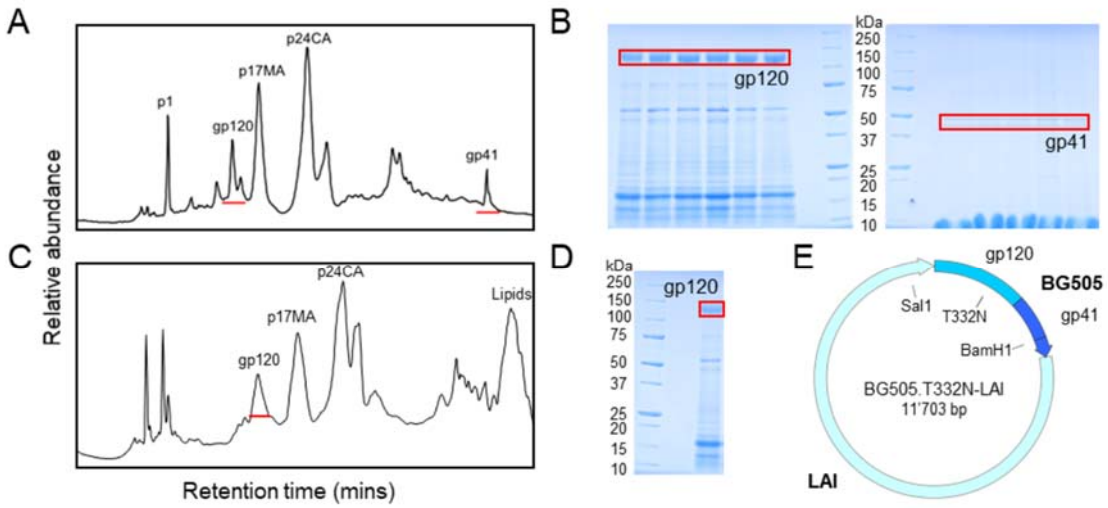


Figure 7.1. Construction of the BG505.T332N-LAI infectious molecular clone and purification of Env proteins. Virion derived-Env proteins were produced by Dr Elena Chertova and Dr Jeffrey Lifson, Frederick National Laboratory for Cancer Research, USA. The HPLC chromatograms and SDS-PAGE data were also provided by them. A) HPLC separation of the BG505.T332N-LAI virus preparation under non-reducing conditions. UV absorbance was measured at 206 nm. Viral protein peaks (identified by subsequent SDS-PAGE gel, protein sequencing and immunoblot analyses) are annotated. B) SDS-PAGE gel of gp120- and gp41-containing fractions, highlighted in red. Excised gel bands were used for subsequent HILIC-UPLC analysis of enzymatically released N-glycans. C) HPLC separation of a second aliquot of BG505.T332N-LAI virus preparation, as above. D) SDS-PAGE gel of pooled gp120 fractions. Approximately 70 μg of HPLC-purified gp120 was used for site-specific glycan analysis. E) Schematic representation of the BG505.T332N-LAI IMC, as described in Medina-Ramírez *et al.*, 2017 [360].

A66-R5 T cell line is justified, as the goal of vaccination is to elicit nAbs that recognise functional Env proteins present on infectious virions produced predominantly in human lymphoid cells. As the IMC was cytopathic, uninfected cells were added to the culture throughout the incubation in order to maintain virus production. After day 54, the culture supernatant was harvested for virion purification by sucrose density-gradient centrifugation, resulting in a 1'000-fold concentrated virus preparation. An aliquot of the virus preparation was disrupted in 8M Guanidine-HCl, and the proteins were separated by HPLC (Figure 7.1A). Fractions containing the gp120 and gp41 subunits were analysed by SDS-PAGE, and the gp120- and gp41-containing gel bands were excised for N-linked glycan analysis (Figure 7.1B). A further aliquot of the virus preparation was then purified, and the pooled gp120 fractions were used for site-specific glycan analysis (~70 µg; Figure 7.1C and D). Unfortunately, the gp41 fraction was insufficient for a similar analysis.

7.2 Relative abundance of oligomannose- and complex-type glycans

The excised gp120- and gp41-containing gel bands (Figure 7.1B) were subjected to in-gel PNGase F digestion. Glycans were fluorescently labelled, and analysed by HILIC-UPLC. The relative amounts of oligomannose- versus complex-type glycans on the virion-derived gp120 and gp41 subunits were compared to those from recombinant BG505 SOSIP.664 trimers produced transiently in HEK cells and purified with the PGT151 bnAb (Table 2.1).

The N-glycans on the virion-derived gp120 proteins were 50% oligomannose-type and 50% complex-type (Figure 7.2A), whereas the gp120 subunits from the SOSIP.664 trimers were enriched for oligomannose-type structures (63%, Figure 7.2B). Similarly, virion-derived gp41 had a lower overall content of oligomannose-type glycans (4%; Figure 7.2C) compared to gp41 from SOSIP.664 trimers (22%; Figure 7.2D). Thus, the N-linked glycans

on viral Env proteins are somewhat more processed than those on the corresponding recombinant trimer.

Furthermore, larger and more elaborate, complex-type glycans were more abundant on the virion-derived gp120 and gp41 subunits, compared to SOSIP.664 (magenta chromatograms, Figure 7.2). This is exemplified by N-glycans containing either α2-3- or α2-6-linked terminal sialic acids. Overall, the extent of gp120 sialylation was higher for BG505.T332N-LAI virions (16%) than it was for BG505 SOSIP.664 trimers (5%), while

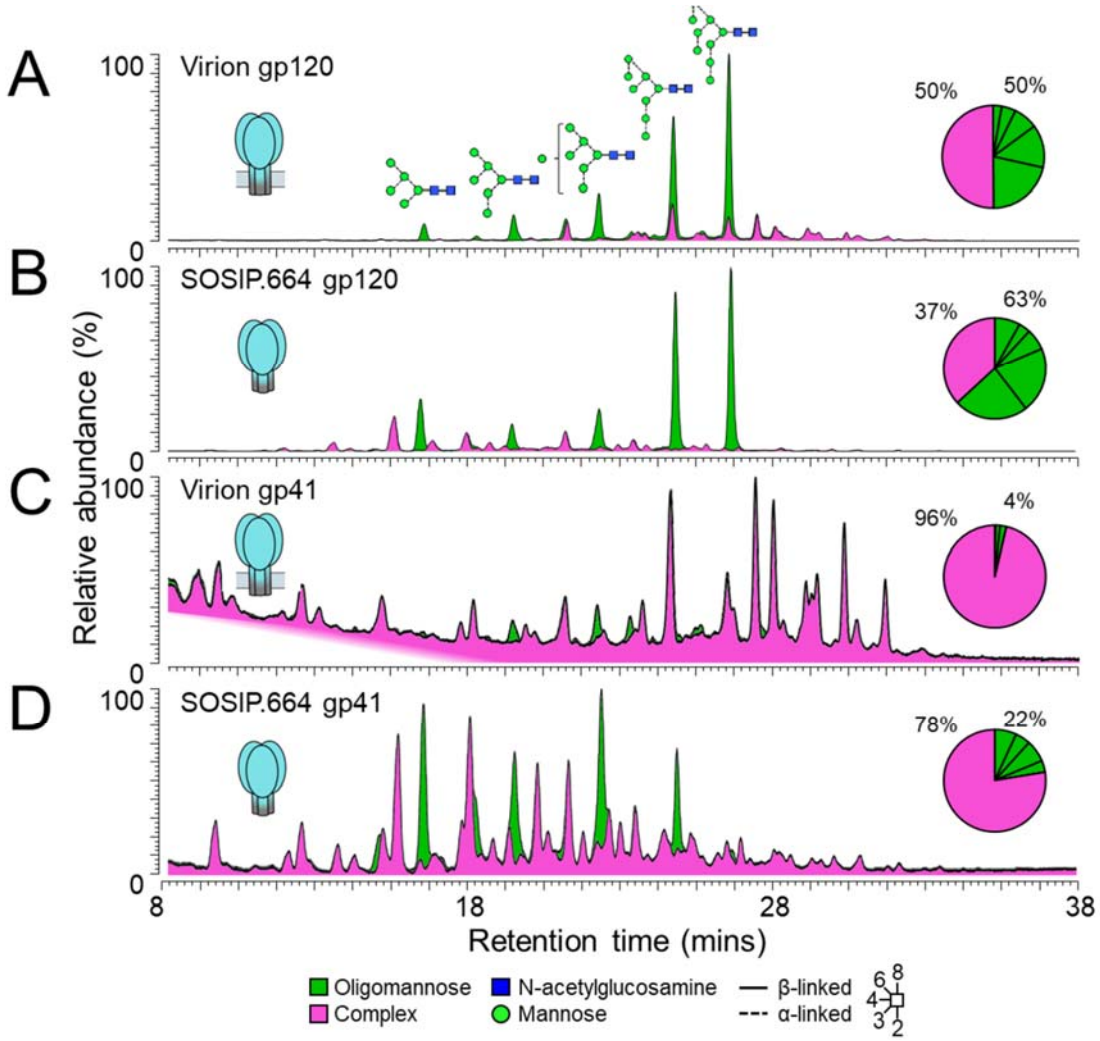


Figure 7.2. The abundance of oligomannose- and complex-type glycans on virion-derived Env and BG505 SOSIP.664. Fluorescently labelled N-linked glycans from the gp120 and gp41 subunits of virion-derived Env (A and C) or BG505 SOSIP.664 produced in HEK cells (B and D) were analysed by HILIC-UPLC. Oligomannose-type glycans (green) were quantified according to their susceptibility to digestion with Endo H (inset pie chart).

the corresponding values for the gp41 subunits were 26% and 8%, respectively. The ratio of α2-6/α2-3 sialic acids on the gp41 subunits was 1.8 for virions and 0.8 for HEK-derived trimers (Table 7.1), a finding consistent with earlier observations that α2-6-linked sialic acid moieties were prevalent on PBMC-derived virion gp41 [333]. Compared to gp41, the gp120 proteins from both Env proteins contained a smaller population of sialic acid-containing glycans (Table 7.1). This observation may reflect the generally lower efficiency of glycan processing on gp120 subunits than on gp41 subunits, irrespective of the Env protein source.

Table 7.1. Quantification of N-linked glycans by glycosidase digestion (% of total glycans).

		gp120		**gp41**	
		Virion-derived	SOSIP.664	Virion-derived	SOSIP.664
Oligomannose	M5	2.7	8.1	0.3	6.9
	M6	4.5	3.7	1.4	4.8
	M7	7.7	6.8	2.0	7.0
	M8	13.6	21.0	0.0	3.7
	M9	21.4	23.5	0.0	0.0
	TOTAL	**49.9**	**63.1**	**3.7**	**22.4**
Sialylated glycans	α2-3	9.6	3.3	9.3	4.2
	α2-6	6.6	2.1	16.9	3.4
	TOTAL	**16.3**	**5.3**	**26.1**	**7.5**

7.3 Divergence and conservation of fine processing at various gp120 glycan sites

To investigate whether trimeric immunogens correctly present key bnAb epitopes at a site-specific level, gp120 derived from virions and BG505 SOSIP.664 trimers were digested with trypsin and chymotrypsin and analysed by LC-ESI MS, as described previously (Figure 7.3). To aid in the assignment of structures, a glycan library was generated by IM-ESI MS analysis of a separate, unlabelled aliquot of enzymatically released N-linked glycans (Appendix D; Prof. David Harvey, University of Oxford, Oxford, UK).

A large number of bnAbs are directed against the oligomannose-type glycans of the outer domain of gp120 (Table 1.1). Comparing the fine processing states of N-linked glycan sites recognised by such bnAbs on the virion-derived trimer with the BG505 SOSIP.664 trimer reveals the overall integrity of oligomannose processing; $Man_9GlcNAc_2$ dominates at the majority of IMP sites on both constructs (Figure 7.3). This finding is consistent with the ability of bnAbs to both neutralise the virus and bind to trimeric immunogens. The only notable deviation in the fine processing of oligomannose-type glycans was seen at N363, where $Man_8GlcNAc_2$ predominated on virion-derived trimers, but $Man_9GlcNAc_2$ did so on the recombinant trimer.

Another region relevant to immunogen design strategies is the trimer apex, which contains highly conserved glycans involved in multiple bnAb epitopes, including PGT145, PG9, and PG16. Compared to recombinant gp120 monomers, the apex region of BG505 SOSIP.664 trimers is enriched for oligomannose-type glycans, attributed to steric constraints imposed by trimerisation [122]. Site-specific analysis was used to assess whether the same influence of trimerisation applied to virion-derived Env. Indeed, the high proportion of oligomannose-type glycans observed at the N160 site is a strong indicator that the same steric restrictions do in fact apply (Figure 7.3), however the distribution of individual oligomannose-type glycans differed slightly.

While the trimer-specific restriction in glycan processing at the conserved N160 site applies to both virion-derived Env and BG505 SOSIP.664 trimers, there were differences at other sites around the apex. Namely, sites on the SOSIP.664 trimer that present a mixed population of oligomannose-type and complex-type glycans (e.g. N197) were more highly processed on the BG505.T332N-LAI virus (Figure 7.3). In addition to elevated oligomannose processing, there was an increase in glycan branching at the N197 and N190 sites (Figure 7.3). The increased processing and branching are consistent with observations from the HILIC-UPLC analyses and are likely to contribute to the ~ 10% increase in

complex-type glycans on virion-derived gp120 compared to the SOSIP.664 trimers (Figure 7.2).

The CD4bs region also contains multiple bnAb epitopes, many of which are shielded or otherwise influenced by the highly conserved N276 glycan [254]. The N276 glycan was more extensively processed on the virion-derived Env than on the BG505 SOSIP.664 trimer

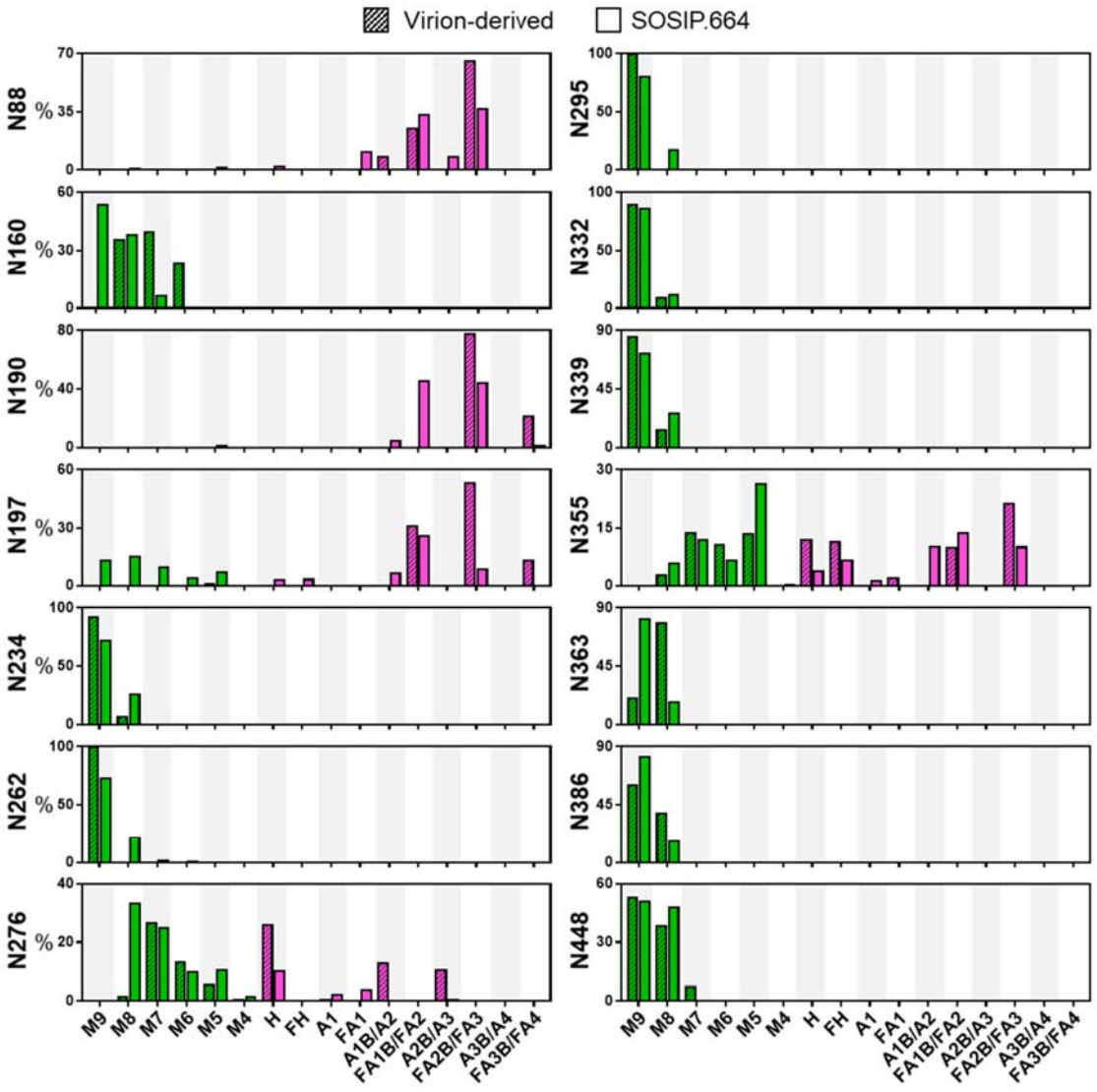


Figure 7.3. Compositional site-specific glycan analysis of virion-derived Env and BG505 SOSIP.664 trimers. Glycosylation profiles of 14 of the 24 PNGS of gp120 derived from virions, produced in a lymphoid cell line (hatched) or BG505 SOSIP.664, produced in HEK cells (plain). Oligomannose-type glycans are coloured in green and complex-/hybrid-type glycans in magenta. Glycans are grouped as shown in Appendix C. Glycan libraries are provided in Appendix D. Virion data were acquired jointly with Joel Allen, University of Southampton, Southampton, UK.

(Figure 7.3). Thus, the N276 site, together with N197, N190, and N462 (not determined for the virion), constitute multiple sources of glycan heterogeneity surrounding the CD4bs. Given that bnAbs targeting this region are able to neutralise the virus, it is hypothesised that the conformational plasticity inherent to glycans means that even large glycans do not preclude the CD4bs. Alternatively, this can be explained by the observation that many CD4bs-targeting bnAbs recognise the conserved *N*-Acetylglucosamine core [254].

The BG505.T332N-LAI virion-derived material contained greater amounts of hybrid-type glycans than were present on BG505 SOSIP.664. This finding was specifically the case for the N276 and N355 sites (Figure 7.3). There is growing interest in hybrid-type glycans because apex-targeting bnAbs, such as PG16, preferentially recognise such structures [257, 288, 290, 400]. However, the hybrid-type glycan populations identified here did not correspond to known bnAb epitopes.

The N88 site, near the base of the trimer, has previously been reported to be occupied exclusively by oligomannose-type glycans on BaL virions [327]. It had been hypothesised that the proximity of the membrane in virion-derived Envs may impose steric constraints on the processing of this site. Here, the N88 glycan was exclusively occupied by complex-type glycans on the virion-derived gp120, but was significantly less processed on the SOSIP.664 trimer (Figure 7.3).

Site-specific glycan compositions could only be determined for 14 of the 24 gp120 PNGS. In order to define the predominant glycan type at the other sites, the remaining glycopeptides were subject to the site-specific classification method by Joel Allen, University of Southampton. The glycan classification data were in broad agreement with the site-specific glycan composition data, in that sites occupied by mixed populations of glycans on the SOSIP.664 trimer were typically occupied by complex-type glycans on the virion-derived material. Mapping the oligomannose-content of individual sites onto a model reveals the regions of elevated processing displayed by the virion derived material (Figure 7.4).

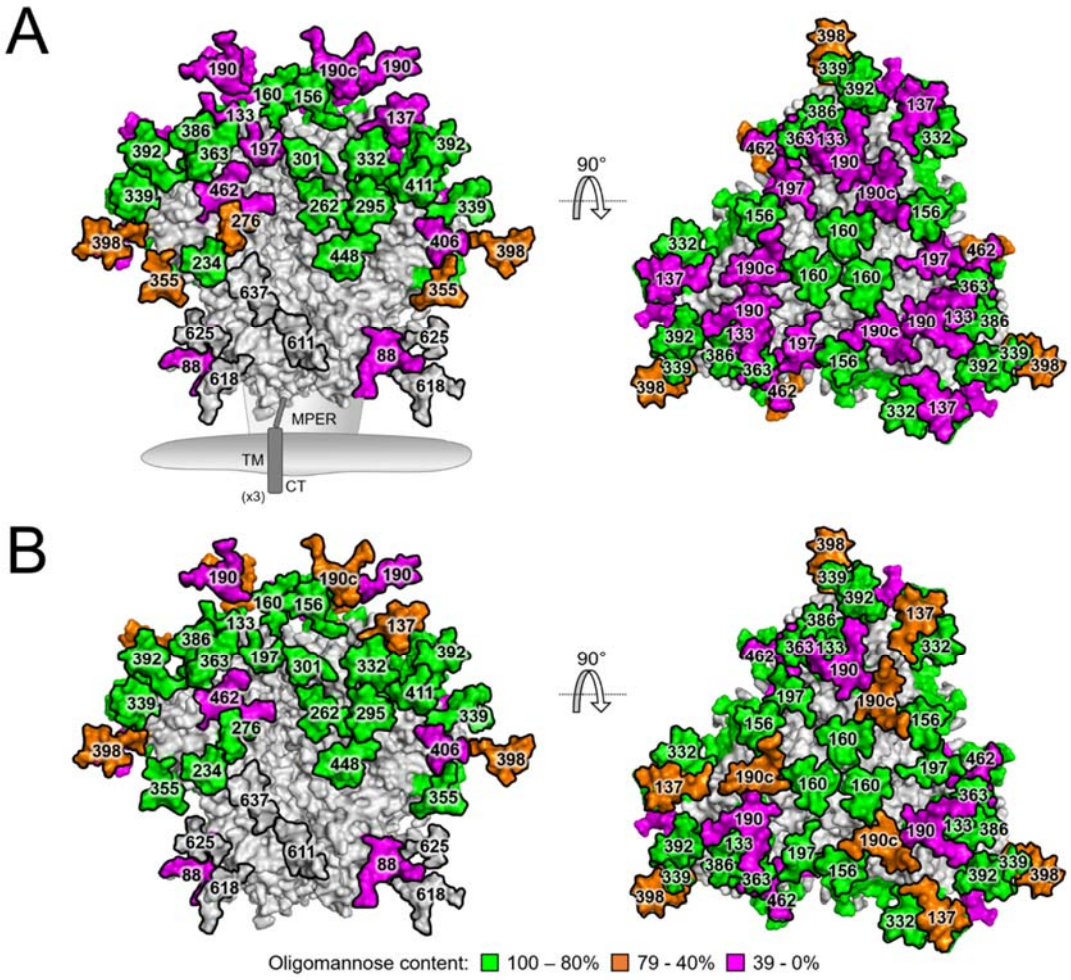


Figure 7.4. Site-specific classification of the predominant glycan type on virion-derived gp120 and BG505 SOSIP.664. A) Model of a fully glycosylated BG505.T332N-LAI Env and B) a BG505 SOSIP.664 trimer, coloured according to the predominant glycan type at each PNGS, as determined by glycosidase-treated glycopeptides. Model constructed according to Behrens *et al.*, 2016 [132]. Glycopeptides were subject to sequential digestion with Endo H and PNGase F (in the presence of O^{18}-labelled water) by Joel Allen, University of Southampton, to generate novel mass signatures for oligomannose-type glycans, complex-type glycans, or unoccupied sites.

7.4 Discussion

This chapter compared the glycosylation of the soluble BG505 SOSIP.664 immunogens with that of a genotype-matched Env, derived from infectious virions. The use of the A66-R5 T cell line for the production of the virus is relevant as a successful vaccine must elicit antibodies capable of recognising virions produced predominantly in human lymphoid cells. However, for the manufacture of vaccines, CHO cells are preferred to HEK cells. Indeed,

BG505 SOSIP.664 has been produced for human trials under cGMP conditions using CHO cells [324]. In this regard, it is notable that the glycan processing of cGMP BG505 SOSIP.664 expressed in CHO cells is highly similar to that expressed in HEK cells [276, 324].

Virion-derived Env contained a greater abundance of complex-type glycans than the corresponding SOSIP.664 trimers. Furthermore, the complex-type glycans present on the virion were generally more branched and sialylated. Site-specific analysis revealed that the elevated glycan processing observed on the virion could largely be attributed to PNGS containing mixed populations of oligomannose- and complex-type glycans on the SOSIP.664 trimers containing predominantly complex-type glycans on the virion. In contrast, oligomannose-type glycan sites were well conserved across both proteins. These results are in agreement with a site-specific glycan analysis of virion-derived Envs by Cao *et al.* [275].

While the difference in complex-type glycan processing may be symptomatic of the producer cell line, which was not controlled for in this study, it may also highlight the potential influence of membrane tethering on glycosylation processing. For example, membrane-bound CD59 also displays increased processing compared to its soluble form [336]. It is possible that membrane tethering may maintain Env in close proximity to membrane-bound glycan processing enzymes, as opposed to soluble constructs which are released into the lumen of the ER, thus influencing the kinetics of egress through the secretory system (Figure 1.10D).

Although both based on the BG505 genotype, the exact amino acid sequence of the virion-derived Env differed from that of the SOSIP.664 trimers, in that the latter contain an additional disulphide bond, a gp41 point mutation, and an optimised furin cleavage site. In Chapter 3, the addition of stabilising mutations was found to subtly restrict the glycosylation processing at some sites. Furthermore, the nucleic acid sequences were not matched, as the SOSIP.664 trimers had undergone a codon optimisation process. This may result in faster

translation and early folding events during Env biosynthesis and subsequently restrict glycan processing to the SOSIP.664 trimers [123]. Future work should explore the contribution of each of these sequence differences to glycosylation processing.

Alternatively, the elevated glycan processing observed on virion-derived Env may be explained by the presence of misfolded, non-functional trimers [123]. The virus preparation was purified by HPLC, while the SOSIP.664 trimers underwent positive selection using a quaternary structure-specific bnAb affinity purification step. However, the absence of complex-type glycans at the N160 'TAMP' site suggests that misfolded proteins were unlikely to have contributed to the analysis [122].

Although not addressed here, the relative occupancy of PNGS has been reported to be lower on SOSIP.664 trimers than on virion-derived material [275, 276]. This is particularly true for sites located on the V1/V2 loops and on gp41. Partial PNGS occupancy has the effect of generating potentially immunodominant holes in the glycan shield on a subpopulation of the trimers [240, 383, 388]. In the case of glycan holes in the highly variable V1 and V2 regions, the likely outcome would be undesirable narrow-specificity nAbs or non-nAbs.

Overall, the glycosylation of virion-derived Env is largely recapitulated by recombinant SOSIP.664 trimers, despite the latter displaying hotspots of elevated glycan processing. This is consistent with the observation that recombinant trimers display the epitopes for numerous glycan-targeting bnAbs that are able to neutralise the corresponding virus, and support their exploration as clinical immunogens [74, 324]. Due to the difficulties associated with eliciting glycan-binding antibodies, it may be advisable to avoid exposing the immune system to glycan epitopes that are not presented on viral Env. Additionally, increasing the occupancy of PNGS on recombinant immunogens may yet prove critical to eliciting a bnAb response. Methods to alter the glycosylation of recombinant Envs to more closely match that of the virus may therefore be required. Lastly, the next generation of

clinical immunogens may also benefit from the development of strategies to boost the immunogenicity of glycan-dependent epitopes, as will be discussed in the next chapter.

8. Future directions: Enhancing the immunogenicity of the glycan shield using yeast

Antigenic mimicry is usually considered fundamental to vaccinology. That is to say, in order to elicit an antibody response against a particular antigen, that antigen must be presented to the immune system. The goal of an effective HIV-1 vaccine is to elicit antibodies capable of neutralising the vast number of circulating strains. Such bnAbs arise in approximately a third of individuals after several years of infection [84-86], and are frequently directed to glycan epitopes on the HIV-1 Env protein [273, 278, 287-289, 297, 299, 300, 304]. The development of recombinant, native-like Env mimics, with glycosylation matching that of the virus (Chapter 7), is therefore likely to form part of an HIV-1 vaccine [215, 275].

However, given the difficulty in eliciting bnAbs during natural infection, it may be unreasonable to assume that mimics of the Env spike will be any more effective [351]. For example, while soluble native-like trimers present the epitopes for numerous bnAbs, they are unable to elicit them in immunisation studies [74, 237]. Similarly, synthetic, multivalent, oligomannose immunogens, designed to recapitulate the clustered glycan epitope of the bnAb, 2G12, are again able to bind the antibody but are unable to generate neutralising antibodies [345]. Thus, there is a disparity between the antigenicity and immunogenicity of current HIV-1 immunogens.

In natural infection, the development of bnAbs is slow, arising only after an extended virus-antibody arms race [80, 81, 83, 198, 255]. There are numerous reasons why the

elicitation of bnAbs in both infection and vaccine settings is challenging, including the high level of somatic hypermutation and unusual antibody architectures typically required for glycan recognition and broad neutralisation (Section 1.6.1) [189, 250, 256-261]. Additionally, it may reflect the fact that glycan-targeting bnAbs have to overcome multiple tolerance checkpoints [194].

Immunological tolerance, if indeed the limiting factor in bnAb development, can be overcome. For example, there is a precedent for tolerance to 'self' glycan structures to be broken through the antigenic mimicry of pathogens. As discussed in Section 1.8.2, in Guillain-Barré syndrome, infection with *C. jejuni* can result in an auto-reactive antibody response directed against the gangliosides of the PNS (Figure 1.12A) [270]. Similarly, the LOS of *Helicobacter pylori* can mimic Lewis blood group antigens, resulting in autoimmune gastritis [401].

In the context of HIV-1, the 2G12 bnAb can cross react with yeast mannans and the *Rhizobium radiobacter* Rv3 LOS; both of which contain α1-2-linked mannose motifs analogous to the 2G12 epitope (Figure 1.12B) [268, 352, 355, 356]. Furthermore, immunisation of rabbits with yeast displaying the α1-2-linked mannose motifs, generated immune sera with similar glycan specificities to 2G12, which were able to weakly neutralise some viruses [353]. Thus, the glycans of yeast are capable of breaking tolerance to 'self' oligomannose-type glycans. Interestingly, the binding of sera to $gp120_{BaL}$ (a strain not neutralised by immune sera) was comparable between the immunised and naïve animals [353]. When $gp120_{BaL}$ was expressed in the presence of kifunensine, however, the sera from the immunised animals displayed much greater affinity. Both kifunensine treated and untreated $gp120_{BaL}$ display dense clusters of oligomannose-glycans. Their glycosylation differs only in that the less tightly packed glycans on untreated gp120 are of the complex-type, while these glycans are retained as $Man_9GlcNAc_2$ on the kifunensine treated gp120. The fact that immune sera preferentially bound to the latter suggests that the elicited

oligomannose-specific antibodies are not capable of recognising these structures in context of the dense oligomannose network presented on the outer domain of gp120 [353]. The next barrier to overcome, therefore, is to break immunological tolerance to such clusters, for example by displaying immunogenic 'non-self' yeast glycans in the context of the Env protein.

