Architecture of the HIV-1 glycan shield

By

Laura Pritchard

Linacre College, Oxford

Submitted for the degree of Doctor of Philosophy

at the

University of Oxford

October 2014

Supervised by

Dr Max Crispin and Dr Mark Wormald

Contents

Architecture of the HIV-1 glycan shield

Laura Pritchard
A thesis submitted for the degree of Doctor of Philosophy
Linacre College, University of Oxford
Michaelmas Term 2014

Abstract

In recent years the glycan shield of the HIV-1 envelope spike (Env) has emerged as a potential target for microbicide and vaccine design. The densely packed glycans on its surface include an intrinsic population of under-processed oligomannose structures, and a number of lectins and broadly neutralising antibodies (bnAbs) have been isolated which are reactive to these 'non-self' glycan structures. The potential value of these agents in therapeutic or vaccine contexts depends upon the prevalence of their glycan targets in nature and their resilience to sequence mutation. Here the prevalence of oligomannose-type glycans on recombinant gp120 was demonstrated across a panel of isolates, revealing subtle cross-clade differences. Alanine scanning of all potential N-glycosylation sites (PNGSs) within a model gp120 demonstrated the overall stability of the oligomannose population, but highlighted regions of glycan clusters where individual glycans act to limit the processing of their neighbours. This was formally demonstrated for the N332 'site of vulnerability', where deletion of nearby glycosylation sites led to altered glycan processing at the N332 site. A panel of N332-dependent bnAbs was screened for their ability to tolerate such changes in glycan processing, with differing results. While some displayed promiscuous binding, others were more sensitive to glycan microheterogeneity. Site-specific glycosylation analysis of the PGT135 epitope revealed that an intolerance of certain glycoforms may explain its limited breadth. While a greater understanding of Env glycan microheterogeneity and bnAb promiscuity is required, these findings reveal insights into the architecture of the HIV-1 glycan shield that suggest it is a conserved and robust target for drug and vaccine design.

Acknowledgements

There are a great many people who have supported me throughout my DPhil to whom I am extremely thankful. I have had the benefit of supervision from several brilliant scientists, and have learnt valuable lessons from them all. I would firstly like to thank Dr Terry Butters for giving me the opportunity to study for a doctorate, and who first mentored me when I arrived at the Institute. I would also like to thank Dr Max Crispin and Dr Mark Wormald for their help and advice in producing this thesis, as well as their immeasurably vast knowledge of glycobiology. Finally I would like to acknowledge Dr Chris Scanlan, who warmly welcomed me into his lab and provided much of the direction for this thesis. I will remember his infectious sense of humour as much as his passion for science.

The Glycobiology Institute has been an incredibly supportive environment to work in, and I am grateful to Professor Raymond Dwek for nurturing that environment and his continual encouragement and engagement with my work. I would also like to thank everyone in the lab, past and present, for all their help and making it a great place to work: Iona, for always being so generous with her time and help; Kavitha, for her wise words and enjoyable discussions; Snezana, for her practical assistance and pragmatic advice; Gemma, for her enthusiasm and hard work; Dr Camille Bonomelli, for her endless knowledge about gp120 glycans; and Dr Ben Yu, for his kindness and help. I would also like to thank Professor Nicole Zitzmann for use of various pieces of equipment, and several members of her lab for their help and advice, in particular Simon Spiro, for his invaluable discussions on HIV and tips with this thesis; Dr Dominic Alonzi, for HPLC advice; and Emma Dixon, for helpful discussions.

I would also like to thank the many collaborators who have contributed to this thesis. Dr Katie Doores (King's College London) contributed the ELISA and neutralisation data presented in chapter six, and provided the envelope sequences used in chapter three. Sergey

Menis and Dan Kulp from Professor William Schief's lab (The Scripps Research Institute) generated the glycosylated trimer model presented in chapter five. Professor John Moore and Dr Albert Cupo (Weill Cornell Medical College) provided the BG505 SOSIP.664, BG505 WT.SEKS and CZA97.012 trimers, and Dr Andrew Ward (Scripps) the BG505.705 trimers, analysed in chapter seven. I am grateful to many people for helping with the site-specific glycosylation work. In Oxford, Dr Holger Kramer was very generous with his time and helped accelerate the work, while Professor David Harvey provided many helpful discussions on mass spectrometry and performed the ion mobility ESI-MS/MS glycan analysis described in chapter four. I am indebted to the scientists at Ludger Ltd who welcomed me and provided invaluable help and training on their instruments. In particular I would like to thank Dr Louise Royle and Dr Daniel Spencer for their oversight of the project, and Dr Li Phing Liew for her practical assistance.

Finally I would like to thank my family and friends who have supported me through the ups and downs of the past three years. I owe my sanity to the residents of Huge Alan, and special thanks to Angie for ploughing through this thesis with a fresh pair of eyes. Finally to George, thank you for your continual support and infinite patience throughout this!

1 Introduction

1.1 The Human Immunodeficiency Virus

1.1.1 Emergence of the HIV/AIDS pandemic

It is over 30 years since the first documented cases of acquired immune deficiency syndrome (AIDS). On June 5th 1981 a report was published by the American Centers for Disease Control and Prevention (CDC) detailing the cases of five, previously healthy, homosexual men in Los Angeles who had presented with a rare lung infection (*Pneumocystis carinii* pneumonia) – two of whom had died. By the end of the year hundreds of similar cases had been reported, where otherwise healthy homosexual men had died from rare opportunistic infections. In early 1983 the CDC reported a high prevalence of similar cases in intravenous drug users and haemophiliacs, and suggested that the cause could be an infectious agent being transmitted through sexual contact or by exposure to infected blood.

The isolation of a novel retrovirus from the lymph node of an AIDS patient by the laboratory of Professor Luc Montagnier provided a likely candidate for the reported cases (Barré-Sinoussi et al., 1983). Referred to as Lymphadenopathy Associated Virus (LAV), the retrovirus could be propagated in cultures of healthy T-lymphocytes and appeared to be a member of the human T-cell leukemia virus (HTLV) family. Months later the laboratory of Dr Robert Gallo independently reported the isolation of retroviruses, named collectively as HTLV-III, from 48 AIDS patients (Gallo et al., 1984). It was soon confirmed that LAV and

HTLV-III were the same virus - now known as human immunodeficiency virus (HIV) – and that it was the causative agent of AIDS.

Since the pandemic began the World Health Organisation (WHO) estimates that over 36 million people have lost their lives to HIV/AIDS. A further 35 million people are currently infected with HIV, and while the number of new infections has declined steadily since 1997, the number still remains high (2.6 million in 2009). Sub-Saharan Africa represents the worst affected region, accounting for two-thirds of those infected, with an average prevalence rate of about 5% in adults. HIV/AIDS remains one of the biggest challenges to public health across the world and represents a considerable social and economic burden (World Health Organisation).

1.1.2 The origin and evolution of HIV

HIV originated from simian immunodeficiency viruses (SIVs) which were transmitted to humans from their natural non-human primate hosts. Based on similarities in genome organisation and phylogenetic analyses, the transmission events can be mapped to central western Africa, where the prevalence of hunting and butchery of non-human primates provides a possible route of transmission (Hahn et al., 2000).

Two distinct types of HIV, HIV-1 and HIV-2, can be distinguished – both of which can cause AIDS. HIV-2 is restricted to regions of West Africa and derived from a simian immunodeficiency virus (SIV) that is harboured in sooty mangabey monkeys (*Cercocebus atys*), known as SIV_{smm}. HIV-1 can be divided into four groups (M, N, O and P) which likely arose from independent primate-to-human transmission events. The M, or 'Major', group accounts for the vast majority of all HIV-1 infections worldwide. It can be further divided into different clades (A, B, C, D, F, G, H, J, and K) based on sequence variation within the *gag* and *env* genes, and furthermore circulating recombinant forms (CRFs) have

been identified which have arisen from recombination between different clades. The individual clades show variable distributions around the world. Clade B is prevalent in Northern America, Latin America, Western Europe and Australia, however it is clade C - predominantly found in Southern Africa - that is responsible for the largest number of infections.

Group M is thought to have derived from an SIV that is harboured in the chimpanzee *Pan troglodytes troglodytes* (SIV_{cpz}) (Gao et al., 1999), and the earliest recorded case of infection comes from a stored plasma sample taken from the Democratic Republic of the Congo in 1959 (Nahmias et al., 1986; Zhu et al., 1998). Infections of HIV-1 within the N ('non-M, non-O') group are very rare, and are mostly limited to Cameroon and its neighbouring countries, although they are believed to also have derived from SIV_{cpz}. Group P was more recently discovered, and has been attributed to transmission of SIV_{gor} from gorillas to humans (Plantier et al., 2009; Vallari et al., 2011). It is currently unclear whether O ('outlier') group viruses were transmitted from gorillas or chimpanzees.

1.1.3 Structure & genome organisation

The HIV-1 genome comprises a single-stranded positive-sense RNA, approximately 9 kB in length (Fig. 1.1), with each virion containing two copies within a capsid core (Fig. 1.2). In common with other lentiviruses, the RNA genome is reverse transcribed into cDNA and integrated into the host cell genome prior to protein expression. Several structural elements exist within the genome to modulate expression of the viral proteins. For example the long terminal repeats (LTRs) at the 5′ and 3′ ends are not only important for integration of the HIV-1 DNA, they also contain regulatory elements that are important for transcriptional activation and polyadenylation (Klaver and Berkhout, 1994). Other elements include the

TAR (trans activation response) region, to which the Tat (trans-activator of transcription) protein binds, and the RRE (Rev response element), to which the Rev protein binds.

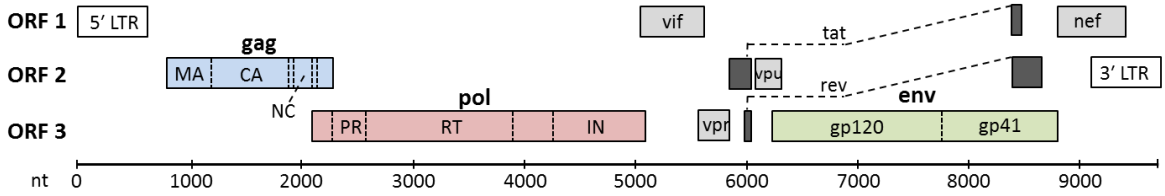


Fig. 1.1. Organisation of the HIV-1 genome. The HIV-1 genome consists of positive, single-stranded RNA and is read in multiple open reading frames (ORFs). The structural proteins (gag; shown in blue) - matrix (MA), capsid (CA) and nucleocapsid (NC) – are cleaved by the viral protease (PR), which also cleaves itself and the other viral enzymes (pol; shown in red) - reverse transcriptase (RT) and integrase (IN). The gp160 precursor produced from env (shown in green) is cleaved by the cellular protease, furin. Regulatory proteins are shown in dark grey and accessory proteins are in light grey. nt – nucleotides. Figure adapted from the Los Alamos HIV sequence database (http://www.hiv.lanl.gov/content/sequence/HIV/MAP/landmark.html).

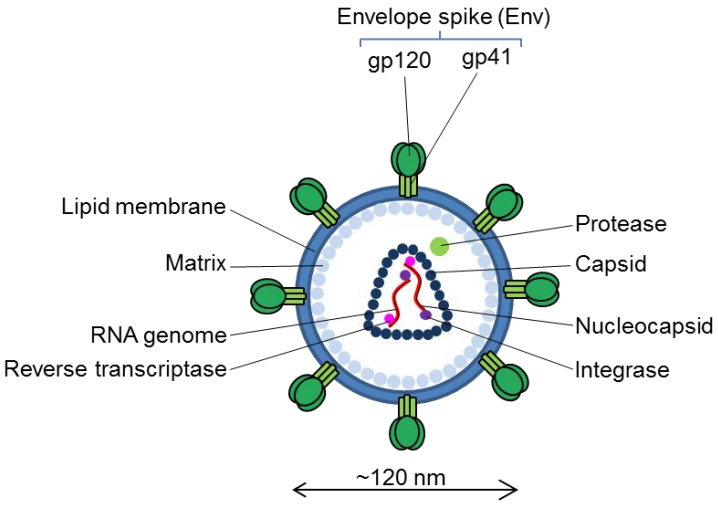


Fig. 1.2. HIV-1 virion structure. The matrix protein provides a lattice beneath the host cell-derived lipid membrane, and interacts with the membrane-bound envelope spike (Env). Within the virion two copies of the RNA genome are coated by nucleocapsid proteins, and are associated with viral cofactors. These complexes are protected within a canonically-shaped, capsid-coated compartment.

Common to other retroviruses, the HIV-1 genome encodes three major gene products, Gag, Pol and Env, across several open reading frames. The Pol polypeptide precursor includes the key HIV-1 enzymes: reverse transcriptase (RT), integrase (IN), an aspartyl protease (PR) and RNase H. The Gag polypeptide precursor includes the major structural proteins which form the underlying virion structure: matrix (MA), capsid (CA),

nucleocapsid (NC) and the accessory protein, p6 (Fig. 1.2). Both the Gag and Pol polypeptide precursors are cleaved by the viral PR. In contrast the gp160 envelope precursor (Env) is cleaved by the cellular protease, furin, during its secretion through the Golgi. The genome also encodes four accessory proteins (Vpr, Vif, Vpu and Nef), and two regulatory proteins (Tat and Rev) (Frankel and Young, 1998).

1.1.4 Lifecycle

An overview of the HIV-1 lifecycle is shown in Fig. 1.3. Infection is initiated through the binding of the envelope spike (Env) to the HIV-1 cellular receptor, CD4, on a target cell (stage 1). Initial attachment to CD4 facilitates interaction with the co-receptor, either CCR5 or CXCR4, which triggers the conformational changes required for membrane fusion (stage 2; see also section 1.3.2). The site of membrane fusion was originally thought to be the plasma membrane, however the actin cortex that lies beneath typically provides a barrier against viruses infecting via this route (Grove and Marsh, 2011). Evidence that co-receptor signalling can activate cofilin, a cellular actin-depolymerizing factor, suggests that HIV-1 may be able to overcome this barrier (Yoder et al., 2008). However several studies have observed HIV-1 entry via receptor-mediated endocytosis, offering an alternative route to bypass this barrier (Miyauchi et al., 2009; van Wilgenburg et al., 2014). Membrane fusion releases the core into the target cell which is followed by uncoating of the capsid (stage 3). The exposed RNA genome is reverse transcribed by viral RT, an error-prone process that frequently introduces mutations into the viral genome (stage 4). A double-stranded hybrid of RNA/DNA is initially produced, before viral RNase H degrades the RNA genome and a second DNA strand is synthesised by RT. The double-stranded DNA remains associated with PR, IN, MA and Vpr, forming the pre-integration complex, which is then imported into

the nucleus where IN integrates the DNA into the host cell genome (stage 5) (Frankel and Young, 1998).

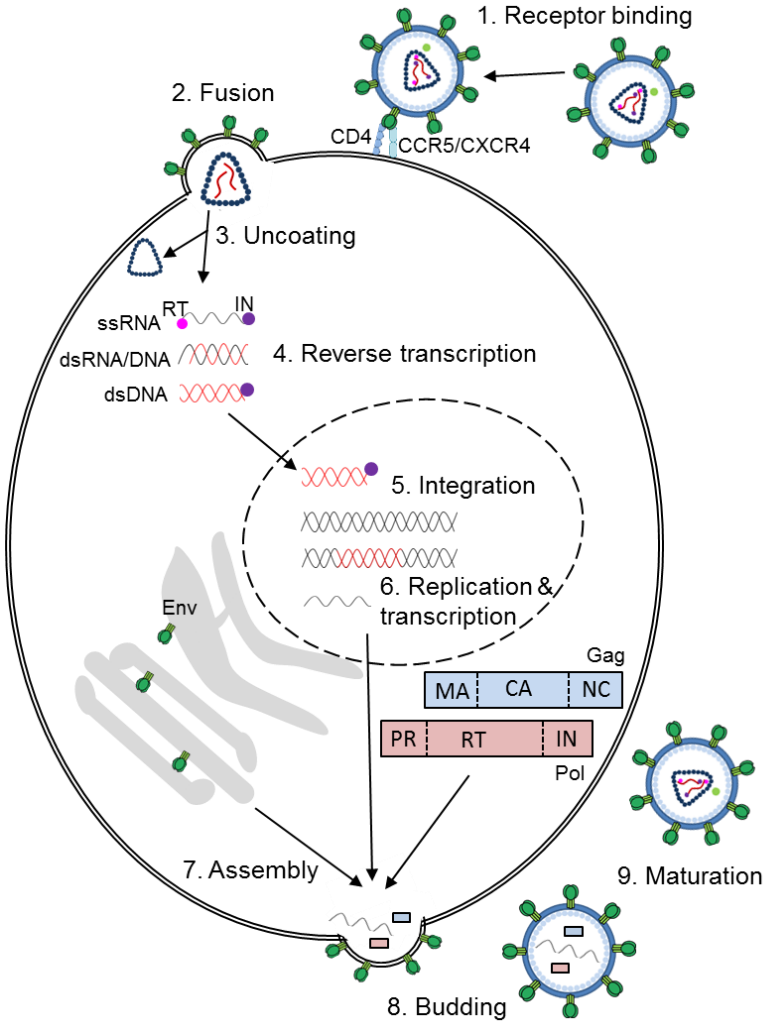


Fig. 1.3. The HIV-1 lifecycle. Traditional view of the HIV-1 lifecycle, see main text for details of individual stages. Note: the location of membrane fusion and capsid coating is currently uncertain.

The timing and precise locations of these events, in particular the uncoating of the viral capsid, have become a matter of debate in recent years. Originally it was thought that uncoating occurred immediately following entry into the cell. However a more recent model assigns essential roles for the capsid in cytosolic transport to the nuclear pore and facilitation of efficient reverse transcription. In this model uncoating is triggered prior to nuclear import only once reverse transcription is complete (Arhel, 2010). The host restriction factor,

TRIM5α, is thought to work by promoting the premature disassembly of the capsid, thereby reducing the level of reverse transcription (Black and Aiken, 2010; Stremlau et al., 2004, 2006).

Once integrated the viral DNA can lie latent for many years, replicating as part of the host cell genome. This gives rise to latent 'reservoirs' of virus, which persist in the face of antiretroviral treatment and thus make it impossible to fully eradicate the virus (Chun et al., 1997; Finzi, 1997). Alternatively, the DNA may be transcribed with the help of the transcriptional activator Tat. In the early stages of infection, spliced mRNA is exported from the nucleus and translated into Tat and Rev. In the later stages the accumulation of Rev, which binds unspliced or singly-spliced mRNAs, promotes the export of full-length mRNA transcripts which encode the Gag, Pol and Env proteins (stage 6). Translation of the Gag, Pol and Env genes occurs in the cytoplasm, with the Env proteins being directed to the ER and secretory pathway via a signal peptide.

While the main site of assembly of new viral particles is accepted to occur on the plasma membrane of $CD4^+$ T lymphocytes (stage 7), the situation is less clear in macrophages (Bieniasz, 2009). Studies have revealed the presence of non-endosomal, neutral pH intracellular compartments within macrophages where the assembly and storage of viral particles has been observed (Deneka et al., 2007; Welsch et al., 2011). The presence of plasma membrane markers within them suggests they derive from invaginations of the plasma membrane, although it is currently unclear what regulates their formation and how HIV-1 assembly might be distributed between these compartments and the plasma membrane (Tan and Sattentau, 2013).

Once the newly synthesised viral genome and proteins have assembled, the virion buds off (stage 8). The human plasma membrane protein tetherin (CD317) acts as a restriction factor to inhibit release of newly formed viral particles, however this is

counteracted by the HIV-1 Vpu protein (Neil et al., 2008). Budded particles remain non-infectious until PR cleaves the Gag polypeptide precursor (maturation; stage 8) (McCune et al., 1988). This triggers conformational rearrangements that allow formation of the capsid core and coating of the viral RNA with nucleocapsid.

1.1.5 Pathogenesis of HIV-1 infection

The cellular tropism of HIV-1 is determined by the distribution of CD4 (present on macrophages, dendritic cells and $CD4^+$ T lymphocytes) as well as the co-receptors, CCR5 or CXCR4. The vast majority of infections are established by CCR5-tropic viruses, even upon transmission of CXCR4-tropic or CCR5/CXCR4 dual-tropic viruses (Margolis and Shattock, 2006; Zhu et al., 1993). Most infections occur via exposure at the genital or rectal mucosa, and thus HIV-1 must penetrate the epithelial walls at these sites to access the lymph tissues beneath. HIV-1 is thought to pass through the small gaps in the epithelial walls, where they encounter dendritic cells (DCs), in particular Langerhans cells (LCs) and $CD4^+$ memory T lymphocytes (Hladik and McElrath, 2008).

Following initial infection HIV-1 disseminates to lymphoid tissues (spleen, blood, bone marrow and lymph nodes; section 1.4.2), and in particular the gut-associated lymphoid tissue (GALT) which has a very high concentration of $CCR5^+$ $CD4^+$ T lymphocytes (Brenchley et al., 2004; Poles et al., 2001; Veazey et al., 2001). This migration has been attributed to an interaction between a region of the gp120 V2 loop with the 'gut-homing receptor', integrin $\alpha_4\beta_7$, on the surface of T lymphocytes (Arthos et al., 2008; Wagner et al., 1996). It was found that gp120 binding triggered LFA-1 activation, a key mediator in the formation of virological synapses, offering a further mechanism of viral spread (Arthos et al., 2008). It has been hypothesised that blocking this $\alpha_4\beta_7$-binding motif represents a mechanism of protection by non-neutralising or weakly-neutralising V2-specific antibodies

(Gorny et al., 2012). However a more recent study could not detect significant binding of a panel of gp120s to $\alpha_4\beta_7$ expressed on activated peripheral blood mononuclear cells (PBMCs) or 293T cells, and thus the significance of this interaction remains unclear (Perez et al., 2014).

The dynamics of HIV-1 infection can be separated into distinct phases (McMichael et al., 2010) (Fig. 1.4). The acute infection phase begins with an eclipse period, where replication is restricted to local mucosal tissues and viral RNA is undetectable in the plasma. Following DC-mediated dissemination of viral particles, plasma viremia increases exponentially with a concomitant decrease in $CD4^+$ T lymphocyte numbers. Arrival of viral particles at GALT and other lymphoid tissues initiates a massive depletion of memory $CCR5^+$ $CD4^+$ T lymphocytes, and it is thought that the effects of this initial depletion are irreversible and ultimately trigger the path towards immune dysregulation (Picker and Watkins, 2005). After several weeks of infection an adaptive immune response is mounted, which is associated with a rallying of $CD4^+$ T lymphocyte numbers and a decrease in plasma viremia to the viral set point. However various mechanisms of immune escape employed by HIV-1 enable repeated evasion of neutralising antibody and cytotoxic T lymphocyte (CTL) responses, which are progressively weakened as further losses of $CD4^+$ T lymphocytes are sustained. In approximately half of infected individuals the latter stages of infection are associated with a switch in co-receptor usage to $CXCR4^+$ $CD4^+$ T lymphocytes, which are particularly abundant in the thymus (Connor et al., 1997; Penn et al., 1999). The latter stages of infection are associated with AIDS-associated opportunistic infections, which ultimately lead to death.

Depletion of T lymphocytes due to productive HIV-1 infection is thought to predominantly occur through apoptosis. Several HIV-1 proteins, including gp120, Tat and Vpr, initiate the intrinsic pathway through the release of cytochrome c from mitochondria.

This pathway is also triggered following the formation of syncytia, which form when Env expressed on the surface of infected cells interacts with CD4 on non-infected cells, triggering membrane fusion and the formation of giant multi-nucleated cells. Apoptosis of T lymphocytes can also occur from HIV-specific immune responses mediated by CTLs, which respond to the presentation of viral antigens on MHC class I molecules by releasing apoptotic-inducing cytotoxic granules (Gougeon, 2003).

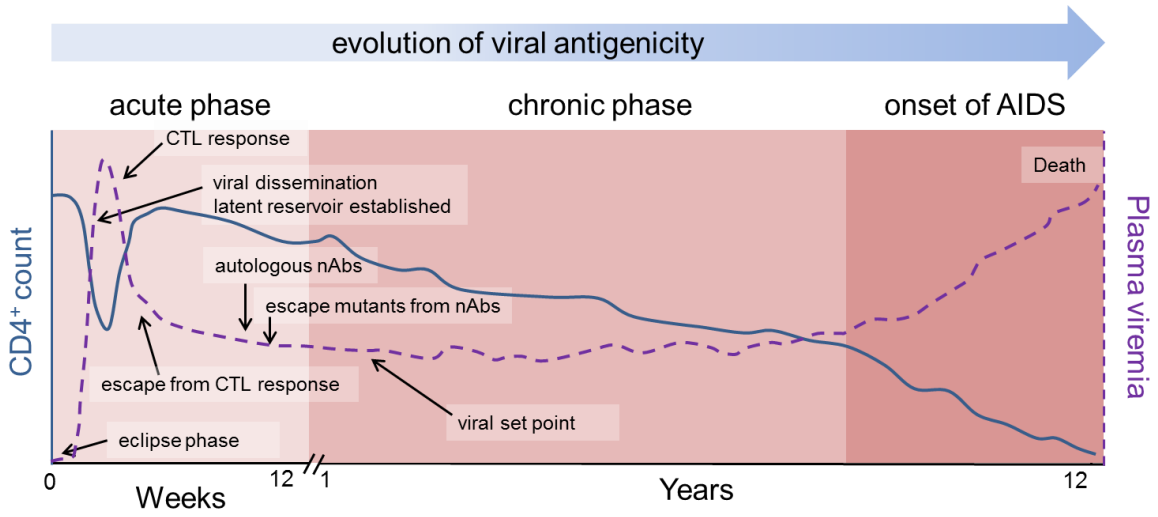


Fig. 1.4. HIV-1 pathogenesis. See main text for details. nAbs, neutralising antibodies; CTL, cytotoxic T lymphocyte; AIDS, acquired immunodeficiency disease.