The following chapter contains preliminary data exploring this concept. Glycan analysis confirms the suitability of yeast for breaking tolerance to oligomannose-type glycans. Additionally, two methods for the generation of glycan-altered Env are discussed: expressing Env in yeast, and using yeast glycosylation enzymes to modify Env *in vitro*.

8.1 Yeast N-glycosylation

Unlike mammalian cells, yeast synthesise glycoproteins bearing only high-mannose glycans (Figure 8.1). Yeast N-glycosylation is initiated in the ER, in much the same manner as mammalian glycosylation, with the transfer of a $Glc_3Man_9GlcNAc_2$ precursor glycan to Asn-X-Ser/Thr (X ≠ proline) motifs on the nascent polypeptide. Following transfer, three glucose residues and one terminal α1-2-mannose residue are removed by glucosidase I, II, and an ER-residing α-mannosidase. The resulting $Man_8GlcNAc_2$ glycoform is then transported to the Golgi apparatus, where N-glycosylation processing differs markedly between yeast and mammals [402].

In mammalian cells, Golgi apparatus-resident mannosidases continue to trim mannose residues to generate $Man_5GlcNAc_2$ or $Man_3GlcNAc_2$ structures. A range of cell-specific glycosyltransferases subsequently rebuild the glycan with the addition of GlcNAc, galactose, sialic acid, and fucose residues (Figure 1.5). In contrast, yeast glycosylation involves the stepwise addition of mannose and mannosylphosphate residues to the $Man_8GlcNAc_2$ structure, catalysed by several transferases (Figure 8.1) [403]. The series of reactions is initiated by Och1, which transfers an α1-6-linked mannose to the α1-3-linked mannose of

the D1-arm. This branch is subsequently elongated by mannan polymerase I and II (M-Pol I and II) with the addition of numerous α1-6-linked mannose residues. The poly-α1-6-linked mannose backbone can be further modified by side-chains comprised of α1-2-linked mannose and mannosylphosphate residues, which can in turn be capped with α1-3-linked mannose residues, transferred by the Mnn1 mannosyltransferase (Figure 8.1; Section 1.8.2).

As per mammalian cells, the precise N-linked glycan structures produced are dependent on the repertoire of glycosidases and glycosyltransferases possessed by the yeast. For example, *Schizosaccharomyces pombe* lacks ER α1-2-mannosidase activity and can also decorate the α1-6-mannose backbone with α1-2-linked galactose residues [404]. In a similar manner, *Kluyveromyces lactis* glycans lack mannosylphosphates, but can incorporate α1-2-linked GlcNAc residues in their side chains [405]. Additionally, *Pichia pastoris* strains lack the Mnn1

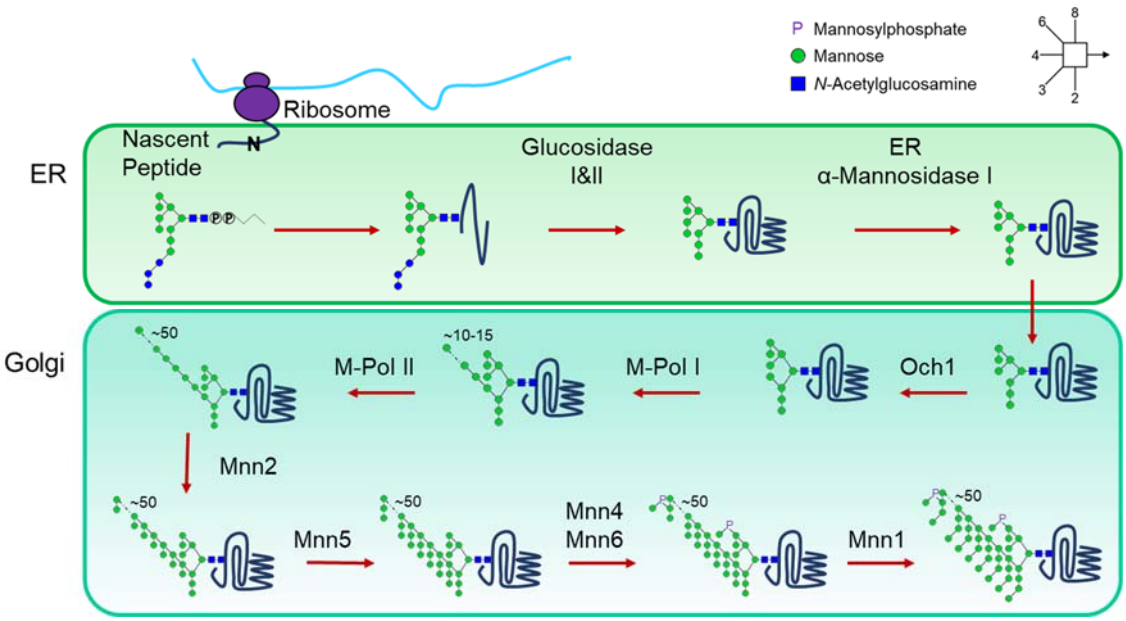


Figure 8.1. Yeast N-glycosylation pathway. The initial N-glycosylation pathway is conserved between yeast and higher eukaryotes. Three glucose residues and one mannose residue (with the exception of *S. pombe*) are trimmed from the $Glc_3Man_9GlcNAc_2$ precursor. The resulting $Man_8GlcNAc_2$-displaying glycoprotein is transferred to the Golgi apparatus where it is modified by the α1-6-mannosyltransferase, Och1, which adds an 'initiating' α1-6-mannose to the α1-3-mannose of the D1-arm. Subsequent elongation of the outer chain is achieved by the sequential action of two protein complexes, mannan polymerase I (M-Pol I; Mnn9, Van1) and mannan polymerase II (M-Pol II; Mnn9, Anp1, Mnn10, Mnn11, Hoc1). The α1-6-mannose backbone is further modified by the addition of α1-2-mannose residues, mannosylphosphate residues (P) and α1-3-mannose caps (with the exception of *P. pastoris*).

mannsosyltransferase responsible for α1-3-linked mannose caps [406]. The size of the high-mannose glycans generated by *P. pastoris* is also much smaller ($Man_{8-14}GlcNAc_2$) than those found in *Saccharomyces cerevisiae*, for example, which can reach up to 50-150 mannose residues [407].

Given the promising results obtained previously with Mnn1 deficient yeast [353], *P. pastoris* was chosen for the expression of Env. To further probe the glycan structures present in *P. pastoris*, N-linked glycans were enzymatically released from cells and analysed by negative-ion IM-ESI MS, with the help of Prof. David Harvey, University of Oxford. Consistent with the literature, the spectrum revealed a series of high-mannose glycans, ranging from $Man_8GlcNAc_2$ to $Man_{15}GlcNAc_2$, with a corresponding series of high-mannose glycans containing a phosphate residue (Figure 8.2A). In order to assign the structures the glycans were fragmented, generating predominantly C-type glycosidic cleavages, A-type cross ring cleavages and D-ions (Appendix B). Figure 8.2B portrays the fragmentation spectrum corresponding to the $Man_9GlcNAc_1$ glycan at m/z 1776.5. This produced D- and D-18- ions at m/z 809 and 791, respectively, indicating that the ninth mannose residue is linked to the D1-arm of $Man_8GlcNAc_2$ and the structure is not the ER-resident $Man_9GlcNAc_2$ isomer [408]. The D- and D-18-ions corresponding to the ER-isomer of $Man_9GlcNAc_2$ could not be detected (Figure 8.2B, indicated with asterisk). Thus, the structure of some *P. pastoris* glycans are 'non-self', confirming the suitability of this expression system for the production of immunogenic Env proteins.

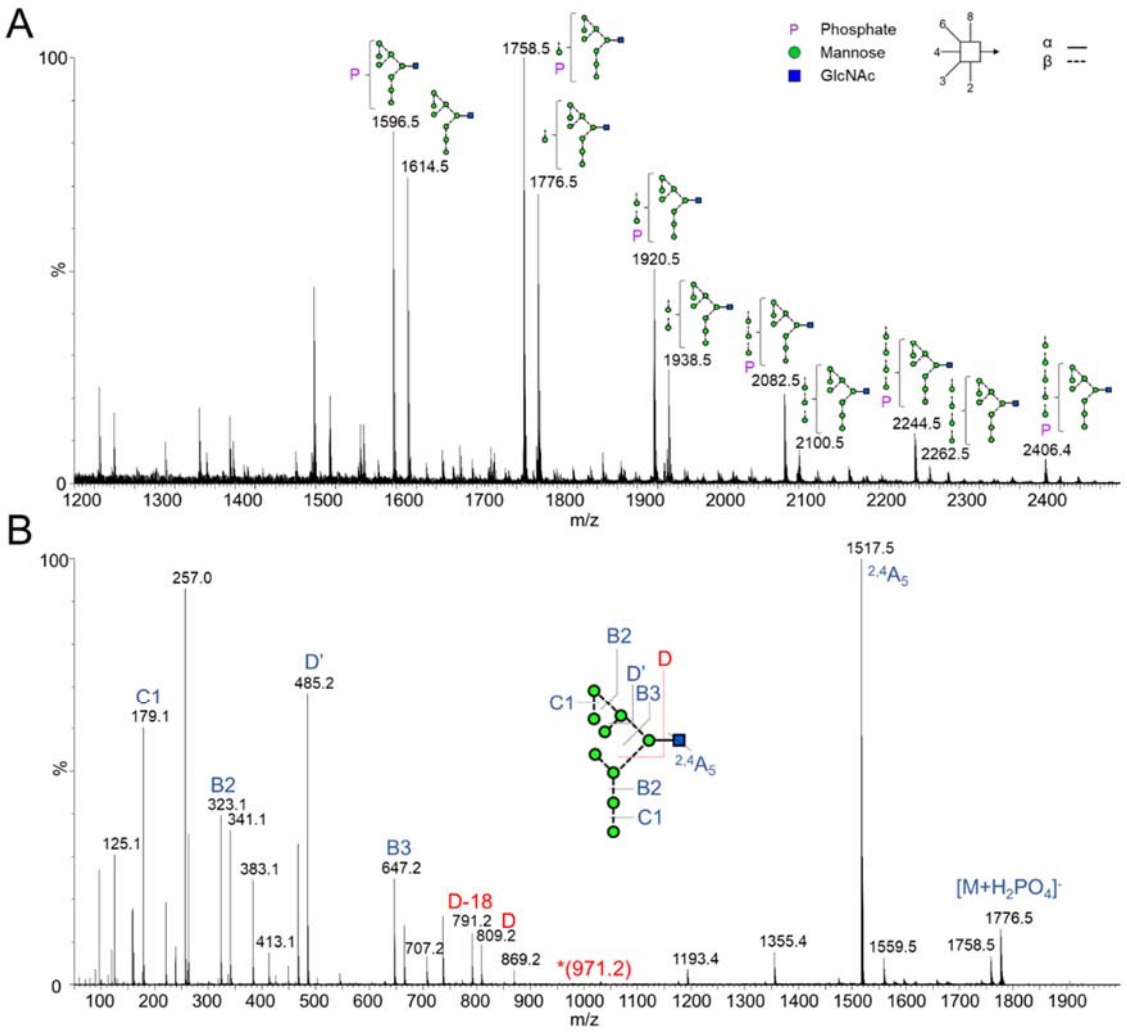


Figure 8.2. N-glycosylation mass spectra of *P. pastoris*. A) N-glycans were released from the surface of *P. pastoris* cells by digestion with Endo H and analysed by negative ion IM-ESI MS, with the help of Prof. David Harvey, University of Oxford. B) Tandem mass spectrum of a *P. pastoris* N-glycan with a molecular ion peak of m/z 1776.5. Diagnostic structural ions annotated in blue and red, according to Appendix B. The characteristic D-ion for the ER-isomer of $Man_9GlcNAc_2$ (m/z 971.2) was not detected in the spectra, indicated with an asterisk.

8.2 *In vivo* modification of Env glycans

The BG505 NFL constructs (provided by Prof. Rich Wyatt, Scripps Research, La Jolla, USA) were chosen for expression in yeast cells. The BG505 NFL sequence is almost identical to that of the BG505 SOSIP.664 trimers, except that the furin cleavage site is replaced by a glycine-serine linker (Figure 3.8). The linker enables the formation of stable, native-like trimers without furin cleavage [234], and thus removes the need to co-express with a furin

encoding plasmid or rely on cleavage by native yeast proteases (for example, Kex2, which cleaves Lys/Arg-Arg motifs).

For expression in *P. pastoris* (X-33 strain), the BG505 NFL constructs were inserted into the pPICLIT plasmid (provided by Prof. Nicole Zitzmann, University of Oxford, UK). The pPICLIT plasmid contains an α-factor signal sequence upstream of a Kex2 cleavage site. The constructs were cloned flush to the cleavage site, to enable secretion of the protein with a native N-terminus. Two constructs were made, one containing a C-terminal His_6-tag, and one without. Transformation of the plasmid into *P. pastoris* was confirmed by PCR, and several transformants were tested for protein expression.

Expression was confirmed by Western blot (Figure 8.3A). The constructs ran slightly higher than trimers expressed in HEK cells, possibly due to hypermannosylation. Several attempts to scale-up expression and purify the proteins were made, including Ni^{2+} affinity (for His_6-tagged proteins) and bnAb affinity chromatography (2G12 and b12), testing multiple transformants and buffer conditions. Figure 8.3C illustrates the difficulties in purifying the proteins using 2G12 affinity chromatography, despite the proteins cross-reacting to 2G12 on a Western blot. Trimers could not be obtained in sufficient quantities for further analysis.

Expression of the BG505 NFL trimers was also attempted in *K. lactis* by Dr Mehul Ganatra, from the laboratory of Dr Chris Taron, Protein Expression & Modification Research division, New England Biolabs, USA. The BG505 NFL construct (containing C-terminal His_6-tag) was transformed into *K. lactis* GG799 cells and a glycan engineered strain, termed Del2, which lacks the Och1 and Mnn9 mannosyltransferases. Additionally, N- or C-terminus chitin binding domain-tagged proteins were transformed into the Del2 cells. The *K. lactis* GG799 cells produce predominantly $Man_{7\text{-}13}GlcNAc_2$ glycans, while the Del2 strain produces predominantly $Man_{9\text{-}11}GlcNAc_2$ glycans (personal communication with Dr Mehul Ganatra). Furthermore, *K. lactis* can incorporate α1-2-linked GlcNAc residues [405].

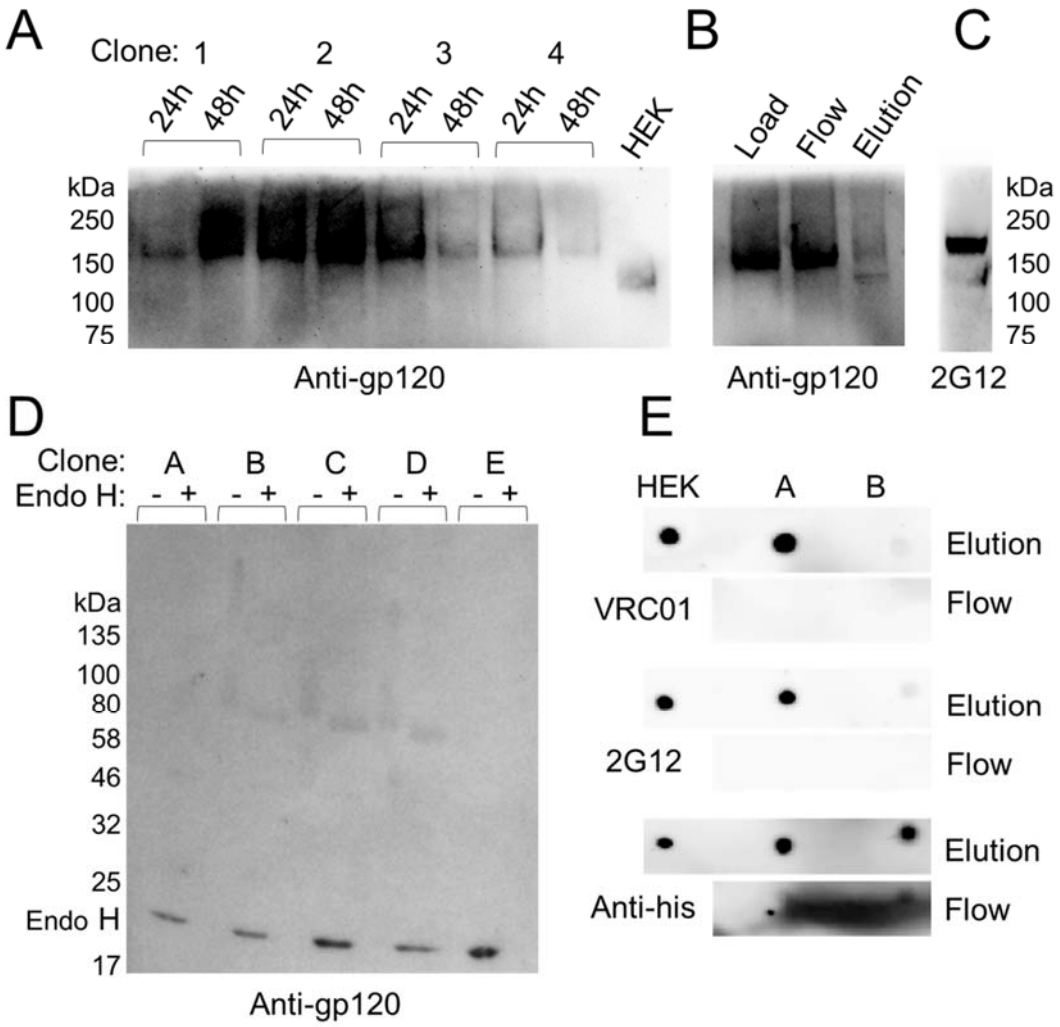


Figure 8.3. Expression and purification of BG505 NFL expressed in yeast. A) Western blot of test expression of four BG505 NFL *P.pastoris* X-33 transformants. Cell culture supernatant was taken at 24h and 48h post methanol induction. Proteins were visualised using an anti-gp120-HRP antibody. A HEK-expressed trimer was included for comparison. B) Western blot of example purification by 2G12 affinity chromatography. The lower band on the elution is likely to be contamination from a previous purification of HEK-expressed trimers. Proteins visualised by anti-gp120-HRP antibody. C) Western blot confirming cross-reactivity of *P. pastoris* expressed BG505 NFL to 2G12. Proteins visualised using 2G12 and anti-human IgG-HRP antibodies. D) Western blot of test expression of four BG505 NFL *K. lactis* transformants. Cell supernatants (-/+ Endo H) were 25× concentrated and visualised using anti-gp120 (goat) and anti-goat IgG-HRP antibodies. The transformants were made, and the Western blot provided, by Dr Mehul Ganatra, from the laboratory of Dr Chris Taron, Protein Expression & Modification Research division, New England Biolabs, USA. Transformants: A, N-terminal His_6-tag in GG799 cell; B, N-terminal His_6-tag in Del2 cells; C, N-terminal chitin binding domain-tag in Del2 cells; D, C-terminal chitin-binding domain-tag in Del2 cells; E, untransformed GG799 cell control. E) Dot blot of test Ni^{2+} affinity purification of BG505 NFL trimers from transformants A and B. Elutions and 25× concentrated flow were visualised with VRC01 or 2G12 and anti-human IgG-HRP antibodies, or anti-his-HRP antibodies. A HEK-expressed trimer is included for comparison.

Expression was confirmed by Western blot (Figure 8.3D; Dr Mehul Ganatra). The proteins ran as a heterogeneous smear, which collapsed to the predicted ~70 kDa when treated with Endo H. The His_6-tagged constructs were expressed and successfully purified by Ni^{2+} affinity chromatography, as confirmed by a dot blot (Figure 8.3E). The GG799 strain appeared to produce higher yields, and moving forward with a large scale expression provided sufficient material for subsequent site-specific glycan analysis.

The Ni^{2+}-purified BG505 NFL proteins were reduced, alkylated and digested with trypsin, as previously described. Analysis by LC-ESI MS confirmed the full-length expression of the construct (Figure 8.4B). The spectra were searched against a glycan library containing $Man_{5\text{-}25}GlcNAc_2$ glycans and the same series containing additional GlcNAc residues. Site-specific glycan compositions were obtained for 12 of the 28 sites (Figure 8.4A), though the fragmentation data varied in quality. It is likely that this is not a complete picture of the site-specific glycan compositions as some expected glycans were missing, for example $Man_{11}GlcNAc_2$ at site N276 (Figure 8.4A). Most of the sites followed a similar trend: a series of high-mannose glycans, ranging from $Man_8GlcNAc_2$, peaking around $Man_{12}GlcNAc_2$, and then rapidly tailing off. Glycans containing further GlcNAc residues were present, though in small quantities.

It is notable that IMP sites (e.g. N332, N363 and N448), which display restricted processing in mammalian expression systems [132], appeared to be processed by the yeast mannosyltransferases. This suggests that the steric constraints which limit mammalian α-mannosidases from accessing their substrate glycans do not apply to yeast mannosyltransferases. This result is encouraging for the production of immunogens with altered oligomannose clusters. The N339 and N355 sites appeared to be the most highly processed sites, despite exhibiting somewhat restricted processing on mammalian cell line expressed trimers [132]. It remains to be determined whether this result is genuine, or an artefact of the comparatively good data quality at these sites.

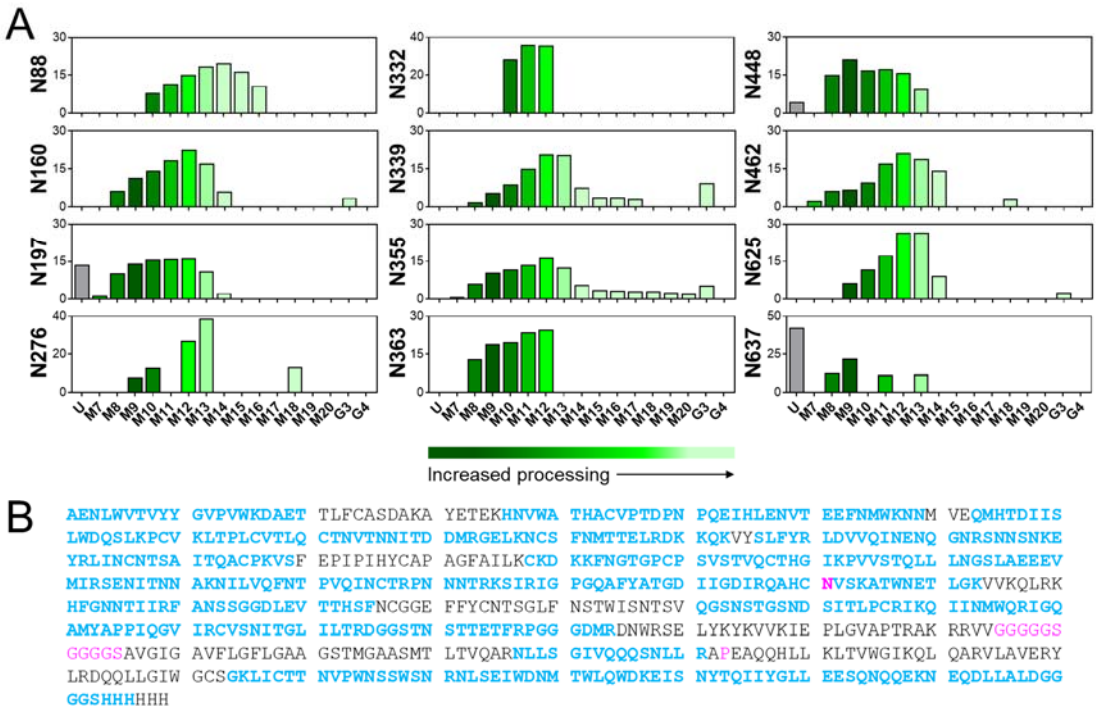


Figure 8.4. Site-specific glycan analysis of Env expressed in yeast. A) BG505 NFL proteins expressed in *K. lactis* GG799 cells were purified by Ni^{2+} affinity chromatography. The elution was reduced, alkylated and digested with trypsin, prior to analysis by LC-ESI MS. Graphs represent the relative abundance at each site (%). U, unoccupied; M7-20, $Man_{7-20}GlcNAc_2$; G3/4, $Man_{10-20}GlcNAc_{3/4}$. B) Amino acid sequence of BG505 NFL trimers. Peptide coverage is highlighted in blue/bold, NFL modifications are highlighted in magenta.

Furthermore, the N197 and N637 sites appeared to have moderately low glycan site occupancy (Figure 8.4A). Should yeast expressed Envs progress to immunogenicity studies, it is likely that strategies to increase glycan site occupancy will be required, in order to cover undesirable protein epitopes.

8.3 *In vitro* modification of Env glycans

An alternative strategy for the production of glycan-modified Env is the enzymatic processing of trimers *in vitro*. This method has the advantage of circumventing the difficulties associated with the expression and purification of Env produced in yeast (and, possibly, PNGS under occupancy).

A proposed strategy is as follows: a soluble Env trimer, displaying predominantly $Man_9GlcNAc_2$ glycans, serves as the starting material. This is readily achievable by expressing the trimer in the presence of kifunensine. The $Man_9GlcNAc_2$ structures are then trimmed *in vitro* by incubation of Env with ER α-mannosidase I. The resulting $Man_8GlcNAc_2$ structures are the preferred substrate for Och1 [409], which, in the presence of GDP-mannose, transfers an α1-6-linked mannose residue to the α1-3-linked mannose of the D1-arm of $Man_8GlcNAc_2$. This initiates the elongation of an α1-6-linked mannose branch by the mannan polymerase I (M-Pol I) complex. The Mnn2 enzyme can then further modify the glycan through the addition of α1-2-linked mannose residues to mannose backbone (Figure 8.5A).

Table 8.1 summarises the progress made towards this project to date. The cloning and expression of the catalytic domain of ER α-mannosidase I, and its activity on recombinant Env proteins, has previously been described by members of the group [313, 353].

The Och1 mannosyltransferase is a type II transmembrane protein, anchored to the Golgi apparatus at its N-terminus. To generate a soluble form of the enzyme, the catalytic domain (residues 38-480) was cloned into the pPICLIT expression vector, as described above, so that the N-terminal region of Och1 contained the α-factor signal sequence and the C-terminus contained a His_6-tag. The pPICLIT-Och1 vector was successfully transformed into *P. pastoris,* and a small-scale, test expression was performed. Expression was confirmed by Western blot (Figure 8.5B). A 48 hour incubation period appeared to provide the highest yields, with expression levels decreasing beyond 72 hours. Large-scale expression and purification by Ni^{2+} affinity chromatography has not yet been successful, though has been reported by another group* [409].

* Note that this paper cloned Och1 residues 31 – 480.

The M-Pol I complex, comprised of Mnn9 and Van1 (1:1), is responsible for the addition of approximately 10 mannose residues to form the α1-6-linked mannose backbone. A previous study failed to express Mnn9 and VanI as separate subunits, but a soluble M-Pol I complex was obtained by simultaneous expression [410]. Thus, to achieve co-expression of both subunits, a sequential transformation approach was undertaken. The catalytic domain of Van1 (residues 92 – 535) was cloned into the pPICLIT vector, so that the N-terminus contained the α-factor signal sequence, and the C-terminus contained a His_6-tag, as described above. Successful transformation into *P. pastoris* was confirmed. The Mnn9 sequence (residues 41 – 395) was amplified for cloning into the pJAN-s1 vector (provided by Dr Karen Polizzi, Imperial College London, UK). The pJAN-s1 vector contains a different resistance marker (Nourseothricin), to allow for the selection of double transformants. The Mnn9 construct was also cloned downstream of the α-factor signal sequence, but was not tagged for purification, as it can be co-purified via the Van1 subunit [410]. Successful cloning of Mnn9 has not yet been confirmed.

The Mnn2 sequence was also truncated prior to the predicted N-terminal transmembrane domain (residues 36 – 597), and was inserted into the pPICLIT vector with α-factor signal sequence and His_6-tag, as before. Cloning in to the pPICLIT vector, and the production of a stable transformed *P. pastrois* cell line has been confirmed, though protein expression is yet to be tested (Table 8.1).

Table 8.1. Progress towards *in vitro* glycan modification.

	Och1	**Van1**	**Mnn9**	**Mnn2**
Cloning into yeast expression vector	✓	✓	✗	✓
Transformation of yeast stable cell lines	✓	✓		✓
Test expression, small scale	✓			
Large scale expression and purification	✗			

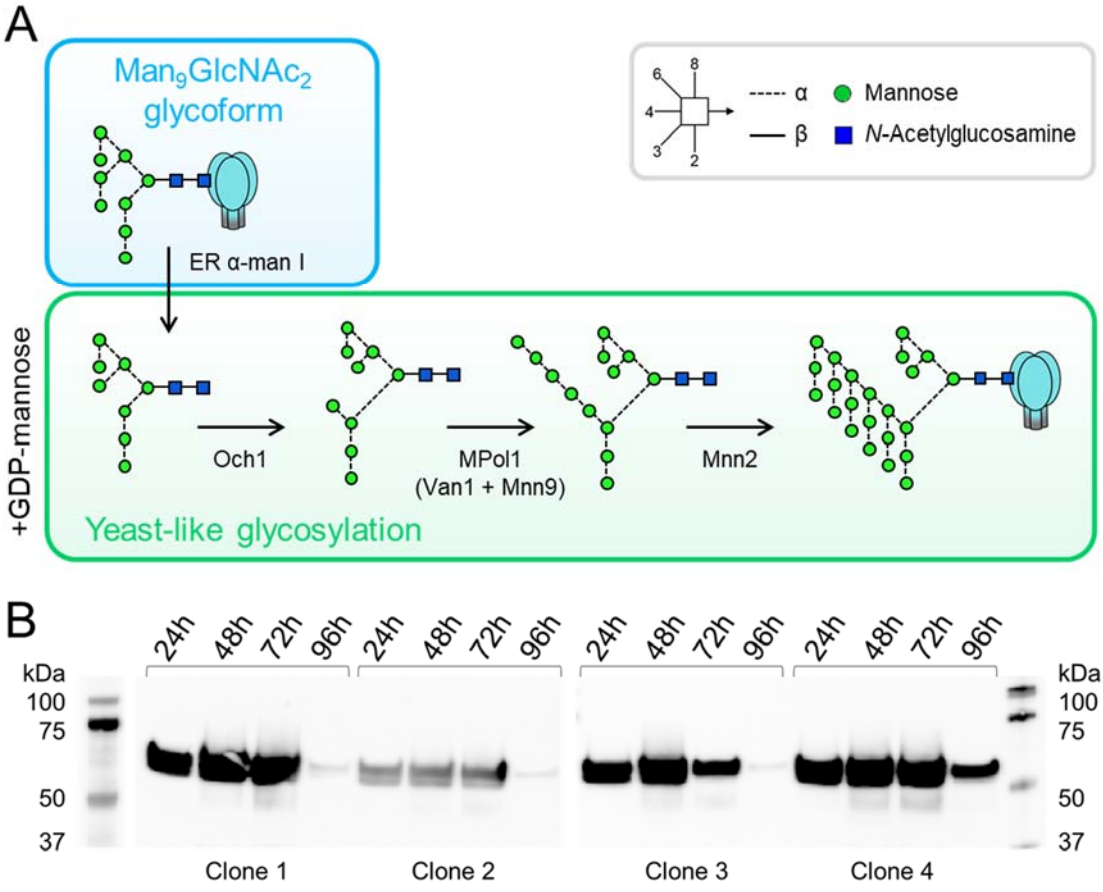


Figure 8.5. Strategy for the modification of Env glycosylation *in vitro.* A) Env trimers expressed with kifunensine are digested with ER α-mannosidase I. The resulting glycoforms are then modified by soluble, recombinant yeast mannosyltransferases, in the presence of GDP-mannose. B) Four transformed *P. pastoris* clones were tested for the expression of Och1 over 96 hours, taking time points every 24 hours. Expression was confirmed by Western blot, using an anti-His-HRP conjugated antibody. Expected MW = 61 kDa.

8.4 Discussion

This chapter presented two possible strategies to produce immunogens displaying clustered, 'non-self' high-mannose glycans, designed to break immunological tolerance to the clusters of 'self' oligomannose-type glycans displayed on HIV-1 Env. The first strategy attempts to exploit the natural N-glycosylation pathway of yeast to modify the Env trimers co-translationally. Preliminary results suggest that this goal is achievable, though further optimisation is required in order to boost expression yields. The second strategy is based on the expression of soluble yeast mannosyltransferases in order to modify Env glycans *in vitro*.

While still in its early stages, this strategy has great potential, given that the activity of many of the enzymes *in vitro* has already been confirmed [313, 353, 409, 410]. Should either of these strategies progress sufficiently, extensive characterisation of the glycosylation, structure, and antigenicity of the resulting trimers will be required.

The IMP glycans are resistant to α-mannosidase I trimming, both *in vivo* and *in vitro* [313]. It is not yet known whether the same steric restrictions apply when these glycans are expressed in yeast. However, the site-specific glycan analysis of BG505 NFL trimers expressed in *K. lactis* presented here, provides evidence that these glycans are somewhat susceptible to processing by yeast mannosyltransferases. Both $Man_8GlcNAc_2$ and the ER-isomer of $Man_9GlcNAc_2$ can act as the substrate for Och1 *in vitro* [409]. The result is promising for the generation of 'non-self' glycan clusters, though it remains to be determined whether such structures are indeed more immunogenic.

Classically, breaking immunological tolerance to self-antigens results in autoimmune disorders [270]. While the glycans on HIV-1 Env are derived from the host, and are thus 'self' antigens, clusters of oligomannose-type glycans are seldom seen on host glycoproteins. Furthermore, the innate immune system already contains anti-mannose responses, for example, the mannose binding lectin [53]. Lastly, no adverse effects were observed following the passive transfer of 2G12 in a clinical setting [411]. Taken together, this suggests that breaking tolerance to the oligomannose-type glycans will not result in damaging autoimmune responses.

9. Concluding remarks

In this thesis, a semi-quantitative, site-specific glycan analysis methodology was employed to investigate the impact of sequence diversity on recombinant Env immunogens [132, 331]. The analysis of a cross-clade panel of native-like trimers revealed remarkable similarities in the composition of the glycan shields, despite differences to the underlying amino acid sequences. Specifically, the mannose patches were a conserved feature of all trimers, as were aspects of the processing of complex-type glycans. Glycan density is known to influence the processing of the oligomannose-type glycans of Env [310]. This concept was further explored by investigating the contribution of individual glycan sites to the fine processing of the surrounding network of glycans.

The above analyses were performed on recombinant, soluble Env constructs. While native-like trimers are good structural and antigenic mimics of virally-presented Env [74, 125], the extent to which their glycan shields match that of infectious virions was unknown. Here, the glycosylation of BG505 SOSIP.664 immunogens was found to be highly similar to that of virion-derived Env. There were, however, regions of elevated glycosylation processing on the viral material. Although both based on the BG505 genotype, it is important to note that the amino acid sequence of the virion-derived Env was not identical to that of the SOSIP.664 trimers; the latter contain an additional disulphide bond, a gp41 point mutation, and an optimised furin cleavage site. Additionally, the nucleic acid sequences were not matched, as the SOSIP.664 trimers had undergone codon optimisation. The addition of

stabilising mutations can subtly restrict the glycosylation processing at some sites, suggesting that there may be a balance between producing stable Env mimics and mimics with glycosylation matching that of the virus. Future work should explore the contribution of such sequence differences to glycosylation processing.

Given the difficulties associated with eliciting glycan-binding bnAbs, it may be disadvantageous to expose the immune system to glycan epitopes that are not present on viral Env [276, 351]. The ability to make targeted alterations to the glycan shields of recombinant Envs may therefore prove beneficial. Here, the site-specific glycan compositions of recombinant Envs could be altered through the use of glycan-dependent bnAbs, either during the expression or purification stages. Thus bnAbs represent a useful tool for manipulating the fine processing of trimers on a site-specific basis.

Lastly, it is important to consider the potential difference between the antigenicity and immunogenicity of a vaccine candidate. The Env glycan shield is inherently immunoquiescent; this is reflected in the very slow development of bnAbs in only a subset of infected individuals. Strategies through which to enhance the immunogenicity of the glycan shield may be required. Accordingly, this thesis built on work by Dunlop *et al.*, exploring the use of 'non-self' high-mannose yeast glycans to break immunological tolerance to the 'self' oligomannose-type glycan clusters observed on Env [268, 352, 353]. While the work is ongoing, it was encouraging to obtain data suggesting that the clustered oligomannose-type glycans of Env are amenable to processing by yeast glycosyltransferases.

List of abbreviations

2-AA	2-aminobenzoic acid
α-Glu I/II	α-glucosidase I or II
α-Man I/II	α-mannosidase I or II
ADCC	Antibody-dependent cellular cytotoxicity
AIDS	Acquired immunodeficiency syndrome
ART	Antiretroviral therapy
BCR	B cell receptor
BMGY	Buffered glycerol-complex medium
BMMY	Buffered methanol-complex medium
bnAbs	Broadly neutralising antibodies
bp	Base pair
C1-C5	Constant regions 1-5
CA	Capsid
cART	Combined antiretroviral therapy
CCR5	C-C chemokine receptor 5
CD4	Cluster of differentiation 4
CD4bs	CD4 binding site
CDR	Complementarity-determining regions
CFG	Centre for Functional Glycomics
cGMP	Current good manufacturing practice
CHO	Chinese hamster ovary
CLR	C-type lectin receptor
CNX/CLN	Calnexin/calreticulin
CRF	Circulating recombinant form
CT	Cytoplasmic tail
CTL	Cytotoxic T lymphocyte (CD8+ T cell)
CXCR4	C-X-C chemokine receptor 4
DC-SIGN	Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin
DTT	Dithiothreitol
EM	Electron microscopy
Endo H	Endoglycosidase H
Env	Envelope spike
ER	Endoplasmic reticulum
FDA	Food & Drug Administration
FP	Fusion peptide
FT	Fucosyltransferase
Fuc	Fucose
Gal	Galactose
GalNAc	*N*-Acetylgalactosamie
GALT	Gut-associated lymphoid tissue
gl-bnAb	germline (encoded)-broadly neutralising antibody
Glc	Glucose
GlcN	Glucosamine
GlcNAc	*N*-Acetylglucosamine
GNL	*Galanthus nivalis* lectin
GnT I	*N*-Acetylglucosaminyltransferase I
HAART	Highly active antiretroviral therapy
HCD	Higher energy collisional dissociation

HCDR3	Third heavy chain complementarity-determining regions
HEK	Human embryonic kidney
Hex	Hexose
HexNAc	*N*-Acetylhexosamine
HILIC-UPLC	Hydrophilic interaction chromatography-Ultra performance liquid chromatography
HIV-1	Human immunodeficiency virus, type 1
HIV-2	Human immunodeficiency virus, type 2
HLA	Human leukocyte antigen
HPLC	High-performance liquid chromatography
HR1/2	Heptad repeat helices 1/2
IAA	Iodacetamide
IM-ESI MS	Ion mobility-electrospray ionisation mass spectrometry
IMC	Infectious molecular clone
IMP	Intrinsic mannose patch
IN	Integrase
kb	Kilobase
kDa	Kilodalton
KDO	2-keto-3-deoxy-D-manno-octulosonic acid
LC-ESI MS	Liquid chromatography-electrospray ionisation mass spectrometry
LOS	Lipooligosaccharide
LRA	Latency reversing agent
LTR	Long terminal repeat
M-Pol I/II	Mannan polymerase I or II
MA	Matrix
MALDI	Matrix assisted laser desorption ionisation
Man	Mannose
MBL	Mannose binding lectin
MPER	Membrane proximal external regions
MW	Molecular weight
nAbs	Neutralising antibodies
NC	Nucleocapsid
Nef	Negative regulatory factor
Neu5Ac	*N*-Acetylneuraminic acid, sialic acid
NFL	Native flexibly linked
OD	Optical density
ORF	Open reading frame
OST	Oligosaccharyltransferase
PBMCs	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PDB	Protein Data Bank
PDI	Protein disulphide isomerase
PNGase F	Peptide-N-glycosidase F
PNGS	Potential N-linked glycosylation site
pp	Percentage point
PR	Protease
PrEP	Pre-exposure prophylaxis
PRR	Pattern recognition receptor
R5 viruses	CCR5-tropic viruses
RP-HPLC	Reversed-phase HPLC (see above)
RT	Reverse transcriptase

SC	Single-chain
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
Siglec-1	Sialic acid-binding Ig-like lectin 1
SIVcpz	Simian immunodeficiency virus, chimpanzees
SIVgor	Simian immunodeficiency virus, gorillas
SOSIP.664	Stabilising mutations, including disulphide bond ('SOS') and Ile to Pro point mutation ('IP'), truncation at residue 664
TAMP	Trimer associated mannose patch
Tat	Trans-activator of transcription
TF	Transmitted/Founder
TM	Transmembrane region
TOF	Time of flight
UFO	Uncleaved pre-fusion optimised
V1-V5	Variable regions/loops 1-5
Vif	Viral infectivity factor;
Vpr	Viral protein R
Vpu	Viral protein U
X4 viruses	CXCR4-tropic viruses
XIC	Extracted ion chromatograms
YCB	Yeast carbon base
YPD	Yeast extract peptone dextrose
YPGal	Yeast extract peptone galactose

Appendix A. Cloning and mutagenesis primers

Table A.1 Primers for PNGS mutagenesis and cloning into yeast expression vectors.

Mutation	Primer (5'-3')
N156A_fwd	GATATGCGCGGAGAACTGAAGGCTTGTTCTTTCAACATGACTAC
N156A_rev	GTAGTCATGTTGAAAGAACAAGCCTTCAGTTCTCCGCGCATATC
N197A_fwd	CAGGCTGATCAATTGCGCCACCAGTGCCATTACAC
N197A_rev	GTGTAATGGCACTGGTGGCGCAATTGATCAGCCTG
N234A_fwd	GTAAAGATAAGAAGTTCGCCGGCACTGGGCCCTGC
N234A_rev	GCAGGGCCCAGTGCCGGCGAACTTCTTATCTTTAC
N295A_fwd	CACCCCCGTCCAGATTGCTTGCACTAGACCTAAC
N295A_rev	GTTAGGTCTAGTGCAAGCAATCTGGACGGGGGTG
N339A_fwd	GTGAGCAAGGCTACATGGGCCGAGACTCTG
N339A_rev	CAGAGTCTCGGCCCATGTAGCCTTGCTCAC
N386A_fwd	GAGGCGAGTTCTTTTACTGTGCCACAAGTGGCCTGTTTAATTC
N386A_rev	GAATTAAACAGGCCACTTGTGGCACAGTAAAAGAACTCGCCTC
N392A_fwd	GTAACACAAGTGGCCTGTTTGCTTCAACTTGGATCAGCAAC
N392A_rev	GTTGCTGATCCAAGTTGAAGCAAACAGGCCACTTGTGTTAC
N411A_fwd	CTAATTCTACCGGCTCTGCCGATAGTATCACACTGCC
N411A_rev	GGCAGTGTGATACTATCGGCAGAGCCGGTAGAATTAG
N448A_fwd	GTGATTCGATGTGTCAGCGCTATCACAGGCCTGATTC
N448A_rev	GAATCAGGCCTGTGATAGCGCTGACACATCGAATCAC
NFL_pPICLIT_fwd	GAAGAAGGGGTATCTCTCGAG**AAAAGA**GCCGAAAATCTGTGGGTG
NFL_pPICLIT_rev	GAGATGAGTTTTTGTTCTAGATTATCAATCCAGGGCCAGCAG
NFL_His_pPICLIT_r	GAGATGAGTTTTTGTTCTAGATCAATGGTGATGGTGGTGATG
Och1_pPICLIT_fwd	GTATCTCTCGAG**AAAAGA**CTGTCCCAATTCTACC
Och1_pPICLIT_rev	GATGATGATGGTCGACCTTGTGACCAGCGTTC
Van1_pPICLIT_fwd	GTATCTCTCGAG**AAAAGA**TCCACTCAAAACGGTAGATC
Van1_pPICLIT_rev	GATGATGATGGTCGACCTCAGACTGCCTTCTC
Mnn2_pPICLIT_fwd	GTATCTCTCGAG**AAAAGA**TCCGTCAAAGAGTACAAAG
Mnn2_pPICLIT_rev	GATGATGATGGTCGACCTTACCAGCGATAGCC
Mnn9_pJAN-s1_fwd	GTGTTTCTTTGGAC**AAGAGA**GGTTTGAACGGTCAATC
Mnn9_pJAN-s1_rev	CCGCGGCCGCGGTCTCGATTAGTGGTTCTCTTCTTCG

Kex2 cleavage site highlighted in bold; restriction sites underlined.

Appendix B. Glycan and glycopeptide fragmentation

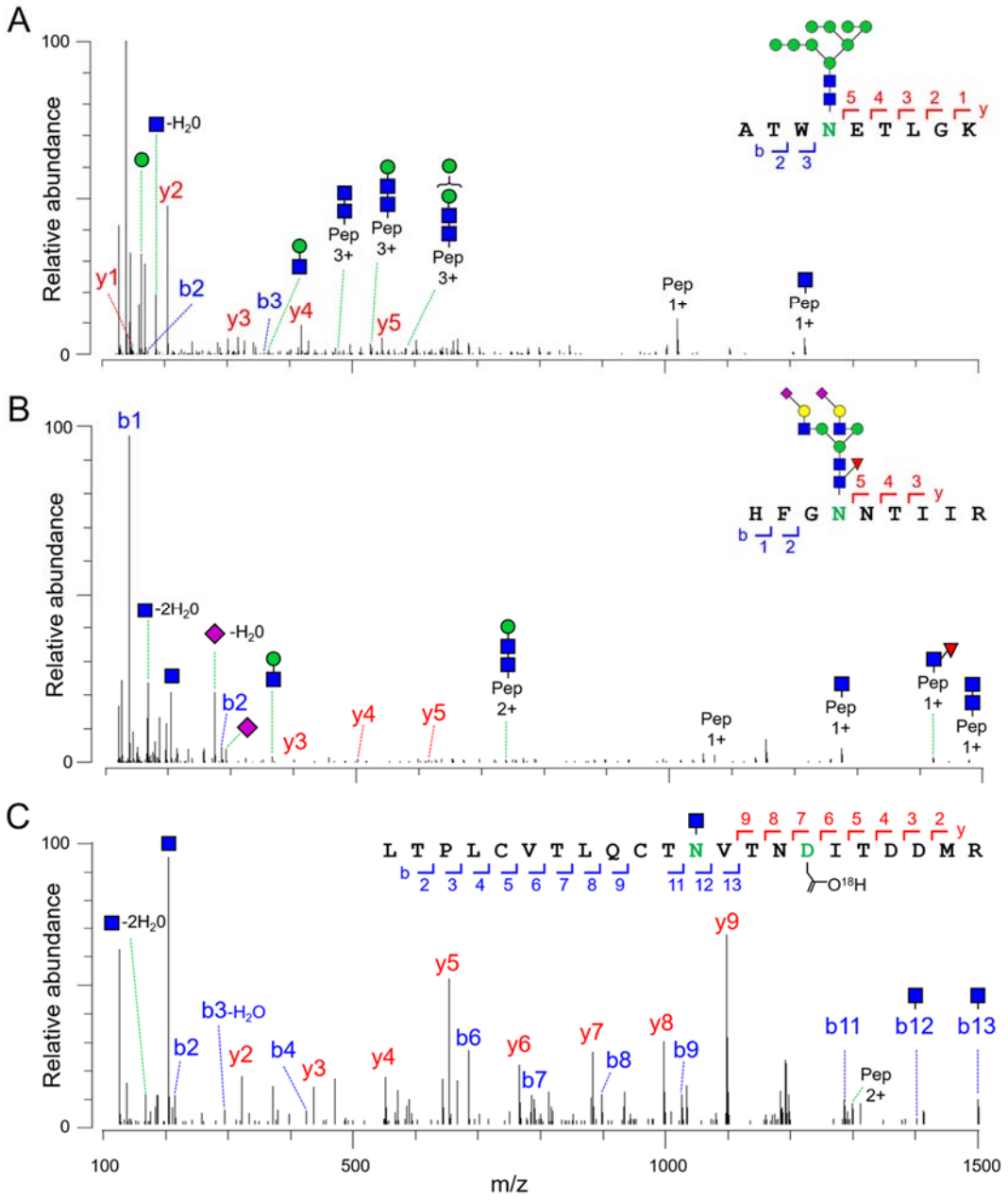


Figure B.1. Example glycopeptide HCD fragmentation spectra. A) Fragmentation of a glycopeptide occupied by a $Man_9GlcNAc_2$ structure. B) Fragmentation of a glycopeptide occupied by a fucosylated bi-antennary glycan. C) Fragmentation of a doubly occupied peptide sequentially treated with Endo H (to cleave oligomannose-type glycans) and PNGase F in the presence of O^{18}-labelled water (to cleave remaining complex-type glycans). The PNGase F reaction coverts the Asn to an O^{18}-labelled Asp. In this example, the untreated glycopeptide therefore contained an oligomannose-type glycan at the N_{12} position and a complex-type glycan at $N(D)_{16}$.

Table B.1. Characteristic oxonium ions from HCD fragmentation of glycopeptides[a].

Oxonium ion (m/z)	Description[b]
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	HexHexNAcHex
657.23	NeuAcHexHexNAc

a. Adapted from Mechref, 2012 [374].
b. 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.

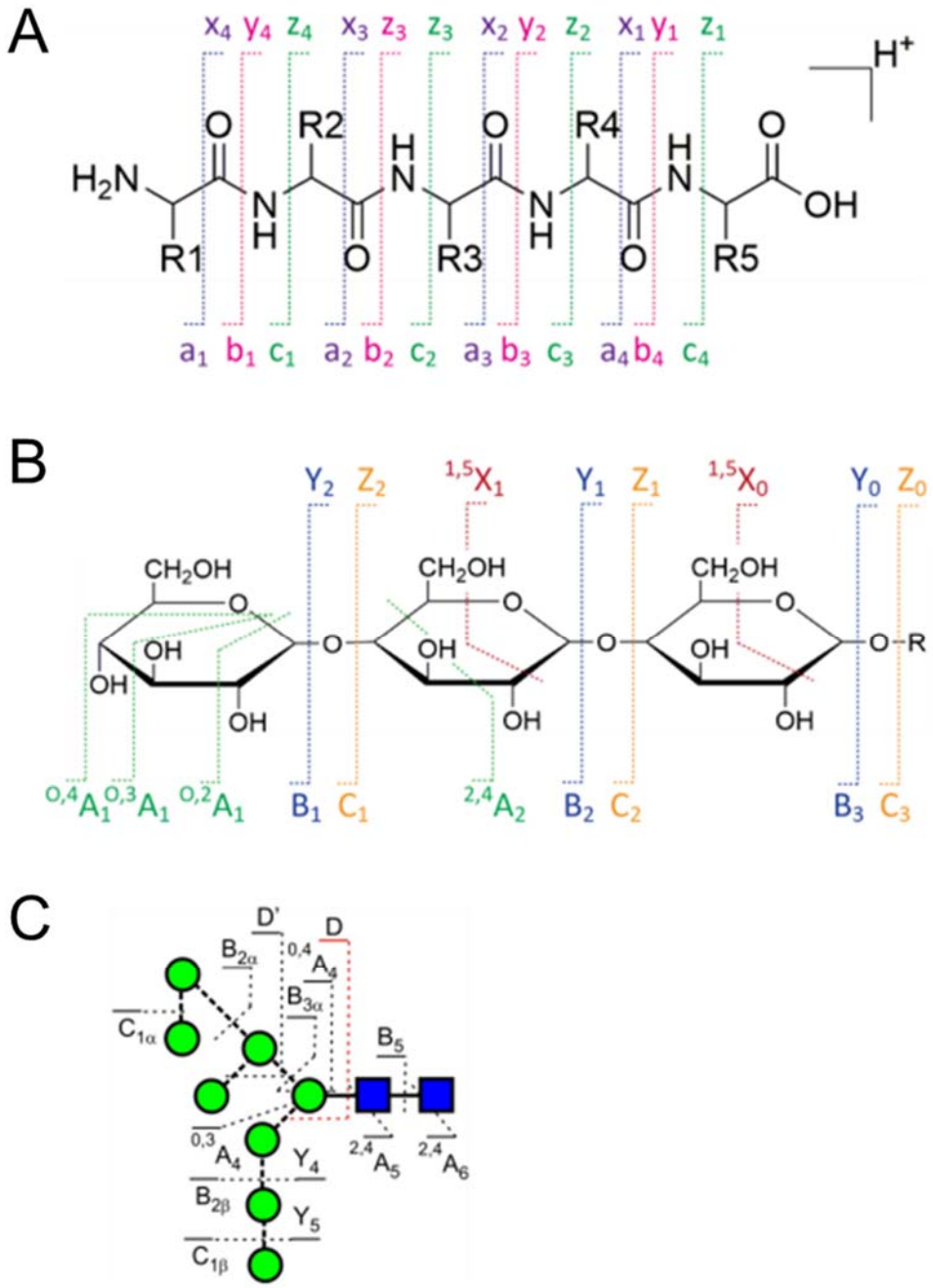


Figure B.2. Fragmentation of glycans and glycopeptides by tandem mass spectrometry. Figure adapted from Behrens, 2017 [366]. A) Ions generated by the fragmentation of the peptide backbone [489]. B) Ions generated by fragmentation of glycans. Subscripts indicate the position relative to the termini, analogous to the system used in peptides, and superscripts indicate cleavages within carbohydrate rings [373]. C) Characteristic ions generated by fragmentation of N-glycans illustrated on $Man_8GlcNAc_2$. The D-ion is highlighted in red [408].

Appendix C. Glycan nomenclature

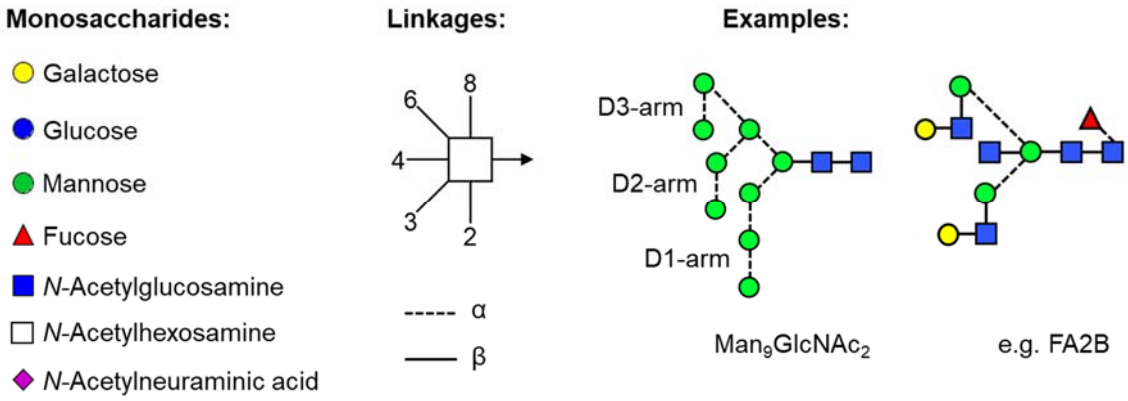


Figure C.1. Glycan nomenclature. Throughout this thesis, the symbolic representation of glycans has followed that of the Consortium for Functional Glycomics, with linkage information based on Harvey *et al.*, 2009 [116, 117]. Numbering of the D1-, D2- and D3-arms is adopted from the system of Vliegenhart *et al.*, 1983 [118].

Table C.1. Classification of glycans for site-specific analysis.

Glycan class	Glycan type
M3-9	$Man_{3\text{-}9}GlcNAc_2$
FM3/4	Fucosylated $Man_{3/4}GlcNAc_2$
H	Hybrid-type
FH	Fucosylated hybrid
A1	Mono-antennary
FA1	Fucosylated mono-antennary
A1B/A2	Mono-antennary with bisecting GlcNAc, or bi-antennary
FA1B/FA2	Fucosylated mono-antennary with bisecting GlcNAc, or fucosylated bi-antennary
A2B/A3	Bi-antennary with bisecting GlcNAc, or tri-antennary
FA2B/FA3	Fucosylated bi-antennary with bisecting GlcNAc, or fucosylated tri-antennary
A3B/A4	Tri-antennary with bisecting GlcNAc, or tetra-antennary
FA3B/FA4	Fucosylated tri-antennary with bisecting GlcNAc, or fucosylated tetra-antennary

Appendix D. Glycan libraries

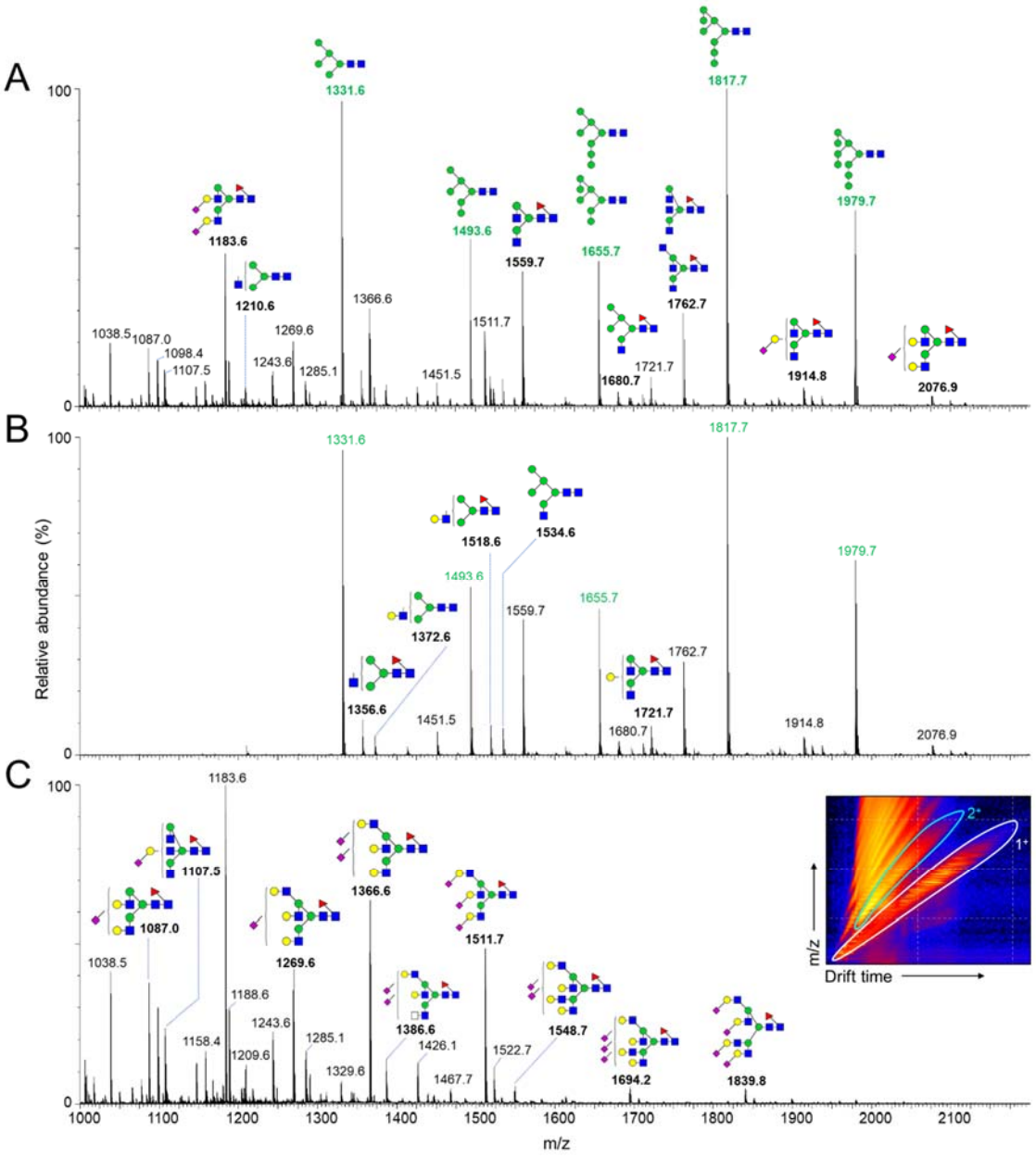


Figure D.1. IM-ESI MS analysis of N-linked glycans from BG505 SOSIP.664 expressed in HEK cells. A) N-linked glycans were enzymatically released from BG505 SOSIP.664 expressed in HEK cells and analysed by negative ion IM-ESI MS, with the help of Prof. David Harvey, University of Oxford. B) Mobility-extracted singly charged negative ions. C) Mobility-extracted doubly charged negative ions. The corresponding ions are encircled in white and blue, respectively, in the inset ion mobility drift plot.