Although important in the acute phase, productive infection causing apoptosis does not appear to be the major reason for widespread T lymphocyte depletion. The chronic depletion of T lymphocytes leading up to the development of AIDS is thought to occur by mechanisms not directly related to viral replication. Non-infected, bystander cells can be destroyed by the extrinsic apoptotic pathway, triggered through the release of viral proteins from nearby infected cells. For example the release of Tat triggers the expression of death receptors and their ligands (CD95 and CD95L) on the surface of neighbouring uninfected cells, increasing their susceptibility to apoptosis by the extrinsic pathway (Gougeon, 2003). However it is widely believed that the biggest cause of T lymphocyte depletion is through

chronic immune activation (Lawn and Butera, 2001). This is supported by observations of sooty mangabeys, infected with non-pathogenic SIV, which display low levels of immune activation and limited T lymphocyte depletion. This is in spite of having high levels of viral replication, suggesting that mechanisms other than viral replication must be responsible for immune activation (Silvestri et al., 2003). It has been suggested that chronic immune activation might arise from circulating bacterial derivatives that are translocated across gut mucosa following damage to GALT during acute infection. One study demonstrated a positive correlation between the levels of circulating LPS and the degree of innate and adaptive immune activation. Furthermore bacterial translocation was not observed in the non-pathogenic SIV infection model, correlating with the observations above (Brenchley et al., 2006). More recently it was shown that in human spleen and tonsil cultures the vast majority of cell death occurred in non-productively infected T lymphocyte, whereby replication was aborted prior to completion of reverse transcription. The products of the abortive infection are recognised by the cellular DNA sensor IFI16, leading to cell death via pyroptosis (Doitsh et al., 2014). Pyroptosis is a highly inflammatory process that involves cell lysis and the release of pro-inflammatory cytokines. Thus it may not only represent a significant cause of direct T lymphocyte death, but the associated inflammatory response likely contributes to chronic immune activation and ongoing T lymphocyte depletion.

1.2 Fighting the HIV/AIDS pandemic

1.2.1 Highly active antiretroviral therapy (HAART) and current treatments

Following the discovery of HIV-1 it took just under four years for the first anti-retroviral drug, azidothymidine (AZT), to be licensed for use (Fischl et al., 1987). Today there are around 20 separate drugs available covering six different drug classes: Table 1.1 outlines the

different classes and their mechanism of action. The demand for the large number and range of drugs available has stemmed from the speed at which HIV-1 can acquire resistance to an individual drug (Clavel and Hance, 2004). This propensity for resistance derives from two features of HIV-1 infection in particular. Firstly, the error-prone method of DNA replication via reverse transcriptase, which has been found to possess a higher mutation rate than other retroviruses (Mansky and Temin, 1995; Preston et al., 1988; Roberts et al., 1988). Secondly, HIV-1 replicates at a very high rate which, when combined with the frequency of mutation, allows the generation of a huge population of viruses, known as quasi-species (Coffin, 1995).

Table 1.1. Classes of anti-retroviral drugs and their mechanisms of action.

Class	Mechanism of action	Examples
Nucleoside and nucleotide reverse transcriptase inhibitors	Directly inhibit reverse transcription of viral RNA into cDNA.	azidothymidine, abacavir, lamivudine, emtricitabine
Non-nucleoside reverse transcriptase inhibitors	Inhibit reverse transcription of viral RNA by allosteric modulation	delavirdine, efavirenz, etravirine, nevirapine
Protease inhibitors	Inhibit proteolytic cleavage of Gag and Gag-Pol precursor polyproteins	tipranivir, ritonavir, darunavir, atazanavir, fosamprenavir
Integrase inhibitors	Inhibit integration of viral DNA into host chromosomes	raltegravir, dolutegravir
Fusion inhibitors	Inhibit gp41-mediated fusion of viral and cell membranes	enfuvirtide
Entry inhibitors	Antagonist of the CCR5 co-receptor	maraviroc

In vitro studies have demonstrated that complete drug resistance to a monotherapy can be observed in a matter of weeks (Clavel and Hance, 2004). Antiretroviral therapy is therefore delivered as a combination therapy (known as highly active antiretroviral therapy, or HAART) of three or more drugs from different classes, with the aim of raising the genetic

barrier to resistance. HAART has proved to be a highly effective strategy and in some regions of the world it has extended life expectancies to near-normal levels. Since 2013 the WHO has also recommended the use of HAART as a preventative measure for certain high-risk groups such as pregnant women, following evidence of its protective capacity (Baeten et al., 2012). Global efforts to increase access to HAART in low- and middle-income countries has made a huge impact, and the WHO estimates that over the last decade it has saved 4.2 million lives and prevented 800,000 new infections in children.

1.2.2 Towards an AIDS vaccine

While the introduction and deliverance of HAART represents a major achievement in the fight against HIV/AIDS, it is far from a perfect solution. HAART is not a cure for HIV-1 infection; it does not clear the persistent latent reservoirs of virus that exist and if treatment is stopped, rapid viral rebound is observed. Antiretroviral drugs are also associated with unpleasant side effects, which are a common reason for failed adherence to HAART regimens. Many of the drugs must be taken with food which can pose a significant obstacle in areas of extreme poverty. HAART is also expensive and access is not universal; of the 26 million people eligible under WHO guidelines, less than 40% are receiving treatment. Effective preventative measures are desperately required to control the HIV/AIDS pandemic in the long-term, which to date have included promotion of condoms, male circumcision, screening of blood products and needle exchange programs. More recently the use of topical microbicides has been investigated (Shattock and Rosenberg, 2012), and a study demonstrating the successful use of a tenofovir vaginal cream has given hope that this approach could be effective (Abdool Karim et al., 2010). However these preventative measures are dependent on human behaviour, which can be unpredictable, and they do not

offer a long-term solution. The gold standard therefore remains the development of a safe and efficacious prophylactic vaccine.

The aim of vaccination is to condition the adaptive immune system to respond rapidly and effectively to infection by a pathogen, through prior exposure to relevant antigens in a non-pathogenic context. Unfortunately, many features of HIV-1 make vaccine development a formidable challenge. The majority of infections are established by a single transmitted virus (Abrahams et al., 2009; Keele et al., 2008), and if infection is permitted the generation of a pool of $CD4^+$ cells harbouring latent virus prevents eradication. There is, therefore, a very narrow window of opportunity for an immune response to intervene, and provision of sterilising immunity is believed to depend upon the presence of high titres of circulating neutralising antibodies (nAbs). Indeed passive transfer of nAbs has been demonstrated to protect rhesus macaques from infection with SIV (Hessell et al., 2009; Mascola et al., 2000; Moldt et al., 2012; Parren et al., 2001; Shingai et al., 2013). However the challenges of inducing high titres of nAbs through vaccination are considerable. Neutralising antibody responses against HIV-1 are targeted to the envelope spike, which varies considerably in both length and sequence (Korber et al., 2001). Furthermore HIV-1 can shield conserved antigenic regions using carbohydrates acquired from the host cell glycosylation machinery, and through the quaternary nature of its structure (Section 1.3).

Other lines of vaccine research have focused on driving a CTL response, since it is known to influence early viral replication events (Ernardin et al., 2005; McMichael et al., 2010; Salazar-Gonzalez et al., 2009). Furthermore long-term controllers of HIV-1 infection have been associated with strong CTL responses (Walker, 2007). Advantages of this approach include the possibility of targeting alternative, more conserved, antigens over the highly variable envelope protein, however CTL responses require presentation of antigens on major histocompatibility complex (MHC) class I molecules, typically requiring cell

infection, and thus do not offer sterilising immunity. Despite this it is hoped that a robust CTL response could control early viremia and prevent irreversible destruction of $CD4^+$ memory T cells, thereby reducing the viral load. In a challenge study of macaques with the SHIV-89.6P isolate (SHIV is a chimera based on the SIV backbone with HIV-1 Env, Tat and Rev genes), a prime/boost vaccination regimen targeting the CTL response did not protect macaques from infection, however vaccinated monkeys demonstrated stronger CTL responses, higher numbers of $CD4^+$ T cells, and low to undetectable levels of viremia (Barouch et al., 2000). A follow up study later found that CTL escape was achieved by mutation of a single nucleotide within a Gag epitope, which led to expanded viral replication and progression to AIDS (Barouch et al., 2002).

To date over 30 different candidate HIV-1 vaccines have been trialled in humans, and a handful of these have been investigated in large scale phase IIB or III studies (McElrath and Haynes, 2010). The first two concurrent major trials, Vax003 and Vax004, involved subunit vaccines comprising recombinant gp120s from the envelope protein of two different strains. AIDSVAX B/B with the adjuvant, alum, was administered in America and in the Netherlands (Flynn et al., 2005), while AIDSVAX B/E with alum was used in Asia (Pitisuttithum et al., 2006). In both cases, no protective effect was observed, and no improvement in $CD4^+$ T cell counts, or levels of viremia, was detected. The next set of trials – the Step trial in the Americas and Australia, and the Phambili trial in South Africa – were focused on eliciting CTL responses and were testing the MRKAd5 HIV-1 gag/pol/nef vaccine, which used a replication-incompetent adenovirus type 5 vector to deliver antigen producing HIV-1 genes into cells. However the studies were stopped early over lack of efficacy and concerns that the vaccine could in fact have caused an increase in rates of HIV-1 infection (Buchbinder et al., 2008; Gray et al., 2010, 2011).

The first glimmer of hope came from a phase III efficacy trial in Thailand. The RV144 trial involved four priming injections using a live recombinant canarypox vector (vCP1521) encoding Gag, $gp120_{93TH023}$ and protease, followed by two booster injections of AIDSVAX B/E. The regimen was found to be moderately protective against HIV-1 infection, with an efficacy rate of 31% (Rerks-Ngarm et al., 2009). However for those who were infected no benefits were observed in terms of $CD4^+$ cell count or viral set point.

Another large study, the HVTN 505 trial, was recently reported (Hammer et al., 2013). This was evaluating the efficacy of a DNA-prime/Ad5-boost regimen in the United States. The DNA prime consisted of 6 plasmids encoding Gag, Pol, Nef and Env proteins from different strains, followed by a rAd5 (recombinant adenovirus type 5) vector boost expressing a Gag-Pol fusion protein and multi-strain Env proteins. This was another trial to be terminated before completion due to lack of efficacy as no protective effects or increase in $CD4^+$ cells was detected. The trail also failed to adequately address the concerns over the potential for adenovirus vaccines to increase rates of infection: 21 participants who received the placebo became infected versus 29 participants who were vaccinated. Opinions differ on the significance of this result, but it has been argued that alternative vectors should be considered (Barouch, 2010; McMichael et al., 2013; Parks et al., 2013).

Since there are no documented cases of natural recovery from HIV-1 infection, the correlates of protection are not clear. Studies are now underway to follow-up on the success of the RV144 trials to determine what immune responses were associated with vaccine-mediated protection. An initial analysis indicated that IgG responses to the V1 and V2 regions of gp120 correlated with protection, while levels of serum IgA seemed to correlate with increased risk of infection (Haynes et al., 2012). The protective effects of IgG have been attributed to antibody-dependent cellular cytotoxicity (ADCC) activity, and the

detrimental effect of IgA appears to derive from competition with IgG for antigen-binding, thereby preventing ADCC (Bonsignori et al., 2012; Tomaras et al., 2013).

1.3 The HIV-1 envelope glycoprotein (Env)

The HIV-1 envelope glycoprotein (Env) mediates infection via its interaction with CD4 and a co-receptor, either CCR5 or CXCR4, on target cells. This critical role, and its exposed location on the virion surface, makes it an important target for drug and vaccine design.

1.3.1 Synthesis and assembly of Env

The functional HIV-1 envelope spike comprises a trimer of non-covalently associated gp120-gp41 heterodimers. The *env* gene is translated as a gp160 precursor and its synthesis is directed to the endoplasmic reticulum (ER) via its signal peptide. $Glc_3Man_9GlcNAc_2$ (Glc, glucose; Man, mannose; GlcNAc, N-acetylglucosamine) oligosaccharide precursors are co-translationally added *en bloc* from dolichol phosphate precursors to free amines on asparagine residues that lie within Asn-X-Ser/Thr-X sequons, where X is any amino acid except proline (Ben-Dor et al., 2004; Kornfeld and Kornfeld, 1985). Depending on the clade and subtype, each gp160 sequence contains between 22-37 putative glycosylation sites (Zhang et al., 2004).

Folding of Env has been demonstrated to be dependent on interactions with the folding chaperones calnexin (CNX) and calreticulin (CLR) (Dettenhofer and Yu, 2001; Li et al., 1996; Otteken and Moss, 1996) (Fig. 1.5). These chaperones interact with mono-glucosylated N-linked oligosaccharides (Ware et al., 1995), which are generated by trimming of the two terminal glucose units on the precursor oligosaccharides by glucosidases I and II. Upon removal of the third and final glucose by glucosidase II, Env faces one of two fates. If it has successfully folded, it will remain deglucosylated and pass to

the Golgi for further processing. If, however, it remains improperly folded the oligosaccharides are reglucosylated by UDP-glucose glucosyl transferase (UGGT) (Helenius and Aebi, 2004); Env is therefore retained within the ER by calnexin and calreticulin for further chaperone-mediated folding. It has been demonstrated that Env spends a considerable amount of time within the ER, presumably due to the complexity of folding (Land et al., 2003; Otteken et al., 1996). This complexity at least partially stems from the requirement to correctly form all ten disulphide bonds for proper folding, which is mediated via protein disulphide isomerase (Earl et al., 1991; Land et al., 2003).

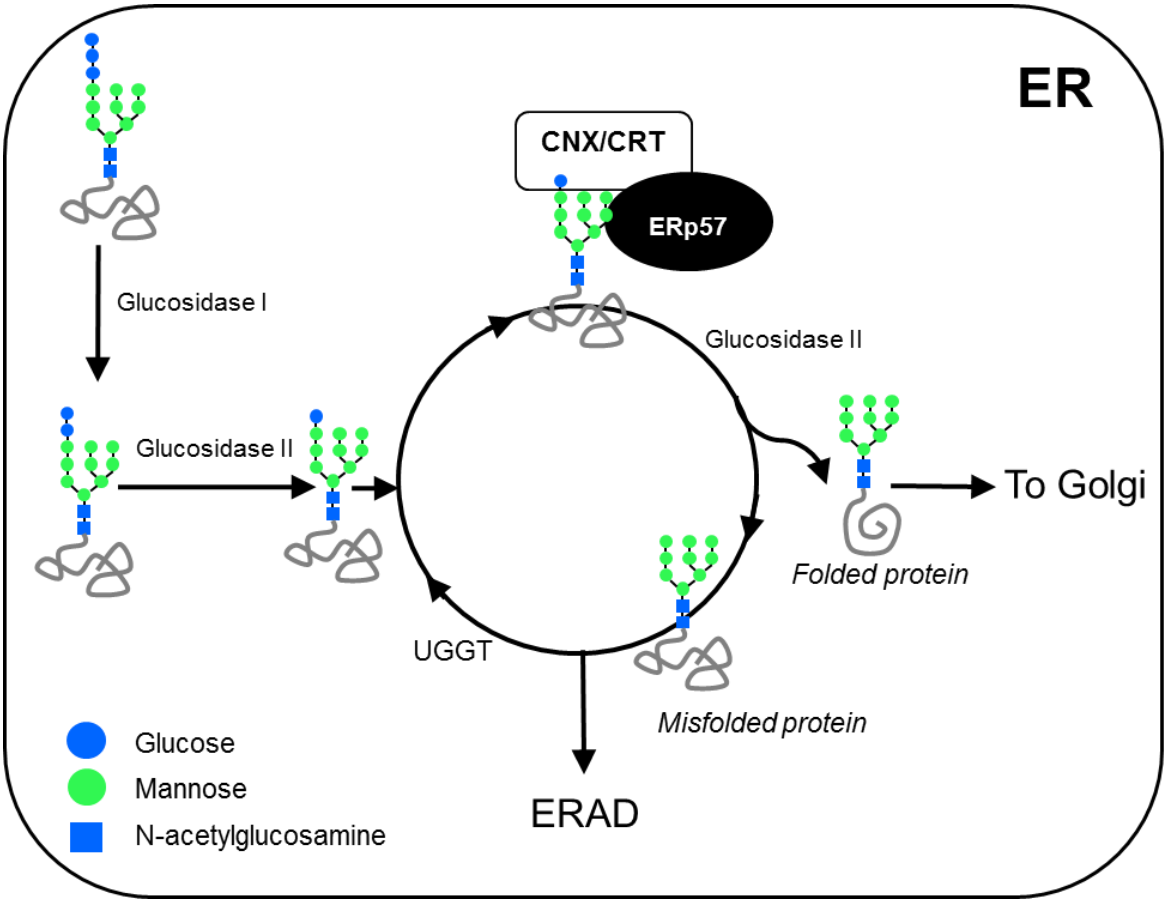


Fig. 1.5. Overview of Env folding via the calnexin (CNX) cycle. $Glc_3Man_9GlcNAc_2$ precursors are added to Env glycan sites by an oligosaccharyltransferase. Two terminal glucose residues are removed by glucosidases I and II, and the monoglucosylated oligosaccharides mediate interaction with CNX/CRT, which associate with the protein disulphide isomerase protein, ERp57. Upon release from CNX, the final glucose residue is removed by glucosidase II. Properly folded Env molecules progress to the Golgi for further glycan processing and proteolytic cleavage, while unfolded or misfolded molecules are reglucosylated by the folding sensor UGGT, initiating another round of chaperone-mediated folding. Molecules that take too long to fold are exposed to components of the ER-associated degradation (ERAD) pathway, which instigate their removal and degradation.

Typically, transit of a glycoprotein through the ER to the Golgi is accompanied by further modifications to the precursor N-linked oligosaccharides (Fig. 1.6A). ER and Golgi α-mannosidases remove up to four mannose residues, before addition of an N-acetlyglucosamine residue to the D1 arm, forming a hybrid-type glycan. Further mannosidase trimming, and the sequential addition of further monosaccharides by an array of glycosyltransferases, result in the formation of diverse hybrid- and complex-type glycan structures (Kornfeld and Kornfeld, 1985). The chemical structures of the most common monosaccharides present in human N-linked glycans are shown in Fig. 6B. This conventional sequence is not observed for all precursor oligosaccharides added to Env, however, despite its lengthy transit time through the secretory pathway. The number of glycan sites on Env is unusually high, and creates a glycan density that is largely impenetrable to the processing enzymes. This lack of access results in many of the oligosaccharides remaining untrimmed, or minimally trimmed, beyond the deglucosylation stage (Bonomelli et al., 2011; Cutalo et al., 2004; Doores et al., 2010a; Go et al., 2009; Leonard et al., 1990; Pabst et al., 2012). This has the effect of producing an 'intrinsic mannose patch' within the outer domain of gp120, which harbours the highest density of glycan sites. Analysis of the glycosylation of virion-derived Env has revealed a greater abundance of oligomannose-type glycans than observed on monomeric gp120, suggesting that oligomerisation of the individual subunits may confer additional steric constraints on the processing enzymes, giving rise to the 'trimer-induced' mannose patch (Bonomelli et al., 2011; Doores et al., 2010a). These ideas are explored in chapter 7.

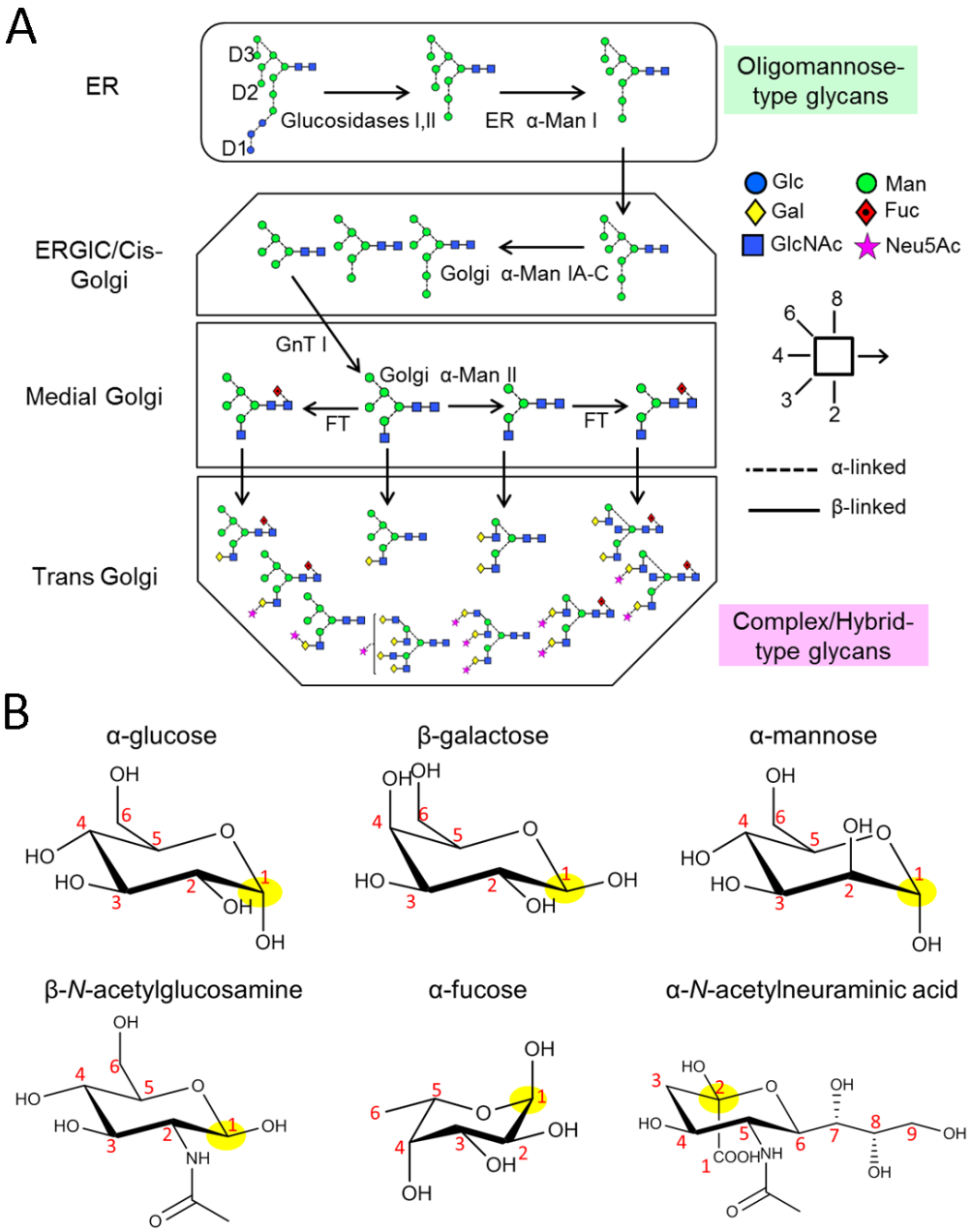


Fig. 1.6. The typical mammalian N-linked glycosylation pathway. (A) Asparagine-linked $Glc_3Man_9GlcNAc_2$ precursors are typically processed in a well characterised sequence. In the ER, glucosidases I and II remove the three terminal glucose residues, while ER α-mannosidase I (ER α-Man I) removes the terminal mannose residue from the D2 arm. Mannosidase trimming continues down to $Man_5GlcNAc_2$ in the ER-Golgi intermediate compartment (ERGIC) and cis-Golgi, through the action of Golgi α-mannosidases IA-C. N-acetylglucosaminyltransferase I (GnT I) then adds a GlcNAc residue to the D1 arm, and extension of this by cell-specific galactosyltransferases, sialyltransferases and fucosyl transferases (FT) produces various hybrid-type glycans. Alternatively, further mannosidase processing down to $GlcNAc_1Man_3GlcNAC_2$ can occur through Golgi α-mannosidase II, and addition of one or more further GlcNAc residues by GnTs II and III can lead to synthesis of a vast array of complex-type glycans. (B) Cyclic chemical structures of the most common monosaccharides present, in their pyranose form, in the precursor and processed N-linked glycans. Monosaccharides can form one of two anomers (α or β) depending on their orientation relative to the anomeric carbon (indicated in yellow). The monosaccharides are typically found in the D form, with the exception of fucose which is found as the L enantiomer. Carbon numbers are labelled in red.

Trimerisation of the individual subunits occurs in the ER, consistent with the early protection of glycan processing observed on viral Env, and is mediated through the ectodomain of gp41 (Earl et al., 1991). A trimer of gp160s therefore passes to the Golgi, where it is cleaved by the furin protease at a designated cleavage site, (RXXR), where X is any amino acid (Hallenberger et al., 1992). Cleavage of the gp160 precursor is essential for the functionality of Env (McCune et al., 1988); uncleaved trimers will interact with CD4 but not undergo fusion with the target cell membrane. Following cleavage, gp120 and gp41 remain associated via non-covalent interactions and pass to the plasma membrane for incorporation into virions. Incorporation into virions is associated with Gag, but the precise mechanism by which this occurs is currently uncertain. However the combination of gp120 shedding, due to the weak interaction gp120-gp41 interaction, and clathrin-mediated endocytosis of Env, results in few functional envelope spikes being incorporated into virions (Checkley et al., 2011).

1.3.2 Structure and function of the Env trimer

The gp120 and gp41 subunits of Env have distinct functions; the extracellular gp120 subunits bind to CD4 and the co-receptor, while the transmembrane gp41 subunits mediate fusion of the viral and target cell membranes. A model for the membrane-fusion process is depicted in Fig. 1.7. Fine atomic resolution structures have been determined for the gp120 and gp41 subunits, giving insights into their mechanisms of action. The overall structure of Env is depicted schematically in Fig 1.8A. Gp120 comprises five constant regions (C1-5), which are highly conserved across different clades and isolates, interspersed with five variable regions (V1-5) which exhibit much greater sequence variation. Once cleaved the gp120 and gp41 subunits remain non-covalently associated, however the weakness of this interaction frequently results in gp120 ‘shedding’ from the virion surface.