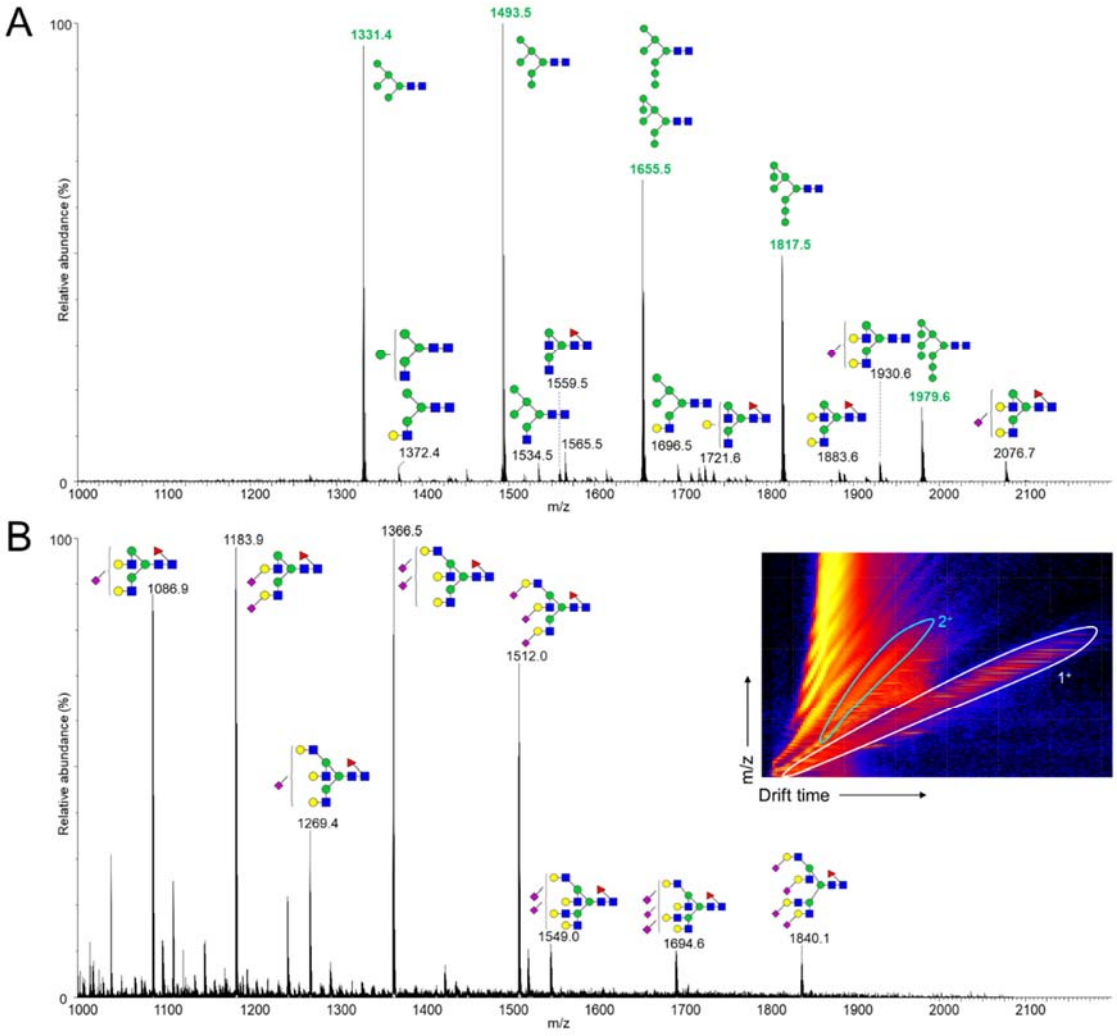


Figure D.2. IM-ESI MS analysis of N-linked glycans from CZA97 SOSIP.664 expressed in CHO cells. N-linked glycans were enzymatically released from CZA97 SOSIP.664 expressed in CHO cells and analysed by negative ion IM-ESI MS, with the help of Yasunori Watanabe, University of Oxford. A) Mobility-extracted singly charged negative ions. B) Mobility-extracted doubly charged negative ions. The corresponding ions are encircled in white and blue, respectively, in the inset ion mobility drift plot.

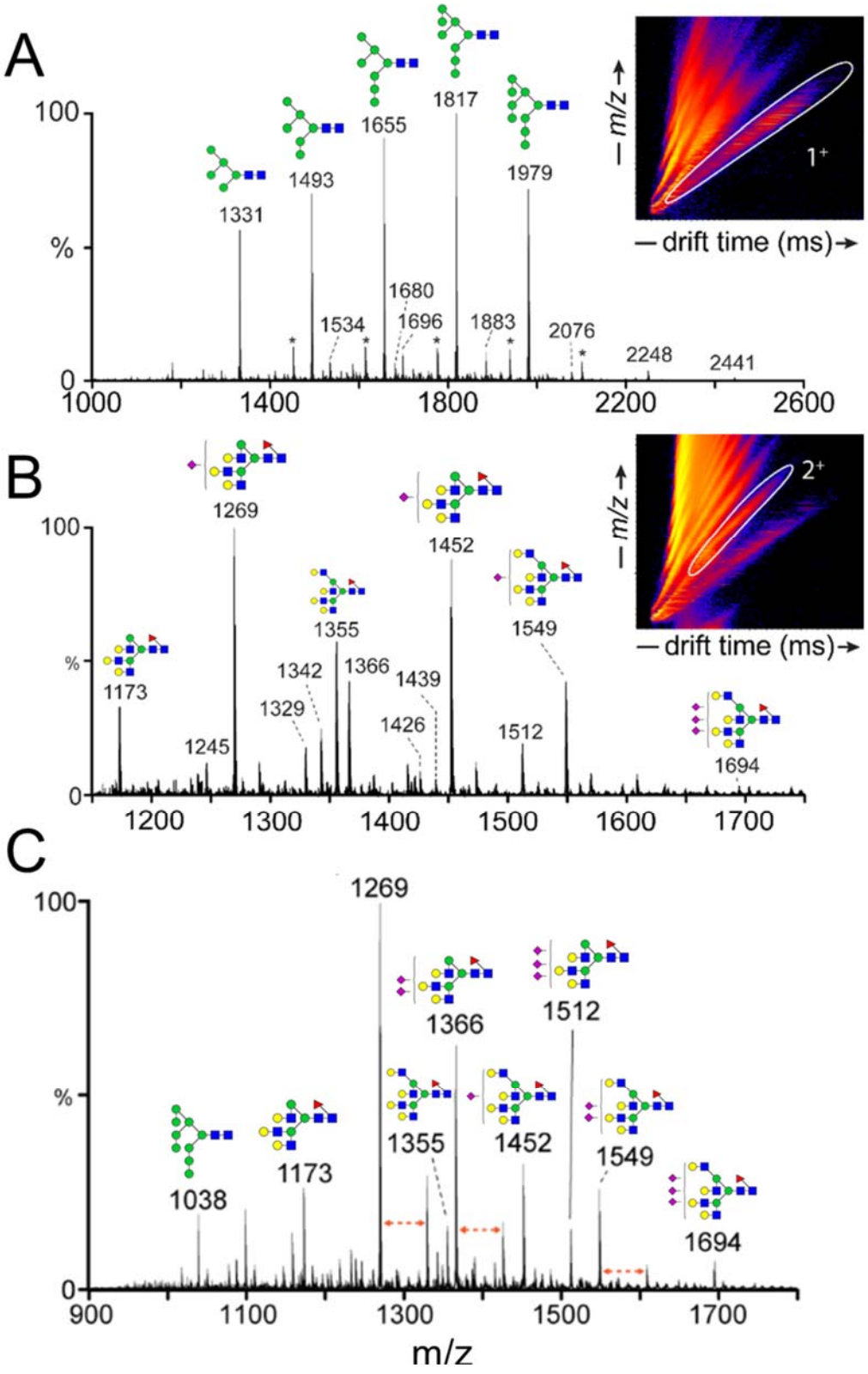
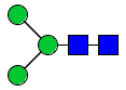
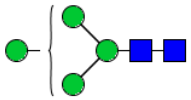
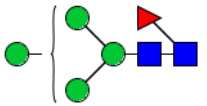
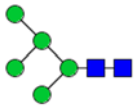
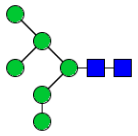
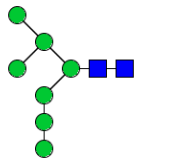
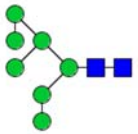
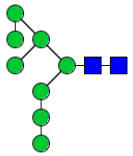
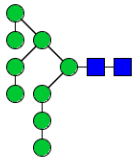
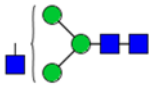
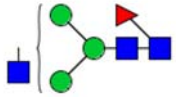


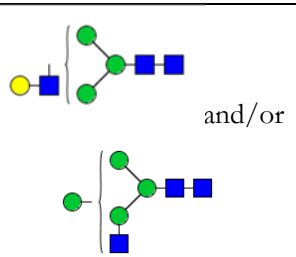
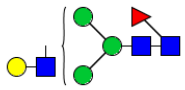
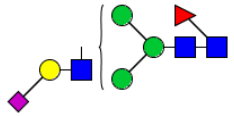
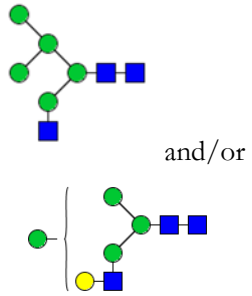
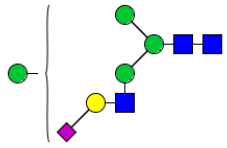
Figure D.3. IM-ESI MS analysis of N-linked glycans from BG505.T332N-LAI virions expressed in A66-R5 T cells. N-linked glycans were enzymatically released from BG505.T332N-LAI virions by in-gel digestion with PNGase F. Glycans were analysed by negative ion IM-ESI MS by Prof. David Harvey, University of Oxford and are published in Struwe *et al.*, 2018 [276]. A) Mobility-extracted singly charged negative ions from BG505.T332N-LAI gp120. The corresponding singly charged ions are encircled in the inset

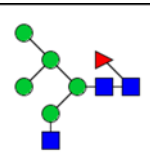
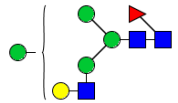
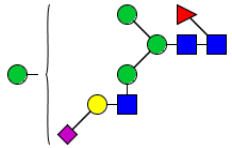
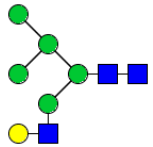
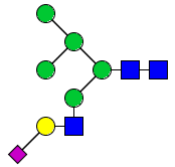
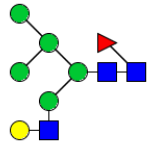
ion mobility drift plot. The series of carbohydrate ions marked with an asterisk is formed by the additional adduction of sodium phosphate to the main oligomannose ions. B) Mobility-extracted doubly charged negative ions from BG505.T332N-LAI gp41. The corresponding doubly charged ions are encircled in the inset ion mobility drift plot. C) Mobility-extracted doubly charged N-glycans from BG505.T332N-LAI gp120. Peaks with additional phosphate adducts are labelled in orange.

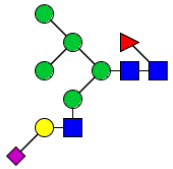
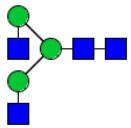
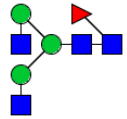
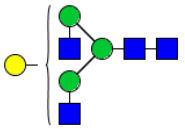
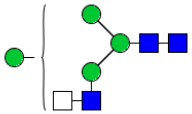
Table D.1. Example glycan library for BG505 SOSIP.664 expressed in HEK cells. Structural assignments are based on tandem mass spectrometry glycan fragmentation data from Behrens *et al.*, 2016 [132], or proposed from the literature. The glycan classification group used for the presentation of site-specific glycan analysis throughout this thesis is given. Where two or more possible isomeric structures could be classed in different groups (e.g. a glycan with an m/z of 1737 could be a bi-antennary or hybrid-type glycan), the most likely/abundant structure is used.

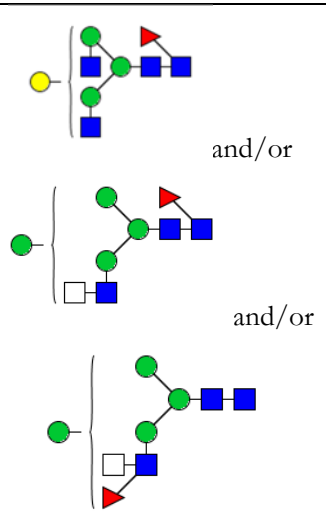
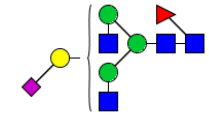
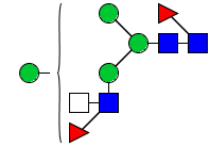
m/z		z	Ion	Composition					Proposed Structure		Glycan class
Obs.	Exp.			Hex NAc	Hex	Fuc	Neu 5Ac	SO4	Behrens *et al.*, 2016	Proposed	
1007.5	1007.3	1	b	2	3	0	0	0	Not present		M3
1169.5	1169.3	1	b	2	4	0	0	0			M4
1315.6	1315.4	1	b	2	4	1	0	0	Not present		FM
1331.6	1331.4	1	b	2	5	0	0	0			M5
1493.6	1493.5	1	b	2	6	0	0	0			M6

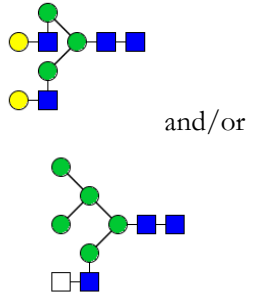
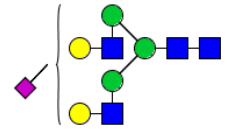
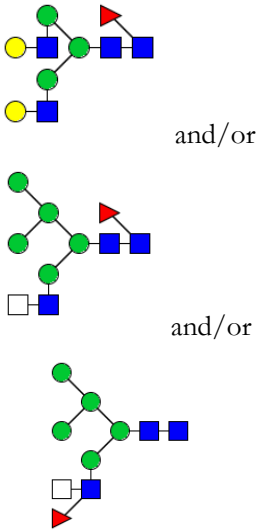
1655.7	1655.5	1	b	2	7	0	0	0	 and/or 		M7
1817.7	1817.6	1	b	2	8	0	0	0			M8
957.4	957.3	2	d	2	8	0	0	0			
1979.7	1979.6	1	b	2	9	0	0	0			M9
1038.5	1038.3	2	d	2	9	0	0	0			
1210.6	1210.4	1	b	3	3	0	0	0			A1
1356.6	1356.4	1	b	3	3	1	0	0			FA1

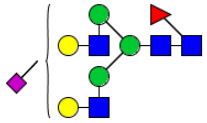
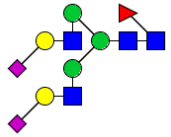
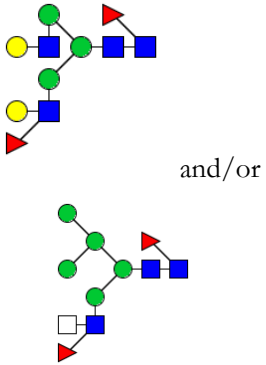
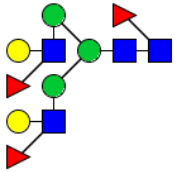
1372.6	1372.4	1	b	3	4	0	0	0			A1
1518.6	1518.5	1	b	3	4	1	0	0			FA1
1711.8	1711.6	1	a	3	4	1	1	0			FA1
1534.6	1534.5	1	b	3	5	0	0	0			H
1727.7	1727.6	1	a	3	5	0	1	0			H

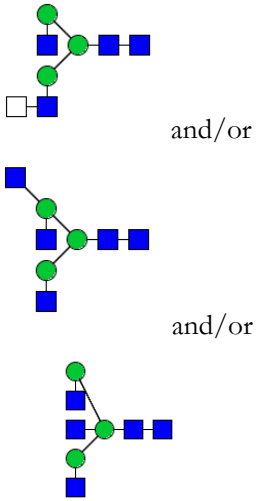
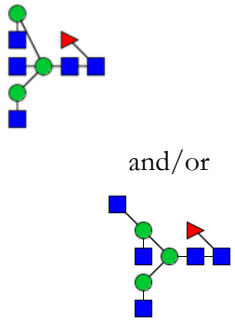
1680.7	1680.5	1	b	3	5	1	0	0	 and/or 		FH
1873.8	1873.7	1	a	3	5	1	1	0			FH
985.5	985.3	2	e	3	5	1	1	0			
1696.7	1696.5	1	b	3	6	0	0	0			H
1889.8	1889.7	1	a	3	6	0	1	0			H
1842.7	1842.6	1	b	3	6	1	0	0			FH
2035.8	2035.7	1	a	3	6	1	1	0			FH
1017.4	1017.4	2	c	3	6	1	1	0			

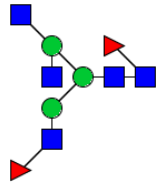
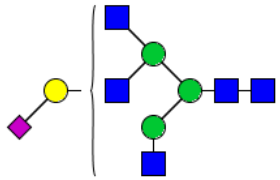
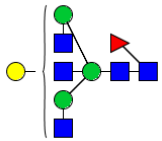
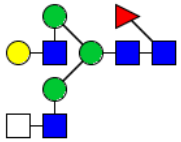
1066.0	1066.3	2	e	3	6	1	1	0			
1413.6	1413.5	1	b	4	3	0	0	0			A2
1559.7	1559.5	1	b	4	3	1	0	0			FA2
1575.7	1575.5	1	b	4	4	0	0	0	 and/or 		A2

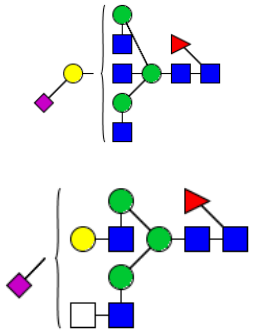
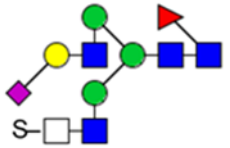
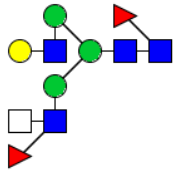
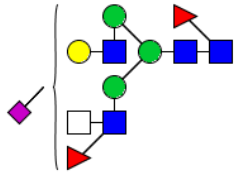
1721.7	1721.6	1	b	4	4	1	0	0			FA2
1914.8	1914.7	1	a	4	4	1	1	0			FA2
1867.8	1867.6	1	b	4	4	2	0	0			FH

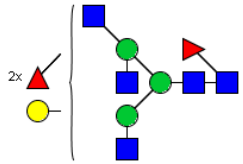
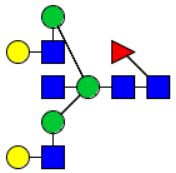
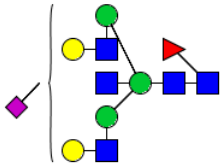
1737.7	1737.6	1	b	4	5	0	0	0	 and/or		A2
1930.8	1930.7	1	a	4	5	0	1	0			A2
1883.7	1883.6	1	b	4	5	1	0	0	 and/or and/or		A2
2076.9	2076.7	1	a	4	5	1	1	0			FA2

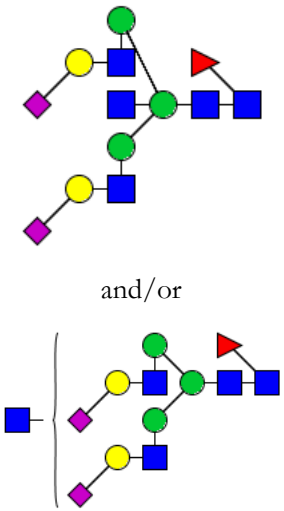
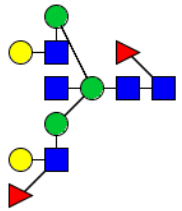
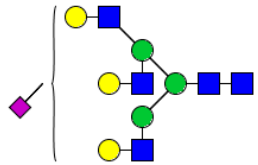
1087.0	1086.9	2	e	4	5	1	1	0			
2367.9	2367.8	1	a	4	5	1	2	0			FA2
1183.6	1183.4	2	c	4	5	1	2	0			
2029.8	2029.7	1	b	4	5	2	0	0	 and/or		FA2
1063.5	1063.3	2	d	4	5	2	0	0			
2175.8	2175.7	1	b	4	5	3	0	0	Not present		FA2

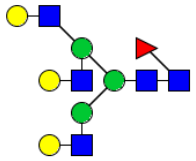
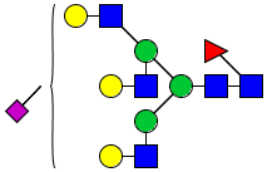
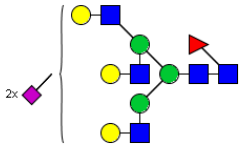
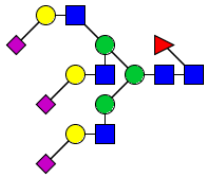
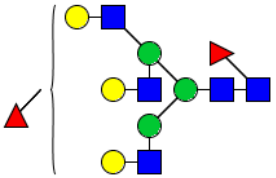
1616.7	1616.5	1	b	5	3	0	0	0	Not fragmented		A2B/3
1762.7	1762.6	1	b	5	3	1	0	0			FA2B/3

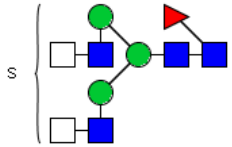
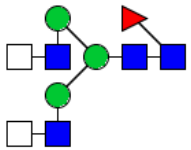
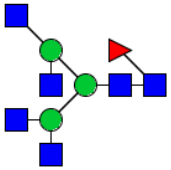
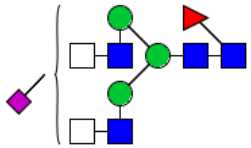
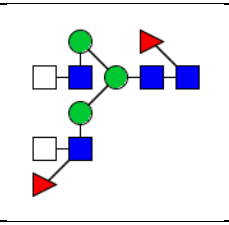
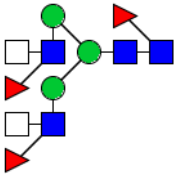
1908.8	1908.7	1	b	5	3	2	0	0	Not present		FA3
1971.8	1971.7	1	a	5	4	0	1	0	Not present		A3
1924.8	1924.6	1	b	5	4	1	0	0	 and/or 		FA2B
1011.0	1010.8	2	d	5	4	1	0	0			
2117.9	2117.7	1	a	5	4	1	1	0			FA2B

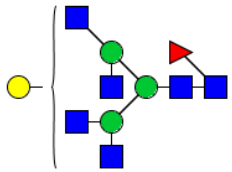
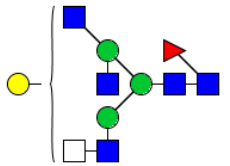
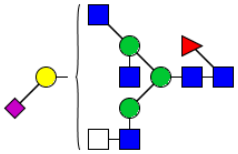
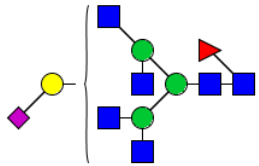
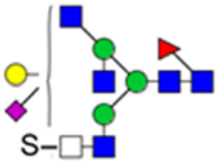
1107.5	1107.4	2	e	5	4	1	1	0			
1098.4	1098.4	2	c	5	4	1	1	1	Not present		FA2
2070.8	2070.7	1	b	5	4	2	0	0			FA2
1084.0	1083.8	2	d	5	4	2	0	0			
2263.9	2263.8	1	a	5	4	2	1	0			FA2
1180.6	1180.4	2	e	5	4	2	1	0			

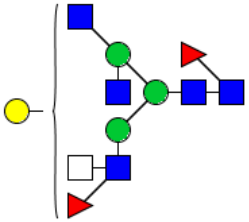
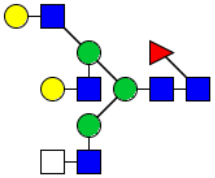
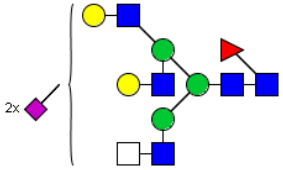
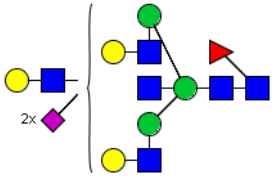
2216.9	2216.8	1	b	5	4	3	0	0	Not present		FA3
2086.8	2086.7	1	b	5	5	1	0	0			FA2B
1092.0	1091.8	2	d	5	5	1	0	0			
2279.9	2279.8	1	a	5	5	1	1	0			FA2B
1188.6	1188.4	2	e	5	5	1	1	0			

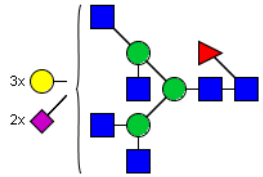
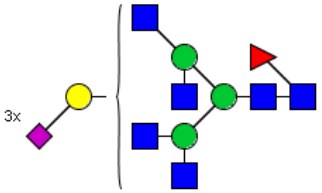
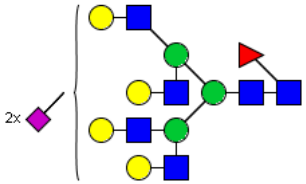
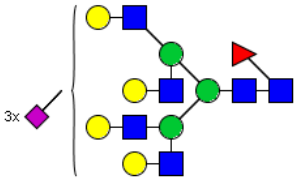
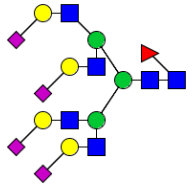
1285.1	1285.0	2	c	5	5	1	2	0	Not present		FA2B/FA3
2232.8	2232.8	1	b	5	5	2	0	0			
1165.0	1164.9	2	d	5	5	2	0	0	Not present		FA2B
2295.8	2295.8	1	a	5	6	0	1	0	Not fragmented		A3
2248.8	2248.7	1	b	5	6	1	0	0			FA3

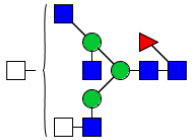
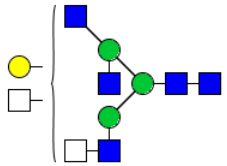
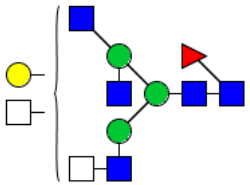
1173.0	1172.9	2	d	5	6	1	0	0			
2441.9	2441.9	1	a	5	6	1	1	0			FA3
1269.6	1269.4	2	e	5	6	1	1	0			
1366.1	1366.0	2	c	5	6	1	2	0			FA3
1511.7	1511.5	2	c	5	6	1	3	0			FA3
1522.7	1522.5	2	c +Na	5	6	1	3	0			
2394.8	2394.8	1	b	5	6	2	0	0	Not present		FA3

1947.8	1947.7	1	a	6	3	1	0	1			FA2
1965.8	1965.7	1	b	6	3	1	0	0			FA2
1031.5	1031.3	2	d	6	3	1	0	0			
2158.9	2158.8	1	a	6	3	1	1	0			FA2
2111.9	2111.7	1	b	6	3	2	0	0			FA2
1104.5	1104.3	2	d	6	3	2	0	0			
2257.8	2257.8	1	b	6	3	3	0	0			FA2

2127.8	2127.7	1	b	6	4	1	0	0	Not fragmented	 and/or 	FA3
1209.6	1209.4	2	e	6	4	1	1	0	Not present	 and/or 	FA4
2399.9	2400.8	1	a	6	4	1	1	1	Not present		FA3

2273.9	2273.8	1	b	6	4	2	0	0	Not present		FA3
2289.9	2289.8	1	b	6	5	1	0	0			FA3
1193.5	1193.4	2	d	6	5	1	0	0			
1386.6	1386.5	2	c	6	5	1	2	0			FA3
1467.7	1467.5	2	c	6	6	1	2	0	Not present		FA3B/FA4

											
1613.2	1613.1	2	c	6	6	1	3	0	Not present		FA4
1548.7	1548.5	2	c	6	7	1	2	0			FA4
1694.2	1694.1	2	c	6	7	1	3	0			FA4
1839.8	1839.6	2	c	6	7	1	4	0	Not present		FA4

1133.0	1132.9	2	d	7	3	1	0	0	Not present		FA3
2184.7	2184.7	1	b	7	4	0	0	0	Not present		A3
2330.9	2330.8	1	b	7	4	1	0	0			FA3
1214.1	1213.9	2	d	7	4	1	0	0			
1451.5		1							Not present	Not determined	
1158.4		2							Not present	Not determined	
1243.6		2							Not present	Not determined	
1329.6		2							Not present	Not determined	
1426.1		2							Not present	Not determined	

a. $[M-H]^-$
b. $[M+H_2PO_4]^-$
c. $[M-H_2]^{2-}$
d. $[M+2H_2PO_4]^{2-}$
e. $[M-H+H_2PO_4]^{2-}$

Appendix E. Env sequences

The following appendix contains the amino acid sequence information for the Env trimers analysed in this thesis (not including PNGS mutants). Sequences were aligned to the HIV-1 reference strain HXB2, using the Clustal Omega multiple sequence alignment tool, and residues are numbered accordingly throughout this thesis. PNGS are highlighted in blue, and stabilising mutations are highlighted in magenta. PNGS not present in the parental strain, e.g. the N332 site in BG505 trimers, are also highlighted in magenta. Table E.1 contains additional information relating to each mutation.

```
HXB2                MRVKEKYQHLWRWGWRWGTMLLG-------------MLMICSATEKLWVTVYYGVPVWKE
ZM197M_NFL_TD+      -------------------MPMGSLQPLATLYLLGMLVASVLAMEQLWVTVYYGVPVWKD
001428_NFL_TD+      -------------------MPMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKD
16055_NFL_TD+       -------------------MPMGSLQPLATLYLLGMLVASVLANGNLWVTVYYGVPVWKD
JR-FL_NFL_TD+       -------------------MPMGSLQPLATLYLLGMLVASVLAVEKLWVTVYYGVPVWKD
B41_SOSIP.v4        -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARA-AKKWVTVYYGVPVWKE
B41_SOSIP.v6        -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARA-AKKWVTVYYGVPVWKE
AMC008_SOSIP.v4     -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGART-EKLWVTVYYGVPVWKE
AMC008_SOSIP.v6     -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGART-EKLWVTVYYGVPVWKE
AMC011_SOSIP.v5     -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARA-EQLWVTVYYGVPVWKE
AMC016_SOSIP.v4     -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAEEELWVTVYYGVPVWKE
CZA97_SOSIP.664     -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGAR-VGNMWVTVYYGVPVWTD
BG505_NFL_TD+       -------------------MPMGSLQPLATLYLLGMLVASVLAAENLWVTVYYGVPVWKD
BG505_SOSIP.v5      -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGAR-AENLWVTVYYGVPVWKD
BG505_SOSIP.664     -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGAR-AENLWVTVYYGVPVWKD
BG505_SOSIP.v4      -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGAR-AENLWVTVYYGVPVWKD
BG505_SOSIP.v6      -------MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGAR-AENLWVTVYYGVPVWKD
BG505_virion        ------------------------------------------AENLWVTVYYGVPVWKD

                    48                                      88                 107
HXB2                ATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNMWKNDMVEQMHED
ZM197M_NFL_TD+      AEATLFCASDAKAYEKEKHNVWATHACVPTDPNPQEIPLGNVTENFNMWKNDMADQMHTD
001428_NFL_TD+      AETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEMVLGNVTENFNMWKNDMVDQMHTD
16055_NFL_TD+       AETTLFCASDAKAYEKEKHNVWATHACVPTDPNPQEMVLENVTENFNMWKNDMVEQMHTD
JR-FL_NFL_TD+       AETTLFCASDAKAYDTEKHNVWATHACVPTDPNPQEVVLENVTEHFNMWKNNMVEQMQTD
B41_SOSIP.v4        ATTTLFCASDAKAYDTKVHNVWATHACVPTDPNPQEIVLGNVTENFNMWKNNMVEQMHED
B41_SOSIP.v6        ACTTLFCASDAKAYDTKVHNVWATCACVPTDPNPQEIVLGNVTENFNMWKNNMVEQMHED
AMC008_SOSIP.v4     ATTTLFCASDAKAYDTEVRNVWATHACVPTDPSPQEVILENVTENFNMWTNNMVEQMHED
AMC008_SOSIP.v6     ACTTLFCASDAKAYDTEVRNVWATHCCVPTDPSPQEVILENVTENFNMWTNNMVEQMHED
AMC011_SOSIP.v5     ATTTLFCASDARAYDTEVRNVWATHCCVPTDPNPQEVVLENVTENFNMWKNNMVEQMHED
AMC016_SOSIP.v4     ATTTLFCASDAKAYDTEVRNVWATHACVPTDPSPQEVVLANVTENFNMWKNNMVEQMHED
CZA97_SOSIP.664     AKTTLFCASDTKAYDREVHNVWATHACVPTDPNPQEIVLENVTENFNMWKNDMVDQMHED
BG505_NFL_TD+       AETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTD
BG505_SOSIP.v5      AETTLFCASDAKAYETKKHNVWATCACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTD
BG505_SOSIP.664     AETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTD
BG505_SOSIP.v4      AETTLFCASDAKAYETKKHNVWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTD
BG505_SOSIP.v6      ACTTLFCASDAKAYETKKHNVWATHCCVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTD
BG505_virion        AETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEEFNMWKNNMVEQMHTD

                    108                 130    137                          155
HXB2                IISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNS-----------SSGRMIMEKGEIK
ZM197M_NFL_TD+      IISLWDQSLKPCVKLTPLCVTLNCSDATSNTTKNATNTNTTSTDNRNATSNDTEMKGEIK
001428_NFL_TD+      VISLWAQSLKPCVKLTPLCVTLECTQVNATQGNTT-----------QVNVTQVNGDEMK
16055_NFL_TD+       VISLWDQSLKPCVKLTPLCVTLECRQVNTTNAT--------------SSVNVTNGEEIK
JR-FL_NFL_TD+       IISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTN-------------DSEGTMERGEIK
B41_SOSIP.v4        IISLWDQSLKPCVKLTPLCVTLNCNNVNTNNTNNSTNAT--------ISDWEKMETGEMK
B41_SOSIP.v6        IISLWDQSLKPCVKLTPLCVTLNCNNVNTNNTNNSTNAT--------ISDWEKMETGEMK
AMC008_SOSIP.v4     IISLWDQSLKPCVKLTPLCVTLNCTNELKNTT----------------KTNNSSWGGEMK
AMC008_SOSIP.v6     IISLWDQSLKPCVKLTPLCVTLNCTNELKNTT----------------KTNNSSWGGEMK
AMC011_SOSIP.v5     IISLWDQSLKPCVKLTPLCVTLNCTDLRNATNTNATNTT--------SSSRGTMEGGEIK
```

```
AMC016_SOSIP.v4      IISLWDQSLKPCVKLTPLCVTLNCTDLGNATDAINRNTT----DAPNSTLRTMEEKGEIK
CZA97_SOSIP.664      IISLWDQSLKPCVKLTPLCVTLHCTNATFKNN-----------------VTNDMNKEIR
BG505_NFL TD+        IISLWDQSLKPCVKLTPLCVTLQCTNVTNN--------------------ITDDMRGELK
BG505_SOSIP.v5       IISLWDQSLKPCVKLTPLCVTLQCTNVTNN--------------------ITDDMRGELK
BG505_SOSIP.664      IISLWDQSLKPCVKLTPLCVTLQCTNVTNN--------------------ITDDMRGELK
BG505_SOSIP.v4       IISLWDQSLKPCVKLTPLCVTLQCTNVTNN--------------------ITDDMRGELK
BG505_SOSIP.v6       IISLWDQSLKPCVKLTPLCVTLQCTNVTNN--------------------ITDDMRGELK
BG505_virion         IISLWDQSLKPCVKLTPLCVTLQCTNVTNN--------------------ITDDMRGELK

                     156 160                                          197      206
HXB2                 NCSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTT---------SYKLTSCNTSVITQACP
ZM197M_NFL TD+       NCTFNITTELRDRKTKVRALFYKLDVVPLEEEKNSSSKNSSYKEYRLISCNTSTCTQACP
001428_NFL TD+       NCSFNTTTELRDKKQKVYALFYRLDLVPLERENRGDSN--SASKYILINCNTSAITQACP
16055_NFL TD+        NCSFNATTELRDKKQKVYALFYRLDIVPLEEERKGNS-----SKYRLINCNTSACTQACP
JR-FL_NFL TD+        NCSFNITTELRDKVQKVYALFYKLDVVPIDNNNT---------SYRLISCDTSVCTQACP
B41_SOSIP.v4         NCSFNVTTSIRDKIKKEYALFYKLDVVPLENKNNIN--NTNITNYRLINCNTSVITQACP
B41_SOSIP.v6         NCSFNVTTSIRDKIKKEYALFYKLDVVPLENKNNIN--NTNITNYRLINCNTSVITQACP
AMC008_SOSIP.v4      NCSFNVTTSIRDKVQKEYALFYKLDIVPIDDDNN-------TSNYRLINCNTSVITQACP
AMC008_SOSIP.v6      NCSFNVTTSIRDKVQKEYALFYKLDIVPIDDDNN-------TSNYRLINCNTSVITQACP
AMC011_SOSIP.v5      NCSFNITTSMRDKVQKEYALFYKLDVVPIKNDNT---------SYRLISCNTSVITQACP
AMC016_SOSIP.v4      NCSFNITTSVRDKMQKEYATFYKLDIVPIDNDNN---------SYRLINCNTSVITQACP
CZA97_SOSIP.664      NCSFNTTTEIRDKKQQGYALFYRPDIVLLKE-NR---NNSNNSEYILINCNASTITQACP
BG505_NFL TD+        NCSFNMTTELRDKKQKVYSLFYRLDVVQINE-NQGNRSNNSNKEYRLINCNTSACTQACP
BG505_SOSIP.v5       NCSFNMTTELRDKKQKVYSLFYRLDVVQINE-NQGNRSNNSNKEYRLINCNTSAITQACP
BG505_SOSIP.664      NCSFNMTTELRDKKQKVYSLFYRLDVVQINE-NQGNRSNNSNKEYRLINCNTSAITQACP
BG505_SOSIP.v4       NCSFNMTTELRDKKQKVYSLFYRLDVVQINE-NQGNRSNNSNKEYRLINCNTSAITQACP
BG505_SOSIP.v6       NCSFNMTTELRDKKQKVYSLFYRLDVVQINE-NQGNRSNNSNKEYRLINCNTSAITQACP
BG505_virion         NCSFNMTTELRDKKQKVYSLFYRLDVVQINE-NQGNRSNNSNKEYRLINCNTSAITQACP

                     207                     230 234    241                  262 266
HXB2                 KVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLA
ZM197M_NFL TD+       KVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCHNVSTVQCTHGIKPVVSTQLLLNGSLA
001428_NFL TD+       KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGSCNNVSTVQCTHGIKPVVSTQLLLNGSLA
16055_NFL TD+        KVTFDPIPIHYCAPAGYAILKCNNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLA
JR-FL_NFL TD+        KISFEPIPIHYCAPAGFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLA
B41_SOSIP.v4         KVSFEPIPIHYCAPAGFAILKCNSKTFNGSGPCTNVSTVQCTHGIRPVVSTQLLLNGSLA
B41_SOSIP.v6         KVSFEPIPIHYCAPAGFAILKCNSKTFNGSGPCTNVSTVQCTHGIRPVVSTQLLLNGSLA
AMC008_SOSIP.v4      KITFEPIPIQFCTPAGFAILKCNNKKFNGKGPCTNVSTVQCTHGIRPVVSTQLLLNGSLA
AMC008_SOSIP.v6      KITFEPIPIQFCTPAGFAILKCNNKKFNGKGPCTNVSTVQCTHGIRPVVSTQLLLNGSLA
AMC011_SOSIP.v5      KVSFEPIPIHYCAPAGFAILKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLA
AMC016_SOSIP.v4      KVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLA
CZA97_SOSIP.664      KVNFDPIPIHYCAPAGYAILKCNNKTFSGKGPCNNVSTVQCTHGIKPVVSTQLLLNGSLA
BG505_NFL TD+        KVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLA
BG505_SOSIP.v5       KVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLA
BG505_SOSIP.664      KVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLA
BG505_SOSIP.v4       KVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLA
BG505_SOSIP.v6       KVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLA
BG505_virion         KVSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLA

                     267      276          289   295   301                      325
HXB2                 EEEVVIRSVNFTDNAKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRAFVTIGK-IGN
ZM197M_NFL TD+       EEEIIIRSENLTDNTKTIIVHLNESVEINCTRPNNYTRKSVRI--GPGQTFFATGEIIGD
001428_NFL TD+       EEEIIIRSENLTDNVKTIIVHLDQSVEIVCTRPNNYTRKSIRI--GPGQTFYATGDIIGN
16055_NFL TD+        EGEIIIRSENLTNNVKTIIVHLNESVEIVCTRPNNYTRKSIRI--GPGQTFYATGDIIGN
JR-FL_NFL TD+        EEEVVIRSDNFTNNAKTIIVQLKESVEINCTRPNNYTRKSIRI--GPGRAFYTTGEIIGD
B41_SOSIP.v4         EEEIVIRSENITDNAKTIIVQLNEAVEINCTRPNNNTRKSIHI--GPGQWFYATGDIIGN
B41_SOSIP.v6         EEEIVIRSENITDNAKTIIVQLNEAVEINCTRPNNNTRKSIHI--GPGRWFYATGDIIGN
AMC008_SOSIP.v4      EEEVIIRSDNFTDNAKTIIVQLNESVVINCTRPNNNTRKSINI--GPGRWFYTTGEIIGD
AMC008_SOSIP.v6      EEEVIIRSDNFTDNAKTIIVQLNESVVINCTRPNNNTRKSINI--GPGRWFYTTGEIIGD
AMC011_SOSIP.v5      EEEVVIRSANFTDNAKIIIVQLNKSVEINCTRPNNNTRKSIHI--GPGRWFYTTGEIIGD
AMC016_SOSIP.v4      EEEIVIRSENFTDNAKTIIVQLNESVEINCTRPNNNTRKSIHI--GPGRWFYTTGQIIGN
CZA97_SOSIP.664      EKEIIIRSENLTDNVKTIIVHLNKSVEIVCTRPNNNTRKSMRI--GPGQTFYATGDIIGD
BG505_NFL TD+        EEEVMIRSENITNNAKNILVQFNTPVQINCTRPNNYTRKSIRI--GPGQAFYATGDIIGD
BG505_SOSIP.v5       EEEVMIRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRI--GPGQWFYATGDIIGD
BG505_SOSIP.664      EEEVMIRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRI--GPGQAFYATGDIIGD
BG505_SOSIP.v4       EEEVMIRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRI--GPGQWFYATGDIIGD
BG505_SOSIP.v6       EEEVMIRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRI--GPGQWFYATGDIIGD
BG505_virion         EEEVMIRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRI--GPGQAFYATGDIIGD

                     326   332    339              356   362                    385
HXB2                 MRQAHCNISRAKWNNTLKQIASKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYC
ZM197M_NFL TD+       IRQAHCNLSKSNWTTTLKRIEKKLKEHFNN-ATIKFESSAGGDLEITTHSFNCRGEFFYC
001428_NFL TD+       IREAHCNISEKKWHEMLRRVSEKLAEHFPN-KTIKFTSSSGGDLEITTHSFNCRGEFFYC
16055_NFL TD+        IRQAYCNISKDDWIRTLQRVGKKLAEHFPR-RIINFTSPAGGDLEITTHSFNCRGEFFYC
JR-FL_NFL TD+        IRQAHCNISRAKWNDTLKQIVIKLREQFEN-KTIVFNHSSGGDPEIVMHSFNCGGEFFYC
B41_SOSIP.v4         IRQAHCNISKARWNETLGQIVAKLEEQFPN-KTIIFNHSSGGDPEIVTHSFNCGGEFFYC
```

```
B41_SOSIP.v6          IRQAHCNISKARWNETLGQIVAKLEEQFPN-KTIIFNHSSGGDPEIVTHSFNCGGEFFYC
AMC008_SOSIP.v4       IRQAHCNLSRTQWNNTLKQIAIKLREQFEN-KTIVFNQSSGGDPEIVMHSFNCGGEFFYC
AMC008_SOSIP.v6       IRQAHCNLSRTQWNNTLKQIAIKLREQFEN-KTIVFNQSSGGDPEIVMHSFNCGGEFFYC
AMC011_SOSIP.v5       IRQAHCNISGTKWNDTLKQIVVKLKEQFGN-KTIVFNHSSGGDPEIVMHSFNCGGEFFYC
AMC016_SOSIP.v4       IRQAHCNISRAKWNNTLHKIVKKLREQFRN-KTIVFKQSSGGDPEIVMHSFNCGGEFFYC
CZA97_SOSIP.664       IRQAYCNISGSKWNETLKRVKEKLQENYNNNKTIKFAPSSGGDLEITTHSFNCRGEFFYC
BG505_NFL_TD+         IRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYC
BG505_SOSIP.v5        IRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYC
BG505_SOSIP.664       IRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYC
BG505_SOSIP.v4        IRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYC
BG505_SOSIP.v6        IRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYC
BG505_virion          IRQAHCNVSKATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYC

                      386   392  397      406                                  445
HXB2                  NSTQLFNSTWFNSTWSTEGSNNTEGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRC
ZM197M_NFL_TD+        NTSGLFNSSLLNDTDGTS---NSTSNATITLPCRIKQIINMWQRVGQCMYASPIAGIITC
001428_NFL_TD+        NTSGLFNSTYMPNGTYMPNG-TNNSNSTIILPCRIKQIINMWQRVGQCMYAPPIAGNITC
16055_NFL_TD+         NTSSLFNSTYNPNDTNSNSS-SSNSSLDITIPCRIKQIINMWQRVGQCMYAPPIEGNITC
JR-FL_NFL_TD+         NSTQLFNSTWNNNT---EGSNNTE-GNTITLPCRIKQIINMWQRVGQCMYAPPIRGQIRC
B41_SOSIP.v4          NTTPLFNSTWNNTR---TDDYPTGGEQNITLQCRIKQIINMWQGVGKAMYAPPIRGQIRC
B41_SOSIP.v6          NTTPLFNSTWNNTR---TDDYPTGGEQNITLQCRIKQIINMWQGVGKAMYAPPIRGQIRC
AMC008_SOSIP.v4       NTTKLFNSTWNDTD---IRGNNTEGNDTITIPCRIKQIVNMWQEVGKAMYAPPIRGQIRC
AMC008_SOSIP.v6       NTTKLFNSTWNDTD---IRGNNTEGNDTITIPCRIKQIVNMWQEVGKAMYAPPIRGQIRC
AMC011_SOSIP.v5       NSTQLFNSTWNDTT---GS----NYTGTIVLPCRIKQIVNMWQEVGKAMYAPPIKGQIRC
AMC016_SOSIP.v4       NSTQLFNSTWYGNE---SSD-NPGVEGNITLPCRIKQIINLWQEVGKAMYAPPIGGQIRC
CZA97_SOSIP.664       NTTRLFNNN------------ATEDETITLPCRIKQIINMWQGVGRAMYAPPIAGNITC
BG505_NFL_TD+         NTSGLFNSTWISNTSV-QGSNSTGSNDSITLPCRIKQIINMWQRIGQCMYAPPIQGVIRC
BG505_SOSIP.v5        NTSGLFNSTWISNTSV-QGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRC
BG505_SOSIP.664       NTSGLFNSTWISNTSV-QGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRC
BG505_SOSIP.v4        NTSGLFNSTWISNTSV-QGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRC
BG505_SOSIP.v6        NTSGLFNSTWISNTSV-QGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRC
BG505_virion          NTSGLFNSTWISNTSV-QGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRC

                      446               463                                  504
HXB2                  SSNITGLLLTRDGGNSN-NESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRR
ZM197M_NFL_TD+        KSNITGLLLTRDGGNKS-AGIETFRPGGGNMKDNWRSELYKYKVVEIKPLGIAPTRAKRR
001428_NFL_TD+        NSNITGLLLVRDGGKN--NNTEIFRPGGGDMRDNWRSELYKYKVVEIKPLGVAPTRAKRR
16055_NFL_TD+         KSNITGLLLVRDGGVES-NETEIFRPGGGDMRNNWRSELYKYKVVEIKPLGIAPTRAKRR
JR-FL_NFL_TD+         SSNITGLLLTRDGGINE-NGTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRAKRR
B41_SOSIP.v4          SSNITGLLLTRDGGRDQ-NGTETFRPGGGNMRDNWRSELYKYKVVKIEPLGIAPTACKRR
B41_SOSIP.v6          SSNITGLLLTRDGGRDQ-NGTETFRPGGGNMRDNWRSELYKYKVVKIEPLGIAPTACKRR
AMC008_SOSIP.v4       SSNITGLLLTRDGGSES-NTTEIFRPGGGDMRDNWRSELYKYKVVRIEPLGVAPTKCKRR
AMC008_SOSIP.v6       SSNITGLLLTRDGGSES-NTTEIFRPGGGDMRDNWRSELYKYKVVRIEPLGVAPTKCKRR
AMC011_SOSIP.v5       SSNITGLILIRDGGKNRSENTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGIAPTKCKRR
AMC016_SOSIP.v4       SSNITGLLLTRDGGNNN-ITTEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKCKRR
CZA97_SOSIP.664       KSNITGLLLVRDGGEDN-KTEEIFRPGGGNMKDNWRSELYKYKVIELKPLGIAPTGCKRR
BG505_NFL_TD+         VSNITGLILTRDGGSTN-STTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRAKRR
BG505_SOSIP.v5        VSNITGLILTRDGGSTN-STTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRR
BG505_SOSIP.664       VSNITGLILTRDGGSTN-STTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRR
BG505_SOSIP.v4        VSNITGLILTRDGGSTN-STTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRR
BG505_SOSIP.v6        VSNITGLILTRDGGSTN-STTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRR
BG505_virion          VSNITGLILTRDGGSTN-STTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRAKRR

                      505                                                    558
HXB2                  VVQREKR------AVGIGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRA
ZM197M_NFL_TD+        VVEGGGGSGGGGSAAGIGAVRRGFLGAAGSTMGAASVMLTVQARQLLSGIVQQQSNLLRA
001428_NFL_TD+        VVEGGGGSGGGGSAVGLGAVRRGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLQA
16055_NFL_TD+         VVEGGGGSGGGGSAVGLGAVRRGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLKA
JR-FL_NFL_TD+         VVQGGGGSGGGGSAVGIGAVRRGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRA
B41_SOSIP.v4          VVQRRRRRR----AVGLGAFILGFLGAAGSTMGAASMALTVQARLLLSGIVQQQNNLLRA
B41_SOSIP.v6          VVQRRRRRR----AVGLGAFILGFLGAAGSTMGAASMALTVQARNLLSGIVQQQNNCLRA
AMC008_SOSIP.v4       VVQRRRRRRAV---GTIGAMFLGFLGAAGSTMGAASITLTVQARLLLSGIVQQQNNLLRA
AMC008_SOSIP.v6       VVQRRRRRRAV---GTIGAMFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQNNCLRA
AMC011_SOSIP.v5       VVQRRRRRRAV---G-IGAVFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRA
AMC016_SOSIP.v4       VVQRRRRRRAV---GTIGAMFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQSNLLRA
CZA97_SOSIP.664       VVERRRRRR----AVGIGAVFLGFLGAAGSTMGAASMTLTVQARQLLSSIVQQQSNLLRA
BG505_NFL_TD+         VVGGGGGSGGGGSAVGIGAVRRGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRA
BG505_SOSIP.v5        VVGRRRRRR----AVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRA
BG505_SOSIP.664       VVGRRRRRR----AVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRA
BG505_SOSIP.v4        VVGRRRRRR----AVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRA
BG505_SOSIP.v6        VVGRRRRRR----AVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNCLRA
BG505_virion          VVG-----------------------------------------------------------

                      559                                                611    618
HXB2                  IEAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKS
ZM197M_NFL_TD+        PEAQQHMLQGTVWGIKQLQTRVLAIERYLKDQQLLGLWGCSGKLICTTAVPWNTSWSNKS
001428_NFL_TD+        PEAQQHLLQGTVWGIKQLQTRVLAIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKS
```

```
16055_NFL_TD+       PEAQQHLLQLGVWGIKQLQTRVLAIERYLKDQQLLGIWGCSGKLICTTAVPWNSSWSNKS
JR-FL_NFL_TD+       PEAQQRMLQLGVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTAVPWNASWSNKS
B41_SOSIP.v4        PEAQQHMLQLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKIICCTNVPWNDSWSNKT
B41_SOSIP.v6        PECQQHMLQLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKIICCTNVPWNDSWSNKT
AMC008_SOSIP.v4     PEAQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTSVPWNTSWSNKS
AMC008_SOSIP.v6     PECQQHLLQLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTSVPWNTSWSNKS
AMC011_SOSIP.v5     PECQQHLLKLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICCTAVPWNTSWSNKS
AMC016_SOSIP.v4     PEAQQHLLKLTVWGIKQLQARVLAVERYLKDQQLLGIWGCSGKLICCTAVPWNASWSNKS
CZA97_SOSIP.664     PEAQQHMLKLTVWGIKQLQTRVLAIERYLKDQQLLGIWGCSGKLICCTNVPWNSSWSNKS
BG505_NFL_TD+       PEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTNVPWNSSWSNRN
BG505_SOSIP.v5      PECQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRN
BG505_SOSIP.664     PEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRN
BG505_SOSIP.v4      PEAQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRN
BG505_SOSIP.v6      PECQQHLLKLTVWGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRN
BG505_virion        ------------------------------------------------------------

                    619   625         637                       664
HXB2                LEQIWNHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELD--------------
ZM197M_NFL_TD+      KDEIWDNMTWMQWDREIGNYTQVIYQLLEVSQNQQEKNENDLLALD--------------
001428_NFL_TD+      LTDIWDNMTWMQWDREVGNYTGIIYRLLEDSQNQQERNEQDLLALD--------------
16055_NFL_TD+       HDEIWGNMTWMQWDREIGNYTNTIYRLLEDSQNQQEQNEKDLLALD--------------
JR-FL_NFL_TD+       LDRIWNNMTWMEWEREIDNYTSEIYTLIEESQNQQEKNEQELLALD--------------
B41_SOSIP.v4        INEIWDNMTWMQWEKEIDNYTQHIYTLLEVSQIQQEKNEQELLELD--------------
B41_SOSIP.v6        INEIWDNMTWMQWEKEIDNYTQHIYTLLEVSQIQQEKNEQELLELD--------------
AMC008_SOSIP.v4     LDKIWNNMTWMEWEREIDNYTSLIYTLLEESQNQQEKNEQELLELD--------------
AMC008_SOSIP.v6     LDKIWNNMTWMEWEREIDNYTSLIYTLLEESQNQQEKNEQELLELD--------------
AMC011_SOSIP.v5     YNQIWNNMTWMEWEREIDNYTSLIYTLIEDSQNQQEKNEQELLELD--------------
AMC016_SOSIP.v4     LDNIWNNMTWMEWEKEISNYTNLIYNLIEESQNQQEKNEQELLELD--------------
CZA97_SOSIP.664     QTDIWNNMTWMEWDREISNYTDTIYRLLEDSQTQQEKNEKDLLALD--------------
BG505_NFL_TD+       LSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD--------------
BG505_SOSIP.v5      LSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD--------------
BG505_SOSIP.664     LSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD--------------
BG505_SOSIP.v4      LSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD--------------
BG505_SOSIP.v6      LSEIWDNMTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD--------------
BG505_virion        ------------------------------------------------------------
```

Table E.1. Native-like trimer stabilising mutations.

Mutation[a]	JR-FL NFL TD+	001428 NFL TD+	ZM197M NFL TD+	16055 NFL TD+	BG505 NFL TD+	BG505 SOSIP.664	BG505 SOSIP.V4	BG505 SOSIP.V5	AMC008 SOSIP.V4	AMC008 SOSIP.V6	AMC011 SOSIP.V5	AMC016 SOSIP.V4	B41 SOSIP.V4	B41 SOSIP.V6	CZA97 SOSIP.664	Stabilising effect	Refs
E47D	y	y	y	y												gp120-gp41 interface stability	[382]
T/K/R49E	y	y	y	y												gp120-gp41 interface stability	[382]
V65K	y	y	y	y												gp120-gp41 interface stability	[382]
E106T	y	y	y	y												gp120-gp41 interface stability	[382]
S164E	y															Variable loop stability	[382]
I/V165L	y	y	y	y												Variable loop stability	[382]
E/A/Q172V	y	y	y													Variable loop stability	[382]
H308R	y															Variable loop stability	[382]
E429R	y	y	y	y												Prebridging sheet stability	[382]
K/R432Q	y	y	y	y												Prebridging sheet stability	[382]
K/A/S500R	y		y	y												gp120-gp41 interface stability	[382]
L543N	y															gp120-gp41 interface stability	[382]
N553S	y															gp120-gp41 interface stability	[382]
G588R	y															gp120-gp41 interface stability	[382]
E662A	y															gp120-gp41 interface stability	[382]
I201C	y	y[b]	y	y	y											Prebridging sheet stability (Intra-gp120 disulphide)	[226,
A433C	y	y	y	y	y												382]
L568G		y	y													Prefusion stability	[412]
T569G	y			y												Prefusion stability	[412]
D/S636G		y	y	y												Prefusion stability	[412]
N302Y	y	y	y	y	y											Variable loop (V3) stability	[412]

Table E.1. Continued.

F519R	y	y	y	y	y											Fusion peptide stability	[412]
L520R	y	y	y	y	y											Fusion peptide stability	[412]
G_4S	y	y	y	y	y											Link gp120-gp41	[234]
N332			y	y	y	y	y	y								Glycan epitope	[74]
N156			y													Glycan epitope	-
N295			y													Glycan epitope	-
A501C						y	y	y	y	y	y	y	y	y	y	Link gp120-gp41 (Intra-protomer disulphide)	[130]
T605C						y	y	y	y	y	y	y	y	y	y		
I559P	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	Prefusion stability	[133, 134]
R_6						y	y	y	y	y	y	y	y	y	y	Optimise furin cleavage	[413]
I/L535M										y		y			y	Enhance trimer formation	[414]
L543Q/N	y									y	y	y		y		Enhance trimer formation	[414]
A316W							y	y	y	y	y	y	y	y		Variable loop (V3) stability	[228]
E64K							y	y					y	y		Impede spontaneous sampling of the CD4-induced conformation	[228]
H66R									y	y	y	y				Impede spontaneous sampling of the CD4-induced conformation	[228]
H72C								y						y		Link gp120-gp41 (Intra-protomer disulphide)	[225]
H564C																	
A73C										y	y					Link gp120-gp41 (Intra-protomer disulphide)	[225]
A561C								y		y	y			y			
E49C										y				y		Link gp120-gp41, trimer stability (Inter-protomer disulphide)	[225]
L555C										y				y			
.664	y	y	y	y	y	y	y	y	y	y	y	y	y	y	y	Solubilise, prevent aggregation	[224]

a. Disulphide pairs highlighted in bold.
b. Sample analysed in chapter four differs from published sequence.

Appendix F. Glycan labelling reaction

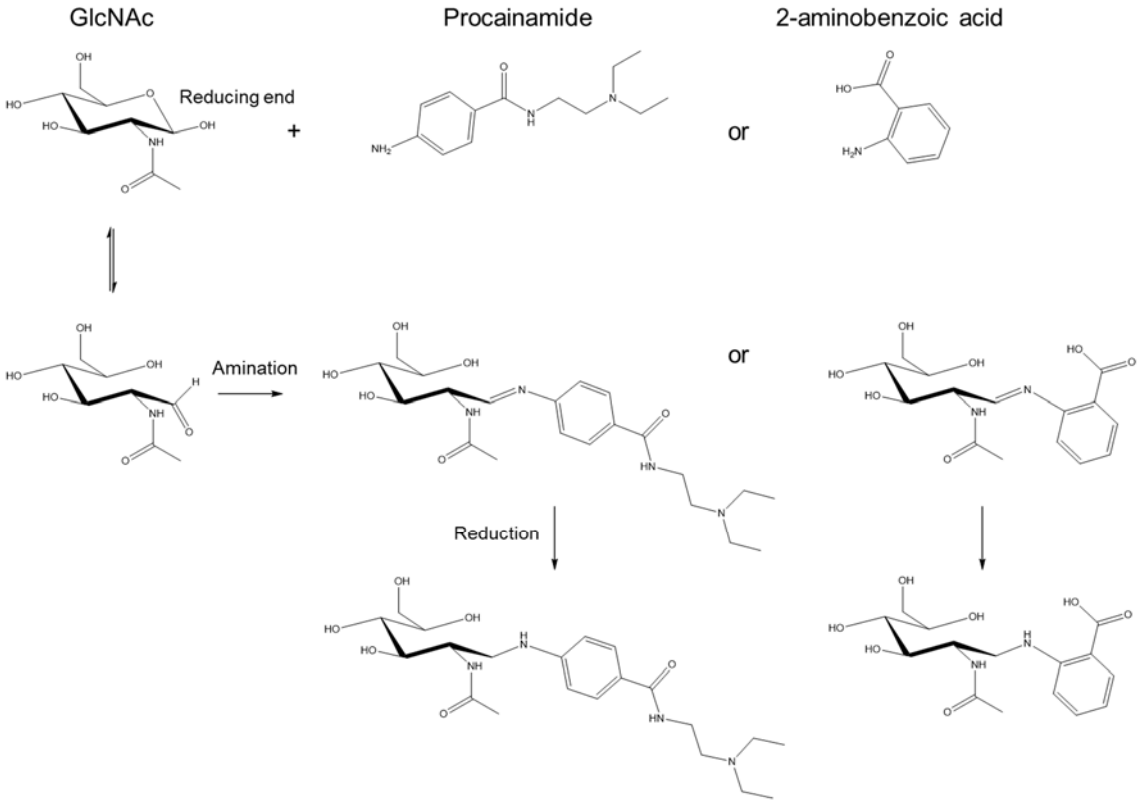


Figure F.1. Fluorescent labelling of enzymatically-released N-linked glycans. N-linked glycans are released from the glycoprotein by in-gel digestion with PNGase F. The reducing end of the terminal *N*-Acetylglucosamine (GlcNAc) can then be labelled using reductive amination. In this reaction, a label containing a primary amine group reacts in a condensation reaction with the aldehyde group of the glycan, resulting in a Schiff base, which is reduced by a reducing agent (sodium cyanoborohydride) to yield a secondary amine [415]. Throughout this thesis, procainamide and 2-aminobenzoic acid (2-AA) have been used to fluorescently label glycans. The respective excitation and emission wavelengths for procainamide and 2-AA are: 310 nm and 370 nm, and 250 nm and 428 nm.