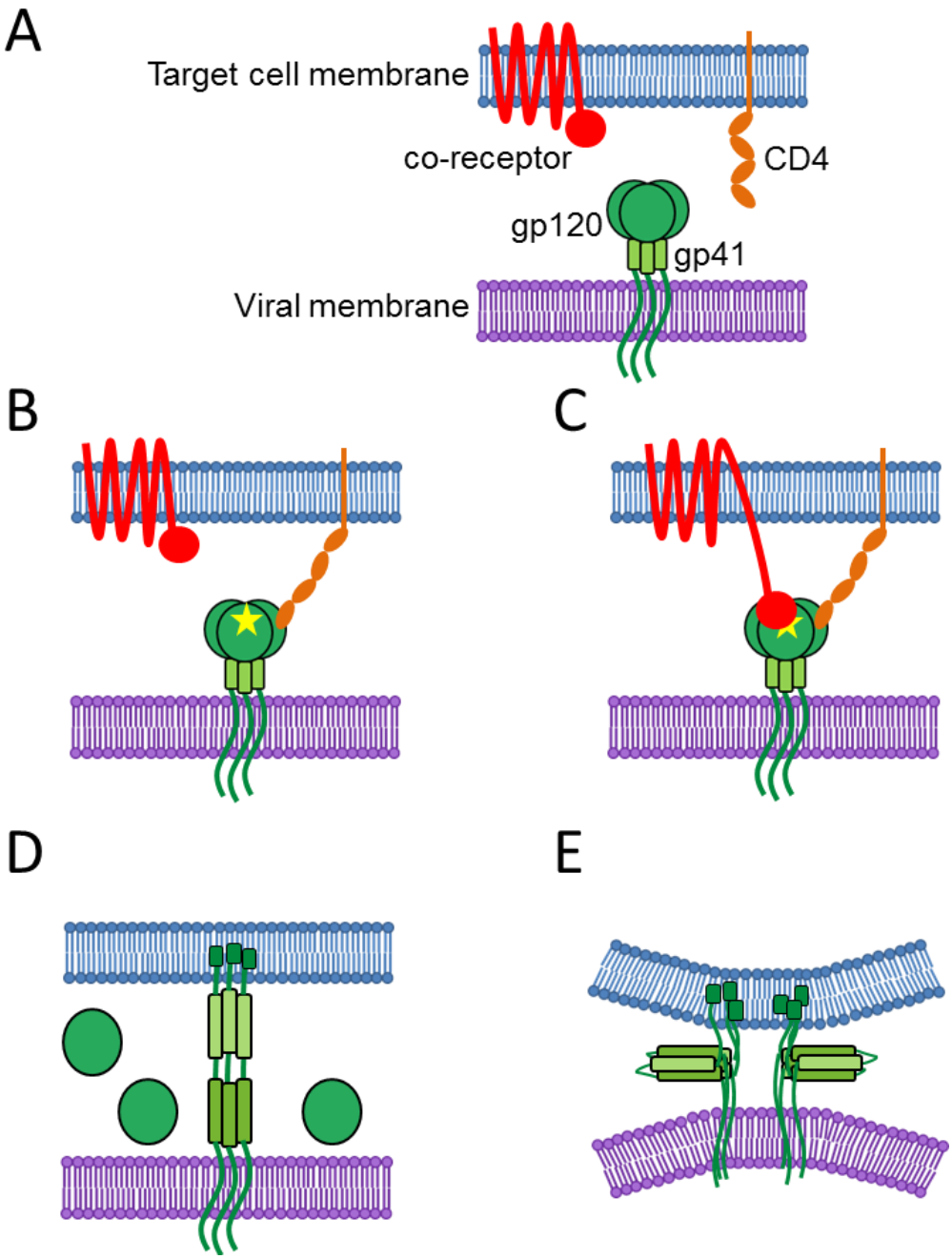


Fig. 1.7. Virus attachment and fusion to target host cell membranes. (A) HIV-1 fusion with target cells is mediated by the envelope protein, a trimer of gp120-gp41heterodimers, which interacts with CD4 and a co-receptor. (B) Fusion is initiated by CD4 binding to gp120, which triggers the formation and stabilisation of the co-receptor binding site (yellow star). (C) The co-receptor, either CCR5 or CXCR4, in turn binds to gp120. (D) Co-receptor binding triggers the N-terminal fusion peptide of gp41 to insert into the target cell membrane, which is associated with shedding of the gp120 subunits. (E) The two helices of each gp41subunit arrange into a hair-pin bend, forming a 6-helix bundle. This draws the membranes together, permitting their fusion.

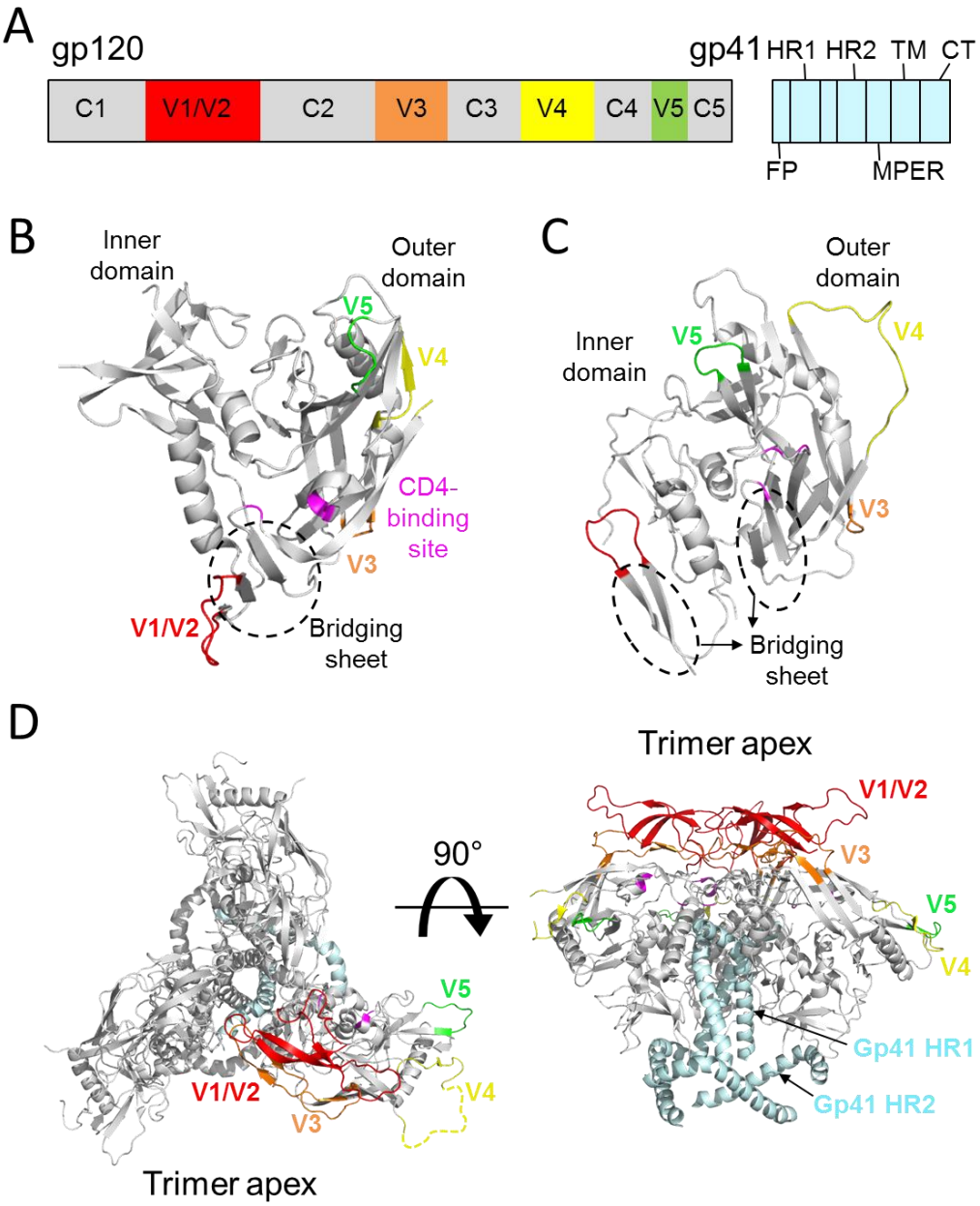


Fig 1.8. Structure of the HIV-1 Env. (A) Schematic illustrating the primary structure of the gp120 and gp41 components. Constant regions of gp120 are in grey, variable regions are coloured. The gp41 peptide is coloured in cyan. FP – fusion peptide, HR1,2 – heptad repeat helices 1 and 2, MPER – membrane proximal external region, TM – transmembrane domain, CT – cytoplasmic tail. (B) Structure of an HIV-1 gp120 core (V1/V2 and V3 loops are replaced by Gly-Ala-Gly tripeptides) in a CD4-liganded conformation. The CD4 binding site (magenta) and bridging sheet (circled), which connects the inner and outer domains, are both formed in the liganded state. Variable loops are coloured according to the scheme in (A). (C) Structure of an unliganded SIV gp120 core. (D) Structure of a soluble BG505 SOSIP.664 gp140 trimer (Julien et al., 2013a), looking down on the trimer apex (left; chain A coloured only) and from the side (right; all chains coloured). Images were made in PyMol using PDB IDs: 1GC1, 2BF1 and 4NCO.

The X-ray crystal structure of a deglycosylated HIV-1 gp120 core in complex with CD4 at 2.5 Å revealed the three-dimensional arrangement of the gp120 domains in the liganded conformation (Fig. 1.8B) (Kwong et al., 1998). The N-terminal end of gp120 forms the inner domain, which is separated from the C-terminal outer domain by a bridging sheet – a four-stranded antiparallel β-sheet that forms upon binding to CD4. The CD4 binding site is a hydrophobic pocket tucked away in the cleft between the inner and outer domains. While binding encompasses interactions with 26 residues on gp120 and 22 on CD4, the most significant interactions occur between Phe43 and Arg59 on CD4 and Asp368, Glu370 and Trp427 of gp120. These three gp120 residues are conserved across all primate immunodeficiency viruses (Kwong et al., 1998). The co-receptor binding site is formed across the bridging sheet and encompasses the adjacent V3 loop, which plays a significant role in determining viral co-receptor tropism (Hwang et al., 1991).

Comparison of the CD4-liganded structure with that of an unbound SIV gp120 core (displaying 70% sequence similarity to the HXB2 gp120 core) revealed that significant conformational changes occur upon CD4 binding (Fig 1.8C) (Chen et al., 2005). The outer domain largely retains the same conformation throughout, while in contrast the inner domain displays a less ordered structure in the unliganded conformation, with individual components capable of independent movement. Upon CD4-binding a more rigid, unified structure is formed, with the two domains slightly tilting away from each other (Chen et al., 2005). These inferred conformational changes are consistent with the high entropic cost associated with CD4-binding (Kwong et al., 2002; Myszka et al., 2000). Upon binding of CD4 and the co-receptor to gp120, membrane fusion is initiated through the action of gp41. The conformational rearrangements expose the gp41 fusion peptide, which inserts into the lipid membrane of the target cell. These rearrangements are thought to be dependent on the action of protein disulphide isomerase in the cellular membrane, which reduces two

disulphide bonds within gp120 (Barbouche et al., 2003). Following insertion of the fusion peptide the HR2 α-helix folds back to lie adjacent with the HR1 α-helix, forming a hairpin-like intermediate. This bundling of the helices brings the viral and cellular membranes together, facilitating their fusion (Doms and Moore, 2000).

For many years structural information on the trimeric complex has been limited to that obtained by low resolution electron microscopy (EM) studies. Electron tomography images of Env on virions at around 20 Å, combined with molecular modelling using the high resolution monomer structures, provided some insight into the trimeric architecture of HIV-1 Env (Liu et al., 2008a; White et al., 2010; Zanetti et al., 2006). These revealed that the V1/V2 region sat at the trimer apex, distal to the viral membrane, contrary to previous suggestions (Chen et al., 2005). Higher resolution cryo-EM studies have been performed using recombinant detergent-stabilised constructs with the cleavage site removed to improve stability (Mao et al., 2012, 2013). However cleavage of the gp160 precursor is known to significantly alter the antigenicity of Env (Herrera et al., 2005; Ringe et al., 2013), and thus it is unknown how representative these constructs are to native trimers. There are also questions over the validity of the method used, which utilised a previously published structure to align and refine the data (Cohen, 2013).

More recently soluble cleaved constructs have been designed, which employ several mutations to facilitate expression and stability. Referred to as 'SOSIP' constructs, they include an artificial disulphide bond to tether the gp120 and gp41 subunits (designated SOS), and an Ile→Pro mutation to improve stability between gp41 subunits (designated IP) (Binley et al., 2000; Klasse et al., 2013; Sanders et al., 2002a; Schülke et al., 2002). These soluble gp140 constructs lack the C-terminal region of gp41, the transmembrane-domain, and most of the membrane proximal external region (MPER). The most advanced SOSIP to date is the BG505 SOSIP.664 construct, which produces trimers that mimic the antigenic

properties of native Env and resemble native trimers by negative stain-EM analysis (Chung et al., 2014; Sanders et al., 2013). The structure of this glycosylated trimer was recently solved using X-ray crystallography and cryo-EM at 4.7 Å and 5.8 Å, respectively (Julien et al., 2013a; Lyumkis et al., 2013). These structures confirmed the orientation modelled using the electron tomography images, with the V1/V2 regions forming the interface at the trimer apex (Fig. 1.8D). Overall the trimers displayed a compact 'mushroom' shape, which contrasts with the presence of a large central pore in the uncleaved structure (Mao et al., 2013). The CD4 binding sites lie sheltered within the inter-protomer clefts and access is further limited by glycans on the V1/V2 and V3 loops.

1.3.3 The antigenic surface of Env

The surface exposed envelope glycoproteins are the sole target for neutralising antibodies (nAbs) against HIV-1. Binding of nAbs to the functional envelope trimer can block infection by several different mechanisms. For example recognition of certain epitopes directly blocks infection by interfering with receptor binding or membrane fusion. Alternatively other nAbs work by aggregating virions and initiating phagocytosis, or by triggering complement-induced virolysis (Zolla-Pazner, 2004). Given the unique role of the envelope glycoproteins in neutralisation, it is unsurprising that they feature several properties designed to protect against recognition by nAbs. These are rooted in the physical characteristics of the Env complex, involving protection at almost every level of its organisation, which gives rise to a unique antigenic profile (Fig. 1.9) (Eggink et al., 2007; Hoxie, 2010; Wyatt et al., 1998).

The oligomeric organisation of the Env complex represents an initial crude barrier to the elicitation of a strong neutralising immune response. Functional Env comprises non-covalently associated trimers of cleaved, gp41-gp120 heterodimers on the virion surface. In this arrangement only the outer regions of the trimer are exposed to antibody

recognition, and are thus said to form the 'neutralising face' of HIV-1 (Fig 1.9) (Wyatt et al., 1998). However effective neutralising antibody responses are rarely observed early on during natural infection and thus fail to prevent HIV-1 infection and dissemination. Instead 'non-neutralising' antibody responses are commonly observed, which are directed against 'non-neutralising' epitopes, i.e. those that are either absent or occluded on the functional trimer. Instead they are present on populations of non-functional Env, which include uncleaved gp160, monomeric gp120, and gp41 'stumps' on the virion surface, the latter two deriving from gp120 shedding (Moore et al., 2006). Epitopes present on these non-functional forms elicit non-neutralising antibody responses which fail to block infection. Furthermore these non-functional forms are thought to be immunodominant, suppressing productive responses against the functional trimer.

Several factors account for the immunodominance of non-neutralising epitopes. One of the major determinants is the presence of an array of glycans that mask much of the neutralising face of the trimer. This 'glycan shield', derived from the host cell glycosylation machinery, is largely recognised as 'self' by the immune system and thus forms the so-called 'silent face' of gp120, protecting potentially immunogenic epitopes, like the CD4 binding site, within the peptide backbone beneath (Fig 1.9). However as discussed in section 1.4.2, it is becoming apparent that the 'glycan shield' may not be as immunologically silent as previously thought, and that some neutralising antibodies are able to recognise and penetrate this shield.

Another major nAb evasion feature is the relative accessibility of conserved and variable regions on Env. The CD4 binding site, which is functionally conserved across isolates, represents a potentially vulnerable site for nAb recognition. However its occluded location within the inter-protomer cleft, with glycans guarding the periphery, reduces the options for nAb approach. Furthermore nAbs that mimic the binding of CD4 typically

induce the corresponding entropic cost of the associated conformational rearrangements (Kwong et al., 2002). Some antibodies have been isolated which bind the CD4bs, however many are only weakly neutralising. One exception is b12, a broadly neutralising antibody that induces a smaller entropic penalty upon binding (Burton et al., 1994; Kwong et al., 2002). More recently the use of engineered Env probes to screen sera from infected individuals has led to the isolation of potent anti-CD4bs bnAbs with neutralisation breadths of up to 90% (Wu et al., 2010). A crystal structure of one such bnAb, VRC01, in complex with gp120 revealed a mode of binding similar to that of CD4, with the exception of a minimal requirement on contacts with the bridging sheet (Zhou et al., 2010). The high potency is therefore thought to arise from the ability to bind with limited conformational rearrangements (Burton and Weiss, 2010). Access to the co-receptor binding site, which is transiently exposed upon CD4 binding, is similarly difficult to target.

In contrast to the 'conformational masking' of these vulnerable conserved regions, the variable sections of gp120 occupy prominent locations at the trimer apex and outer domains, providing immunodominant distractions that divert the immune response away from less adaptable epitopes. Early antibody responses contain high titres of autologous neutralising antibodies targeted against variable regions, in particular the V1/V2 loops. However viral evolution, involving amino acid deletions and substitutions and the shifting of glycosylation sites, drives viral escape (Moore et al., 2010).

Finally the low number of Env spikes on virion surfaces is thought to represent another barrier of protection against antibody recognition. HIV-1 typically contains 7-14 spikes per virion, a much lower number compared to other viruses (Chertova et al., 2002; Chojnacki et al., 2012; Zhu et al., 2006). A low density of spikes on the virion surface reduces the opportunity for bivalent binding by nAbs, which would increase avidity effects and therefore potency (Klein and Bjorkman, 2010). Stimulation of B cells, which may

depend upon cross-linking of multiple B cell receptors, may also be restricted (Liu et al., 2010). However recent data suggests that clustering of Env molecules occurs upon proteolytic cleavage of Gag (maturation), which may help improve infectivity (Chojnacki et al., 2012). Clustering of Env may represent a window of opportunity for nAbs to bind bivalently.

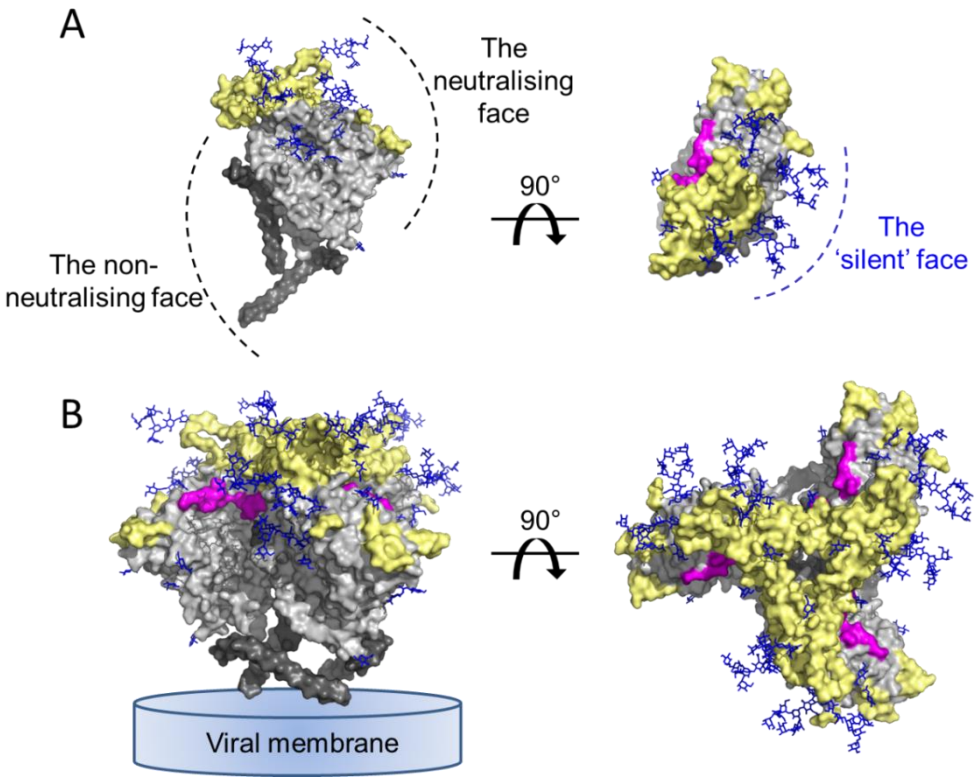


Fig. 1.9. The antigenic surface of HIV-1 Env. The surface of gp120/gp41 can be divided into different antigenic regions. The neutralising and non-neutralising faces are shown in the context of the monomer and trimer. Glycans cover much of the neutralising surface, creating an immunologically 'silent' face. Gp120s are in light grey, gp41 ecto domains are in dark grey. Gp120 variable regions are coloured in yellow. The CD4 binding site is coloured in magenta. Glycans are depicted as blue sticks. Images were generated in PyMol and were based on PDB ID 4NCO (Julien et al., 2013a).

The extensive protection of HIV-1 Env obstructs the production of effective immune responses, and the majority of HIV-1 infected individuals produce restricted, strain-specific (autologous) antibody responses. The first nAbs appear within a few months of infection, typically targeting the V1/V2 regions of Env, and drive viral evolution through changes to

amino acid sequences and the positions of glycosylation sites. While nAb responses increase in potency over time, only approximately 25% of individuals develop potent antibodies capable of neutralising a broad range HIV-1 strains. Why only a subset of individuals develop broadly neutralising antibodies (bnAbs) is unknown, but such responses typically require several years and are associated with chronic antigen exposure (Overbaugh and Morris, 2012). The isolation of potent bnAbs that target HIV-1 glycans is discussed in section 1.4.3.

1.4 Targeting HIV-1 glycans for drug and vaccine design

The envelope spike is critical for HIV-1 function; it must mediate cell infection while trying to evade host immune responses. The glycans which coat the spike are pivotal in achieving this, and thus represent an attractive target for drug and vaccine design.

1.4.1 Conservation and characterisation of Env glycans

The glycans of HIV-1 fulfil a diverse array of essential functions across the virus lifecycle. This functional dependence imposes constraints on the extent to which glycosylation sites can be deleted or shifted; indeed it has been shown that despite years of viral evolution, the average number of glycosylation sites remains relatively constant (Zhang et al., 2004). The glycosylation sites of gp120 display varying degrees of conservation, suggesting that greater functional constraints may exist for some sites (Fig. 1.10). Broadly speaking, the most highly conserved glycosylation sites lie within the inner domain of gp120, spanning the C1, V1/V2, C2 and V3 regions.

Site-specific glycosylation analyses have helped define the locations of complex- and oligomannose-type glycans on gp120 (Cutalo et al., 2004; Go et al., 2008, 2009, 2011, 2014; Leonard et al., 1990; Pabst et al., 2012; Zhu et al., 2000). On recombinant monomeric gp120 these analyses have suggested that the sites within the C1 and variable regions are

occupied by complex-type glycans, while sites on the outer domain are predominantly occupied by oligomannose-type glycans. The glycosylation pattern of functional envelope trimers is less well characterized, as complete site-specific mass spectrometric analysis of an Env trimer is currently lacking. However as described earlier, global glycan analysis of gp120 isolated from PBMC-derived virus revealed an elevation in the abundance of oligomannose-type glycans compared to monomeric gp120, indicating that additional steric constraints that exist within a trimer context could lead to further underprocessing of the glycans (Bonomelli et al., 2011; Doores et al., 2010a). This suggests that some sites that are designated as 'complex' type on the monomer will likely contain oligomannose-type glycans on the trimer.

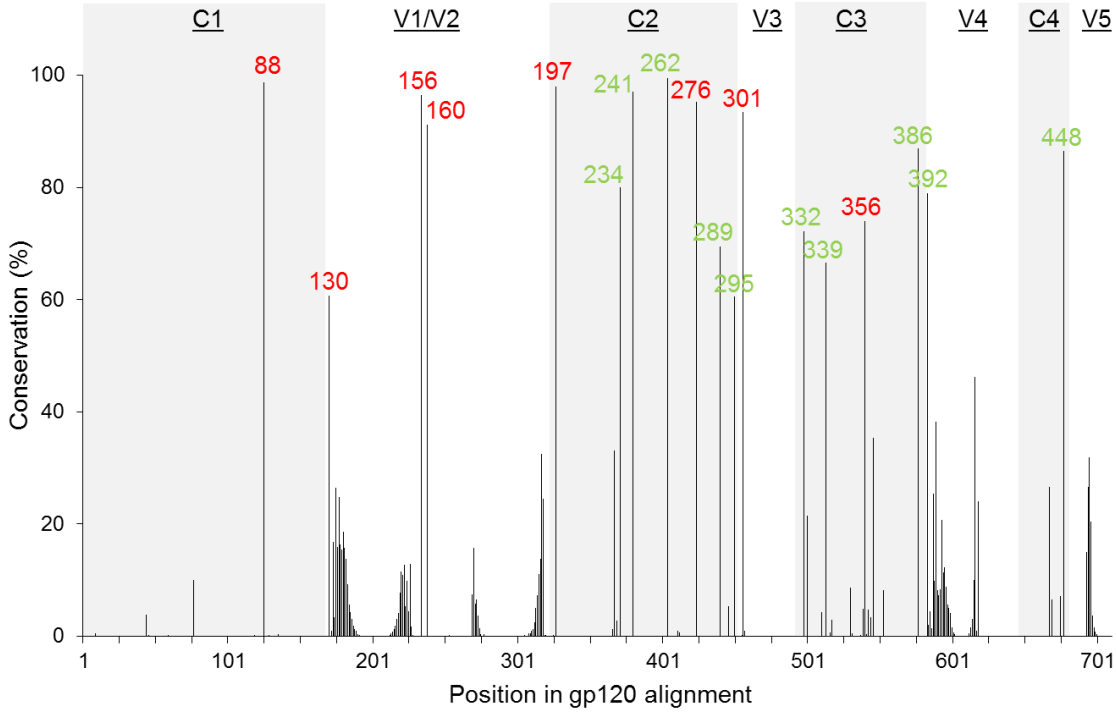


Fig. 1.10. Conservation of gp120 glycosylation sites. Percentage conservation of the individual potential N-linked glycosylation sites (PNGSs) was based on analysis of over 4000 aligned Env sequences from the Los Alamos HIV sequence database (http://www.hiv.lanl.gov/). Sites displaying greater than 50% conservation are indicated, and are allocated as putative complex-type (red) or oligomannose-type (green) based on previous reports (Leonard et al., 1990; Cutalo et al., 2004; Pabst et al., 2012; Go et al., 2013).

While the median number of N-linked glycosylation sites within gp120 is 24 (Korber et al., 2001), gp41 contains on average just 4 potential N-glycosylation sites, with the sites at N611, N616, N625 and N637 each exhibiting a high degree of conservation (>90%). Site-specific glycosylation analyses of gp41 glycans derived from recombinant gp140 constructs have provided conflicting results, and it is not clear whether all sites are occupied by complex-type glycans, or if the sites at N625 and N637 contain predominantly oligomannose-type glycans (Go et al., 2014; Guttman et al., 2014; Pabst et al., 2012). These analyses also suggest that compared to gp120 glycosylation sites, which display a high degree of occupancy, some gp41 PNGSs may be unoccupied or partially occupied (Go et al., 2014; Pabst et al., 2012).

HIV-1 Env also contains O-linked glycans, however the technical challenges associated with isolating them (Mariño et al., 2010) mean they have been poorly characterized in comparison with N-linked glycans (Bernstein et al., 1994; Hansen et al., 1992). The first signs of O-linked glycosylation derived from studies with monoclonal antibodies specific for O-linked glycans; antibodies targeting the Tn antigen (a N-acetylgalactosamine monosaccharide attached to serine or threonine) or sialyl-Tn antigen could neutralise multiple strains of HIV-1 and HIV-2 and prevent syncytium formation *in vitro* (Hansen et al., 1991, 1990). Predictions of the number of O-linked glycosylation sites vary, ranging between 1-8 sites for HIV-1, and are dependent on strain (Bernstein et al., 1994; Hansen et al., 1998). Recent site-specific glycosylation analysis of recombinant Env molecules expressed in CHO or 293T cells were found to contain O-linked glycans only at the T499 site in the C5 region (Go et al., 2013, 2014). This site was also found to be occupied with Gal-GalNAc glycans on gp120 derived from virus grown in a T-cell line (Yang et al., 2014). Evolution of serine and threonine residues, predicted to harbour O-linked glycans, within the V1 domain of SIV gp120 were attributed to viral escape from

neutralising antibodies (Chackerian et al., 1997), however it remains unclear if such a role exists for O-linked glycans on HIV-1.

1.4.2 Functions of HIV-1 glycans

The functions of HIV-1 glycans can broadly be separated into three main categories: protein folding and stability; protection from neutralising antibody responses; and interaction with cellular receptors. Each are discussed below.

Glycans in protein folding and stabilisation

As described in section 1.3.1, Env glycans are important for folding within the ER as they mediate interactions with folding chaperones via the CNX/CRT cycle (Fig. 1.5). Such a role for achieving proper conformation is supported by observations that non-glycosylated gp120, produced by treatment with tunicamycin, fails to bind CD4, whereas deglycosylated gp120, produced by enzymatic treatment, retains binding capability (Li et al., 1993). It is difficult to delineate which glycans may directly mediate interactions with CNX/CRT, however it would be expected that such a critical role would translate to a high degree of conservation (Fig. 1.10). Insights can also be obtained from mutagenesis studies which examine the effect of glycan site deletions on infectivity, although infectivity incorporates a number of requirements other than productive folding.

A recent study systematically dissected the role of each glycosylation site of a clade FE isolate in viral infectivity, highlighting a number of sites within the V1/V2 and C2 domains that made major contributions (Wang et al., 2013). These included N156, N160, N197 and N262, which are all over 90% conserved (Fig. 1.10). However another study on a clade B isolate revealed infectivity was decreased upon deletion of the N332, N386 and N262 sites, while in another isolate only loss of the N262 glycan had an effect (Lavine et al., 2012). Infectivity depends upon a number of factors, and the differences in results could

reflect isolate-specific factors. However the common requirement for the N262 is consistent with a number of other studies which suggest this glycan is critical for effective Env function (François and Balzarini, 2011; Sterjovski et al., 2007; Willey et al., 1988). Loss of N262 dramatically reduced gp120 incorporation into virions (François and Balzarini, 2011; Wang et al., 2013), which could arise from misfolding of Env preventing its efficient secretion.

Loss of certain glycans has also been found to influence the tropism of HIV-1 through modulation of local protein conformation. For example loss of the N197 glycan was found to be sufficient to mediate CD4-independent infection through changes to the V1/V2 loop conformation (Kolchinsky et al., 2001). Similarly a study in SIV found that loss of the SIV N173 glycan, which corresponds to N160 on HIV-1 gp120, facilitated CD4-independent cell-to-cell transmission between macrophages (Yen et al., 2014). Such influences of protein conformation also likely explain observed roles in determining co-receptor tropism. For example loss of the N301 glycan is associated with a change from CCR5 to CXCR4 tropism (Clevestig et al., 2006; Pollakis et al., 2001; Tsuchiya et al., 2013). V1/V2 glycans have also been demonstrated to allosterically modulate the gp120-gp41 association site (Drummer et al., 2013), while computational modelling was recently employed to analyse the effect of glycosylation on V3 dynamics (Wood et al., 2013).

Protective role of glycans

The role of glycans in protecting against neutralising antibody responses has been well documented (Reitter et al., 1998; Wei et al., 2003). Transmitter-founder viruses typically have lower numbers of glycosylation sites, with the differences mostly mapping to the V1/V2 region sites (Chohan et al., 2005; Derdeyn et al., 2004; Li et al., 2006; Liu et al., 2008b). This suggests that fewer glycosylation sites in the V1/V2 region may confer an

advantage to viral fitness, however with the onset of neutralising antibody responses a balance must be reached between maintaining infectivity and protecting vulnerable epitopes.

Protection of the vulnerable CD4 binding site is particularly important and several glycans have been identified which fulfil this role, mostly mapping to regions immediately adjacent to the binding site (Fig. 1.11). N197, and the equivalent glycan on SIV gp120, have been shown to modulate CD4 accessibility to the binding site as well as sensitivity to CD4-binding site antibodies (Pikora et al., 2005; Utachee et al., 2010). The proximal glycans at the N386 and N276 sites have also been demonstrated to play a critical role in conferring resistance to CD4-binding site antibodies (Duenas-Decamp et al., 2008; Koch et al., 2003; Sanders et al., 2008).

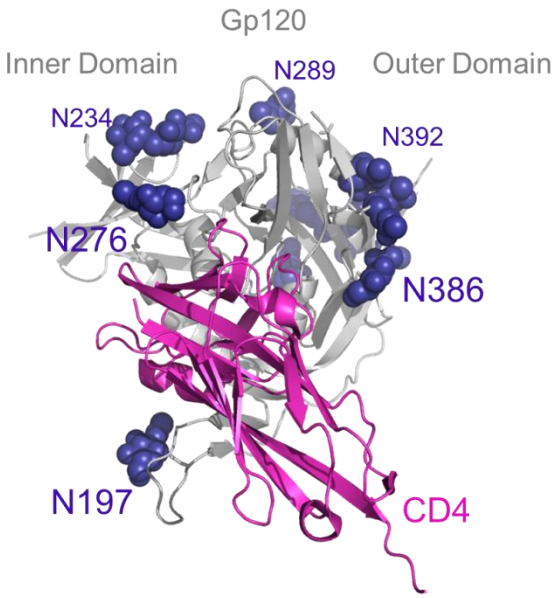


Fig. 1.11. N-linked glycans proximal to the CD4-binding site help protect against CD4-directed neutralising antibodies. Gp120 is shown in grey and CD4 in magenta. Glycans are blue spheres. Image was made in PyMol using PDB ID: 1GC1 (Kwong et al., 1998).

The V3 N301 glycan has also been shown to play an important role in shielding both the immunogenic V3 loop and the CD4 binding site from neutralising antibodies (Back et al., 1994; Koch et al., 2003; Schønning et al., 1996). Since the N301 glycan is not directly adjacent to the CD4 binding site its mechanism of protection is not clear. While it could

feasibly mediate the orientations by which nAbs can approach the binding site, an indirect role has also been proposed, whereby removal of the N301 glycan impacts the conformation and ability of the nearby V1/V2 region to effectively shield against neutralisation (Koch et al., 2003).

Mutagenesis studies of the four highly conserved glycosylation sites within gp41 have indicated that, depending on the strain examined, none are individually essential for infectivity (Johnson et al., 2001; Wang et al., 2013). While simultaneous removal of all four sites rendered gp41 incapable of membrane-fusion, perhaps indicating some functional role in folding, the high conservation of the glycans may instead reflect important roles in epitope shielding; indeed removal of the glycosylation sites increased recognition by gp120-specific antibodies (Fenouillet et al., 1993).

Interaction with cellular receptors

The glycans of HIV-1 interact with an array of cellular receptors and consequently play a significant role in mediating viral infectivity. One of the best characterised interactions occurs with the Dendritic Cell-Specific Intercellular adhesion molecule-3 (ICAM)-Grabbing Non-integrin (DC-SIGN) receptor on DCs in the mucosal tissues. Typically during infection DCs endocytose pathogens in order to present processed antigens on MHC class II molecules to $CD4^+$ T cells. However interaction of HIV-1 with DC-SIGN triggers the endocytosis of virions into non-lysosomal compartments, allowing them to remain intact as DCs migrate to secondary lymphoid tissues. Upon arrival they mediate *trans* infection of $CD4^+$ $CCR5^+$ T cells (Geijtenbeek et al., 2000a; Kwon et al., 2002). DC-SIGN is a C-type lectin with primary specificity for oligomannose-type glycans (Feinberg et al., 2001; Geijtenbeek et al., 2000b; Hong et al., 2002), although it can also bind blood group antigens via an alternative binding mode (Guo et al., 2004). The high affinity interaction for HIV-1 is

thought to derive from a multivalent interaction with clusters of oligomannose-type glycans on gp120, achievable due to the tetrameric nature of DC-SIGN (Mitchell et al., 2001). Interestingly, while the interaction of DC-SIGN with HIV-1 is mediated by oligomannose-type glycans, an increase in oligomannose content was demonstrated to lead to increased viral degradation and presentation to $CD4^+$ T cells, and reduced transmission to target cells, suggesting that there may be a compromise between initial uptake and downstream pathways (van Montfort et al., 2011). A number of other C-type lectin receptors also interact with oligomannose-residues on gp120 (Figdor et al., 2002), including Langerin found on the surface of Langerhans cells (section 1.2.3). Gp120 binding to Langerin triggers internalisation of virions into Birbeck granules (Stambach and Taylor, 2003; Valladeau et al., 2000). However the role of this interaction in mediating infection of $CD4^+$ T lymphocytes has been controversial, with differing opinions on whether Langerin either promotes or protects $CD4^+$ T lymphocytes from infection (Hladik and McElrath, 2008; de Witte et al., 2007). However more recent evidence has demonstrated a role for LCs in viral dissemination and infection of $CD4^+$ T lymphocytes following vaginal infection (Ballweber et al., 2011).

Gp120 binding to DC-SIGN and other C-type lectin receptors has consequences beyond internalisation of virions. For example it has been demonstrated that a mannose-dependent interaction with receptors on monocyte-derived DCs triggered the release of the immunosuppressive cytokine, IL-10, in a strain-dependent manner (Shan et al., 2007). Such an interaction may contribute to the poor immunogenicity of HIV-1 Env as a vaccine candidate, and indeed it has been shown that enzymatic removal of gp120 mannose residues, or occlusion using Griffithsin, improves the immunogenicity of recombinant gp120 (Banerjee et al., 2009, 2012). Further immunosuppression as a consequence of the DC-SIGN-gp120 interaction is thought to occur through destruction of DCs. Gp120 binding

sensitises DCs to undergo accelerated apoptosis in response to ligands such as CD40, lipopolysaccharide, TNF-α or IL-1β (Chen et al., 2013). Furthermore DC-SIGN may promote productive infection of DCs through stabilisation of the CD4-gp120 interaction (Hijazi et al., 2011).

It is not just oligomannose-type glycans on gp120 that modulate the infectivity of HIV-1. The precise role of sialic acid residues in particular remains to be fully understood. For example desialylation of gp120 with neuraminidase was observed to increase viral infectivity (Hu et al., 1996), however another recent study demonstrated that binding of gp120 to sialic acid-binding immunoglobulin-like lectin (SIGLEC)-1 facilitated infection of both CCR5- and CXCR4-tropic viruses in macrophages (Zou et al., 2011). Meanwhile binding of galactose residues on gp120 by galectin-1 has been shown to increase infectivity (Ouellet et al., 2005; St-Pierre et al., 2011). Galectin-1 is a soluble lectin that is abundant in the thymus and lymph nodes, and its binding to gp120 is thought to help mediate the gp120-CD4 interaction on target cells (Sato et al., 2012).

Further roles for HIV-1 glycans likely await discovery. For example a recent study analysed the glycan sites carried by viruses that are transmitted from mother-to-child, and found that transmitted viruses more commonly carried a glycan site at N339, but lacked the glycan site at N295 (Baan et al., 2012). This lack of a glycan at N295 has previously been noted in infected infants (Wolinsky et al., 1992), however the associated mechanism remains unknown. Evidence also suggests that Env glycans may play a role in modulating antigen presentation to $CD4^+$ T cells (Botarelli et al., 1991; Li et al., 2008).

1.4.3 The broadly neutralising antibody response to HIV-1 glycans

While the glycans coating HIV-1 Env were originally viewed as a glycan ‘shield’, protecting against host immune responses, in recent years a number of broadly neutralising antibodies

(bnAbs) have been isolated that directly target these shielding glycans (Chuang et al., 2013; Falkowska et al., 2014; Huang et al., 2014; Mouquet et al., 2012; Scheid et al., 2011; Walker et al., 2009a, 2011a; Wu et al., 2010). Indeed many of the emerging sites of vulnerability on the envelope spike involve gp120 and gp41 glycans (Fig. 1.12). Many of these antibodies display extraordinary breadth and potency, and have been demonstrated to protect against high-dose SHIV challenge in macaques (Moldt et al., 2012). Structural determination of their epitopes is currently in progress as part of a reverse engineering approach to immunogen design (Burton et al., 2012). These antibodies also display promise therapeutically; passive transfer of a combination of these bnAbs was shown to suppress viral load in macaques and humanised mouse models (Klein et al., 2012; Shingai et al., 2013). The glycan-dependent epitopes of these bnAbs are discussed below.

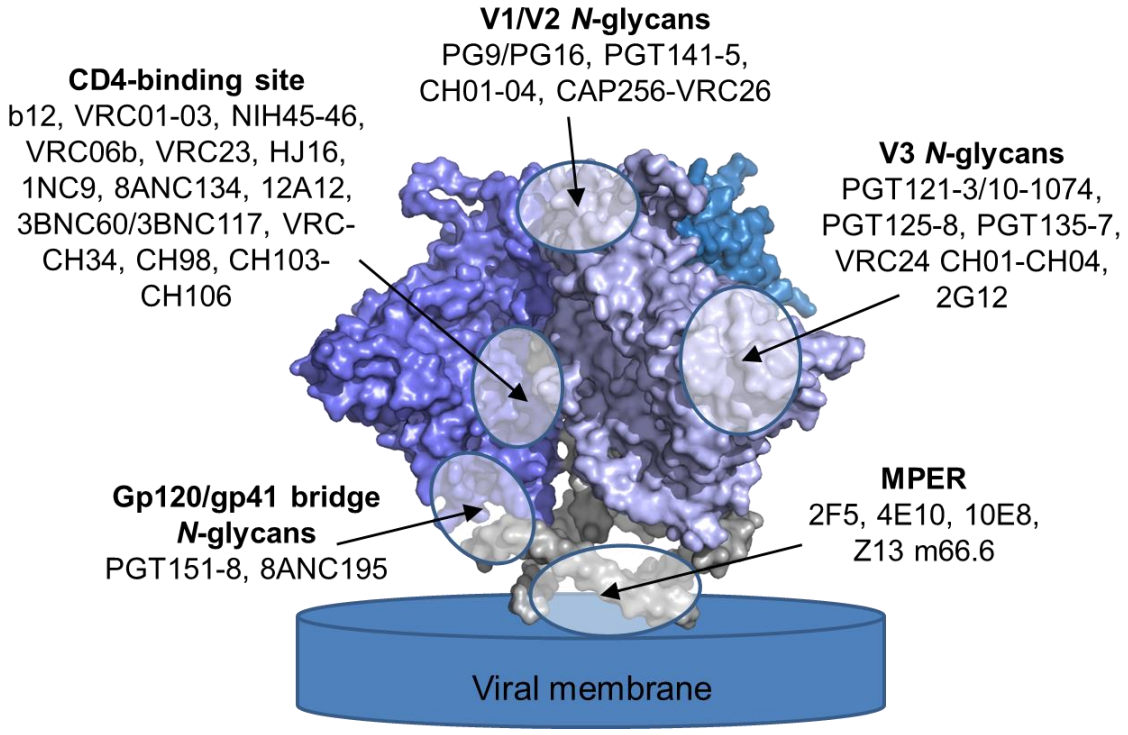


Fig. 1.12. Sites of vulnerability on the HIV-1 envelope spike. Many of the recently isolated bnAbs are targeted to epitopes involving N-linked glycans. Figure was made in PyMol using PDB ID: 4NCO and was adapted from (Mouquet, 2014).

2G12 – a carbohydrate-dependent epitope

The first antibody discovered to possess broadly neutralising activity against HIV-1 glycans was 2G12 (Binley et al., 2004; Sanders et al., 2002b; Scanlan et al., 2002; Trkola et al., 1996). 2G12 was isolated from a chronically infected individual and neutralised 41% of a 93-virus panel with an IC_{50} of <50 µg/ml (Binley et al., 2004). The involvement of glycans was first hypothesised following mutagenesis studies, which revealed the importance of the N295, N392, N332, N386 and N448 glycan sites for binding (Trkola et al., 1996). Treatment with a panel of glycosidases, and competition experiments with various monosaccharides, confirmed a glycan-dependent epitope involving α1-2-linked mannose residues (Sanders et al., 2002b; Scanlan et al., 2002).

Antibody-carbohydrate interactions are typically lower affinity compared to antibody-protein binding. The surprising potency of 2G12 was rationalised by a crystal structure of a 2G12 $(Fab)_2$ complex, which showed how 2G12 adopts an unusual domain-exchanged configuration, whereby the variable domains of the heavy chains swap over to form an extended antigen-binding surface, comprising the conventional V_L/V_H primary combining sites as well as a novel V_H/V_H' binding interface (Fig. 1.13). This configuration allows 2G12 to make multivalent interactions with the closely clustered oligomannose glycans on gp120, thereby increasing avidity without the entropic penalties associated with equivalent bivalent binding by typical IgG (Calarese et al., 2003). The importance of the domain-exchanged configuration in contributing avidity effects is demonstrated by observations that an equivalent Y-shaped (non-domain-exchanged) 2G12 IgG could bind monomeric gp120, presumably through its primary V_L/V_H binding sites, but could not neutralise HIV-1 pseudovirus (Doores et al., 2010b).

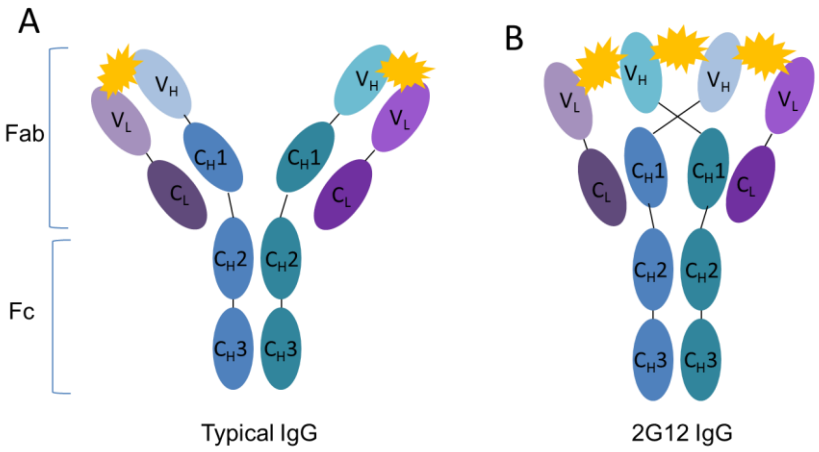


Fig. 1.13. Domain-exchanged conformation of 2G12. The arrangement of domains within a typical (A) and 2G12 (B) IgG are depicted schematically. Typical IgG forms two primary combining sites on independent arms, while the domain-exchange of the V_H domains forms an additional secondary binding interface. Heavy chains are shown in blue, light chains are shown in purple. Antigen binding sites are yellow stars. C – constant, V – variable, H – heavy, L – light. Adapted from (Wu et al., 2013).

The crystal structure of 2G12 in complex with $Man_9GlcNAc_2$ revealed the presence of four bound glycans; one at each of the primary combining sites and two at the V_H/V_H' site (Calarese et al., 2003). At the primary combining sites the vast majority of contacts (85%) were to the terminal Manα1-2Man residues of the D1 arm of $Man_9GlcNAc_2$, involving 12 hydrogen bonds and 48 van der Waals interactions, although it has been shown that 2G12 can also bind to Manα1-2Man residues on the D3 arm (Calarese et al., 2005). Additional interactions were also made with the innermost mannose residue of the D1 arm and the core mannose. The V_H/V_H' interface was observed to form a relatively shallow binding site, primarily contacting the terminal mannose of the D2 arm of $Man_9GlcNAc_2$, with eight to nine hydrogen bonds and 22-26 van der Waals interactions being made with the two $Man_9GlcNAc_2$ glycans at this site.

A negative-stain single particle EM reconstruction of 2G12 $(Fab)_2$ in complex with a BG505 SOSIP.664 trimer, combined with modelling using the high resolution crystal structures (Calarese et al., 2003; Julien et al., 2013a), enabled identification of the likely binding sites of 2G12 on the viral surface (Murin et al., 2014). The EM reconstruction

showed three 2G12 $(Fab)_2$, one bound to each protomer, with one arm of the Fab binding closer to the apex trimer and the second Fab binding lower down on the gp120 outer domain. This arrangement suggested that the primary combining site of the upper Fab was likely occupied by the N392 glycan, with the lower Fab contacting the glycan at N295. Meanwhile the V_H/V_H' binding site appeared to contact the glycans at N332 and N339. These results broadly correspond to mutagenesis studies that have highlighted critical glycan sites for 2G12 binding (Murin et al., 2014; Scanlan et al., 2002; Trkola et al., 1996), although it was previously thought that the N332 glycan was a primary binding site (Calarese et al., 2003). Implicated glycans that are not observed to directly contact 2G12, for example at the N386 and N448 sites, are thought to play indirect, strain-dependent roles in binding.

Numerous studies have attempted to elicit 2G12-like bnAbs via immunisation with constructs bearing α1-2-linked mannose residues in different immunogenic contexts (Agrawal-Gamse et al., 2011; Doores et al., 2010c; Dunlop et al., 2010; Kong et al., 2010; Luallen et al., 2008; Ni et al., 2006). While many have resulted in elicitation of nAbs that are capable of binding monomeric gp120, none have elicited nAbs that contain a domain-exchanged configuration or are capable of effectively neutralising virus. Given the requirement of domain-exchange for effective neutralisation (Calarese et al., 2003; Doores et al., 2010b), successful immunisation strategies will likely require targeted stimulation to drive domain-exchange, which requires surprisingly few mutations from germline precursors (Huber et al., 2010).

As a target for vaccine design 2G12 is limited in scope by its dependence on the N295 glycan site, which is lacking in most clade C isolates (Binley et al., 2004; Zhang et al., 2004). While other bnAbs have now been described which offer broader and more potent protection against HIV-1, 2G12 still offers promise, both in prophylactic (Baba et al., 2000;

Hessell et al., 2009; Mascola et al., 2000) and therapeutic contexts (Mehandru et al., 2007; Trkola et al., 2005). Furthermore a dimeric form of 2G12 has recently been described which displays greater potency and mediates stronger ADCC responses compared to the monomeric form (Klein et al., 2010; West et al., 2009). This is presumably through increased valency of binding translating to increased avidity effects, although it is unclear if the additional interactions are within the same trimer or across neighbouring Env molecules (Murin et al., 2014; Wu et al., 2013).