Appendix G. List of publications and conferences

Publications*

- **Seabright GE**†, Cottrell CA†, van Gils MJ, D'addabo A, Harvey DJ, Behrens AJ, Allen J, Watanabe Y, Maker A, Vasiljevic S, de Val N, Sanders RW, Ward AB, Crispin M. *Structure*. 2019. Networks of HIV-1 envelope glycans maintain antibody epitopes in the face of glycan additions and deletions. (Submitted).

- **Seabright GE**, Doores KJ, Burton DR, Crispin M. *J Mol Biol*. 2019. Protein and glycan mimicry in HIV vaccine design.

- Ringe RP, Pugach P, Cottrell CA, LaBranche CC, **Seabright GE**, Ketas TJ, Ozorowski G, Kumar S, Schorcht A, van Gils MJ, Crispin M, Montefiori DC, Wilson IA, Ward AB, Sanders RW, Klasse PJ, Moore JP. *J Virol*. 2018. 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.

- Struwe WB†, Chertova E†, Allen JD†, **Seabright GE**, Watanabe Y, Harvey DJ, Medina-Ramirez M, Roser JD, Smith R, Westcott D, Keele BF, Bess JW Jr, Sanders RW, Lifson JD, Moore JP, Crispin M. *Cell Rep*. 2018. Site-Specific Glycosylation of Virion-Derived HIV-1 Env Is Mimicked by a Soluble Trimeric Immunogen.

- Watanabe Y, Raghwani J, Allen JD, **Seabright GE**, Li S, Moser F, Huiskonen JT, Strecker T, Bowden TA, Crispin M. *Proc Natl Acad Sci U S A*. 2018. Structure of the Lassa virus glycan shield provides a model for immunological resistance.

- Watanabe Y, Vasiljevic S, Allen JD, **Seabright GE**, Duyvesteyn HME, Doores KJ, Crispin M, Struwe WB. *Anal Chem*. 2018. Signature of Antibody Domain Exchange by Native Mass Spectrometry and Collision-Induced Unfolding.

- Harvey DJ, **Seabright GE**, Vasiljevic S, Crispin M, Struwe WB. *J Am Soc Mass Spectrom*. 2018. Isomer Information from Ion Mobility Separation of High-Mannose Glycan Fragments.

- Behrens AJ, **Seabright GE**, Crispin M. Chemical Biology of Glycoproteins, Book Chapter. 2017. Targeting Glycans of HIV Envelope Glycoproteins for Vaccine Design.

- Behrens AJ, Vasiljevic S, Pritchard LK, Harvey DJ, Andev RS, Krumm SA, Struwe WB, Cupo A, Kumar A, Zitzmann N, **Seabright GE**, Kramer HB, Spencer DI, Royle L, Lee JH, Klasse PJ, Burton DR, Wilson IA, Ward AB, Sanders RW, Moore JP, Doores KJ, Crispin M. *Cell Rep*. 2016. Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein.

*Between commencing DPhil (October 2015) and the time of submitting this thesis (May 2019).
†The authors contributed equally.

Conferences

04/2016	EAVI2020 1st PhD training course, Stockholm, Sweden (oral presentation).
11/2016	EAVI2020 Annual meeting and 2nd PhD training course, Barcelona, Spain.
01/2017	CHAVI-ID Annual retreat, La Jolla, USA (2016 Chris Scanlan Memorial Travel Award Winner).
02/2017	EAVI2020 Weekly research call (teleconference, oral presentation).
11/2017	EAVI2020 Annual meeting and 4th PhD training course, Barcelona, Spain.
01/2018	EAVI2020 Weekly research call (teleconference, oral presentation).
07/2018	Immunoshape International Symposium on Glycoimmunology, San Sebastian, Spain.
10/2018	EAVI2020 Annual meeting and 6th PhD training course, Madrid, Spain (oral and poster presentation).
03/2019	EAVI2020 Weekly research call (teleconference, oral presentation).

References

[1] World Health Organisation. HIV/AIDS Fact sheets. 2018. https://www.who.int/news-room/fact-sheets/detail/hiv-aids.

[2] Sharp PM, Hahn BH. Origins of HIV and the AIDS pandemic. Cold Spring Harb Perspect Med. 2011;1:a006841.

[3] Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V. Evolutionary and immunological implications of contemporary HIV-1 variation. British medical bulletin. 2001;58:19-42.

[4] Buonaguro L, Tornesello ML, Buonaguro FM. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. J Virol. 2007;81:10209-19.

[5] Hemelaar J, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. AIDS. 2006;20:W13-23.

[6] Wolfs TF, Zwart G, Bakker M, Goudsmit J. HIV-1 genomic RNA diversification following sexual and parenteral virus transmission. Virology. 1992;189:103-10.

[7] Zhang LQ, MacKenzie P, Cleland A, Holmes EC, Brown AJ, Simmonds P. Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. J Virol. 1993;67:3345-56.

[8] Wolinsky SM, Wike CM, Korber BT, Hutto C, Parks WP, Rosenblum LL, et al. Selective transmission of human immunodeficiency virus type-1 variants from mothers to infants. Science. 1992;255:1134-7.

[9] Shaw GM, Hunter E. HIV transmission. Cold Spring Harb Perspect Med. 2012;2.

[10] Smyth RP, Davenport MP, Mak J. The origin of genetic diversity in HIV-1. Virus research. 2012;169:415-29.

[11] Cuevas JM, Geller R, Garijo R, Lopez-Aldeguer J, Sanjuan R. Extremely High Mutation Rate of HIV-1 In Vivo. PLoS Biol. 2015;13:e1002251.

[12] Zhu P, Liu J, Bess J, Jr., Chertova E, Lifson JD, Grise H, et al. Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature. 2006;441:847-52.

[13] Lasky LA, Groopman JE, Fennie CW, Benz PM, Capon DJ, Dowbenko DJ, et al. Neutralization of the AIDS retrovirus by antibodies to a recombinant envelope glycoprotein. Science. 1986;233:209-12.

[14] Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci U S A. 2008;105:7552-7.

[15] Salazar-Gonzalez JF, Salazar MG, Keele BF, Learn GH, Giorgi EE, Li H, et al. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. J Exp Med. 2009;206:1273-89.

[16] Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell. 1996;86:367-77.

[17] Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, et al. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. Nature. 1996;382:722-5.

[18] Hu WS, Hughes SH. HIV-1 reverse transcription. Cold Spring Harb Perspect Med. 2012;2.

[19] Roberts JD, Bebenek K, Kunkel TA. The accuracy of reverse transcriptase from HIV-1. Science. 1988;242:1171-3.

[20] Desimmie BA, Delviks-Frankenberrry KA, Burdick RC, Qi D, Izumi T, Pathak VK. Multiple APOBEC3 restriction factors for HIV-1 and one Vif to rule them all. J Mol Biol. 2014;426:1220-45.

[21] Campbell EM, Hope TJ. HIV-1 capsid: the multifaceted key player in HIV-1 infection. Nat Rev Microbiol. 2015;13:471-83.

[22] Peng K, Muranyi W, Glass B, Laketa V, Yant SR, Tsai L, et al. Quantitative microscopy of functional HIV post-entry complexes reveals association of replication with the viral capsid. Elife. 2014;3:e04114.

[23] Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. Cell. 2002;110:521-9.

[24] Han Y, Lassen K, Monie D, Sedaghat AR, Shimoji S, Liu X, et al. Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. J Virol. 2004;78:6122-33.

[25] Karn J, Stoltzfus CM. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. Cold Spring Harb Perspect Med. 2012;2:a006916.

[26] Nguyen DH, Hildreth JE. Evidence for budding of human immunodeficiency virus type 1 selectively from glycolipid-enriched membrane lipid rafts. J Virol. 2000;74:3264-72.

[27] Ono A, Freed EO. Plasma membrane rafts play a critical role in HIV-1 assembly and release. Proc Natl Acad Sci U S A. 2001;98:13925-30.

[28] Neil SJ, Zang T, Bieniasz PD. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature. 2008;451:425-30.

[29] Perez-Caballero D, Zang T, Ebrahimi A, McNatt MW, Gregory DA, Johnson MC, et al. Tetherin inhibits HIV-1 release by directly tethering virions to cells. Cell. 2009;139:499-511.

[30] Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li C, Wabwire-Mangen F, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. The New England journal of medicine. 2000;342:921-9.

[31] Hladik F, McElrath MJ. Setting the stage: host invasion by HIV. Nat Rev Immunol. 2008;8:447-57.

[32] Hladik F, Sakchalathorn P, Ballweber L, Lentz G, Fialkow M, Eschenbach D, et al. Initial events in establishing vaginal entry and infection by human immunodeficiency virus type-1. Immunity. 2007;26:257-70.

[33] Norvell MK, Benrubi GI, Thompson RJ. Investigation of microtrauma after sexual intercourse. J Reprod Med. 1984;29:269-71.

[34] Khanna KV, Whaley KJ, Zeitlin L, Moench TR, Mehrazar K, Cone RA, et al. Vaginal transmission of cell-associated HIV-1 in the mouse is blocked by a topical, membrane-modifying agent. J Clin Invest. 2002;109:205-11.

[35] Dimitrov DS, Willey RL, Sato H, Chang LJ, Blumenthal R, Martin MA. Quantitation of human immunodeficiency virus type 1 infection kinetics. J Virol. 1993;67:2182-90.

[36] Guadalupe M, Reay E, Sankaran S, Prindiville T, Flamm J, McNeil A, et al. Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. J Virol. 2003;77:11708-17.

[37] Veazey RS, DeMaria M, Chalifoux LV, Shvetz DE, Pauley DR, Knight HL, et al. Gastrointestinal tract as a major site of CD4+ T cell depletion and viral replication in SIV infection. Science. 1998;280:427-31.

[38] Cummins NW, Badley AD. Making sense of how HIV kills infected CD4 T cells: implications for HIV cure. Mol Cell Ther. 2014;2:20.

[39] Cummins NW, Badley AD. Mechanisms of HIV-associated lymphocyte apoptosis: 2010. Cell Death Dis. 2010;1:e99.

[40] Sodroski J, Goh WC, Rosen C, Campbell K, Haseltine WA. Role of the HTLV-III/LAV envelope in syncytium formation and cytopathicity. Nature. 1986;322:470-4.

[41] Gougeon ML. Apoptosis as an HIV strategy to escape immune attack. Nat Rev Immunol. 2003;3:392-404.

[42] Scarlatti G, Tresoldi E, Bjorndal A, Fredriksson R, Colognesi C, Deng HK, et al. In vivo evolution of HIV-1 co-receptor usage and sensitivity to chemokine-mediated suppression. Nat Med. 1997;3:1259-65.

[43] Connor RI, Sheridan KE, Ceradini D, Choe S, Landau NR. Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. J Exp Med. 1997;185:621-8.

[44] Tersmette M, Gruters RA, de Wolf F, de Goede RE, Lange JM, Schellekens PT, et al. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential HIV isolates. J Virol. 1989;63:2118-25.

[45] Tersmette M, Lange JM, de Goede RE, de Wolf F, Eeftink-Schattenkerk JK, Schellekens PT, et al. Association between biological properties of human immunodeficiency virus variants and risk for AIDS and AIDS mortality. Lancet. 1989;1:983-5.

[46] Simon V, Ho DD. HIV-1 dynamics in vivo: implications for therapy. Nat Rev Microbiol. 2003;1:181-90.

[47] Izquierdo-Useros N, Lorizate M, Puertas MC, Rodriguez-Plata MT, Zangger N, Erikson E, et al. Siglec-1 is a novel dendritic cell receptor that mediates HIV-1 trans-infection through recognition of viral membrane gangliosides. PLoS Biol. 2012;10:e1001448.

[48] Zou Z, Chastain A, Moir S, Ford J, Trandem K, Martinelli E, et al. Siglecs facilitate HIV-1 infection of macrophages through adhesion with viral sialic acids. PLoS One. 2011;6:e24559.

[49] van Kooyk Y, Rabinovich GA. Protein-glycan interactions in the control of innate and adaptive immune responses. Nature immunology. 2008;9:593-601.

[50] Manches O, Frleta D, Bhardwaj N. Dendritic cells in progression and pathology of HIV infection. Trends Immunol. 2014;35:114-22.

[51] Geijtenbeek TB, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GC, Middel J, et al. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell. 2000;100:587-97.

[52] de Witte L, Nabatov A, Pion M, Fluitsma D, de Jong MA, de Gruijl T, et al. Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. Nat Med. 2007;13:367-71.

[53] Ji X, Gewurz H, Spear GT. Mannose binding lectin (MBL) and HIV. Mol Immunol. 2005;42:145-52.

[54] Ebenbichler CF, Thielens NM, Vornhagen R, Marschang P, Arlaud GJ, Dierich MP. Human immunodeficiency virus type 1 activates the classical pathway of complement by direct C1 binding through specific sites in the transmembrane glycoprotein gp41. J Exp Med. 1991;174:1417-24.

[55] Spear GT, Takefman DM, Sullivan BL, Landay AL, Zolla-Pazner S. Complement activation by human monoclonal antibodies to human immunodeficiency virus. J Virol. 1993;67:53-9.

[56] Haurum JS, Thiel S, Jones IM, Fischer PB, Laursen SB, Jensenius JC. Complement activation upon binding of mannan-binding protein to HIV envelope glycoproteins. AIDS. 1993;7:1307-13.

[57] Pruenster M, Wilflingseder D, Banki Z, Ammann CG, Muellauer B, Meyer M, et al. C-type lectin-independent interaction of complement opsonized HIV with monocyte-derived dendritic cells. Eur J Immunol. 2005;35:2691-8.

[58] Bajtay Z, Speth C, Erdei A, Dierich MP. Cutting edge: productive HIV-1 infection of dendritic cells via complement receptor type 3 (CR3, CD11b/CD18). J Immunol. 2004;173:4775-8.

[59] Yu Q, Yu R, Qin X. The good and evil of complement activation in HIV-1 infection. Cell Mol Immunol. 2010;7:334-40.

[60] Wilson JD, Ogg GS, Allen RL, Davis C, Shaunak S, Downie J, et al. Direct visualization of HIV-1-specific cytotoxic T lymphocytes during primary infection. AIDS. 2000;14:225-33.

[61] Turnbull EL, Wong M, Wang S, Wei X, Jones NA, Conrod KE, et al. Kinetics of expansion of epitope-specific T cell responses during primary HIV-1 infection. J Immunol. 2009;182:7131-45.

[62] Borrow P, Lewicki H, Wei X, Horwitz MS, Peffer N, Meyers H, et al. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nat Med. 1997;3:205-11.

[63] Rousseau CM, Daniels MG, Carlson JM, Kadie C, Crawford H, Prendergast A, et al. HLA class I-driven evolution of human immunodeficiency virus type 1 subtype c proteome: immune escape and viral load. J Virol. 2008;82:6434-46.

[64] Brumme ZL, Brumme CJ, Heckerman D, Korber BT, Daniels M, Carlson J, et al. Evidence of differential HLA class I-mediated viral evolution in functional and accessory/regulatory genes of HIV-1. PLoS Pathog. 2007;3:e94.

[65] Betts MR, Nason MC, West SM, De Rosa SC, Migueles SA, Abraham J, et al. HIV nonprogressors preferentially maintain highly functional HIV-specific CD8+ T cells. Blood. 2006;107:4781-9.

[66] Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, et al. A whole-genome association study of major determinants for host control of HIV-1. Science. 2007;317:944-7.

[67] Collins KL, Chen BK, Kalams SA, Walker BD, Baltimore D. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. Nature. 1998;391:397-401.

[68] Mann JK, Chopera D, Omarjee S, Kuang XT, Le AQ, Anmole G, et al. Nef-mediated down-regulation of CD4 and HLA class I in HIV-1 subtype C infection: association with disease progression and influence of immune pressure. Virology. 2014;468-470:214-25.

[69] Tomaras GD, Yates NL, Liu P, Qin L, Fouda GG, Chavez LL, 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. J Virol. 2008;82:12449-63.

[70] Parren PW, Gauduin MC, Koup RA, Poignard P, Sattentau QJ, Fisicaro P, et al. Erratum to "Relevance of the antibody response against human immunodeficiency virus type 1 envelope to vaccine design". Immunology letters. 1997;58:125-32.

[71] Parren PW, Burton DR, Sattentau QJ. HIV-1 antibody--debris or virion? Nat Med. 1997;3:366-7.

[72] Sattentau QJ, Moore JP. Human immunodeficiency virus type 1 neutralization is determined by epitope exposure on the gp120 oligomer. J Exp Med. 1995;182:185-96.

[73] Moore PL, Crooks ET, Porter L, Zhu P, Cayanan CS, Grise H, et al. Nature of nonfunctional envelope proteins on the surface of human immunodeficiency virus type 1. J Virol. 2006;80:2515-28.

[74] Sanders RW, Derking R, Cupo A, Julien JP, Yasmeen A, de Val N, 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. PLoS Pathog. 2013;9:e1003618.

[75] Pancera M, Wyatt R. Selective recognition of oligomeric HIV-1 primary isolate envelope glycoproteins by potently neutralizing ligands requires efficient precursor cleavage. Virology. 2005;332:145-56.

[76] Neil SJ, McKnight A, Gustafsson K, Weiss RA. HIV-1 incorporates ABO histo-blood group antigens that sensitize virions to complement-mediated inactivation. Blood. 2005;105:4693-9.

[77] Arendrup M, Hansen JE, Clausen H, Nielsen C, Mathiesen LR, Nielsen JO. Antibody to histo-blood group A antigen neutralizes HIV produced by lymphocytes from blood group A donors but not from blood group B or O donors. AIDS. 1991;5:441-4.

[78] Rong R, Bibollet-Ruche F, Mulenga J, Allen S, Blackwell JL, Derdeyn CA. Role of V1V2 and other human immunodeficiency virus type 1 envelope domains in resistance to autologous neutralization during clade C infection. J Virol. 2007;81:1350-9.

[79] Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proc Natl Acad Sci U S A. 2003;100:4144-9.

[80] Liao HX, Lynch R, Zhou T, Gao F, Alam SM, Boyd SD, et al. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. Nature. 2013;496:469-76.

[81] Bonsignori M, Liao HX, Gao F, Williams WB, Alam SM, Montefiori DC, et al. Antibody-virus co-evolution in HIV infection: paths for HIV vaccine development. Immunol Rev. 2017;275:145-60.

[82] Wibmer CK, Bhiman JN, Gray ES, Tumba N, Abdool Karim SS, Williamson C, et al. Viral escape from HIV-1 neutralizing antibodies drives increased plasma neutralization breadth through sequential recognition of multiple epitopes and immunotypes. PLoS Pathog. 2013;9:e1003738.

[83] Doria-Rose NA, Landais E. Coevolution of HIV-1 and broadly neutralizing antibodies. Curr Opin HIV AIDS. 2019.

[84] van Gils MJ, Euler Z, Schweighardt B, Wrin T, Schuitemaker H. Prevalence of cross-reactive HIV-1-neutralizing activity in HIV-1-infected patients with rapid or slow disease progression. AIDS. 2009;23:2405-14.

[85] Doria-Rose NA, Klein RM, Manion MM, O'Dell S, Phogat A, Chakrabarti B, et al. Frequency and phenotype of human immunodeficiency virus envelope-specific B cells from patients with broadly cross-neutralizing antibodies. J Virol. 2009;83:188-99.

[86] Hraber P, Seaman MS, Bailer RT, Mascola JR, Montefiori DC, Korber BT. Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection. AIDS. 2014;28:163-9.

[87] Euler Z, van Gils MJ, Bunnik EM, Phung P, Schweighardt B, Wrin T, et al. Cross-reactive neutralizing humoral immunity does not protect from HIV type 1 disease progression. The Journal of infectious diseases. 2010;201:1045-53.

[88] Julg B, Liu PT, Wagh K, Fischer WM, Abbink P, Mercado NB, et al. Protection against a mixed SHIV challenge by a broadly neutralizing antibody cocktail. Sci Transl Med. 2017;9.

[89] Mascola JR, Stiegler G, VanCott TC, Katinger H, Carpenter CB, Hanson CE, et al. Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. Nat Med. 2000;6:207-10.

[90] Julg B, Tartaglia LJ, Keele BF, Wagh K, Pegu A, Sok D, et al. Broadly neutralizing antibodies targeting the HIV-1 envelope V2 apex confer protection against a clade C SHIV challenge. Sci Transl Med. 2017;9.

[91] Shingai M, Donau OK, Plishka RJ, Buckler-White A, Mascola JR, Nabel GJ, et al. Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques. J Exp Med. 2014;211:2061-74.

[92] Saunders KO, Pegu A, Georgiev IS, Zeng M, Joyce MG, Yang ZY, et al. Sustained Delivery of a Broadly Neutralizing Antibody in Nonhuman Primates Confers Long-Term Protection against Simian/Human Immunodeficiency Virus Infection. J Virol. 2015;89:5895-903.

[93] Mascola JR, Lewis MG, Stiegler G, Harris D, VanCott TC, Hayes D, et al. Protection of Macaques against pathogenic simian/human immunodeficiency virus 89.6PD by passive transfer of neutralizing antibodies. J Virol. 1999;73:4009-18.

[94] Baba TW, Liska V, Hofmann-Lehmann R, Vlasak J, Xu W, Ayehunie S, et al. Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection. Nat Med. 2000;6:200-6.

[95] Parren PW, Marx PA, Hessell AJ, Luckay A, Harouse J, Cheng-Mayer C, et al. Antibody protects macaques against vaginal challenge with a pathogenic R5 simian/human immunodeficiency virus at serum levels giving complete neutralization in vitro. J Virol. 2001;75:8340-7.

[96] Hessell AJ, Poignard P, Hunter M, Hangartner L, Tehrani DM, Bleeker WK, et al. Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. Nat Med. 2009;15:951-4.

[97] Pegu A, Yang ZY, Boyington JC, Wu L, Ko SY, Schmidt SD, et al. Neutralizing antibodies to HIV-1 envelope protect more effectively in vivo than those to the CD4 receptor. Sci Transl Med. 2014;6:243ra88.

[98] Hessell AJ, Rakasz EG, Tehrani DM, Huber M, Weisgrau KL, Landucci G, et al. Broadly neutralizing monoclonal antibodies 2F5 and 4E10 directed against the human immunodeficiency virus type 1 gp41 membrane-proximal external region protect against

mucosal challenge by simian-human immunodeficiency virus SHIVBa-L. J Virol. 2010;84:1302-13.

[99] Gautam R, Nishimura Y, Pegu A, Nason MC, Klein F, Gazumyan A, et al. A single injection of anti-HIV-1 antibodies protects against repeated SHIV challenges. Nature. 2016;533:105-9.

[100] Parren PW, Ditzel HJ, Gulizia RJ, Binley JM, Barbas CF, 3rd, Burton DR, et al. Protection against HIV-1 infection in hu-PBL-SCID mice by passive immunization with a neutralizing human monoclonal antibody against the gp120 CD4-binding site. AIDS. 1995;9:F1-6.

[101] Gauduin MC, Parren PW, Weir R, Barbas CF, Burton DR, Koup RA. Passive immunization with a human monoclonal antibody protects hu-PBL-SCID mice against challenge by primary isolates of HIV-1. Nat Med. 1997;3:1389-93.

[102] Pietzsch J, Gruell H, Bournazos S, Donovan BM, Klein F, Diskin R, et al. A mouse model for HIV-1 entry. Proc Natl Acad Sci U S A. 2012;109:15859-64.

[103] Deruaz M, Moldt B, Le KM, Power KA, Vrbanac VD, Tanno S, et al. Protection of Humanized Mice From Repeated Intravaginal HIV Challenge by Passive Immunization: A Model for Studying the Efficacy of Neutralizing Antibodies In Vivo. The Journal of infectious diseases. 2016;214:612-6.

[104] Stoddart CA, Galkina SA, Joshi P, Kosikova G, Long BR, Maidji E, et al. Efficacy of broadly neutralizing monoclonal antibody PG16 in HIV-infected humanized mice. Virology. 2014;462-463:115-25.

[105] Sun M, Li Y, Yuan Z, Lu W, Kang G, Fan W, et al. VRC01 antibody protects against vaginal and rectal transmission of human immunodeficiency virus 1 in hu-BLT mice. Archives of virology. 2016;161:2449-55.

[106] Veselinovic M, Neff CP, Mulder LR, Akkina R. Topical gel formulation of broadly neutralizing anti-HIV-1 monoclonal antibody VRC01 confers protection against HIV-1 vaginal challenge in a humanized mouse model. Virology. 2012;432:505-10.

[107] Mellquist JL, Kasturi L, Spitalnik SL, Shakin-Eshleman SH. The amino acid following an asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. Biochemistry. 1998;37:6833-7.

[108] Vance BA, Wu W, Ribaudo RK, Segal DM, Kearse KP. Multiple dimeric forms of human CD69 result from differential addition of N-glycans to typical (Asn-X-Ser/Thr) and atypical (Asn-X-cys) glycosylation motifs. J Biol Chem. 1997;272:23117-22.

[109] Land A, Zonneveld D, Braakman I. Folding of HIV-1 envelope glycoprotein involves extensive isomerization of disulfide bonds and conformation-dependent leader peptide cleavage. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2003;17:1058-67.

[110] Leonard CK, Spellman MW, Riddle L, Harris RJ, Thomas JN, Gregory TJ. 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. J Biol Chem. 1990;265:10373-82.

[111] Gallina A, Hanley TM, Mandel R, Trahey M, Broder CC, Viglianti GA, et al. Inhibitors of protein-disulfide isomerase prevent cleavage of disulfide bonds in receptor-bound glycoprotein 120 and prevent HIV-1 entry. J Biol Chem. 2002;277:50579-88.

[112] Carbohydrate Active Enzymes Database. Homo sapiens genome. http://www.cazy.org/e358.html.

[113] Hallenberger S, Bosch V, Angliker H, Shaw E, Klenk HD, Garten W. Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. Nature. 1992;360:358-61.

[114] McCune JM, Rabin LB, Feinberg MB, Lieberman M, Kosek JC, Reyes GR, et al. Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. Cell. 1988;53:55-67.

[115] Ringe RP, Sanders RW, Yasmeen A, Kim HJ, Lee JH, Cupo A, et al. Cleavage strongly influences whether soluble HIV-1 envelope glycoprotein trimers adopt a native-like conformation. Proc Natl Acad Sci U S A. 2013;110:18256-61.

[116] Harvey DJ, Merry AH, Royle L, Campbell MP, Dwek RA, Rudd PM. Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. Proteomics. 2009;9:3796-801.

[117] Consortium for Functional Glycomics. Nomenclature. http://www.functionalglycomics.org/static/consortium/Nomenclature.shtml.

[118] Johannes F.G. Vliegenthart LD, Herman van Halbeek. High-Resolution, 1H-Nuclear Magnetic Resonance Spectroscopy as a Tool in the Structural Analysis of Carbohydrates Related to Glycoproteins. In: R. Stuart Tipson DH, editor. Advances in Carbohydrate Chemistry and Biochemistry: Academic Press; 1983. p. 209-374.

[119] Watanabe Y, Bowden TA, Wilson IA, Crispin M. Exploitation of glycosylation in enveloped virus pathobiology. Biochim Biophys Acta Gen Subj. 2019.

[120] Thomas G. Furin at the cutting edge: from protein traffic to embryogenesis and disease. Nat Rev Mol Cell Biol. 2002;3:753-66.

[121] Bass J, Turck C, Rouard M, Steiner DF. Furin-mediated processing in the early secretory pathway: sequential cleavage and degradation of misfolded insulin receptors. Proc Natl Acad Sci U S A. 2000;97:11905-9.

[122] Behrens AJ, Harvey DJ, Milne E, Cupo A, Kumar A, Zitzmann N, et al. Molecular Architecture of the Cleavage-Dependent Mannose Patch on a Soluble HIV-1 Envelope Glycoprotein Trimer. J Virol. 2017;91.

[123] Pritchard LK, Vasiljevic S, Ozorowski G, Seabright GE, Cupo A, Ringe R, et al. Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers. Cell Rep. 2015;11:1604-13.

[124] Checkley MA, Luttge BG, Freed EO. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. J Mol Biol. 2011;410:582-608.

[125] Ward AB, Wilson IA. The HIV-1 envelope glycoprotein structure: nailing down a moving target. Immunol Rev. 2017;275:21-32.

[126] Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature. 1998;393:648-59.

[127] Huang CC, Lam SN, Acharya P, Tang M, Xiang SH, Hussan SS, et al. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. Science. 2007;317:1930-4.

[128] Pancera M, Majeed S, Ban YE, Chen L, Huang CC, Kong L, et al. Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. Proc Natl Acad Sci U S A. 2010;107:1166-71.

[129] Kwon YD, Finzi A, Wu X, Dogo-Isonagie C, Lee LK, Moore LR, et al. Unliganded HIV-1 gp120 core structures assume the CD4-bound conformation with regulation by quaternary interactions and variable loops. Proc Natl Acad Sci U S A. 2012;109:5663-8.

[130] Binley JM, Sanders RW, Clas B, Schuelke N, Master A, Guo Y, et al. A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. J Virol. 2000;74:627-43.

[131] Lee JH, de Val N, Lyumkis D, Ward AB. Model Building and Refinement of a Natively Glycosylated HIV-1 Env Protein by High-Resolution Cryoelectron Microscopy. Structure. 2015;23:1943-51.

[132] Behrens AJ, Vasiljevic S, Pritchard LK, Harvey DJ, Andev RS, Krumm SA, et al. Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein. Cell Rep. 2016;14:2695-706.