Dual protein-glycan dependent bnAbs

Advances in methods for screening memory B cells from infected HIV-1 individuals have led to the isolation of a new collection of bnAbs, displaying significantly broader and more potent activity against HIV-1 (Falkowska et al., 2014; Scheid et al., 2011; Walker et al., 2009a, 2011a; Wu et al., 2010). While some of these antibodies recognise conventional protein epitopes, in particular the CD4 binding site (Scheid et al., 2011; Wu et al., 2010), the majority of them have been demonstrated to include both protein and glycan components within their epitopes (Blattner et al., 2014; Falkowska et al., 2014; Garces et al., 2014; Kong et al., 2013; Mouquet et al., 2012; Pancera et al., 2013; Pejchal et al., 2011; Walker et al., 2009a; Wu et al., 2010).

As more of these antibodies are isolated and characterised, a growing number of sites of vulnerability within the glycan shield are being identified (Table 1.12). The epitopes of several antibodies, including PGT121-3, PGT124, PGT125-8, PGT130-1 and PGT135-7, overlap with the 2G12 footprint on the outer domain and are focused around the glycan at N332 – thus referred to as a ‘supersite’ of vulnerability (Kong et al., 2013; Walker et al., 2011). Glycan array studies with PGT128 revealed a similar glycan specificity to 2G12 for $Man_9GlcNAc_2$ and $Man_8GlcNAc_2$ glycans containing a complete D1 arm. However a crystal

structure of PGT128 in complex with $Man_9GlcNAc_2$ revealed subtle differences in their modes of recognition; while 2G12 makes the majority of its contacts with the terminal Manα1-2Man residues of the D1 arm, burying 220 $Å^2$ of the terminal disaccharide, PGT128 buries 394 $Å^2$ of the $Man_9GlcNAc_2$ by reaching along the length of the glycan. A crystal structure of PGT128 in complex with an engineered outer domain (eODmV3) revealed that in the context of gp120 the primary binding site is N332. A secondary binding interaction is also made with the N301 glycan. In contrast to recognition at the N332 site, the main interactions are with the core pentasaccharide of the N301 glycan, which is common to all N-linked glycan structures, and thus there is likely limited glycan specificity at this secondary binding site. The remaining component of the PGT128 epitope comprises the C-terminal end of the V3 loop; residues I323 and D325 makes hydrogen bonds and van der Waals interactions with the CDR H3 loop of PGT128 (Pejchal et al., 2011).

PGT135 has also been revealed to display a preference for oligomannose-type glycans. Glycan array studies showed a preference for $Man_{7-9}GlcNAc_2$ glycans, and pseudoviruses produced in HEK 293S $GnTI^{-/-}$ cells (and thus all contained only oligomannose-type glycans) were neutralized more potently (Kong et al., 2013). A crystal structure of PGT135 in complex with gp120 revealed the primary binding site to be the glycan at N392, with 546 $Å^2$ of the occupying $Man_8GlcNAc_2$ glycan buried. The N332 glycan occupied the secondary binding site, with 365 $Å^2$ of its surface buried, while a much smaller interaction was also observed with the chitobiose core of the N386 glycan (100 $Å^2$ buried surface area) (Kong et al., 2013).

In contrast to the strict oligomannose specificity of 2G12, PGT128 and PGT135, the binding of PGT121 appears to be more promiscuous. Glycan array experiments demonstrated binding to complex-type glycans but not oligomannose-type glycans, however PGT121 could still neutralise pseudovirus expressed in HEK 293S $GnTI^{-/-}$ cells or in the

presence of kifunensine (Kif) (Julien et al., 2013b; Mouquet et al., 2012; Sok et al., 2014). Crystal structures have been obtained demonstrating binding of PGT121 to a sialylated, biantennary complex-type glycan (Julien et al., 2013b; Mouquet et al., 2012), however mutagenesis studies suggest the N332 glycan is the primary binding site (Mouquet et al., 2012; Walker et al., 2011), which is thought to be occupied by oligomannose-type glycans (Leonard et al., 1990; Pabst et al., 2012; Zhu et al., 2000). The apparently conflicting results might be explained by a recent study that demonstrated the promiscuity of some N332-dependent bnAbs in the absence of N332 (Sok et al., 2014). It was observed that in viruses lacking the N332 site neutralisation could still occur if the N136/N137 (on the V1/V2 loop) remained, however if both sites were removed neutralisation was abolished. While viruses containing wholly oligomannose-type glycans could be neutralised in the presence of the N332 glycan, neutralisation didn't occur in the presence of N136/N137 only. These results are consistent with a promiscuous mode of binding by PGT121, whereby sufficient affinity can be achieved either via an oligomannose-type glycan at the N332 site, or via a complex-type glycan at the N136/N137 site. Roles for both of these glycans were supported by an EM reconstruction of the somatically related PGT122 antibody in complex with the BG505 SOSIP.664 trimer, which also implicated minor roles for further neighbouring glycans (Julien et al., 2013a; Sok et al., 2014). Higher-resolution structures of PGT121 in complex with gp120 are needed to structurally rationalise these apparent glycan specificities.

Glycans of the V1/V2 region also form the epitope for the quaternary-dependent somatic variants, PG9 and PG16 (Doores and Burton, 2010; McLellan et al., 2011; Pancera et al., 2010a, 2013; Walker et al., 2009a). Mutagenesis studies consistently identified a critical role for the glycan at the N160 site, and a variable requirement for N156 or N173 glycans depending on strain (Doores and Burton, 2010; Walker et al., 2009a). Neutralisation was found to be abolished for pseudovirus produced in the presence of Kif, and reduced for

pseudovirus produced in HEK 293S GnTI$^{-/-}$ cells. In contrast neutralisation, particularly for PG16, slightly increased for pseudovirus produced in the presence of swainsonine, which leads to the production of hybrid-type glycans (Doores and Burton, 2010). High-resolution crystal structures of PG9 and PG16 in complex with scaffolded V1/V2 regions help explain these results (McLellan et al., 2011; Pancera et al., 2013). PG9 was observed to make extensive contacts with a $Man_5GlcNAc_2$ glycan at the N160 site (1,150 Å^2 buried), primarily with the protein-proximal GlcNAc residue and the two outermost Man residues. Inspection of the binding pocket revealed that larger glycan structures would clash with the light chain of PG9, while smaller structures would not stretch across the binding site (McLellan et al., 2011). Significant interactions were also observed with a $Man_5GlcNAc_2$ glycan at the N156 site, helping to form an epitope that is 70% glycan-dependent. In the PG16-V1/V2 structure a sialylated hybrid-type glycan was observed at the secondary binding site, in this case N173. Interactions were identified with the terminal sialic acid residue, a dependence that was supported by reduced neutralisation against neuraminidase-treated virus (Pancera et al., 2013). These glycan preferences, for a $Man_5GlcNAc_2$ at the N160 site and a sialylated glycan at the secondary site (N156 or N173), were supported by studies with synthetic glycopeptides (Amin et al., 2013). The PGT141-145 family of antibodies share many of the hallmarks of PG9 and PG16, including a dependence on the N160 glycan and a lack of neutralisation against Kif-treated pseudovirus, suggesting they may bind with similar modes and glycan specificities as PG9 and PG16. The CH01-CH04 and CAP256-VRC26.01–12 lineage of antibodies also display a dependence on the N160 residue, however their glycan dependence and specificity have not been fully elucidated (Bonsignori et al., 2011; Doria-Rose et al., 2014).

Table 1.2. Sites of bnAb vulnerability within the HIV-1 glycan shield.

Region	Antibody	Breadth[§] %	Median IC_{50} μg/ml[†]	Glycan epitope	Predicted glycan specificity	Refs
V1/V2	**PG9**	79 (162)	0.22	N160 N156 N173	Man5/ Hybrid	(Julien et al., 2013c; McLellan et al., 2011; Walker et al., 2009)
	PG16	73 (162)	0.15	N160 N156 N173	Man5/ Hybrid	(Pancera et al., 2013; Walker et al., 2009)
	PGT145 PGT141-5	78 (162)	0.29	N160	Nd	(Walker et al., 2011)
Outer domain	**2G12**	32 (162)	2.38	N392, N295, N332, N339	Mannose	(Calarese et al., 2003; Murin et al., 2014)
	PGT121 PGT121-4 10-1074	70 (162)	0.03	N332, N136	Mannose/ Complex	(Garces et al., 2014; Julien et al., 2013b; Mouquet et al., 2012; Pancera et al., 2014; Sok et al., 2014; Walker et al., 2011)
	PGT128 PGT125-8 PGT130-1	72 (162)	0.02	N332, N301, N295	Mannose	(Pejchal et al., 2011; Walker et al., 2011)
	PGT135 PGT135-7	33 (162)	0.17	N392, N332, N295, N386	Mannose	(Kong et al., 2013; Walker et al., 2011)
gp120/gp41 interface	**35O22**	62 (182)	0.033	N88, N230, N241, N625	Mannose	(Huang et al., 2014; Pancera et al., 2014)
	PGT151 PGT151-8	66% (117)	0.008	N611 N637	Complex	(Blattner et al., 2014; Falkowska et al., 2014)
	8ANC195	70% (118)	3.7	N276 N234 N637	Mannose	(Chuang et al., 2013; Scharf et al., 2014; West et al., 2013)

[§]Percentage of viruses neutralised with an IC_{50}<50 μg/ml. Number in brackets indicates size of virus panel.

[†]Median IC_{50} of viruses neutralised with an IC_{50}<50 μg/ml.

Prototypic antibodies from each related antibody family are highlighted in bold.

Nd – not determined

More recently new glycan-dependent epitopes have been identified on and around gp41. The PGT151-158 family of antibodies are cleavage-dependent, binding to a novel

epitope on gp41 (Blattner et al., 2014; Falkowska et al., 2014). Alanine scanning mutagenesis identified the N611 and N637 sites as being critical for binding. Neutralisation of glycan-modified pseudoviruses indicated an intolerance of oligomannose-type glycans, while glycan array data revealed a preference for tetra- and, to a lesser extent, tri-antennary complex-type glycans. Sialic acid was tolerated only when adjoined with an α2,3-linkage and not an α2,6-linkage (Blattner et al., 2014; Falkowska et al., 2014). An EM reconstruction of PGT151 in complex with a cell-surface membrane-associated JR-FL trimer revealed the PGT151 epitope spans the gp120-gp41 interface, explaining its cleavage-dependence (Blattner et al., 2014). The modelling also suggested possible roles for nearby gp120 glycans, including at the N448 and N262 sites, however this has yet to be confirmed by mutagenesis or higher resolution structural characterisation.

The epitopes for two further unrelated antibodies also appear to span the gp41/gp120 interface. Binding by the bnAb 35O22 was found to be dependent on N88, N230, N241 and N625A by mutagenesis, and a role for these glycans was supported by an EM reconstruction and modelling of 35O22 in complex with a BG505 SOSIP.664 trimer (Huang et al., 2014). Meanwhile the bnAb 8ANC195 is dependent upon the glycans at N276 and N234 (Chuang et al., 2013; West et al., 2013), with EM modelling also suggesting a possible role for N637 (Scharf et al., 2014). Neutralisation for both antibodies was increased when glycan processing on virions was restricted to oligomannose-type glycans, and a high-resolution structure of 8ANC195 in complex with a gp120 core revealed an extensive interaction with a $Man_8GlcNAc_2$ glycan (1,645 $Å^2$ buried) at the N234 site (Scharf et al., 2014). Meanwhile interactions with the N276 glycan were primarily made to the core $Man_3GlcNAc_2$ pentasaccharide, suggesting there could be a greater degree of binding tolerance at this site.

1.4.4 Antiviral carbohydrate-binding agents

HIV-1 glycans are also the target of various lectins and other carbohydrate-binding agents (CBAs) isolated from plants or prokaryotes. Investigations are being carried out into their potential preventative or therapeutic uses as a number have been found to display anti-HIV-1 activity (Balzarini, 2006). In particular such agents could have the potential to effectively be employed as topical microbicides (Koharudin and Gronenborn, 2014).

The majority of anti-HIV CBAs that have been characterised to date are specific for mannose residues in one or more linkages. One of the first CBAs to be characterised was cyanovirin-N (CV-N), a small 11 kDa protein comprising two highly similar domains, A and B. Isolated from the cyanobacterium *Nostoc ellipsosporum*, CV-N was found to inhibit HIV-1 replication *in vitro* at low nanomolar concentrations (Boyd et al., 1997a). Viruses treated with CV-N lose infectivity, with CV-N blocking both CD4 and co-receptor binding (Dey et al., 2000; Esser et al., 1999; Mori and Boyd, 2001). Analysis of the CV-N structure revealed two binding sites for glycans, with a specificity for $Man_8GlcNAc_2$ and $Man_9GlcNac_2$ structures (Bewley and Otero-Quintero, 2001; Bolmstedt et al., 2001). The natural quaternary nature of CV-N remains uncertain (Fig. 1.14); crystal structures capture a domain-exchanged dimer (Botos et al., 2002), however this is unstable at physiological conditions and monomeric CV-N has been characterised in solution (Bewley, 2001). Artificial stabilisation of dimeric CV-N configurations has been demonstrated to confer increased potency, and thus potentially offers greater therapeutic value (Keeffe et al., 2011; Kelley et al., 2002). In a challenge study, topical application of CV-N protected 15 of 18 macaques from vaginal challenge with pathogenic SHIV89.6P, while 8 of 8 placebo-treated macaques were infected (Tsai et al., 2004). Protection was similarly demonstrated in the context of an intrarectal challenge (Tsai et al., 2003).

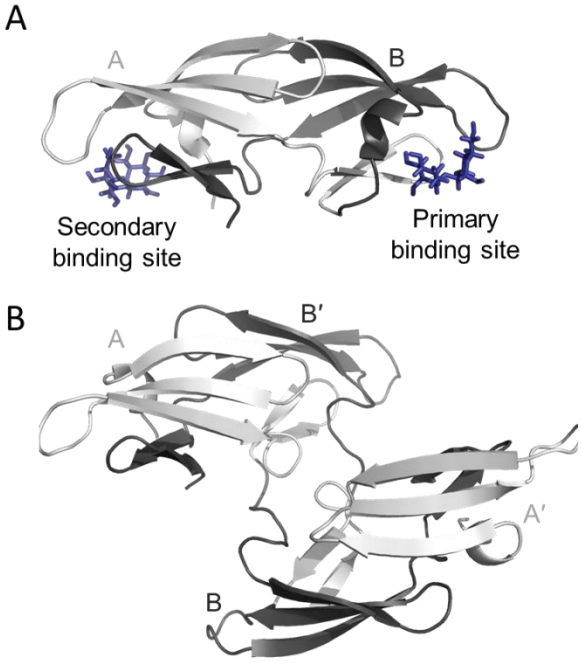


Fig. 1.14. Organisation of monomeric and dimeric CV-N. (A) NMR solution structure of monomeric CV-N with a Manα(1-2)Man disaccharide (blue sticks) occupying each of the binding sites. (B) NMR solution structure of a domain-swapped dimer. The 'A' domains are coloured in light grey and the 'B' domains in dark grey. Image was made in PyMol using PDB ID: 1IIY and 1J4V.

A number of other mannose-specific lectins have also been found to possess *in vitro* anti-HIV activity (Balzarini, 2006). These include *Galanthus nivalis* agglutininin (GNA), *Listera ovate* agglutinin (LOA), *Hippeastrum* hybrid agglutinin (HHA), *Microcystis viridis* lectin (MVL), *Scytonema varium* lectin (SVN), Actinohivin lectin (AH) (Hoorelbeke et al., 2010), *Oscillatoria agardhii* agglutinin (OAA) (Férir et al., 2014), BanLec (isolated from *Musa acuminata* bananas) (Swanson et al., 2010), and *Griffithsia* sp. lectin (GRFT) (Mori et al., 2005). While all show specificity for mannose-linked residues, their preferences for different branch arms and residue linkages differ (Fig. 1.15) (Koharudin and Gronenborn, 2014). Of the lectins characterised to date, GRFT has shown remarkable promise as a topical microbicide; it inactivates HIV-1 at picomolar concentrations, can be synthesised in a cost-effective manner, and lacks the mitogenic activity associated with other lectins like CV-N (Mori et al., 2005; O'Keefe et al., 2009).

Studies aimed at determining the resistance profile of these lectins frequently reveal that deletion of multiple glycosylation sites (typically those of predicted oligomannose, within the outer domain) are required for HIV-1 to obtain resistance, suggesting that the lectins may possess multiple modes of binding to gp120 glycans, offering a higher genetic barrier to resistance (Alexandre et al., 2010, 2013; Balzarini et al., 2004, 2005, 2006; Hu et al., 2007; Huang et al., 2011b; Witvrouw et al., 2005). Given the functional role of glycans in hiding neutralising epitopes beneath, it has been proposed that one therapeutic strategy could involve treatment with CBAs, which would force the deletion of shielding glycans and thereby increase the likelihood of recognition by neutralising antibodies (Balzarini, 2005).

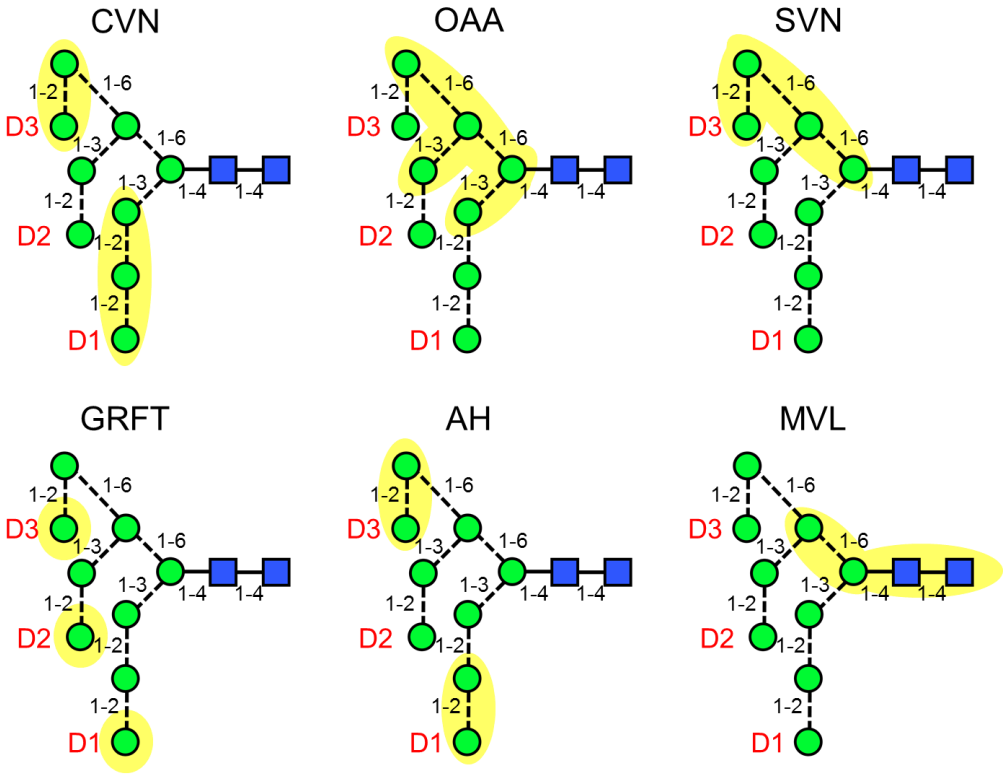


Fig. 1.15. Carbohydrate specificities of mannose-specific lectins. Schematic of $Man_9GlcNAc_2$ is shown with mannose residues as green circles and N-acetylglucosamine residues as blue squares. The three arms are labelled D1, D2 and D3. Residues recognised by mannose-specific lectins are highlighted in yellow (Koharudin and Gronenborn, 2014). CVN, cyanovirin-N; OAA, *Oscillatoria agardhii* agglutinin; SVN, *Scytonema varium* lectin; GRFT, *Griffithsia* sp. lectin; AH, Actinovirin lectin; MVL, *Microcystis viridis* lectin.

1.4.5 Inhibitors of glycosylation

A number of glycosylation inhibitors have been explored for their potential use as antiretroviral agents against HIV-1. Given the functional dependence of gp120 on its glycosylation, it is perhaps unsurprising that inhibitors of the ER glucosidases such as castanospermine and 1-deoxynojirimycin (DNJ), which interfere with the early events of glycan processing, demonstrate anti-HIV activity (Gruters et al., 1987; Tyms et al., 1987). However in contrast deoxymannojirimycin (DMJ), which inhibits Golgi α-mannosidase I, does not possess antiviral activity (Gruters et al., 1987). Searches for more potent derivatives of these natural plant alkaloids led to the discovery of N-butyl-deoxynojirimycin (NB-DNJ), which displays potent activity at non-cytotoxic concentrations (Fleet et al., 1988; Karpas et al., 1988; Ratner, 1992). However side-effects relating from inhibition of gut glycosidases have limited their development in humans. Enhanced delivery using ER-targeted liposomes, as well as the synthesis of more potent derivatives, is currently being explored as a route to reducing the required effective dose, thereby minimising unwanted side effects (Miller et al., 2012; Pollock et al., 2008).

The precise mechanism of action of NB-DNJ and related imino sugars remains unclear. Treatment dramatically alters the glycosylation profile of gp120 (Harvey et al., 2011; Karlsson et al., 1993), potentially affecting receptor binding, although it was demonstrated that NB-DNJ-treated gp120 could undergo CD4-binding, albeit not fusion (Karlsson et al., 1993; Papandréou et al., 2002). Screening with a panel of conformation-dependent antibodies revealed disturbances to the V1/V2 region, which were suggested to explain the observed inhibition of entry (Fischer et al., 1996). Despite the current unknowns regarding the precise antiviral mechanism, an antiretroviral strategy that involves targeting a host protein rather than HIV-1 itself is attractive as it is associated with a high genetic

barrier to resistance. Indeed no resistance mutations were identified following a 15-week *in vitro* culture of HIV-1 in the presence of NB-DNJ (Pollock et al., 2008).

An alternative therapeutic approach that has been proposed involves combining glycosidase inhibitors, to engineer particular glycoforms, with the use of carbohydrate-binding agents. For example combining treatment with 1-deoxymannojirimycin, which stalls glycan processing at the oligomannose stage, and mannose-specific CBAs has been shown to produce a synergistic antiviral effect allowing more potent neutralisation (Balzarini, 2007; Hart et al., 2003). Neuraminidase treatment to remove sialic acid was also found to improve neutralisation by mannose-binding lectin (MBL) (Hart et al., 2003).

1.5 Architecture of the HIV-1 glycan shield

The glycans of HIV-1 play critical roles in the viral lifecycle. These functional roles demand their preservation, limiting the degree to which they can be mutated and offering a window of vulnerability for targeting the virus. The antiviral activity of several glycosylation inhibitors and carbohydrate-binding agents has been known for a while, and more recently the isolation of numerous glycan-reactive bnAbs has suggested HIV-1 glycans could be a promising target for vaccine design. It is therefore necessary to understand the factors that determine the architecture of the HIV-1 glycan shield, and to understand how it may be subject to sequence mutation. This is true in particular for the population of oligomannose-type glycans, which are thought to arise from the high density of glycans, and thus may be sensitive to glycosylation site mutation.

Chapter three explores the variation of the glycan population across different HIV-1 isolates, assessing the relative abundances of oligomannose- and complex-type glycans

across different clades, in an attempt to dissect how the presence or absence of different glycosylation sites may influence glycan processing.

Chapter four describes the development of a method for quantitative site-specific glycosylation analysis of recombinant gp120. This was used to demonstrate the variable degree of glycan microheterogeneity found at individual glycosylation sites across different locations on a model gp120.

Chapter five extends the analysis of the role of individual glycosylation sites in glycan processing, by using alanine scanning to determine the contribution of each glycosylation site to the formation of oligomannose-type glycans. The method for site-specific glycosylation analysis was then utilised to explore the effects of individual glycosylation site-deletions on glycan processing at a site-specific level.

In chapter six the ability of a panel of N332-dependent bnAbs to tolerate glycosylation site-deletions was examined. Site-specific glycosylation analysis of one of these antibodies, PGT135, was performed to investigate how glycan microheterogeneity may provide a mechanism of viral immune escape from glycan-reactive bnAbs.

Finally in chapter 7, the glycan profiles of different candidate immunogens are analysed, providing insights into the glycan processing of trimeric Env.

2 Materials and Methods

2.1 Materials

Enzymes for cloning and glycan digests were purchased from New England Biolabs (Ipswich, MA, USA), unless otherwise specified. Tissue Culture reagents were purchased from Invitrogen™ (Paisley, UK). Solvents were from VWR International and were HPLC grade. Chemicals were from Sigma unless otherwise specified. Water was MilliQ grade.

The pHLsec vector carrying $gp120_{BaL}$ was originally obtained from Dr Guillaume Stewart-Jones (University of Oxford). The following plasmids were obtained from the laboratory of Dennis Burton (Scripps, CA, USA): pcDNA vector carrying $gp120_{JRCSF}$, $gp120_{92TH021}$, $gp120_{92RW020}$ and $gp120_{BG505}$; pCMV vector carrying $gp120_{C22}$, $gp120_{94UG103}$. All other gp120 envelope plasmids were provided by Dr Katie Doores (King's College London).

Primers were synthesised by Thermofisher (Waltham, MA, USA). AmpliTaq Gold PCR Master Mix was obtained from Applied Biosystems (Carlsbad, CA, USA). EndoFree Plasmid Maxi and Mega kits were purchased from Qiagen (Crawley, UK), and NucleoSpin Plasmid and NucleoSpin Gel and PCR clean-up kits were from Macherey-Nagel (Düren, Germany). DNA 1 kB ladder and loading buffer were from Bioline (London). GelRed was obtained from Cambridge Biosciences (Cambridge, UK). DEPC-treated water was from Ambion®. XL1-Blue competent cells were from Stratagene (La Jolla, CA, USA). DNA sequencing was performed by Source Bioscience (Department of Biochemistry, University of Oxford).

Pre-cast NuPAGE 4-12% Bis-Tris polyacrylamide gels, NuPAGE loading buffer, 3-morpholinoethane-sulphonic acid (MES) buffer, HRP-conjugated anti-HIS antibody and

BenchMark His-tagged protein standard were purchased from Invitrogen™. Precision Plus Kaleidoscope Standard and Trans-Blot Turbo Mini PVDF Transfer Packs were obtained from Biorad (Hemel Hempstead, UK). Vivapsin 10- and 50-kDa cut-off concentrators were purchased from Sartorius (Goettingen, Germany). HisTrap FF crude purification columns were from GE Healthcare (Waukesha, WI, USA).

Spe-ed Amide-2 and empty SPE cartridges were obtained from Applied Separations (Allentown, PA, USA). 2-AB labelling kits, stock ammonium formate buffer (pH 4.4) and 2-AA labelled glucose homopolymer ladder were from Ludger Ltd (Culham, UK).

2.2 Cloning of gp120 constructs

2.2.1 Plasmid constructs

Gp120 and gp160 open reading frames were inserted into the pHLsec vector (Ariescu citation) using the Age1 (3') and Kpn1 (5') restriction sites. Sequences were amplified by PCR using primers containing the above restriction sites (for sequences, see Appendix I). PCR reactions were analysed by electrophoresis with a 1% agarose gel, and target PCR products were recovered. Restriction digestion was then performed according to manufacturer's instructions. Where necessary, BsiW1 was used as an alternative to Kpn1 at the 5' site. The empty pHLsec vector was prepared by restriction digestion with Age1 and Kpn1, followed by treatment with Antarctic phosphatase for 45 min at 37ºC. Where the BsiW1 site was introduced into the insert, the vector was propagated in methylation-deficient E. coli cells and digested with Age1 and Acc651. PCR products were ligated into the vector using a ratio of 7:1 (v:v) using T4 DNA ligase, followed by transformation into XL-1 Blue competent cells, according to manufacturer's instructions. Cultures were plated on to LB agar plates containing carbenicillin (50 µg / ml), and incubated at 37ºC overnight. Individual colonies were isolated, cultured and sequenced.

2.2.2 Site-directed mutagenesis

Site-directed mutagenesis was achieved using a two-step overlap PCR method. Firstly, two separate PCR reactions were performed: one using the 3' flanking primer and the 5' primer bearing the mutation, and a second using the 5' flanking primer and the 3' primer bearing the mutation. PCR products were analysed by electrophoresis with a 1% agarose gel, and target PCR products recovered. Secondly, PCR products from the two reactions were then mixed with the two flanking primers to produce full length inserts containing the mutation. Following electrophoresis and gel extraction, inserts were digested with the appropriate restriction enzymes, ligated into pHLsec vector and transformed in XL-1 Blue cells as described above.

2.3 Recombinant gp120 expression

Recombinant gp120s were expressed transiently from Human Embryonic Kidney (HEK) 293T cells via transfection with polyethylenimine (PEI). HEK 293T cells were cultured in DMEM, supplemented with 10% FCS, 100 units/ml penicillin, 100 µg/ml streptomycin and 2 mM glutamine, and maintained at 37ºC and 5% CO_2. For transfection, for every 50 ml of culture media, 65 µg DNA was incubated with 225 µg PEI in 15 ml of serum-free DMEM for 15 min, with mixing. Spent culture media was removed from seeded cells and replaced with 35 ml of fresh, serum-free media. The transfection mixture was then added, and antibiotics added to the final concentrations stated above. Cells were incubated for 3 days, following which culture media was harvested, clarified by centrifugation (300 *g* for 5 min) and sterile filtered through a 0.22 µM membrane.

2.4 Glycoprotein purification

2.4.1 Nickel-affinity purification

A HisTrap HP column was used. The column was pre-equilibrated with buffer A (20 ml; 50 mM sodium phosphate, 0.3 M NaCl, pH 7.5), followed by application of culture media, which was adjusted to pH 7.5 and contained 20 mM imidazole. The column was then washed with buffer B (20 ml; 50 mM sodium phosphate, 20 mM imidazole, 0.3 M NaCl, pH 7.5) and the gp120 eluted with buffer C (50 mM sodium phosphate, 0.3 M imidazole, 0.3 M NaCl, pH 7.5). Eluted fractions were pooled, concentrated using 50 kDa cut-off centrifugal filters, and buffer exchanged into phosphate buffered saline (PBS). Protein concentration was determined using a Nanodrop™ 1000 spectrophotometer (Thermo Scientific).

2.4.2 Size exclusion chromatography

For site-specific glycosylation analysis, monomeric gp120 was purified by size exclusion chromatography. His-purified gp120 was concentrated to 1 ml or less, and injected onto a Superdex 200 Prep grade column (GE healthcare), pre-equilibrated with PBS. Collected fractions were analysed by SDS-PAGE, and those containing monomeric gp120 were pooled and concentrated. Protein concentration was determined using a Nanodrop™ 1000 spectrophotometer.

2.5 SDS-PAGE and Western blotting

Gp120 samples (approximately 5-10 µg monomer) were analysed by SDS-PAGE. Samples were mixed with SDS loading buffer and incubated for 10 min at 70°C, before loading on to a 4-12% Bis-tris NuPage gel. Electrophoresis was carried out in MES buffer for 1 h at 180 V. Gels were then either stained with coomassie brilliant blue, or blotted on to PVDF membrane for immunostaining. Blotting was carried out at 1.3A / 25V for 10 min using a Trans Turbo Blotter (Biorad). Membranes were blocked in 5% dry skimmed milk (with

0.05% Tween 20) for 1 h at room temperature or overnight at 4°C, and then probed with HRP-conjugated anti-HIS antibody (1/10,000 dilution in PBS/1% milk). Following washing, detection was carried out using ECL substrate, according to manufacturer's instructions. A Fujifilm Las-1000 Intelligent DarkBoxII system was used for visualisation by chemiluminescence.

2.6 Enzymatic glycan digestion

2.6.1 In-gel Peptide N-glycosidase F (PNGase F) digestion of glycoproteins

Coomassie-stained gel bands were excised and destained with 200 mM ammonium bicarbonate in 40% acetonitrile, before drying under vacuum. Gel pieces were rehydrated in 50 mM sodium phosphate buffer, pH 7.5, and incubated with PNGase F (1 µl) for 16 h at 37°C. Released glycans were extracted from the gel matrix by 3x washing steps with water.

2.6.2 Glycosidase digestions of free glycans

The abundance of oligomannose glycans was measured by digestion with Endoglycosidase H (Endo H). Each glycan sample was digested with Endo H in duplicate. 2-AA labelled glycans were resuspended in water and digested with Endo H in a total volume of 20 µl, for 16 h at 37°C, according to manufacturer's instructions. Digested glycans were purified using a PVDF protein-binding membrane plate prior to HILIC-UPLC analysis. Exoglycosidase digestions and clean-up for glycan sequencing were carried out in the same way using the following enzymes: Neuraminidase from *Clostridium perfringens* (P0720), α-L-Fucosidase from Bovine Kidney (F5884), β1,4-Galactosidase from *Streptococcus pneumoniae* (E-BG07, QA Bio), β-N-Acetylglucosaminidase from *Streptococcus pneumonia* (E-GL01, QA Bio) and α(1-2,3,6)-mannosidase from Jack Bean (E-AM01, QA Bio).

2.7 Fluorescent Labelling of glycans

PNGase F-released oligosaccharides were fluorescently labelled with 2-aminobenzoic acid (2-AA). The labelling mixture comprised 2-AA (30 mg/ml) and sodium cyanoborohydride (45 mg/ml) dissolved in a solution of sodium acetate trihydrate (4%, w/v) and boric acid (2%, w/v) in methanol. The labelling mixture (80 µl) was added to each sample, which was then vortexed and incubated at 80°C for 1 h. Following cooling, labelled oligosaccharides were purified using Spe-ed Amide 2 columns, pre-equilibrated with acetonitrile (2 ml). Before loading, 97% acetonitrile (v/v) (1 ml) was added to each sample. Loaded samples were then washed with 95% acetonitrile (v/v) (2 ml) and eluted with water (1.5 ml).

2.8 Hydrophilic interaction liquid chromatography (HILIC)-ultra-performance liquid chromatography (UPLC)

Fluorescently labelled glycans were separated by HILIC-UPLC using a 2.1 mm × 10 mm Acquity BEH glycan column (1.7 µm particle size) (Waters, Elstree, UK) using a Waters ACQUITY UPLC® instrument. The following gradient was run: time = 0 min (*t*=0): 22% A, 78% B (flow rate of 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; *t*=46.5: 22% A, 78% B (0.5 ml/min), where solvent A was 50 mM ammonium formate, pH 4.4, and solvent B was acetonitrile. Fluorescence was measured using an excitation wavelength of 250 nm and a detection wavelength of 428 nm. Glucose units were calculated from a standard curve generated using a glucose homopolymer external calibration standard. Data processing was performed using Empower 3 software.

2.9 Glycopeptide analysis

Glycopeptide analysis was performed in collaboration with Ludger Ltd (Abingdon, UK).

2.9.1 Protease digestions

For site analysis of Ni^{2+}-NTA/SEC purified $gp120_{BaL}$ in chapters four and six, protease digestions were performed in solution. For site analysis of Ni^{2+}-NTA-only purified monomeric $gp120_{BaL}$ wild-type and mutant N332 glycopeptides in chapter five, the trypsin digestion was performed in-gel, following resolution by SDS-PAGE. In both cases reduction was carried out by addition of 10 mM dithiothreitol in 100 mM ammonium bicarbonate and incubation at 65°C for 30 min. Alkylation was subsequently performed by addition of 50 mM iodoacetamide in 100 mM ammonium bicarbonate, with incubation at room temperature for 50 min. Trypsin (Promega) was then added at a concentration of 12.5 µg/ml and incubated at 37°C for 16 h. Alternatively chymotrypsin was added at a concentration of 1 µg/ml and incubated at 25°C for 16 h. Glycopeptides were eluted from the in-gel digestions with alternate washes of water and acetonitrile and filtered through a 0.45 µM membrane (Whatman). Glycopeptide samples were dried in a SpeedVac concentrator.

2.9.2 RP-HPLC fractionation of glycopeptides

After resuspension in 0.1% TFA, the glycopeptide pool was fractionated by reverse phase-HPLC using a Jupiter C18 5 µm 250 × 4.5 mm column (300 Å pore size) (Phenomenex, Macclesfield, UK) and a Dionex U3000 LC system. The following gradient was run at a flow rate of 1 ml/min, with fractions being collected every minute for 90 min: time = 0 min (*t*=0): 95% A, 5% B; *t*= 5: 95% A, 5% B; *t*=90: 10% A, 90% B; *t*=95: 10% A, 90% B; *t*=97: 95% A, 5% B; *t*= 120: 95% A, 5% B, where solvent A was 50 mM ammonium formate, pH 4.4, and solvent B was 50% acetonitrile. UV absorbance was detected at 214 and 280 nm.

2.9.3 MALDI-MS

Glycopeptide fractions were desalted using C18 ZipTip® pipette tips, according to manufacturer's instructions, before spotting onto a ground steel target plate in a solution of

10 mg/ml 2,5-Dihydrozybenzoic acid (DHB) in 50% acetonitrile. MALDI-MS was performed using an Autoflex™ Speed MALDI-TOF/TOF instrument (Bruker), operated in positive ion mode. MS/MS was performed to confirm the peptide identity.

2.9.4 LC-ESI-MS/MS of deglycosylated peptides

Deglycosylated peptide fractions were diluted in 0.1% TFA and approximately 1 pmol loaded onto a Thermo PepMap RSLC C18 column (75 µm × 150 mm, 2 µm particle size, 100 Å pore size). A gradient from 2% to 55% acetonitrile was run over 18 min at a flow rate of 350 µl/min, with 0.1% formic acid as the aqueous solvent. ESI-MS/MS was performed using an AmaZon Speed ETD instrument (Bruker), operated in positive ion mode. ETD was performed with methane gas. Maximum ETD precursor was 1500 m/z and maximum accumulation time 10 ms. For CID amplification of fragmentation was from 80 to 150%. 3 precursor ions would be selected for MS/MS and released after 0.2 min. For general MS the instrument ICC Target was set at 200,000 with a maximum accumulation time of 50 ms. The target mass was set to 1200. Data processing was performed using Bruker DataAnalysis software.

2.9.5 HILIC-UPLC of glycans

Glycans were fluorescently labelled with 2-AB using a LudgerTag™ 2-AB labelling kit and purified using LudgerClean T1 cartridges (Ludger Ltd), according to manufacturer's instructions. Fluorescently labelled glycans were separated by HILIC-UPLC using a 2.1 mm × 150 mm ACQUITY UPLC® BEH-Glycan column (1.7 µm particle size) (Waters, Elstree, UK) at 40°C using a Waters ACQUITY UPLC® H class instrument. The following gradient was run at a flow rate of 0.4 ml/min: time = 0 min (*t*=0): 28% A, 72% B; *t*= 54: 48% A, 52% B; *t*=57: 100% A, 0% B; *t*=58: 100% A, 0% B; *t*=59: 28% A, 72% B; *t*= 60: 28% A, 72% B, where solvent A was 50 mM ammonium formate, pH 4.4, and solvent B was acetonitrile.

Fluorescence was measured using an excitation wavelength of 330 nm and a detection wavelength of 420 nm. Data processing was performed using Empower 2 software.

2.10 ELISAs

Enzymes-linked immunosorbent assays (ELISAs) were performed in collaboration with Dr Katie Doores (King's College, London). Microtitre ELISA plates (Corning) were coated with mouse anti-HIS capture antibody (2 µg/mL in PBS) (4A12E4, Life Technologies) overnight at 4°C. Plates were washed five times with a solution of PBS containing 0.05% Tween 20 (v/v) and then blocked for 1 h at room temperature with 5% non-fat milk in PBS + 0.05% Tween (blocking buffer). Gp120 was then added at a concentration of 2 µg/ml in blocking buffer, followed by incubation for 1 h at room temperature. Plates were washed (×5) before addition of primary antibodies (b12, b17, 2G12, PGT121, PGT128 & PGT135). A 1:5 dilution series was used starting at 20 µg/ml or 40 µg/ml in blocking buffer. Incubation was carried out at room temperature for 2 h. Plates were then washed (×5) and alkaline phosphatase-conjugated goat anti-human Fab secondary antibody (Thermo Scientific Pierce) was added as a 1:1000 dilution in blocking buffer. Plates were washed (×5) and then AP substrate (50 µL / well) was added. Once colour had developed the OD at 405 nm was measured.

2.11 Structural modelling of glycosylated Env

Structural modelling was performed by Dan Kulp and Sergey Menis from Professor William Schief's lab at the Scripps Research Institute, CA. A model of the BaL gp120 trimer was constructed using the atomic-level resolution structure of a BG505 SOSIP.664 gp140 trimer (PDB id: 4NCO) (Julien et al., 2013a). First, atomic clashes present in the crystal structure were relieved and missing side-chains rebuilt, by executing 1,000 symmetric ROSETTA-fixbb simulations, selecting the lowest scoring model, and then running a constrained

ROSETTA-relax simulation. Next, the BaL sequence was modelled onto the gp120 trimer structure, by running 1,000 ROSETTA-remodel trajectories (Huang et al., 2011a), which included building new conformations for the variable loops, and then filtering the models by structural metrics. The best 15 models were considered as an ensemble of possible BaL gp120 trimer structures, as each model was in the top 33% of the 3 structural metrics Ramachandran (rama), attractive portion of the Lennard-Jones potential (fa_atr) and the total ROSETTA score. Each N-linked glycosylation motif for each of the 15 models was decorated with $Man_8GlcNAc_2$ glycans at sites of predicted oligomannose-type glycans and with $Man_5GlcNAc_2$ glycans at remaining sites. GlycanRelax (Pancera et al., 2010b) was used to approximate the conformational behaviour of glycans in a glycoprotein context. For each model, 10 separate GlycanRelax trajectories of 10,000 cycles of MonteCarlo trials were carried out. Each glycan on the gp120 was allowed to move independently throughout the GlycanRelax minimization. A single low energy model was then selected.

2.12 Neutralisation assays

Neutralisation assays were performed by Dr Katie Doores. To produce pseudoviruses, plasmids encoding Env were co-transfected with an Env-deficient genomic backbone plasmid (pSG3ΔEnv) in a 1:2 ratio with the transfection reagent PEI (1 mg/mL, 1:3 PEI:total DNA, Polysciences) into HEK 293T cells. Glycosylation site deletions in the full-length envelope were made using site-directed mutagenesis and the sequences were verified by DNA sequencing. Pseudoviruses were harvested 72 h post transfection for use in neutralisation assays (Montefiori, 2005). Neutralizing activity was assessed using a single round replication pseudovirus assay with TZM-bl target cells, as described previously (Walker et al., 2011). Briefly, TZM-bl cells were seeded in a 96-well flat bottom plate at a concentration of 20,000 cells/well. The serially diluted antibody/virus mixture, which was

pre-incubated for 1 h, was then added to the cells and luminescence was quantified after 72 h following infection via lysis and addition of Bright-Glo™ Luciferase substrate (Promega). To determine IC_{50} values, serial dilutions of mAbs were incubated with wild-type virus or mutant virus and the dose-response curves were fitted using nonlinear regression.

2.13 Production of trimeric constructs

Soluble cleaved BG505 SOSIP.664 gp140 trimers, soluble uncleaved BG505 WT.SEKS gp140 trimers, and soluble uncleaved CZA97.012 gp140 trimers were provided by Dr Albert Cupo and Professor John Moore (Cornell University, New York). BG505 SOSIP.664 gp140 trimers were expressed in stable Flp-In™ 293T and CHO cells, and purified by 2G12-affinity chromatography, followed by SEC, as described previously (Chung et al., 2014). Trimers were alternatively purified by PGT145 or PGT151 in the same manner. BG505 WT.SEKS gp140 trimers were expressed in 293T cells and 2G12/SEC-purified as described previously (Ringe et al., 2013). CZA97.012 gp140 trimers were expressed in 293T cells and Ni^{2+}-NTA/SEC purified as described previously (Kovacs et al., 2012). Membrane-bound BG505.705 trimers, produced in 293F cells and purified by PGT151 affinity and SEC as described previously, (Blattner et al., 2014), were provided by Dr Andrew Ward (The Scripps Research Institute, CA).

3 Oligomannose-type glycans are a conserved feature of HIV-1 gp120

HIV-1 gp120 is unique in the extent of its glycosylation, with approximately half of its mass being attributed to a dense coat of N-linked glycans (Lasky et al., 1986). Previous studies have revealed that HIV-1 gp120 contains processed complex-type glycans, as well as a population of under-processed oligomannose-type glycans, referred to as the 'intrinsic mannose patch' (Bonomelli et al., 2011; Doores et al., 2010a; Go et al., 2008; Leonard et al., 1990; Zhu et al., 2000). Oligomannose-type glycans are rarely found on secreted mammalian glycoproteins, and their presence on gp120 has been attributed to the unusually high density of glycosylation sites that cover the surface of the molecule. This density is thought to sterically impede access by glycosidases and glycosyltransferases, leading to under-processing of the glycans and a divergence from host-cell glycosylation (Bonomelli et al., 2011; Doores et al., 2010a). This divergence offers scope for the immune system to target these 'non-self' glycan epitopes, and the isolation of glycan-reactive bnAbs highlights the vulnerability of this region of HIV-1 (Blattner et al., 2014; Falkowska et al., 2014; Huang et al., 2014; Sanders et al., 2002b; Scanlan et al., 2002; Scharf et al., 2014; Walker et al., 2009, 2011a). Furthermore a number of mannose-specific carbohydrate-binding agents have been demonstrated to possess antiviral activity (Alexandre et al., 2010; Boyd et al., 1997b).

The success of any vaccine or antiviral strategy depends upon the prevalence of the target in nature. In this chapter, the glycan profiles of a cross-clade panel of 26 gp120s were quantitatively characterised to determine the prevalence of oligomannose-type glycans. A population of oligomannose-type glycans was observed in every gp120 tested, although subtle differences in the size and nature of the glycans could be identified.

3.1 Expression and purification of recombinant HIV-1 gp120s

A cross-clade pane of 26 recombinant gp120s was assembled, covering clade A (*n = 7*), clade B (*n = 8*), clade C (*n = 5*), CRF01_AE (*n = 3*) and clade G (*n = 3*). Gp160 sequences were obtained from Dr Katie Doores (King's College, London), and open reading frames corresponding to gp120 were cloned into the pHLsec expression vector (Aricescu et al., 2006). The cloning of $gp120_{BaL}$ has previously been described (Dunlop et al., 2010). Gp120s were expressed transiently in HEK 293T cells, and were purified by immobilised metal affinity chromatography to avoid any glycan-biasing that could occur with lectin affinity purification.

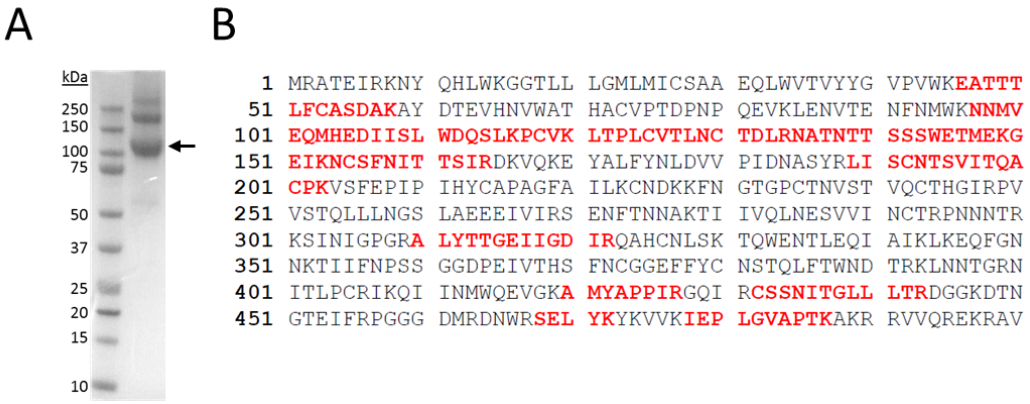


Fig. 3.1. Expression of recombinant, monomeric gp120. (A) SDS-PAGE analysis of nickel affinity-purified $gp120_{BaL}$ under non-reducing conditions. The band corresponding to the monomer is indicated by the arrow. (B) Results of peptide mass fingerprinting: the gel band corresponding to the gp120 monomer was excised and digested with PNGase F and trypsin, before analysis by LC-MS/MS. A query of the data verified the gel band to contain gp120 – the peptide coverage is shown in red. The relatively low coverage is likely caused by a mismatch between the BaL and query sequences. LC-MS/MS was performed by Dr Holger Kramer.

When analysed by SDS-PAGE under non-reducing conditions, the recombinant gp120s were observed to run in both monomeric and higher-order forms (Fig. 3.1A shows $gp120_{BaL}$ as an example). The higher molecular weight species were confirmed to be gp120 by Western Blot, and could be collapsed down to the monomeric band when treated with DTT (data not shown), suggesting they are disulphide-linked aggregates of gp120 subunits.

All glycan analysis was performed on the monomeric species, excised directly from the SDS-PAGE gel. To verify the identity of the band, and to check for any contaminants, the gel band corresponding to the monomer was subjected to an in-gel tryptic digest (Shevchenko et al., 2006), and the peptide fragments were then analysed by LC-MS/MS. A query of the data using the MASCOT search engine confirmed the band to contain gp120 and no other contaminating glycoproteins (Fig. 3.1B).

3.2 HILIC-UPLC analysis of gp120 glycans

N-linked glycans were released directly from gel bands using PNGase F. Following fluorescent labelling by reductive amination with 2-aminobenzoic acid (2-AA), the glycans were analysed by hydrophilic interaction liquid chromatography-ultra performance liquid chromatography (HILIC-UPLC). Profiles from a representative of each clade are shown in Fig. 3.2. Glycan profiles from all strains tested, and their protein sequences, are given in Appendix II, dataset A.

Several features were evident upon comparison of the glycan profiles. Firstly, there was a considerable degree of heterogeneity within each strain tested. Each peak corresponds to at least one unique glycan structure, and in some profiles approximately 50 different peaks could be distinguished. As some peaks contained multiple species, the overall heterogeneity is even higher. Secondly, a substantial population of oligomannose-type glycans could be identified in all profiles.