[133] Sanders RW, Vesanen M, Schuelke N, Master A, Schiffner L, Kalyanaraman R, et al. Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. J Virol. 2002;76:8875-89.

[134] Sanders RW, Busser E, Moore JP, Lu M, Berkhout B. Evolutionary repair of HIV type 1 gp41 with a kink in the N-terminal helix leads to restoration of the six-helix bundle structure. AIDS research and human retroviruses. 2004;20:742-9.

[135] Julien JP, Cupo A, Sok D, Stanfield RL, Lyumkis D, Deller MC, et al. Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science. 2013;342:1477-83.

[136] Lyumkis D, Julien JP, de Val N, Cupo A, Potter CS, Klasse PJ, et al. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. Science. 2013;342:1484-90.

[137] Lee JH, Ozorowski G, Ward AB. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. Science. 2016;351:1043-8.

[138] Crooks ET, Osawa K, Tong T, Grimley SL, Dai YD, Whalen RG, et al. Effects of partially dismantling the CD4 binding site glycan fence of HIV-1 Envelope glycoprotein trimers on neutralizing antibody induction. Virology. 2017;505:193-209.

[139] Wu L, Gerard NP, Wyatt R, Choe H, Parolin C, Ruffing N, et al. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. Nature. 1996;384:179-83.

[140] Hwang SS, Boyle TJ, Lyerly HK, Cullen BR. Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. Science. 1991;253:71-4.

[141] Doms RW, Moore JP. HIV-1 membrane fusion: targets of opportunity. J Cell Biol. 2000;151:F9-14.

[142] U.S. Food & Drug Administration. Antiretroviral drugs used in the treatment of HIV infection. 2018. https://www.fda.gov/patients/hiv-treatment/antiretroviral-drugs-used-treatment-hiv-infection

[143] Iyidogan P, Anderson KS. Current perspectives on HIV-1 antiretroviral drug resistance. Viruses. 2014;6:4095-139.

[144] Shafer RW, Vuitton DA. Highly active antiretroviral therapy (HAART) for the treatment of infection with human immunodeficiency virus type 1. Biomed Pharmacother. 1999;53:73-86.

[145] Antiretroviral Therapy Cohort C. Survival of HIV-positive patients starting antiretroviral therapy between 1996 and 2013: a collaborative analysis of cohort studies. Lancet HIV. 2017;4:e349-e56.

[146] Johnson LF, Mossong J, Dorrington RE, Schomaker M, Hoffmann CJ, Keiser O, et al. Life expectancies of South African adults starting antiretroviral treatment: collaborative analysis of cohort studies. PLoS Med. 2013;10:e1001418.

[147] Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, et al. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. The New England journal of medicine. 2012;367:399-410.

[148] Harrigan PR, Whaley M, Montaner JS. Rate of HIV-1 RNA rebound upon stopping antiretroviral therapy. AIDS. 1999;13:F59-62.

[149] Klein F, Halper-Stromberg A, Horwitz JA, Gruell H, Scheid JF, Bournazos S, et al. HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. Nature. 2012;492:118-22.

[150] Horwitz JA, Halper-Stromberg A, Mouquet H, Gitlin AD, Tretiakova A, Eisenreich TR, et al. HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. Proc Natl Acad Sci U S A. 2013;110:16538-43.

[151] Shingai M, Nishimura Y, Klein F, Mouquet H, Donau OK, Plishka R, et al. Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. Nature. 2013;503:277-80.

[152] Barouch DH, Whitney JB, Moldt B, Klein F, Oliveira TY, Liu J, et al. Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys. Nature. 2013;503:224-8.

[153] Caskey M, Klein F, Nussenzweig MC. Broadly neutralizing anti-HIV-1 monoclonal antibodies in the clinic. Nat Med. 2019.

[154] Caskey M, Klein F, Lorenzi JC, Seaman MS, West AP, Buckley N, et al. Corrigendum: Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. Nature. 2016;535:580.

[155] Lynch RM, Boritz E, Coates EE, DeZure A, Madden P, Costner P, et al. Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. Sci Transl Med. 2015;7:319ra206.

[156] Ledgerwood JE, Coates EE, Yamshchikov G, Saunders JG, Holman L, Enama ME, et al. Safety, pharmacokinetics and neutralization of the broadly neutralizing HIV-1 human monoclonal antibody VRC01 in healthy adults. Clin Exp Immunol. 2015;182:289-301.

[157] Caskey M, Schoofs T, Gruell H, Settler A, Karagounis T, Kreider EF, et al. Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. Nat Med. 2017;23:185-91.

[158] Keizer RJ, Huitema AD, Schellens JH, Beijnen JH. Clinical pharmacokinetics of therapeutic monoclonal antibodies. Clin Pharmacokinet. 2010;49:493-507.

[159] Schoofs T, Klein F, Braunschweig M, Kreider EF, Feldmann A, Nogueira L, et al. HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1. Science. 2016;352:997-1001.

[160] Lu CL, Murakowski DK, Bournazos S, Schoofs T, Sarkar D, Halper-Stromberg A, et al. Enhanced clearance of HIV-1-infected cells by broadly neutralizing antibodies against HIV-1 in vivo. Science. 2016;352:1001-4.

[161] Kim Y, Anderson JL, Lewin SR. Getting the "Kill" into "Shock and Kill": Strategies to Eliminate Latent HIV. Cell host & microbe. 2018;23:14-26.

[162] Deeks SG. HIV: Shock and kill. Nature. 2012;487:439-40.

[163] Xu W, Li H, Wang Q, Hua C, Zhang H, Li W, et al. Advancements in Developing Strategies for Sterilizing and Functional HIV Cures. Biomed Res Int. 2017;2017:6096134.

[164] Abner E, Jordan A. HIV "shock and kill" therapy: In need of revision. Antiviral Res. 2019;166:19-34.

[165] Hutter G, Nowak D, Mossner M, Ganepola S, Mussig A, Allers K, et al. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. The New England journal of medicine. 2009;360:692-8.

[166] Gupta RK, Abdul-Jawad S, McCoy LE, Mok HP, Peppa D, Salgado M, et al. HIV-1 remission following CCR5Delta32/Delta32 haematopoietic stem-cell transplantation. Nature. 2019.

[167] Kordelas L, Verheyen J, Beelen DW, Horn PA, Heinold A, Kaiser R, et al. Shift of HIV tropism in stem-cell transplantation with CCR5 Delta32 mutation. The New England journal of medicine. 2014;371:880-2.

[168] Symons J, Vandekerckhove L, Hutter G, Wensing AM, van Ham PM, Deeks SG, et al. Dependence on the CCR5 coreceptor for viral replication explains the lack of rebound of CXCR4-predicted HIV variants in the Berlin patient. Clin Infect Dis. 2014;59:596-600.

[169] Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol. 2013;31:230-2.

[170] Wang W, Ye C, Liu J, Zhang D, Kimata JT, Zhou P. CCR5 gene disruption via lentiviral vectors expressing Cas9 and single guided RNA renders cells resistant to HIV-1 infection. PLoS One. 2014;9:e115987.

[171] Xu L, Yang H, Gao Y, Chen Z, Xie L, Liu Y, et al. CRISPR/Cas9-Mediated CCR5 Ablation in Human Hematopoietic Stem/Progenitor Cells Confers HIV-1 Resistance In Vivo. Mol Ther. 2017;25:1782-9.

[172] Bella R, Kaminski R, Mancuso P, Young WB, Chen C, Sariyer R, et al. Removal of HIV DNA by CRISPR from Patient Blood Engrafts in Humanized Mice. Mol Ther Nucleic Acids. 2018;12:275-82.

[173] Deal CE, Balazs AB. Vectored antibody gene delivery for the prevention or treatment of HIV infection. Curr Opin HIV AIDS. 2015;10:190-7.

[174] Johnson PR, Schnepp BC, Zhang J, Connell MJ, Greene SM, Yuste E, et al. Vector-mediated gene transfer engenders long-lived neutralizing activity and protection against SIV infection in monkeys. Nat Med. 2009;15:901-6.

[175] Saayman S, Ali SA, Morris KV, Weinberg MS. The therapeutic application of CRISPR/Cas9 technologies for HIV. Expert opinion on biological therapy. 2015;15:819-30.

[176] International AIDS Vaccine Initiative. Database of Preventative HIV Vaccine Candidates. https://www.iavi.org/trials-database/trials.

[177] Pitisuttithum P, Gilbert P, Gurwith M, Heyward W, Martin M, van Griensven F, 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. The Journal of infectious diseases. 2006;194:1661-71.

[178] Flynn NM, Forthal DN, Harro CD, Judson FN, Mayer KH, Para MF, et al. Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. The Journal of infectious diseases. 2005;191:654-65.

[179] Buchbinder SP, Mehrotra DV, Duerr A, Fitzgerald DW, Mogg R, Li D, 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. Lancet. 2008;372:1881-93.

[180] Gray GE, Allen M, Moodie Z, Churchyard G, Bekker LG, Nchabeleng M, 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. The Lancet Infectious diseases. 2011;11:507-15.

[181] Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, et al. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. The New England journal of medicine. 2009;361:2209-20.

[182] Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. The New England journal of medicine. 2012;366:1275-86.

[183] Hammer SM, Sobieszczyk ME, Janes H, Karuna ST, Mulligan MJ, Grove D, et al. Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. The New England journal of medicine. 2013;369:2083-92.

[184] Pham HT, Mesplede T. The latest evidence for possible HIV-1 curative strategies. Drugs Context. 2018;7:212522.

[185] Plotkin SA. Immunologic correlates of protection induced by vaccination. The Pediatric infectious disease journal. 2001;20:63-75.

[186] Bachmann MF, Rohrer UH, Kundig TM, Burki K, Hengartner H, Zinkernagel RM. The influence of antigen organization on B cell responsiveness. Science. 1993;262:1448-51.

[187] McCoy LE, Quigley AF, Strokappe NM, Bulmer-Thomas B, Seaman MS, Mortier D, et al. Potent and broad neutralization of HIV-1 by a llama antibody elicited by immunization. J Exp Med. 2012;209:1091-103.

[188] Sok D, Le KM, Vadnais M, Saye-Francisco KL, Jardine JG, Torres JL, et al. Rapid elicitation of broadly neutralizing antibodies to HIV by immunization in cows. Nature. 2017;548:108-11.

[189] Burton DR, Hangartner L. Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design. Annu Rev Immunol. 2016;34:635-59.

[190] McGuire AT, Glenn JA, Lippy A, Stamatatos L. Diverse recombinant HIV-1 Envs fail to activate B cells expressing the germline B cell receptors of the broadly neutralizing anti-HIV-1 antibodies PG9 and 447-52D. J Virol. 2014;88:2645-57.

[191] Hoot S, McGuire AT, Cohen KW, Strong RK, Hangartner L, Klein F, et al. Recombinant HIV envelope proteins fail to engage germline versions of anti-CD4bs bNAbs. PLoS Pathog. 2013;9:e1003106.

[192] Xiao X, Chen W, Feng Y, Zhu Z, Prabakaran P, Wang Y, et al. Germline-like predecessors of broadly neutralizing antibodies lack measurable binding to HIV-1 envelope glycoproteins: implications for evasion of immune responses and design of vaccine immunogens. Biochem Biophys Res Commun. 2009;390:404-9.

[193] Doores KJ, Huber M, Le KM, Wang SK, Doyle-Cooper C, Cooper A, et al. 2G12-expressing B cell lines may aid in HIV carbohydrate vaccine design strategies. J Virol. 2013;87:2234-41.

[194] Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, Nussenzweig MC. Predominant autoantibody production by early human B cell precursors. Science. 2003;301:1374-7.

[195] Rudd PM, Dwek RA. Glycosylation: heterogeneity and the 3D structure of proteins. Critical reviews in biochemistry and molecular biology. 1997;32:1-100.

[196] Cohen M. Notable Aspects of Glycan-Protein Interactions. Biomolecules. 2015;5:2056-72.

[197] Wyatt R, Kwong PD, Desjardins E, Sweet RW, Robinson J, Hendrickson WA, et al. The antigenic structure of the HIV gp120 envelope glycoprotein. Nature. 1998;393:705-11.

[198] Moore PL, Gray ES, Wibmer CK, Bhiman JN, Nonyane M, Sheward DJ, et al. Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. Nat Med. 2012;18:1688-92.

[199] Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, et al. Antibody neutralization and escape by HIV-1. Nature. 2003;422:307-12.

[200] Dacheux L, Moreau A, Ataman-Onal Y, Biron F, Verrier B, Barin F. Evolutionary dynamics of the glycan shield of the human immunodeficiency virus envelope during natural infection and implications for exposure of the 2G12 epitope. J Virol. 2004;78:12625-37.

[201] Chohan B, Lang D, Sagar M, Korber B, Lavreys L, Richardson B, et al. Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1-V2 loop sequences occurs during transmission of certain genetic subtypes and may impact viral RNA levels. J Virol. 2005;79:6528-31.

[202] Derdeyn CA, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham SA, et al. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science. 2004;303:2019-22.

[203] Liu Y, Curlin ME, Diem K, Zhao H, Ghosh AK, Zhu H, et al. Env length and N-linked glycosylation following transmission of human immunodeficiency virus Type 1 subtype B viruses. Virology. 2008;374:229-33.

[204] Wagh K, Kreider EF, Li Y, Barbian HJ, Learn GH, Giorgi E, et al. Completeness of HIV-1 Envelope Glycan Shield at Transmission Determines Neutralization Breadth. Cell Rep. 2018;25:893-908 e7.

[205] McCaffrey RA, Saunders C, Hensel M, Stamatatos L. N-linked glycosylation of the V3 loop and the immunologically silent face of gp120 protects human immunodeficiency virus type 1 SF162 from neutralization by anti-gp120 and anti-gp41 antibodies. J Virol. 2004;78:3279-95.

[206] Wang W, Zirkle B, Nie J, Ma J, Gao K, Chen XS, et al. N463 Glycosylation Site on V5 Loop of a Mutant gp120 Regulates the Sensitivity of HIV-1 to Neutralizing Monoclonal Antibodies VRC01/03. Journal of acquired immune deficiency syndromes. 2015;69:270-7.

[207] Koch M, Pancera M, Kwong PD, Kolchinsky P, Grundner C, Wang L, et al. Structure-based, targeted deglycosylation of HIV-1 gp120 and effects on neutralization sensitivity and antibody recognition. Virology. 2003;313:387-400.

[208] Li Y, Cleveland B, Klots I, Travis B, Richardson BA, Anderson D, et al. Removal of a single N-linked glycan in human immunodeficiency virus type 1 gp120 results in an enhanced ability to induce neutralizing antibody responses. J Virol. 2008;82:638-51.

[209] Back NK, Smit L, De Jong JJ, Keulen W, Schutten M, Goudsmit J, et al. An N-glycan within the human immunodeficiency virus type 1 gp120 V3 loop affects virus neutralization. Virology. 1994;199:431-8.

[210] Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature. 2014;514:455-61.

[211] Liang Y, Guttman M, Williams JA, Verkerke H, Alvarado D, Hu SL, et al. Changes in Structure and Antigenicity of HIV-1 Env Trimers Resulting from Removal of a Conserved CD4 Binding Site-Proximal Glycan. J Virol. 2016;90:9224-36.

[212] Crispin M, Bowden TA. Antibodies expose multiple weaknesses in the glycan shield of HIV. Nat Struct Mol Biol. 2013;20:771-2.

[213] Gilbert PB, Peterson ML, Follmann D, Hudgens MG, Francis DP, Gurwith M, 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. The Journal of infectious diseases. 2005;191:666-77.

[214] VanCott TC, Bethke FR, Burke DS, Redfield RR, Birx DL. Lack of induction of antibodies specific for conserved, discontinuous epitopes of HIV-1 envelope glycoprotein by candidate AIDS vaccines. J Immunol. 1995;155:4100-10.

[215] Sanders RW, Moore JP. Native-like Env trimers as a platform for HIV-1 vaccine design. Immunol Rev. 2017;275:161-82.

[216] Yang X, Florin L, Farzan M, Kolchinsky P, Kwong PD, Sodroski J, et al. Modifications that stabilize human immunodeficiency virus envelope glycoprotein trimers in solution. J Virol. 2000;74:4746-54.

[217] Frey G, Peng H, Rits-Volloch S, Morelli M, Cheng Y, Chen B. A fusion-intermediate state of HIV-1 gp41 targeted by broadly neutralizing antibodies. Proc Natl Acad Sci U S A. 2008;105:3739-44.

[218] Ringe RP, Yasmeen A, Ozorowski G, Go EP, Pritchard LK, Guttman M, et al. Influences on the Design and Purification of Soluble, Recombinant Native-Like HIV-1 Envelope Glycoprotein Trimers. J Virol. 2015;89:12189-210.

[219] Go EP, Cupo A, Ringe R, Pugach P, Moore JP, Desaire H. Native Conformation and Canonical Disulfide Bond Formation Are Interlinked Properties of HIV-1 Env Glycoproteins. J Virol. 2015;90:2884-94.

[220] Yasmeen A, Ringe R, Derking R, Cupo A, Julien JP, Burton DR, et al. Differential binding of neutralizing and non-neutralizing antibodies to native-like soluble HIV-1 Env trimers, uncleaved Env proteins, and monomeric subunits. Retrovirology. 2014;11:41.

[221] Lewis DJ, Fraser CA, Mahmoud AN, Wiggins RC, Woodrow M, Cope A, et al. Phase I randomised clinical trial of an HIV-1(CN54), clade C, trimeric envelope vaccine candidate delivered vaginally. PLoS One. 2011;6:e25165.

[222] Spearman P, Lally MA, Elizaga M, Montefiori D, Tomaras GD, McElrath MJ, et al. A trimeric, V2-deleted HIV-1 envelope glycoprotein vaccine elicits potent neutralizing antibodies but limited breadth of neutralization in human volunteers. The Journal of infectious diseases. 2011;203:1165-73.

[223] Cosgrove CA, Lacey CJ, Cope AV, Bartolf A, Morris G, Yan C, et al. Comparative Immunogenicity of HIV-1 gp140 Vaccine Delivered by Parenteral, and Mucosal Routes in Female Volunteers; MUCOVAC2, A Randomized Two Centre Study. PLoS One. 2016;11:e0152038.

[224] Klasse PJ, Depetris RS, Pejchal R, Julien JP, Khayat R, Lee JH, et al. Influences on trimerization and aggregation of soluble, cleaved HIV-1 SOSIP envelope glycoprotein. J Virol. 2013;87:9873-85.

[225] Torrents de la Pena A, Julien JP, de Taeye SW, Garces F, Guttman M, Ozorowski G, et al. Improving the Immunogenicity of Native-like HIV-1 Envelope Trimers by Hyperstabilization. Cell Rep. 2017;20:1805-17.

[226] Kwon YD, Pancera M, Acharya P, Georgiev IS, Crooks ET, Gorman J, et al. Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. Nat Struct Mol Biol. 2015;22:522-31.

[227] Guenaga J, de Val N, Tran K, Feng Y, Satchwell K, Ward AB, et al. Well-ordered trimeric HIV-1 subtype B and C soluble spike mimetics generated by negative selection display native-like properties. PLoS Pathog. 2015;11:e1004570.

[228] de Taeye SW, Ozorowski G, Torrents de la Pena A, Guttman M, Julien JP, van den Kerkhof TL, et al. Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes. Cell. 2015;163:1702-15.

[229] Julien JP, Lee JH, Ozorowski G, Hua Y, Torrents de la Pena A, de Taeye SW, et al. Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens. Proc Natl Acad Sci U S A. 2015;112:11947-52.

[230] Pugach P, Ozorowski G, Cupo A, Ringe R, Yasmeen A, de Val N, et al. A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. J Virol. 2015;89:3380-95.

[231] Joyce MG, Georgiev IS, Yang Y, Druz A, Geng H, Chuang GY, et al. Soluble Prefusion Closed DS-SOSIP.664-Env Trimers of Diverse HIV-1 Strains. Cell Rep. 2017;21:2992-3002.

[232] Rutten L, Lai YT, Blokland S, Truan D, Bisschop IJM, Strokappe NM, et al. A Universal Approach to Optimize the Folding and Stability of Prefusion-Closed HIV-1 Envelope Trimers. Cell Rep. 2018;23:584-95.

[233] Ahmed S, Shrivastava T, Kumar N, Ozorowski G, Ward AB, Chakrabarti BK. Stabilization of a soluble, native-like trimeric form of an efficiently cleaved Indian HIV-1 clade C envelope glycoprotein. J Biol Chem. 2017;292:8236-43.

[234] Sharma SK, de Val N, Bale S, Guenaga J, Tran K, Feng Y, et al. Cleavage-independent HIV-1 Env trimers engineered as soluble native spike mimetics for vaccine design. Cell Rep. 2015;11:539-50.

[235] Georgiev IS, Joyce MG, Yang Y, Sastry M, Zhang B, Baxa U, et al. Single-Chain Soluble BG505.SOSIP gp140 Trimers as Structural and Antigenic Mimics of Mature Closed HIV-1 Env. J Virol. 2015;89:5318-29.

[236] Kong L, He L, de Val N, Vora N, Morris CD, Azadnia P, et al. Uncleaved prefusion-optimized gp140 trimers derived from analysis of HIV-1 envelope metastability. Nat Commun. 2016;7:12040.

[237] Sanders RW, van Gils MJ, Derking R, Sok D, Ketas TJ, Burger JA, et al. HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers. Science. 2015;349:aac4223.

[238] Feng Y, Tran K, Bale S, Kumar S, Guenaga J, Wilson R, et al. Thermostability of Well-Ordered HIV Spikes Correlates with the Elicitation of Autologous Tier 2 Neutralizing Antibodies. PLoS Pathog. 2016;12:e1005767.

[239] Torrents de la Pena A, de Taeye SW, Sliepen K, LaBranche CC, Burger JA, Schermer EE, et al. Immunogenicity in Rabbits of HIV-1 SOSIP Trimers from Clades A, B, and C, Given Individually, Sequentially, or in Combination. J Virol. 2018;92.

[240] Klasse PJ, LaBranche CC, Ketas TJ, Ozorowski G, Cupo A, Pugach P, et al. Sequential and Simultaneous Immunization of Rabbits with HIV-1 Envelope Glycoprotein SOSIP.664 Trimers from Clades A, B and C. PLoS Pathog. 2016;12:e1005864.

[241] Cheng C, Pancera M, Bossert A, Schmidt SD, Chen RE, Chen X, et al. Immunogenicity of a Prefusion HIV-1 Envelope Trimer in Complex with a Quaternary-Structure-Specific Antibody. J Virol. 2015;90:2740-55.

[242] Aldon Y, McKay PF, Allen J, Ozorowski G, Felfodine Levai R, Tolazzi M, et al. Rational Design of DNA-Expressed Stabilized Native-Like HIV-1 Envelope Trimers. Cell Rep. 2018;24:3324-38 e5.

[243] NCT03699241. A Randomized, Double-blinded, Placebo-controlled, Dose-escalation Phase 1 Clinical Trial to Evaluate the Safety and Immunogenicity of Recombinant HIV Envelope Protein BG505 SOSIP.664 gp140 Vaccine, Adjuvanted, in Healthy, HIV-1 Uninfected Adults.

[244] Crispin M, Ward AB, Wilson IA. Structure and Immune Recognition of the HIV Glycan Shield. Annu Rev Biophys. 2018.

[245] Wibmer CK, Moore PL, Morris L. HIV broadly neutralizing antibody targets. Curr Opin HIV AIDS. 2015;10:135-43.

[246] Zhou T, Zheng A, Baxa U, Chuang GY, Georgiev IS, Kong R, et al. A Neutralizing Antibody Recognizing Primarily N-Linked Glycan Targets the Silent Face of the HIV Envelope. Immunity. 2018;48:500-13 e6.

[247] Calarese DA, Lee HK, Huang CY, Best MD, Astronomo RD, Stanfield RL, et al. Dissection of the carbohydrate specificity of the broadly neutralizing anti-HIV-1 antibody 2G12. Proc Natl Acad Sci U S A. 2005;102:13372-7.

[248] Trkola A, Purtscher M, Muster T, Ballaun C, Buchacher A, Sullivan N, et al. Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. J Virol. 1996;70:1100-8.

[249] Murin CD, Julien JP, Sok D, Stanfield RL, Khayat R, Cupo A, et al. Structure of 2G12 Fab2 in complex with soluble and fully glycosylated HIV-1 Env by negative-stain single-particle electron microscopy. J Virol. 2014;88:10177-88.

[250] Calarese DA, Scanlan CN, Zwick MB, Deechongkit S, Mimura Y, Kunert R, et al. Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. Science. 2003;300:2065-71.

[251] Sanders RW, Venturi M, Schiffner L, Kalyanaraman R, Katinger H, Lloyd KO, et al. The mannose-dependent epitope for neutralizing antibody 2G12 on human immunodeficiency virus type 1 glycoprotein gp120. J Virol. 2002;76:7293-305.

[252] Scanlan CN, Pantophlet R, Wormald MR, Saphire EO, Calarese D, Stanfield R, et al. The carbohydrate epitope of the neutralizing anti-HIV-1 antibody 2G12. Adv Exp Med Biol. 2003;535:205-18.

[253] Scanlan CN, Pantophlet R, Wormald MR, Ollmann Saphire E, Stanfield R, Wilson IA, et al. The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of alpha1-->2 mannose residues on the outer face of gp120. J Virol. 2002;76:7306-21.

[254] Zhou T, Georgiev I, Wu X, Yang ZY, Dai K, Finzi A, et al. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science. 2010;329:811-7.

[255] Doria-Rose NA, Schramm CA, Gorman J, Moore PL, Bhiman JN, DeKosky BJ, et al. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. Nature. 2014;509:55-62.

[256] Lee JH, Andrabi R, Su CY, Yasmeen A, Julien JP, Kong L, et al. A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic beta-Hairpin Structure. Immunity. 2017;46:690-702.

[257] Pejchal R, Walker LM, Stanfield RL, Phogat SK, Koff WC, Poignard P, et al. Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. Proc Natl Acad Sci U S A. 2010;107:11483-8.

[258] Pejchal R, Doores KJ, Walker LM, Khayat R, Huang PS, Wang SK, et al. A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. Science. 2011;334:1097-103.

[259] Kong L, Lee JH, Doores KJ, Murin CD, Julien JP, McBride R, et al. Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. Nat Struct Mol Biol. 2013;20:796-803.

[260] Julien JP, Sok D, Khayat R, Lee JH, Doores KJ, Walker LM, et al. Broadly neutralizing antibody PGT121 allosterically modulates CD4 binding via recognition of the HIV-1 gp120 V3 base and multiple surrounding glycans. PLoS Pathog. 2013;9:e1003342.

[261] Irimia A, Sarkar A, Stanfield RL, Wilson IA. Crystallographic Identification of Lipid as an Integral Component of the Epitope of HIV Broadly Neutralizing Antibody 4E10. Immunity. 2016;44:21-31.

[262] Zemlin M, Klinger M, Link J, Zemlin C, Bauer K, Engler JA, et al. Expressed murine and human CDR-H3 intervals of equal length exhibit distinct repertoires that differ in their amino acid composition and predicted range of structures. J Mol Biol. 2003;334:733-49.

[263] Klein F, Diskin R, Scheid JF, Gaebler C, Mouquet H, Georgiev IS, et al. Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. Cell. 2013;153:126-38.

[264] Sok D, Laserson U, Laserson J, Liu Y, Vigneault F, Julien JP, et al. The effects of somatic hypermutation on neutralization and binding in the PGT121 family of broadly neutralizing HIV antibodies. PLoS Pathog. 2013;9:e1003754.

[265] Georgiev IS, Rudicell RS, Saunders KO, Shi W, Kirys T, McKee K, et al. Antibodies VRC01 and 10E8 neutralize HIV-1 with high breadth and potency even with Ig-framework regions substantially reverted to germline. J Immunol. 2014;192:1100-6.

[266] Haynes BF, Fleming J, St Clair EW, Katinger H, Stiegler G, Kunert R, et al. Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. Science. 2005;308:1906-8.

[267] Liu M, Yang G, Wiehe K, Nicely NI, Vandergrift NA, Rountree W, et al. Polyreactivity and autoreactivity among HIV-1 antibodies. J Virol. 2015;89:784-98.

[268] Dunlop DC, Ulrich A, Appelmelk BJ, Burton DR, Dwek RA, Zitzmann N, et al. Antigenic mimicry of the HIV envelope by AIDS-associated pathogens. AIDS. 2008;22:2214-7.

[269] Haynes BF, Moody MA, Verkoczy L, Kelsoe G, Alam SM. Antibody polyspecificity and neutralization of HIV-1: a hypothesis. Hum Antibodies. 2005;14:59-67.

[270] Komagamine T, Yuki N. Ganglioside mimicry as a cause of Guillain-Barre syndrome. CNS & neurological disorders drug targets. 2006;5:391-400.

[271] Shivatare VS, Shivatare SS, Lee CD, Liang CH, Liao KS, Cheng YY, et al. Unprecedented Role of Hybrid N-Glycans as Ligands for HIV-1 Broadly Neutralizing Antibodies. J Am Chem Soc. 2018;140:5202-10.

[272] Amin MN, McLellan JS, Huang W, Orwenyo J, Burton DR, Koff WC, et al. Synthetic glycopeptides reveal the glycan specificity of HIV-neutralizing antibodies. Nat Chem Biol. 2013;9:521-6.

[273] Pancera M, Shahzad-Ul-Hussan S, Doria-Rose NA, McLellan JS, Bailer RT, Dai K, et al. Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. Nat Struct Mol Biol. 2013;20:804-13.

[274] Cao L, Diedrich JK, Kulp DW, Pauthner M, He L, Park SR, et al. Global site-specific N-glycosylation analysis of HIV envelope glycoprotein. Nat Commun. 2017;8:14954.

[275] Cao L, Pauthner M, Andrabi R, Rantalainen K, Berndsen Z, Diedrich JK, et al. Differential processing of HIV envelope glycans on the virus and soluble recombinant trimer. Nat Commun. 2018;9:3693.

[276] Struwe WB, Chertova E, Allen JD, Seabright GE, Watanabe Y, Harvey DJ, et al. Site-Specific Glycosylation of Virion-Derived HIV-1 Env Is Mimicked by a Soluble Trimeric Immunogen. Cell Rep. 2018;24:1958-66 e5.

[277] Sok D, Doores KJ, Briney B, Le KM, Saye-Francisco KL, Ramos A, et al. Promiscuous glycan site recognition by antibodies to the high-mannose patch of gp120 broadens neutralization of HIV. Sci Transl Med. 2014;6:236ra63.

[278] Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien JP, et al. Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature. 2011;477:466-70.