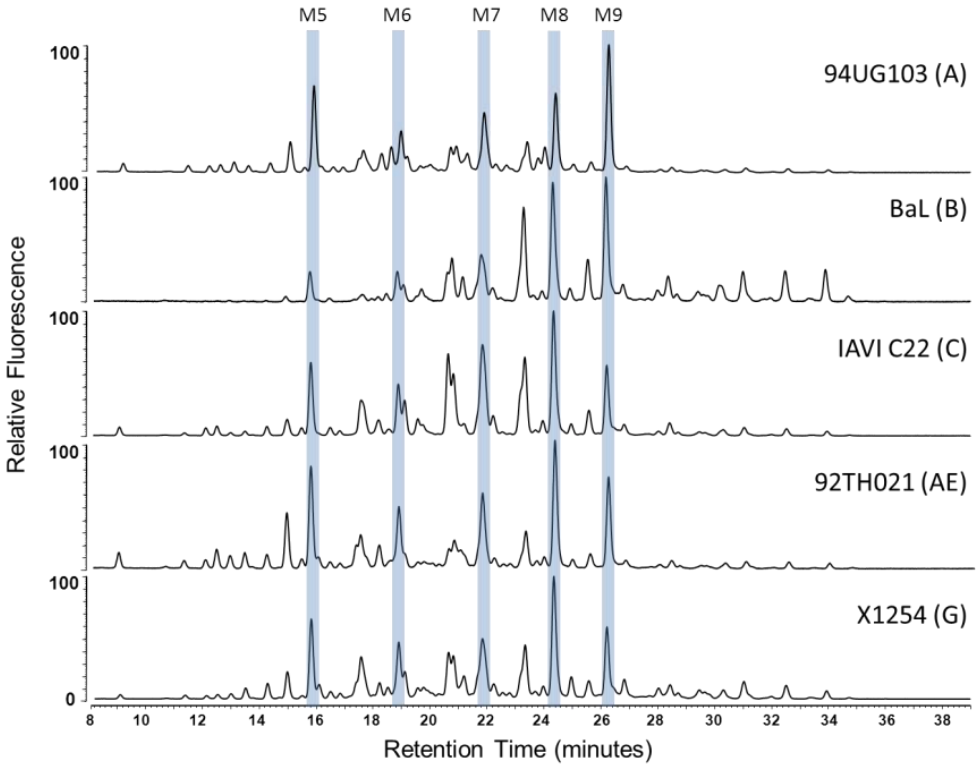


Fig. 3.2. HILIC-UPLC analysis of recombinant gp120. N-linked glycans were released from recombinant gp120s in-gel using PNGase F, labelled with 2-AA and resolved by HILIC-UPLC. Profiles are shown for representative isolates from different clades. Peaks corresponding to oligomannose-type glycans are highlighted in blue: M5, $Man_5GlcNAc_2$; M6, $Man_6GlcNAc_2$; M7, $Man_7GlcNAc_2$; M8, $Man_8GlcNAc_2$; M9, $Man_9GlcNAc_2$.

3.3 Endoglycosidase H digestion of glycan profiles

To accurately quantitate the abundances of the individual oligomannose-type glycans, endoglycosidase H (Endo H) was used, which selectively cleaves oligomannose-type (and some hybrid) glycans (Trimble and Tarentino, 1991). Total pools of 2-AA-labelled PNGase F-released glycans were digested with Endo H to remove oligomannose-type glycans from the spectra (Fig. 3.3). Following normalisation of the two profiles using the later-eluting sialylated structures (approximately 27 min onwards), which are untouched by Endo H, subtraction of the digested profile from the undigested profile enabled calculation of the abundances of individual oligomannose species (Fig. 3.4).

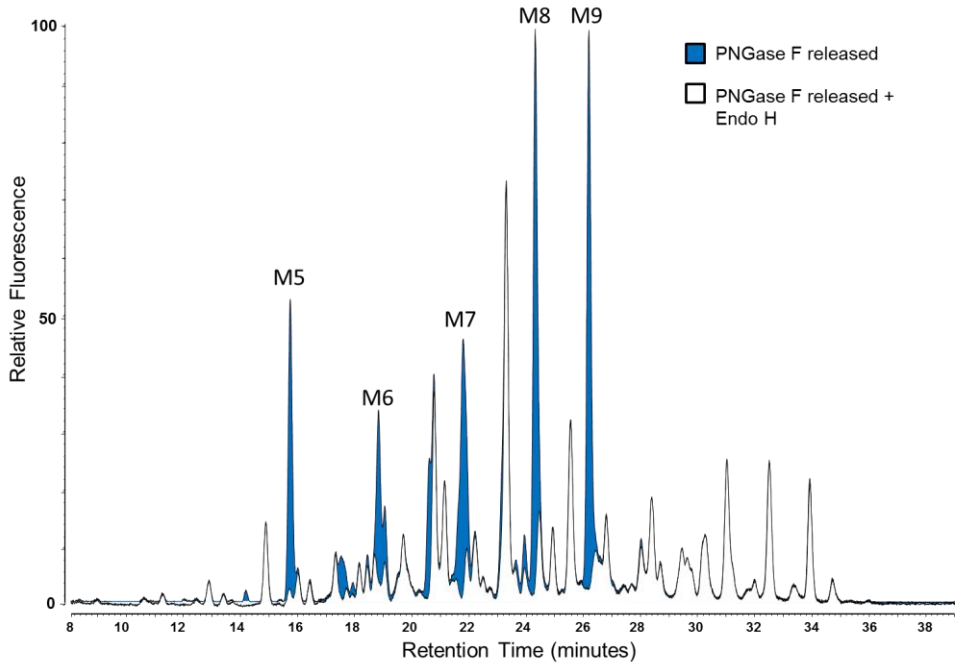


Fig. 3.3. Endoglycosidase H digestion of gp120 glycans. To quantitate the abundance of oligomannose-type glycans, total PNGase F-released glycan pools were digested with Endo H in duplicate. Corresponding peaks were integrated before and after digestion. The HILIC-UPLC profile of a PNGase F-released total (undigested) glycan pool (for $gp120_{BaL}$) is coloured in blue, and is overlaid with the Endo H-digested profile in white.

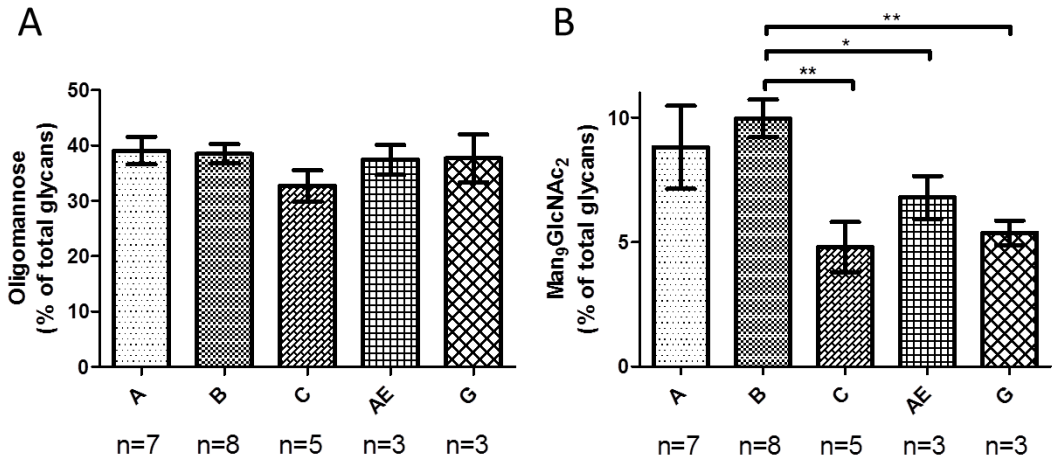


Fig. 3.4. Abundances of oligomannose-type glycans within a cross-clade panel of gp120s. Abundances of $Man_{5\text{-}9}GlcNAc_2$ (A), and $Man_9GlcNAc_2$ only (B), were measured from a cross-clade panel of recombinant gp120s by digestion with Endo H (see text). The number of isolates analysed per clade (n) is indicated. Asterisks represent p values in an unpaired two-tailed Student's t-test: *$p<0.05$, **$p<0.01$. Only significant ($p<0.05$) results are indicated. Raw data is listed in Appendix II, dataset B.

A full breakdown of the abundances of oligomannose-type glycans for each isolate is given in Appendix II, dataset B. Within this panel, the overall size of the oligomannose population varied between 24% and 47% (mean average 37%, standard deviation 6%). Clade C isolates appeared to contain slightly smaller overall populations, although this result was not determined to be significant (Fig. 3.4A). More noticeable differences were observed when comparing individual oligomannose-type glycan species, in particular the least processed $Man_9GlcNAc_2$ species (Fig. 3.4B). On average Clade B isolates were found to possess significantly higher proportions of $Man_9GlcNAc_2$ compared with clades C, G and CRF01_AE. Clade A isolates also appeared to have slightly higher oligomannose content relative to clades C, G and CRF01_AE, however this result was not significant.

3.4 Correlation of oligomannose abundance with outer domain glycosylation sites.

To investigate the differences observed in the abundances of oligomannose-type glycans, the number and distribution of potential N-glycosylation sites (PNGSs) within each isolate was analysed. No significant correlation was found between the overall abundance of oligomannose-type glycans and the total number of PNGSs (Fig. 3.5A). Given the hypothesised role of density in contributing to oligomannose persistence, the data were re-analysed looking at the number of sites present solely on the outer domain, which contains the highest density of glycan sites. In this case linear regression analysis showed a significant positive correlation ($p = 0.015$) between the number of PNGSs on the outer domain and the abundance of oligomannose-type glycans (Fig. 3.5B).

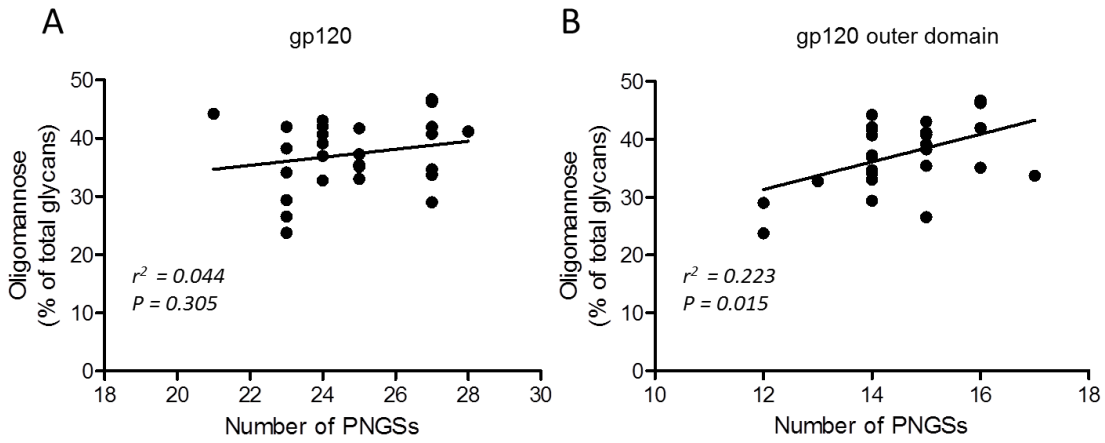


Fig. 3.5. Correlation between the number of PNGSs and abundance of oligomannose-type glycans. The total abundance of oligomannose-type glycans measured from each gp120 isolate was plotted against the total number of potential N-glycosylation sites (PNGSs) **(A)**, or the number of PNGSs within the outer domain (residues 252-482) only **(B)**. Correlations were assessed using linear regression analysis, performed using GraphPad Prism software.

Different clades are associated with the presence or absence of certain glycosylation sites (Fig. 3.6A), for example clade C isolates frequently lack the glycan at N295, whereas CRF01_AE isolates typically have a glycan at N334 rather than N332 (Travers, 2012). To investigate whether individual PNGSs could be associated with higher levels of oligomannose-type glycans, the data were grouped and analysed according to the presence or absence of individual sites. No sites were found to be correlated with a significantly higher overall abundance of oligomannose (data not shown). However when the presence of each site was compared with the levels of $Man_9GlcNAc_2$ only, two sites – N295 and N392 – were found to be significantly associated with higher levels of $Man_9GlcNAc_2$ (Fig. 3.6B). Isolates carrying the N332 glycan site also appeared to show higher levels of $Man_9GlcNAc_2$, however this result was not significant. The analysis also revealed that the presence of the site at N442, located at the periphery of the outer domain, was associated with lower levels of $Man_9GlcNAc_2$. It is unclear if this observation has a structural basis or, more likely, derives from its higher association with clade C and G isolates which more frequently lack

the N295 PNGS (Fig. 3.6A). No other sites were found to be significantly associated with higher or lower abundances of oligomannose.

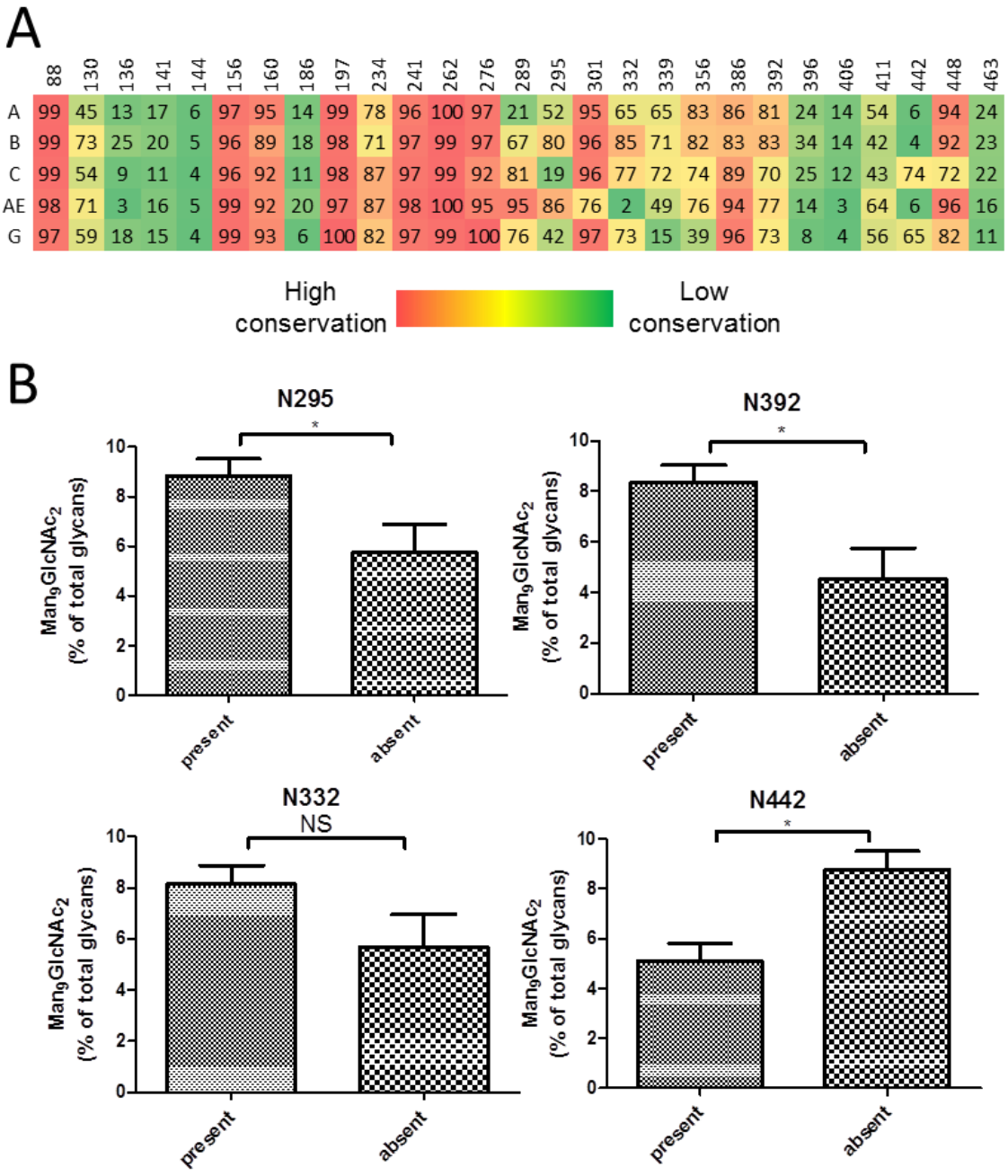


Fig. 3.6. Abundance of $Man_9GlcNAc_2$ in the presence and absence of individual PNGSs. (A) Conservation (%) of individual PNGSs across different clades. Sequence data from the 2013 filtered web alignment from the Los Alamos HIV sequence database (http://www.hiv.lanl.gov/) was used for analysis: clade A, *n = 226*; clade B, *n = 1502*; clade C, *n = 1031*; CRF01_AE, *n = 356*; clade G, *n = 71*. (B) Data were grouped according to the presence or absence of individual PNGSs, and compared using a student t test (two-tailed).* represents $p \leq 0.05$. NS: not significant. All other sites showed no significant association with levels of $Man_9GlcNAc_2$.

3.5 Conclusions

The data presented here represent the first quantitative comparison of glycan compositions from an extensive cross-clade panel of monomeric gp120s. Purification of the gp120s using nickel affinity chromatography avoided the possible bias that can be introduced through lectin affinity chromatography, and analysis by HILIC-UPLC enabled accurate quantitation of the whole glycan population, which includes neutral and sialylated glycans. This analysis revealed that a population of oligomannose-type glycans is a conserved feature of all isolates tested and is independent of clade. Larger populations of oligomannose-type glycans were found to be associated with a greater number of PNGSs within the outer domain specifically, supporting the hypothesis that the formation of oligomannose-type glycans is a consequence of high density in this region (Bonomelli et al., 2011; Doores et al., 2010a; Scanlan et al., 2007).

While overall levels of oligomannose-type glycans were consistent across different clades, differences could be detected at the level of individual oligomannose-type species. The abundance of $Man_9GlcNAc_2$ displayed the greatest variability, with clade B isolates possessing the highest abundance. This is consistent with the high conservation of the N295 and N392 sites within clade B, since both of these sites were associated with higher levels of $Man_9GlcNAc_2$. However this analysis is limited by the relatively small sample size, which makes it very difficult to delineate site-specific associations with $Man_9GlcNAc_2$ levels from broad clade-associative effects. This is highlighted by the association of the N442 site with lower levels of $Man_9GlcNAc_2$. This site is located on the outer domain and has been implicated in formation of the 2G12 epitope (Gray et al., 2007), suggesting it either contains α1,2-mannose glycans ($Man_8GlcNAc_2$ and $Man_9GlcNAc_2$), or helps limit glycan processing at nearby sites that are contacted by 2G12. Its observed association with lower levels of $Man_9GlcNAc_2$ is therefore likely to derive from a general association with clade C and G

isolates which have lower levels of $Man_9GlcNAc_2$, presumably due to the absence of sites like N295. A more direct approach to understanding the contribution of individual sites to oligomannose formation, using alanine-scanning mutagenesis, is taken in chapter 5.

Despite the considerable heterogeneity of glycosylation that was evident in the glycan profiles from the different isolates, a population of oligomannose-glycans represents a conserved feature of HIV-1 gp120 that could be targeted in a vaccine or therapeutic context. The next chapter aims to explore whether the heterogeneity in glycosylation is evenly distributed across the molecule, or if there are site-specific differences.

4 Variable microheterogeneity of gp120 glycosylation sites.

Recombinant gp120 bears both complex and oligomannose-type glycans, however the distribution of these two classes across gp120 sites is not uniform. Site-specific glycosylation studies have revealed that the majority of sites containing oligomannose-type glycans lie within the outer domain of gp120, which harbours the highest density of sites (Cutalo et al., 2004; Go et al., 2008, 2009, 2011, 2013; Leonard et al., 1990; Pabst et al., 2012; Zhu et al., 2000). This correlates with the concept of density-dependent processing, where inaccessibility of glycosidases and glycosyltransferases to these glycan sites results in their under-processing (Doores et al., 2010a). For vaccine design, valuable information could be gained from an understanding of the distribution of different glycans across gp120, and in particular knowledge of the different glycan structures and their abundances at sites directly targeted by bnAbs. This information would complement structural studies, which reveal the molecular interactions between bnAbs and their glycan epitopes, offering insights into the levels of selectivity displayed by different bnAbs (Blattner et al., 2014; Falkowska et al., 2014; Huang et al., 2014; Kong et al., 2013; Mouquet et al., 2012; Pejchal et al., 2011; Scharf et al., 2014; Walker et al., 2009, 2011a).

Site-specific glycosylation analysis of gp120 is technologically challenging due to the high number of glycan sites and the diversity of glycans present, and different approaches have been employed (Cutalo et al., 2004; Go et al., 2013; Leonard et al., 1990; Pabst et al., 2012; Zhu et al., 2000). Strategies typically start with a proteolytic digest of the recombinant material, for example using trypsin or a combination of proteases. The digestion mixture of glycopeptides and peptides is then either directly analysed by

ESI-LC-MS/MS, or alternatively the sample may be first subjected to a fractionation step in order to reduce the complexity of the subsequent analysis. During ESI-LC-MS/MS, fragmentation is routinely performed using collision-induced dissociation (CID). While this method causes effective fragmentation of glycosidic linkages, providing information on the glycan composition, fragmentation of the peptide backbone is rarely observed (Wuhrer et al., 2007). Peptide identity can instead be obtained using the complementary electron-transfer dissociation (ETD) fragmentation method, which preferentially causes cleavage of the peptide bonds. MALDI-MS has also been effectively employed to characterise gp120 glycopeptides following fractionation of a protease digest (Go et al., 2008). Fragmentation of glycopeptides by MALDI produces characteristic fragment ions which provide the peptide mass and the glycan composition, however diagnostic fragmentation of the peptide backbone is unreliable and dependent on the amino acid composition and matrix used (Stephens et al., 2004; Wuhrer et al., 2007). Both methods can also be hindered by the poor intrinsic ionisation properties of glycopeptides relative to peptides, and thus assignment of the peptide identity is commonly achieved by MS/MS analysis of PNGase F-treated deglycosylated peptides.

While a number of studies have achieved almost complete coverage of recombinant gp120 glycan sites (Go et al., 2008, 2009, 2011, 2013, 2014; Leonard et al., 1990; Pabst et al., 2012; Zhu et al., 2000), limitations of the methodology used often result in uncertainties regarding the abundances of different glycan structures (that is, the microheterogeneity) at a given site. Measured intensities are influenced by the ionisation efficiencies of glycopeptides, which are dependent on their glycan chemical composition (Zaia, 2010). Direct comparison of neutral and sialylated glycopeptides in particular, which differ significantly in their ionisation properties, must therefore be interpreted with caution. Furthermore the presence of contaminating peptides or other glycopeptides can lead to ion

suppression effects, distorting observed intensities (Zaia, 2010). One approach to overcome this problem is to isolate all glycoforms derived from one site, typically by reverse-phase HPLC, and analyse them collectively (Go et al., 2008; Pabst et al., 2012).

While quantitation by MS methods are subject to possible bias introduced by the ionisation properties of a glycopeptide, an alternative approach is to release the glycans from the glycopeptide and analyse them by HILIC-UPLC. Quantitation, based on detection of a fluorescent tag, is therefore independent of the nature of the glycoforms (Anumula and Dhume, 1998). In this chapter this method was developed and applied to the site-specific glycosylation analysis of recombinant $gp120_{BaL}$ to quantitatively assess the microheterogeneity of glycosylation at different sites.

4.1 Global heterogeneity of monomeric gp120.

Recombinant, monomeric $gp120_{BaL}$, which contains 23 potential N-glycosylation sites (PNGSs; Fig. 4.1A) was used for analysis. The gp120 was expressed in 293T cells and purified by nickel- and size exclusion-chromatography (SEC) (Fig. 4.1B,C). To determine the overall glycan population, glycans were released from gp120 using PNGase F, labelled with 2-amino benzamide (2-AB), and then resolved using HILIC-UPLC (Fig. 4.2). A fraction of the glycans were retained unlabelled, and analysed using ion mobility ESI-MS/MS (performed by Professor David Harvey). A mass table of glycan structures found is given in Appendix III (Table S3.1). These data were used to help assign the HILIC-UPLC glycan chromatogram in parallel with exoglycosidase sequencing of the glycans (Fig. 4.2; Table 4.1).

Analysis of the total glycan pool by ion mobility ESI-MS/MS, which has remarkable sensitivity (Harvey et al., 2011), revealed a very high level of heterogeneity, with at least 65 different structures being identified (Appendix III, Table S3.1). The same number could not

be resolved by HILIC-UPLC due to the co-elution of certain glycans and a lower sensitivity of detection (Fig. 4.2), however a considerable degree of heterogeneity was still evident. Table 4.1 lists the glycan structures identified by exoglycosidase sequencing.

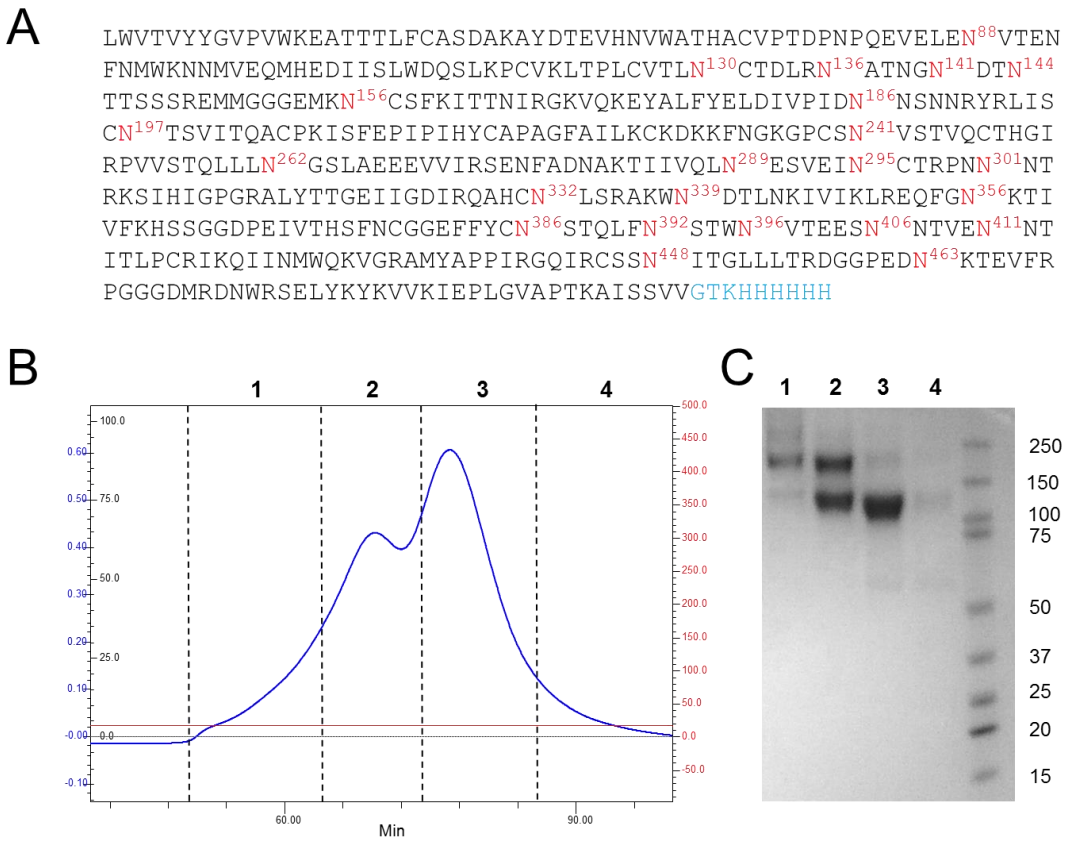


Fig. 4.1. Purification of monomeric gp120$_{BaL}$. (A) The sequence of the mature gp120$_{BaL}$ protein (without the signal peptide) is given. PNGSs are highlighted in red and numbered according to the HXB2 reference strain. The C-terminal hexahistidine tag and associated cloning site is shown in blue. (B) Size exclusion chromatogram showing elution of gp120 monomers (fraction 3) and higher-order oligomers (fractions 1 and 2). (C) Coomassie-stained SDS-PAGE of collected fractions. The monomeric species, collected in fraction 3, was used for analysis.