[279] Pritchard LK, Spencer DI, Royle L, Vasiljevic S, Krumm SA, Doores KJ, et al. Glycan Microheterogeneity at the PGT135 Antibody Recognition Site on HIV-1 gp120 Reveals a Molecular Mechanism for Neutralization Resistance. J Virol. 2015;89:6952-9.

[280] Stewart-Jones GB, Soto C, Lemmin T, Chuang GY, Druz A, Kong R, et al. Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. Cell. 2016;165:813-26.

[281] Mouquet H, Scharf L, Euler Z, Liu Y, Eden C, Scheid JF, et al. Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. Proc Natl Acad Sci U S A. 2012;109:E3268-77.

[282] Barnes CO, Gristick HB, Freund NT, Escolano A, Lyubimov AY, Hartweger H, et al. Structural characterization of a highly-potent V3-glycan broadly neutralizing antibody bound to natively-glycosylated HIV-1 envelope. Nat Commun. 2018;9:1251.

[283] Gristick HB, von Boehmer L, West AP, Jr., Schamber M, Gazumyan A, Golijanin J, et al. Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. Nat Struct Mol Biol. 2016;23:906-15.

[284] Bonsignori M, Hwang KK, Chen X, Tsao CY, Morris L, Gray E, 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. J Virol. 2011;85:9998-10009.

[285] Sok D, van Gils MJ, Pauthner M, Julien JP, Saye-Francisco KL, Hsueh J, et al. Recombinant HIV envelope trimer selects for quaternary-dependent antibodies targeting the trimer apex. Proc Natl Acad Sci U S A. 2014;111:17624-9.

[286] Cale EM, Gorman J, Radakovich NA, Crooks ET, Osawa K, Tong T, et al. Virus-like Particles Identify an HIV V1V2 Apex-Binding Neutralizing Antibody that Lacks a Protruding Loop. Immunity. 2017;46:777-91 e10.

[287] Walker LM, Phogat SK, Chan-Hui PY, Wagner D, Phung P, Goss JL, et al. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science. 2009;326:285-9.

[288] McLellan JS, Pancera M, Carrico C, Gorman J, Julien JP, Khayat R, et al. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature. 2011;480:336-43.

[289] Doores KJ, Burton DR. Variable loop glycan dependency of the broad and potent HIV-1-neutralizing antibodies PG9 and PG16. J Virol. 2010;84:10510-21.

[290] Andrabi R, Su CY, Liang CH, Shivatare SS, Briney B, Voss JE, et al. Glycans Function as Anchors for Antibodies and Help Drive HIV Broadly Neutralizing Antibody Development. Immunity. 2017;47:1004.

[291] Landais E, Murrell B, Briney B, Murrell S, Rantalainen K, Berndsen ZT, et al. HIV Envelope Glycoform Heterogeneity and Localized Diversity Govern the Initiation and Maturation of a V2 Apex Broadly Neutralizing Antibody Lineage. Immunity. 2017;47:990-1003 e9.

[292] Falkowska E, Ramos A, Feng Y, Zhou T, Moquin S, Walker LM, et al. PGV04, an HIV-1 gp120 CD4 binding site antibody, is broad and potent in neutralization but does not induce conformational changes characteristic of CD4. J Virol. 2012;86:4394-403.

[293] Balla-Jhagjhoorsingh SS, Corti D, Heyndrickx L, Willems E, Vereecken K, Davis D, et al. The N276 glycosylation site is required for HIV-1 neutralization by the CD4 binding site specific HJ16 monoclonal antibody. PLoS One. 2013;8:e68863.

[294] Wibmer CK, Gorman J, Anthony CS, Mkhize NN, Druz A, York T, et al. Structure of an N276-Dependent HIV-1 Neutralizing Antibody Targeting a Rare V5 Glycan Hole Adjacent to the CD4 Binding Site. J Virol. 2016;90:10220-35.

[295] Zhou T, Lynch RM, Chen L, Acharya P, Wu X, Doria-Rose NA, et al. Structural Repertoire of HIV-1-Neutralizing Antibodies Targeting the CD4 Supersite in 14 Donors. Cell. 2015;161:1280-92.

[296] Freund NT, Horwitz JA, Nogueira L, Sievers SA, Scharf L, Scheid JF, et al. A New Glycan-Dependent CD4-Binding Site Neutralizing Antibody Exerts Pressure on HIV-1 In Vivo. PLoS Pathog. 2015;11:e1005238.

[297] Scharf L, Scheid JF, Lee JH, West AP, Jr., Chen C, Gao H, et al. Antibody 8ANC195 reveals a site of broad vulnerability on the HIV-1 envelope spike. Cell Rep. 2014;7:785-95.

[298] Kong L, Torrents de la Pena A, Deller MC, Garces F, Sliepen K, Hua Y, et al. Complete epitopes for vaccine design derived from a crystal structure of the broadly neutralizing antibodies PGT128 and 8ANC195 in complex with an HIV-1 Env trimer. Acta Crystallogr D Biol Crystallogr. 2015;71:2099-108.

[299] Falkowska E, Le KM, Ramos A, Doores KJ, Lee JH, Blattner C, et al. Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. Immunity. 2014;40:657-68.

[300] Blattner C, Lee JH, Sliepen K, Derking R, Falkowska E, de la Pena AT, et al. Structural delineation of a quaternary, cleavage-dependent epitope at the gp41-gp120 interface on intact HIV-1 Env trimers. Immunity. 2014;40:669-80.

[301] Andrabi R, Voss JE, Liang CH, Briney B, McCoy LE, Wu CY, et al. Identification of Common Features in Prototype Broadly Neutralizing Antibodies to HIV Envelope V2 Apex to Facilitate Vaccine Design. Immunity. 2015;43:959-73.

[302] Kong R, Xu K, Zhou T, Acharya P, Lemmin T, Liu K, et al. Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. Science. 2016;352:828-33.

[303] van Gils MJ, van den Kerkhof TL, Ozorowski G, Cottrell CA, Sok D, Pauthner M, et al. An HIV-1 antibody from an elite neutralizer implicates the fusion peptide as a site of vulnerability. Nat Microbiol. 2016;2:16199.

[304] Huang J, Kang BH, Pancera M, Lee JH, Tong T, Feng Y, et al. Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. Nature. 2014;515:138-42.

[305] Kong L, Sheppard NC, Stewart-Jones GB, Robson CL, Chen H, Xu X, et al. Expression-system-dependent modulation of HIV-1 envelope glycoprotein antigenicity and immunogenicity. J Mol Biol. 2010;403:131-47.

[306] Mizuochi T, Spellman MW, Larkin M, Solomon J, Basa LJ, Feizi T. Carbohydrate structures of the human-immunodeficiency-virus (HIV) recombinant envelope glycoprotein gp120 produced in Chinese-hamster ovary cells. Biochem J. 1988;254:599-603.

[307] Zhu X, Borchers C, Bienstock RJ, Tomer KB. Mass spectrometric characterization of the glycosylation pattern of HIV-gp120 expressed in CHO cells. Biochemistry. 2000;39:11194-204.

[308] Cutalo JM, Deterding LJ, Tomer KB. Characterization of glycopeptides from HIV-I(SF2) gp120 by liquid chromatography mass spectrometry. J Am Soc Mass Spectrom. 2004;15:1545-55.

[309] Go EP, Liao HX, Alam SM, Hua D, Haynes BF, Desaire H. Characterization of host-cell line specific glycosylation profiles of early transmitted/founder HIV-1 gp120 envelope proteins. J Proteome Res. 2013;12:1223-34.

[310] Pritchard LK, Spencer DI, Royle L, Bonomelli C, Seabright GE, Behrens AJ, et al. Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. Nat Commun. 2015;6:7479.

[311] Coss KP, Vasiljevic S, Pritchard LK, Krumm SA, Glaze M, Madzorera S, et al. HIV-1 Glycan Density Drives the Persistence of the Mannose Patch within an Infected Individual. J Virol. 2016;90:11132-44.

[312] Bonomelli C, Doores KJ, Dunlop DC, Thaney V, Dwek RA, Burton DR, et al. The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade. PLoS One. 2011;6:e23521.

[313] Doores KJ, Bonomelli C, Harvey DJ, Vasiljevic S, Dwek RA, Burton DR, et al. Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. Proc Natl Acad Sci U S A. 2010;107:13800-5.

[314] Geyer H, Holschbach C, Hunsmann G, Schneider J. Carbohydrates of human immunodeficiency virus. Structures of oligosaccharides linked to the envelope glycoprotein 120. J Biol Chem. 1988;263:11760-7.

[315] Go EP, Herschhorn A, Gu C, Castillo-Menendez L, Zhang S, Mao Y, et al. Comparative Analysis of the Glycosylation Profiles of Membrane-Anchored HIV-1 Envelope Glycoprotein Trimers and Soluble gp140. J Virol. 2015;89:8245-57.

[316] Go EP, Ding H, Zhang S, Ringe RP, Nicely N, Hua D, et al. Glycosylation Benchmark Profile for HIV-1 Envelope Glycoprotein Production Based on Eleven Env Trimers. J Virol. 2017;91.

[317] Pabst M, Chang M, Stadlmann J, Altmann F. Glycan profiles of the 27 N-glycosylation sites of the HIV envelope protein CN54gp140. Biological chemistry. 2012;393:719-30.

[318] Go EP, Hewawasam G, Liao HX, Chen H, Ping LH, Anderson JA, et al. Characterization of glycosylation profiles of HIV-1 transmitted/founder envelopes by mass spectrometry. J Virol. 2011;85:8270-84.

[319] Go EP, Chang Q, Liao HX, Sutherland LL, Alam SM, Haynes BF, et al. Glycosylation site-specific analysis of clade C HIV-1 envelope proteins. J Proteome Res. 2009;8:4231-42.

[320] Go EP, Irungu J, Zhang Y, Dalpathado DS, Liao HX, Sutherland LL, et al. Glycosylation site-specific analysis of HIV envelope proteins (JR-FL and CON-S) reveals major differences in glycosylation site occupancy, glycoform profiles, and antigenic epitopes' accessibility. J Proteome Res. 2008;7:1660-74.

[321] Sarkar A, Bale S, Behrens AJ, Kumar S, Sharma SK, de Val N, et al. Structure of a cleavage-independent HIV Env recapitulates the glycoprotein architecture of the native cleaved trimer. Nat Commun. 2018;9:1956.

[322] Zhou TQ, Doria-Rose NA, Cheng C, Stewart-Jones GBE, Chuang GY, Chambers M, et al. Quantification of the Impact of the HIV-1-Glycan Shield on Antibody Elicitation. Cell Reports. 2017;19:719-32.

[323] Behrens AJ, Crispin M. Structural principles controlling HIV envelope glycosylation. Curr Opin Struct Biol. 2017;44:125-33.

[324] Dey AK, Cupo A, Ozorowski G, Sharma VK, Behrens AJ, Go EP, et al. cGMP production and analysis of BG505 SOSIP.664, an extensively glycosylated, trimeric HIV-1 envelope glycoprotein vaccine candidate. Biotechnology and bioengineering. 2018;115:885-99.

[325] Yolitz J, Schwing C, Chang J, Van Ryk D, Nawaz F, Wei D, et al. Signal peptide of HIV envelope protein impacts glycosylation and antigenicity of gp120. Proc Natl Acad Sci U S A. 2018;115:2443-8.

[326] McCoy LE, Falkowska E, Doores KJ, Le K, Sok D, van Gils MJ, et al. Incomplete Neutralization and Deviation from Sigmoidal Neutralization Curves for HIV Broadly Neutralizing Monoclonal Antibodies. PLoS Pathog. 2015;11:e1005110.

[327] Panico M, Bouche L, Binet D, O'Connor MJ, Rahman D, Pang PC, et al. Mapping the complete glycoproteome of virion-derived HIV-1 gp120 provides insights into broadly neutralizing antibody binding. Sci Rep. 2016;6:32956.

[328] Behrens AJ, Kumar A, Medina-Ramirez M, Cupo A, Marshall K, Cruz Portillo VM, et al. Integrity of Glycosylation Processing of a Glycan-Depleted Trimeric HIV-1 Immunogen Targeting Key B-Cell Lineages. J Proteome Res. 2018;17:987-99.

[329] Watanabe Y, Raghwani J, Allen JD, Seabright GE, Li S, Moser F, et al. Structure of the Lassa virus glycan shield provides a model for immunological resistance. Proc Natl Acad Sci U S A. 2018;115:7320-5.

[330] Crispin M, Doores KJ. Targeting host-derived glycans on enveloped viruses for antibody-based vaccine design. Curr Opin Virol. 2015;11:63-9.

[331] Behrens AJ, Struwe WB, Crispin M. Glycosylation profiling to evaluate glycoprotein immunogens against HIV-1. Expert review of proteomics. 2017;14:881-90.

[332] Willey RL, Shibata R, Freed EO, Cho MW, Martin MA. Differential glycosylation, virion incorporation, and sensitivity to neutralizing antibodies of human immunodeficiency virus type 1 envelope produced from infected primary T-lymphocyte and macrophage cultures. J Virol. 1996;70:6431-6.

[333] Pritchard LK, Harvey DJ, Bonomelli C, Crispin M, Doores KJ. Cell- and Protein-Directed Glycosylation of Native Cleaved HIV-1 Envelope. J Virol. 2015;89:8932-44.

[334] Schnaar RL. Glycobiology simplified: diverse roles of glycan recognition in inflammation. J Leukoc Biol. 2016;99:825-38.

[335] Rantalainen K, Berndsen ZT, Murrell S, Cao L, Omorodion O, Torres JL, et al. Co-evolution of HIV Envelope and Apex-Targeting Neutralizing Antibody Lineage Provides Benchmarks for Vaccine Design. Cell Rep. 2018;23:3249-61.

[336] Wheeler SF, Rudd PM, Davis SJ, Dwek RA, Harvey DJ. Comparison of the N-linked glycans from soluble and GPI-anchored CD59 expressed in CHO cells. Glycobiology. 2002;12:261-71.

[337] Li Y, Luo L, Thomas DY, Kang CY. Control of expression, glycosylation, and secretion of HIV-1 gp120 by homologous and heterologous signal sequences. Virology. 1994;204:266-78.

[338] Li Y, Bergeron JJ, Luo L, Ou WJ, Thomas DY, Kang CY. Effects of inefficient cleavage of the signal sequence of HIV-1 gp 120 on its association with calnexin, folding, and intracellular transport. Proc Natl Acad Sci U S A. 1996;93:9606-11.

[339] Li Y, Luo L, Thomas DY, Kang CY. The HIV-1 Env protein signal sequence retards its cleavage and down-regulates the glycoprotein folding. Virology. 2000;272:417-28.

[340] Snapp EL, McCaul N, Quandte M, Cabartova Z, Bontjer I, Kallgren C, et al. Structure and topology around the cleavage site regulate post-translational cleavage of the HIV-1 gp160 signal peptide. Elife. 2017;6.

[341] Asmal M, Hellmann I, Liu W, Keele BF, Perelson AS, Bhattacharya T, et al. A signature in HIV-1 envelope leader peptide associated with transition from acute to chronic infection impacts envelope processing and infectivity. PLoS One. 2011;6:e23673.

[342] Guttman M, Garcia NK, Cupo A, Matsui T, Julien JP, Sanders RW, et al. CD4-induced activation in a soluble HIV-1 Env trimer. Structure. 2014;22:974-84.

[343] Go EP, Hua D, Desaire H. 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. J Proteome Res. 2014;13:4012-27.

[344] Stansell E, Panico M, Canis K, Pang PC, Bouche L, Binet D, et al. Gp120 on HIV-1 Virions Lacks O-Linked Carbohydrate. PLoS One. 2015;10:e0124784.

[345] Wang LX. Synthetic carbohydrate antigens for HIV vaccine design. Curr Opin Chem Biol. 2013;17:997-1005.

[346] Doores KJ, Fulton Z, Huber M, Wilson IA, Burton DR. Antibody 2G12 recognizes di-mannose equivalently in domain- and nondomain-exchanged forms but only binds the HIV-1 glycan shield if domain exchanged. J Virol. 2010;84:10690-9.

[347] Aussedat B, Vohra Y, Park PK, Fernandez-Tejada A, Alam SM, Dennison SM, et al. Chemical synthesis of highly congested gp120 V1V2 N-glycopeptide antigens for potential HIV-1-directed vaccines. J Am Chem Soc. 2013;135:13113-20.

[348] Alam SM, Dennison SM, Aussedat B, Vohra Y, Park PK, Fernandez-Tejada A, et al. Recognition of synthetic glycopeptides by HIV-1 broadly neutralizing antibodies and their unmutated ancestors. Proc Natl Acad Sci U S A. 2013;110:18214-9.

[349] Alam SM, Aussedat B, Vohra Y, Meyerhoff RR, Cale EM, Walkowicz WE, et al. Mimicry of an HIV broadly neutralizing antibody epitope with a synthetic glycopeptide. Sci Transl Med. 2017;9.

[350] Cai H, Zhang R, Orwenyo J, Giddens J, Yang Q, LaBranche CC, et al. Multivalent Antigen Presentation Enhances the Immunogenicity of a Synthetic Three-Component HIV-1 V3 Glycopeptide Vaccine. ACS Cent Sci. 2018;4:582-9.

[351] Scanlan CN, Offer J, Zitzmann N, Dwek RA. Exploiting the defensive sugars of HIV-1 for drug and vaccine design. Nature. 2007;446:1038-45.

[352] Dunlop DC. Antigenic mimicry as a strategy for AIDS vaccine design. University of Oxford. 2009.

[353] Dunlop DC, Bonomelli C, Mansab F, Vasiljevic S, Doores KJ, Wormald MR, et al. Polysaccharide mimicry of the epitope of the broadly neutralizing anti-HIV antibody, 2G12, induces enhanced antibody responses to self oligomannose glycans. Glycobiology. 2010;20:812-23.

[354] Agrawal-Gamse C, Luallen RJ, Liu B, Fu H, Lee FH, Geng Y, et al. Yeast-elicited cross-reactive antibodies to HIV Env glycans efficiently neutralize virions expressing exclusively high-mannose N-linked glycans. J Virol. 2011;85:470-80.

[355] Clark BE, Auyeung K, Fregolino E, Parrilli M, Lanzetta R, De Castro C, et al. A bacterial lipooligosaccharide that naturally mimics the epitope of the HIV-neutralizing antibody 2G12 as a template for vaccine design. Chem Biol. 2012;19:254-63.

[356] Stanfield RL, De Castro C, Marzaioli AM, Wilson IA, Pantophlet R. Crystal structure of the HIV neutralizing antibody 2G12 in complex with a bacterial oligosaccharide analog of mammalian oligomannose. Glycobiology. 2015;25:412-9.

[357] Pantophlet R, Trattnig N, Murrell S, Lu N, Chau D, Rempel C, et al. Bacterially derived synthetic mimetics of mammalian oligomannose prime antibody responses that neutralize HIV infectivity. Nat Commun. 2017;8:1601.

[358] Bhiman JN, Anthony C, Doria-Rose NA, Karimanzira O, Schramm CA, Khoza T, et al. Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. Nat Med. 2015;21:1332-6.

[359] Stamatatos L, Pancera M, McGuire AT. Germline-targeting immunogens. Immunol Rev. 2017;275:203-16.

[360] Medina-Ramirez M, Garces F, Escolano A, Skog P, de Taeye SW, Del Moral-Sanchez I, et al. Design and crystal structure of a native-like HIV-1 envelope trimer that engages multiple broadly neutralizing antibody precursors in vivo. J Exp Med. 2017;214:2573-90.

[361] McGuire AT, Hoot S, Dreyer AM, Lippy A, Stuart A, Cohen KW, et al. Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies. J Exp Med. 2013;210:655-63.

[362] Jardine J, Julien JP, Menis S, Ota T, Kalyuzhniy O, McGuire A, et al. Rational HIV immunogen design to target specific germline B cell receptors. Science. 2013;340:711-6.

[363] Steichen JM, Kulp DW, Tokatlian T, Escolano A, Dosenovic P, Stanfield RL, et al. HIV Vaccine Design to Target Germline Precursors of Glycan-Dependent Broadly Neutralizing Antibodies. Immunity. 2016;45:483-96.

[364] Escolano A, Steichen JM, Dosenovic P, Kulp DW, Golijanin J, Sok D, et al. Sequential Immunization Elicits Broadly Neutralizing Anti-HIV-1 Antibodies in Ig Knockin Mice. Cell. 2016;166:1445-58 e12.

[365] Burton DR. What Are the Most Powerful Immunogen Design Vaccine Strategies? Reverse Vaccinology 2.0 Shows Great Promise. Cold Spring Harbor perspectives in biology. 2017;9.

[366] Behrens AJ. Fine structure of the HIV-1 glycan shield. University of Oxford. 2017.

[367] Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 2014;42:W252-8.

[368] Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics. 2011;27:343-50.

[369] Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 2006;22:195-201.

[370] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010;66:486-501.

[371] Emsley P, Crispin M. Structural analysis of glycoproteins: building N-linked glycans with Coot. Acta Crystallogr D Struct Biol. 2018;74:256-63.

[372] Wyatt R, Sodroski J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. Science. 1998;280:1884-8.

[373] Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconjugate Journal. 1988;5:397-409.

[374] Mechref Y. Use of CID/ETD mass spectrometry to analyze glycopeptides. Curr Protoc Protein Sci. 2012;Chapter 12:Unit 12 1 1-1.

[375] Wada Y. Glycan Profiling: Label-Free Analysis of Glycoproteins. In: Kohler JJ, Patrie SM, editors. Mass Spectrometry of Glycoproteins: Methods and Protocols. Totowa, NJ: Humana Press; 2013. p. 245-53.

[376] Wada Y, Azadi P, Costello CE, Dell A, Dwek RA, Geyer H, et al. Comparison of the methods for profiling glycoprotein glycans--HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. Glycobiology. 2007;17:411-22.

[377] Harvey DJ, Scarff CA, Crispin M, Scanlan CN, Bonomelli C, Scrivens JH. MALDI-MS/MS with traveling wave ion mobility for the structural analysis of N-linked glycans. J Am Soc Mass Spectrom. 2012;23:1955-66.

[378] Pucic M, Knezevic A, Vidic J, Adamczyk B, Novokmet M, Polasek O, et al. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. Mol Cell Proteomics. 2011;10:M111 010090.

[379] Munro JB, Gorman J, Ma X, Zhou Z, Arthos J, Burton DR, et al. Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. Science. 2014;346:759-63.

[380] Schiffner T, de Val N, Russell RA, de Taeye SW, de la Pena AT, Ozorowski G, et al. Chemical Cross-Linking Stabilizes Native-Like HIV-1 Envelope Glycoprotein Trimer Antigens. J Virol. 2016;90:813-28.

[381] Yang L, Sharma SK, Cottrell C, Guenaga J, Tran K, Wilson R, et al. Structure-Guided Redesign Improves NFL HIV Env Trimer Integrity and Identifies an Inter-Protomer Disulfide Permitting Post-Expression Cleavage. Frontiers in immunology. 2018;9:1631.

[382] Guenaga J, Dubrovskaya V, de Val N, Sharma SK, Carrette B, Ward AB, et al. Structure-Guided Redesign Increases the Propensity of HIV Env To Generate Highly Stable Soluble Trimers. J Virol. 2015;90:2806-17.

[383] McCoy LE, van Gils MJ, Ozorowski G, Messmer T, Briney B, Voss JE, et al. Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies. Cell Rep. 2016;16:2327-38.

[384] Zhang M, Gaschen B, Blay W, Foley B, Haigwood N, Kuiken C, et al. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. Glycobiology. 2004;14:1229-46.

[385] Kong L, Wilson IA, Kwong PD. Crystal structure of a fully glycosylated HIV-1 gp120 core reveals a stabilizing role for the glycan at Asn262. Proteins. 2015;83:590-6.

[386] Wang W, Nie J, Prochnow C, Truong C, Jia Z, Wang S, et al. A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. Retrovirology. 2013;10:14.

[387] Crooks ET, Grimley SL, Cully M, Osawa K, Dekkers G, Saunders K, et al. Glycoengineering HIV-1 Env creates 'supercharged' and 'hybrid' glycans to increase neutralizing antibody potency, breadth and saturation. PLoS Pathog. 2018;14:e1007024.

[388] Crooks ET, Tong T, Chakrabarti B, Narayan K, Georgiev IS, Menis S, et al. Vaccine-Elicited Tier 2 HIV-1 Neutralizing Antibodies Bind to Quaternary Epitopes Involving

Glycan-Deficient Patches Proximal to the CD4 Binding Site. PLoS Pathog. 2015;11:e1004932.

[389] Voss JE, Andrabi R, McCoy LE, de Val N, Fuller RP, Messmer T, et al. Elicitation of Neutralizing Antibodies Targeting the V2 Apex of the HIV Envelope Trimer in a Wild-Type Animal Model. Cell Rep. 2017;21:222-35.

[390] Klasse PJ, Ketas TJ, Cottrell CA, Ozorowski G, Debnath G, Camara D, et al. Epitopes for neutralizing antibodies induced by HIV-1 envelope glycoprotein BG505 SOSIP trimers in rabbits and macaques. PLoS Pathog. 2018;14:e1006913.

[391] Ringe RP, Pugach P, Cottrell CA, LaBranche CC, Seabright GE, Ketas TJ, 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. J Virol. 2019;93.

[392] Ma BJ, Alam SM, Go EP, Lu X, Desaire H, Tomaras GD, et al. Envelope deglycosylation enhances antigenicity of HIV-1 gp41 epitopes for both broad neutralizing antibodies and their unmutated ancestor antibodies. PLoS Pathog. 2011;7:e1002200.

[393] Watanabe Y, Vasiljevic S, Allen JD, Seabright GE, Duyvesteyn HME, Doores KJ, et al. Signature of Antibody Domain Exchange by Native Mass Spectrometry and Collision-Induced Unfolding. Anal Chem. 2018;90:7325-31.

[394] Sanders RW, van Anken E, Nabatov AA, Liscaljet IM, Bontjer I, Eggink D, et al. The carbohydrate at asparagine 386 on HIV-1 gp120 is not essential for protein folding and function but is involved in immune evasion. Retrovirology. 2008;5:10.

[395] Chuang GY, Zhou J, Acharya P, Rawi R, Shen CH, Sheng Z, et al. Structural Survey of Broadly Neutralizing Antibodies Targeting the HIV-1 Env Trimer Delineates Epitope Categories and Characteristics of Recognition. Structure. 2019;27:196-206 e6.

[396] Lemmin T, Soto C, Stuckey J, Kwong PD. Microsecond Dynamics and Network Analysis of the HIV-1 SOSIP Env Trimer Reveal Collective Behavior and Conserved Microdomains of the Glycan Shield. Structure. 2017;25:1631-9 e2.

[397] Binley JM, Ban YE, Crooks ET, Eggink D, Osawa K, Schief WR, et al. Role of complex carbohydrates in human immunodeficiency virus type 1 infection and resistance to antibody neutralization. J Virol. 2010;84:5637-55.

[398] Wu X, Yang ZY, Li Y, Hogerkorp CM, Schief WR, Seaman MS, et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science. 2010;329:856-61.

[399] Tran EE, Borgnia MJ, Kuybeda O, Schauder DM, Bartesaghi A, Frank GA, et al. Structural mechanism of trimeric HIV-1 envelope glycoprotein activation. PLoS Pathog. 2012;8:e1002797.

[400] Shivatare VS, Shivatare SS, Lee CD, Liang CH, Liao KS, Cheng YY, et al. Unprecedented Role of Hybrid N-Glycans as Ligands for HIV-1 Broadly Neutralizing Antibodies. J Am Chem Soc. 2018.

[401] Appelmelk BJ, Simoons-Smit I, Negrini R, Moran AP, Aspinall GO, Forte JG, et al. Potential role of molecular mimicry between Helicobacter pylori lipopolysaccharide and host Lewis blood group antigens in autoimmunity. Infect Immun. 1996;64:2031-40.

[402] Wildt S, Gerngross TU. The humanization of N-glycosylation pathways in yeast. Nat Rev Microbiol. 2005;3:119-28.

[403] Kukuruzinska MA, Bergh ML, Jackson BJ. Protein glycosylation in yeast. Annu Rev Biochem. 1987;56:915-44.

[404] Gemmill TR, Trimble RB. Overview of N- and O-linked oligosaccharide structures found in various yeast species. Biochim Biophys Acta. 1999;1426:227-37.

[405] Smith WL, Nakajima T, Ballou CE. Biosynthesis of yeast mannan. Isolation of Kluyveromyces lactis mannan mutants and a study of the incorporation of N-acetyl-D-glucosamine into the polysaccharide side chains. J Biol Chem. 1975;250:3426-35.

[406] Cregg JM, Vedvick TS, Raschke WC. Recent advances in the expression of foreign genes in Pichia pastoris. Biotechnology (N Y). 1993;11:905-10.

[407] Grinna LS, Tschopp JF. Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, Pichia pastoris. Yeast. 1989;5:107-15.

[408] Harvey DJ. Fragmentation of negative ions from carbohydrates: part 2. Fragmentation of high-mannose N-linked glycans. J Am Soc Mass Spectrom. 2005;16:631-46.

[409] Kitajima T, Chiba Y, Jigami Y. Saccharomyces cerevisiae alpha1,6-mannosyltransferase has a catalytic potential to transfer a second mannose molecule. FEBS J. 2006;273:5074-85.

[410] Rodionov D, Romero PA, Berghuis AM, Herscovics A. Expression and purification of recombinant M-Pol I from Saccharomyces cerevisiae with alpha-1,6 mannosylpolymerase activity. Protein Expr Purif. 2009;66:1-6.

[411] Armbruster C, Stiegler GM, Vcelar BA, Jager W, Koller U, Jilch R, et al. Passive immunization with the anti-HIV-1 human monoclonal antibody (hMAb) 4E10 and the hMAb combination 4E10/2F5/2G12. J Antimicrob Chemother. 2004;54:915-20.

[412] Guenaga J, Garces F, de Val N, Stanfield RL, Dubrovskaya V, Higgins B, et al. Glycine Substitution at Helix-to-Coil Transitions Facilitates the Structural Determination of a Stabilized Subtype C HIV Envelope Glycoprotein. Immunity. 2017;46:792-803 e3.

[413] Binley JM, Sanders RW, Master A, Cayanan CS, Wiley CL, Schiffner L, et al. Enhancing the proteolytic maturation of human immunodeficiency virus type 1 envelope glycoproteins. J Virol. 2002;76:2606-16.

[414] Dey AK, David KB, Klasse PJ, Moore JP. Specific amino acids in the N-terminus of the gp41 ectodomain contribute to the stabilization of a soluble, cleaved gp140 envelope glycoprotein from human immunodeficiency virus type 1. Virology. 2007;360:199-208.

[415] Ruhaak LR, Zauner G, Huhn C, Bruggink C, Deelder AM, Wuhrer M. Glycan labeling strategies and their use in identification and quantification. Anal Bioanal Chem. 2010;397:3457-81.