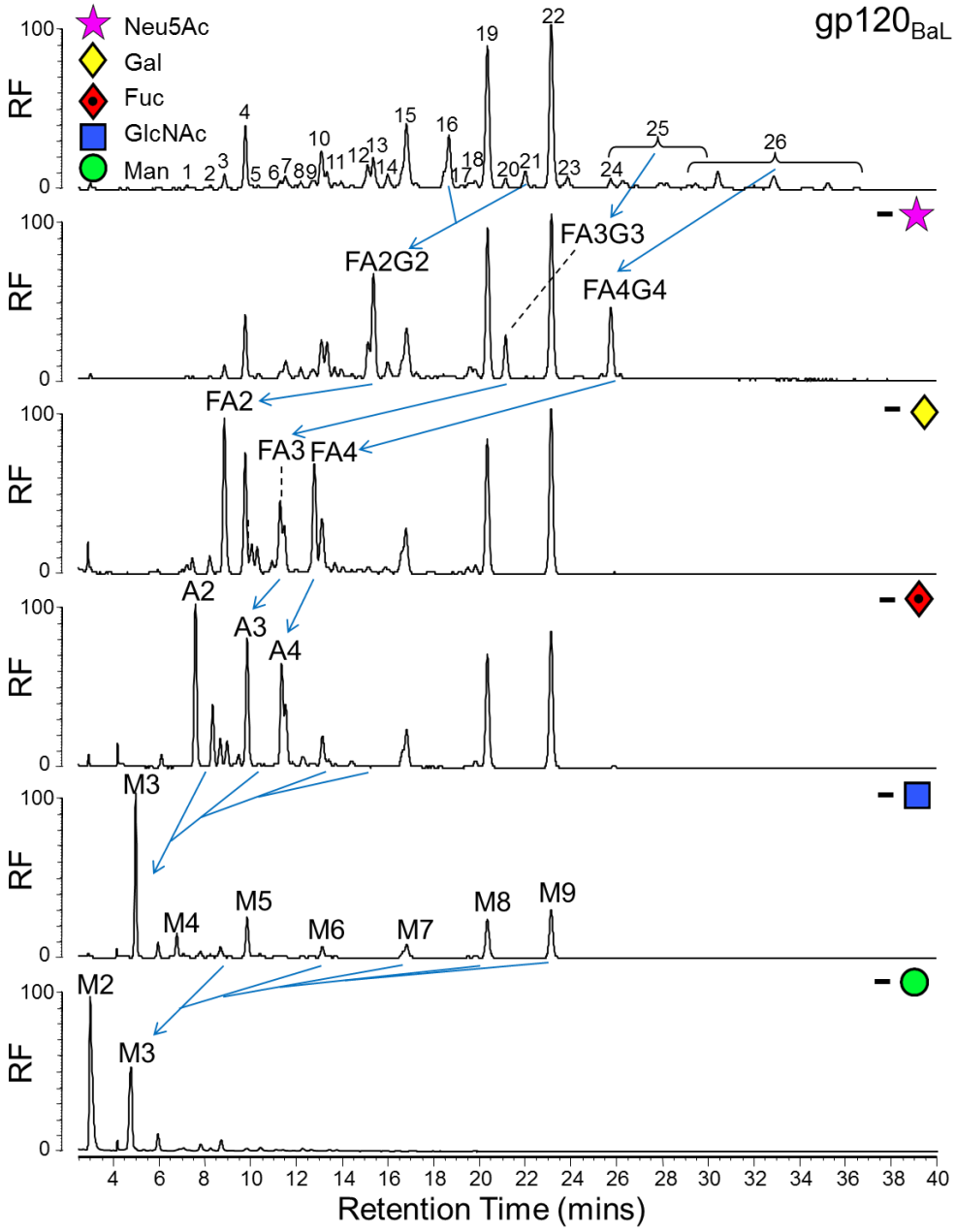


Fig. 4.2. Sequencing of gp120$_{BaL}$ glycans by exoglycosidase digestion. Structural assignment of HILIC-UPLC peaks was performed by sequential digestion with a panel of exoglycosidases. Top panel shows the undigested gp120$_{BaL}$ glycan profile. Panels below represent sequential digestion with the following exoglycosidases: Neuraminidase from *A. ureafaciens*, α-L-Fucosidase from Bovine Kidney, β1,4-Galactosidase from *Streptococcus pneumoniae*, β-N-Acetylglucosaminidase from *Streptococcus pneumonia* and α(1-2,3,6)-mannosidase from Jack Bean. Arrows indicate the digestion product. Neu5Ac, N-acetylneuraminic acid; Gal, galactose; Fuc, fucose; GlcNAc, N-acetylglucosamine; Man, mannose. Glycan structures are represented using Oxford glycan nomenclature (Harvey et al., 2009).

Table 4.1. Glycan structures identified from recombinant $gp120_{BaL}$ by exoglycosidase sequencing and HILIC-UPLC analysis.

Peak	Retention Time (min)	Glucose Units (GU)	Major structure	Minor structures
1	7.27	5.34	FA1	A2
2	8.27	5.68	M4A1	
3	8.90	5.89	FA2	
4	9.80	6.17	M5	FA2B, A3′[§]
5	10.09	6.25	FA3	FM4A1
6	11.33	6.60	A4	
7	11.54	6.65	M5A1	FA3′[§]
8	12.22	6.83	M4A1G1	
9	12.80	6.98	FA4	
10	13.12	7.06	M6	FA4B
11	13.37	7.12	FM4A1G1	
12	15.15	7.57	M5A1G1	A2BG2
13	15.39	7.63	FA2G2	
14	16.01	7.79	FA2BG2	
15	17.24	8.00	M7	
16	18.69	8.46	FA2G2S1	M5A1G1S1
17	19.25	8.60	FA2BG2S1	
18	19.60	8.68	A3′G3[§]	M8
19	20.37	8.87	M8	FA3G3
20	21.16	9.07	FA3’G3[§]	
21	22.02	9.29	FA2G2S2	
22	23.16	9.58	M9	
23	23.89	9.76	FA3′G3S1[§]	
24	25.76	10.25	FA4G4	
25	-	-	Tri-antennaries	
26	-	-	Tetra-antennaries	

[§]A3′ structures are isomers of A3

4.2 Variable microheterogeneity of individual glycan sites

To understand how the glycan microheterogeneity of individual sites compared with the global diversity of glycosylation, site-specific glycan analysis of $gp120_{BaL}$ was performed. The strategy for analysis is represented schematically in Fig. 4.3. Gp120 was reduced, alkylated with iodoacetamide, and then digested with trypsin or chymotrypsin. The digest

was then fractionated by reverse phase-high performance liquid chromatography (RP-HPLC) over a 90-minute gradient, with 90 fractions being collected per digest. A small portion of each fraction was retained and analysed directly by MALDI-MS in order to obtain glycopeptide masses. The remainder of the fraction was subjected to PNGase F digestion to release the glycans from the glycopeptides. The digest was then analysed by ESI-LC-MS/MS in order to obtain the identity of the deglycosylated peptides, before fluorescent labelling of the glycans with 2-AB and HILIC-UPLC analysis.

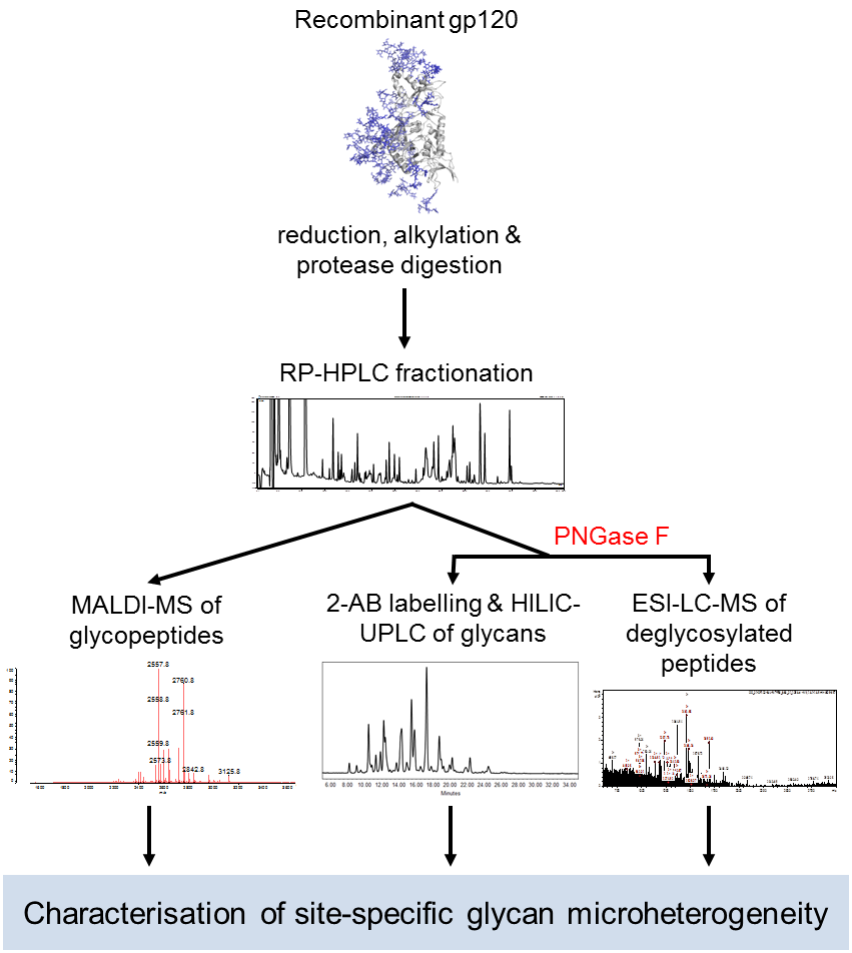


Fig. 4.3. Schematic representation of the methodology used for site-specific glycan analysis.

Data obtained from an example site, N197, is shown in Fig. 4.4. The three approaches: ESI-MS/MS of deglycosylated peptides, HILIC-UPLC of released glycans, and MALDI-MS of glycopeptides, provide complementary information that facilitate assignment and quantitation of the glycoforms at a glycan site. For example at the N197 site

approximately 23% of the glycans were of the oligomannose-type compared to 77% complex-type. The data also demonstrate the advantages of HILIC-UPLC for comparing complex glycan mixtures accurately; the FA2G2S1 glycan represents the second most abundant species, as determined by HILIC-UPLC, however the intensity of the equivalent glycopeptide is much lower by MALDI-MS. Furthermore no multiply sialylated structures could be detected by MALDI-MS despite being present in the HILIC-UPLC chromatogram.

This combinatorial approach was used to obtain HILIC-UPLC glycan profiles of six gp120 sites contained within tryptic glycopeptides (Fig. 4.5). Glycopeptides containing further sites were identified in other tryptic and chymotryptic fractions, however co-elution of these glycopeptides during RP-HPLC fractionation prevented the acquisition of distinct glycan profiles. All glycopeptides identified by MALDI-MS, and HILIC-UPLC glycan profiles of all fractions tested, are given in Appendix III. Of the six sites analysed, considerable variation in the extent of microheterogeneity was evident with some sites containing a much greater distribution of glycan structures (Fig. 4.5; Table 4.1). For example the N463 site, located in a relatively accessible location on the V5 loop, contained approximately 50 different glycan structures with the most predominant structure, FA2, accounting for just 10% of the total glycan population. This is in contrast with the N332 site, located centrally within the outer domain, where $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ together account for 73% of the glycan population. The N130 site, which has previously been reported to contain only complex-type glycans (Cutalo et al., 2004; Go et al., 2009; Zhu et al., 2000), was found to contain approximately equal levels of oligomannose-type and complex-type glycans.

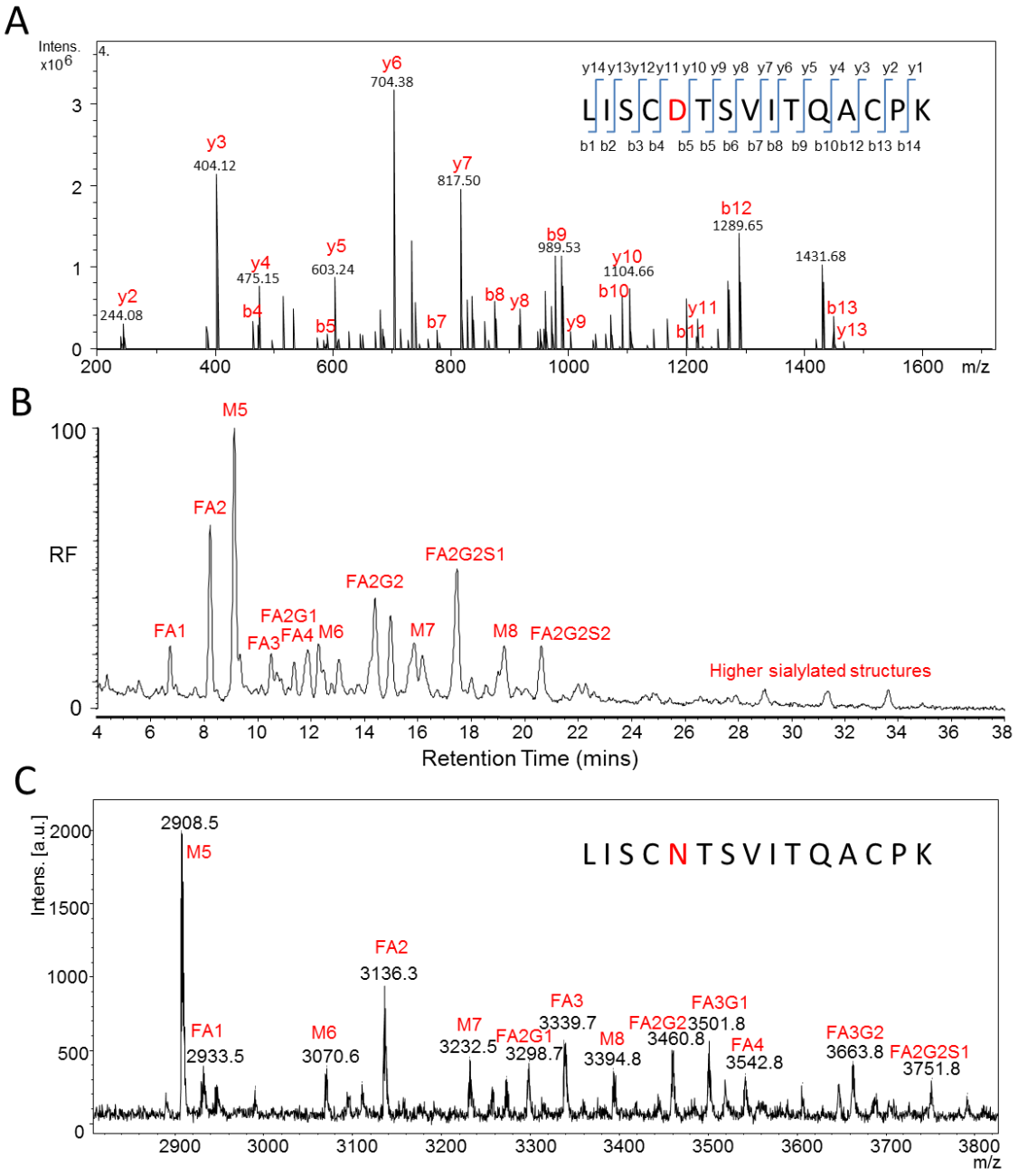


Fig. 4.4. Identification and glycoform characterisation of an example site, N197. (A) ESI-LC-MS/MS spectrum of the deglycosylated tryptic peptide LISCDTSVITQACPK (deglycosylation by PNGase F causes conversion of N→D. Precursor mass was 847.4, $[M_{pep}+2]^{2+}$. Cysteine residues were modified with carbamidomethyl (+57). (B) HILIC-UPLC chromatogram of glycans released from the LISCNTSVITQACPK glycopeptides (following 2-AB labelling). (C) MALDI-MS spectrum of fraction containing LISCNTSVITQACPK glycopeptides. Observed masses correspond to $[M+H]^+$. RF – relative fluorescence.

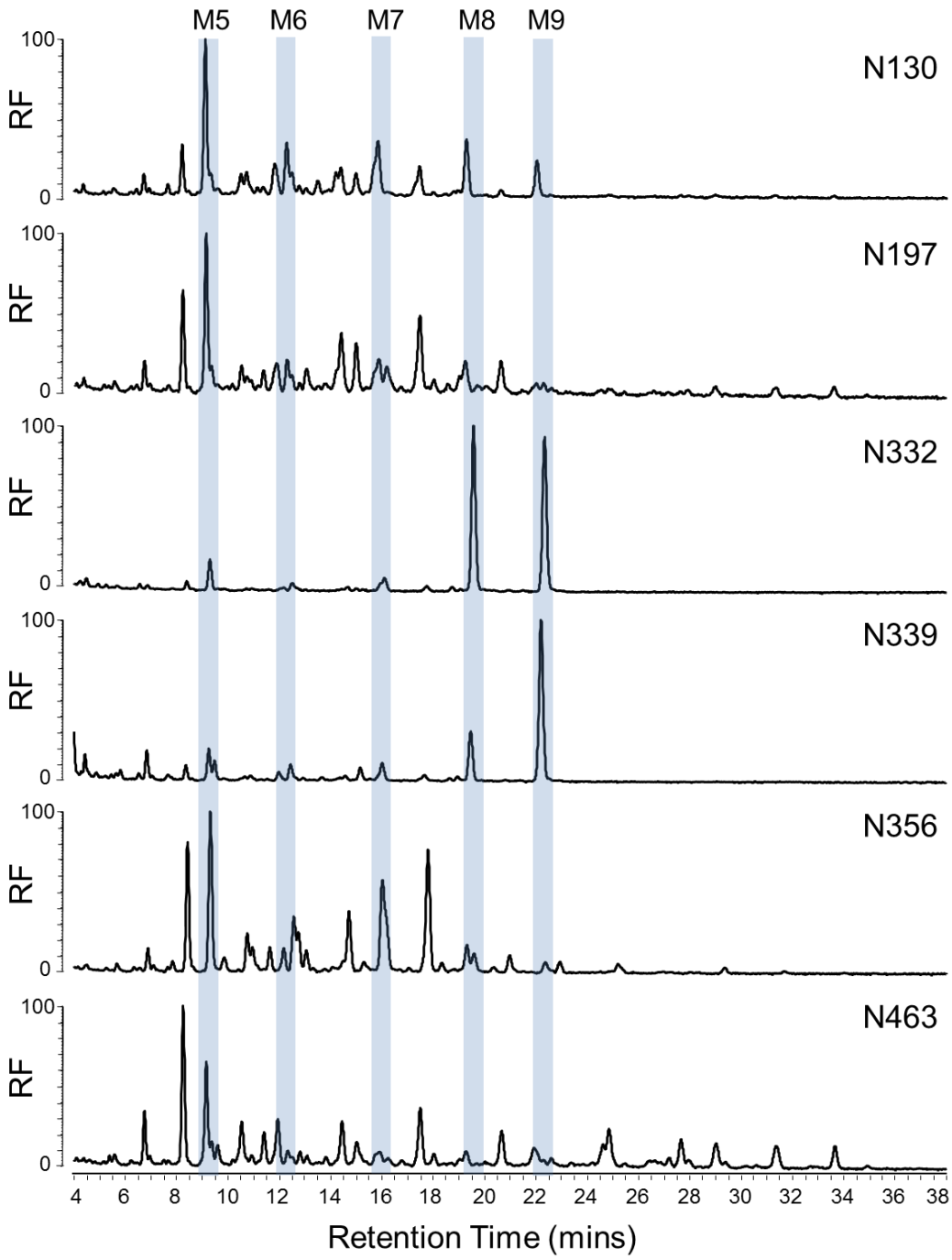


Fig. 4.5. HILIC-UPLC chromatograms of glycans present at six different gp120 sites. N-linked glycans were released from isolated glycopeptides using PNGase F, 2-AB labelled, and resolved by HILIC-UPLC. Where glycopeptides eluted over two fractions the relevant fractions were pooled and re-analysed. Fractions used were as follows: N130, T47 and T48; N197, T34 and T35; N332, T17 and T18; N339, T21 and T22; N356, T19 and T20; N463, T24 and T25. Profiles of individual fractions are shown in Appendix III, Fig. S3.1. Oligomannose-type glycans are highlighted in blue. RF – relative fluorescence.

Table 4.2. Abundances of glycan structures found at gp120 sites by HILIC-UPLC.

Site	Abundance (% of total glycans)						
	M5	M6	M7	M8	M9	**Total M5-M9**	**Total Complex**
N130	16	7	10	7	5	**45**	**55**
N197	11	3	5	4	1	**23**	**77**
N332	6	2	5	37	36	**86**	**14**
N339	6	4	3	11	38	**62**	**38**
N356	12	7	13	2	2	**36**	**64**
N463	9	2	2	2	0	**14**	**86**

4.3 Conclusions

The strategy for site-specific glycosylation analysis outlined here represents a sensitive method for assessing the microheterogeneity of glycosylation at gp120 glycan sites. HILIC-UPLC analysis of released glycans allows an unbiased quantitation of neutral and sialylated glycans, which is not always possible with MS methods. However a major challenge with this method lies in achieving complete coverage, and further method development would be required to cover all sites. Here glycan profiles for six sites (N130, N197, N332, N339, N356 and N463) were obtained from a tryptic digest, with a further two sites (N156 and N448) identified by MALDI-MS alongside co-eluting glycopeptides (Appendix III). Given the distribution of trypsin target residues (Lys and Arg) the N88 and N186 sites were theoretically resolvable (Fig. 4.1A), however their predicted glycopeptide masses were larger than for other sites, perhaps explaining why none were detected by MALDI-MS. The remaining sites were not resolvable through tryptic digest, and thus a separate chymotryptic digest of the $gp120_{BaL}$ was performed. Chymotrypsin cleaves aromatic amino acids (Phe, Tyr and Trp) preferentially, but also has modest activity against Leu, Ile and Met. This broader specificity facilitated identification of ten sites by MALDI-MS, five of them new (N88, N186, N262, N386 and N392), bringing the total coverage to

13 out of 23 sites. However while the chymotryptic digest allowed a greater number of sites to be resolved, it concomitantly increased the number of glycopeptides produced which in turn led to an increased number of glycopeptides co-eluting. It therefore wasn't possible to obtain HILIC-UPLC glycan profiles for any isolated sites using chymotrypsin.

There is therefore a trade-off between increasing the number of glycopeptides produced, so as to resolve more sites, and the risk of multiple glycopeptides co-eluting. Higher specificity enzymes like trypsin offer more predictable behaviour, facilitating data analysis and leading to fewer missed cleavages. One option is to extend their specificity by chemically modifying other amino acids, for example it has been shown that modification of cysteine residues with 2-bromoethylamine converts them to a trypsin substrate (Lindley, 1956). For $gp120_{BaL}$ this would theoretically enable additional resolution of the N241 and N301 sites. Alternatively other high specificity enzymes like Asp-N, which cleaves Asp and Glu residues (Ingrosso et al., 1989), or Glu-C, could be employed. However given the problem of co-elution of glycopeptides, an alternative strategy could involve multiple rounds of RP-HPLC fractionation, whereby glycopeptides carrying multiple sites are digested with a second protease and the digestion product is re-fractionated. In this way the complexity of the sample is reduced with each digest, and as many digests as required could be performed to resolve all sites. Improvements to the RP-HPLC fractionation gradient could also be investigated.

Despite the difficulties in obtaining complete coverage of gp120, this method represents a powerful tool for investigating bnAb epitopes, and targeted digestion strategies could be employed for sites of interest. Comparison of the sites identified here revealed a considerable variation in the degree of microheterogeneity at individual sites, with some being dominated by only a couple of oligomannose structures, compared to highly heterogeneous distributions of complex-type glycans at others. The presence of

oligomannose-type glycans is thought to arise from steric factors which prevent full access by glycosidases and glycosyltransferases (Doores et al., 2010a), and these steric factors are likely also responsible for the restricted microheterogeneity observed at sites like N332 and N339. This more limited microheterogeneity may contribute to the immunogenicity of these sites since it would correspond to a higher concentration of a particular epitope. This is supported by the observation that a number of the glycan-reactive bnAbs characterised to date display a binding preference for oligomannose-type glycans (Kong et al., 2013; Sanders et al., 2002b; Scanlan et al., 2002; Walker et al., 2009, 2011). However in terms of maintaining breadth of protection, an important consideration is the degree to which microheterogeneity of glycoforms can be tolerated by bnAbs; data reported here and elsewhere indicate that all gp120 sites display some level of microheterogeneity, and the degree to which this can be tolerated will affect the efficacy of bnAbs. This idea of antibody promiscuity in glycan recognition is explored further in chapter 6.

Another important consideration for drug and vaccine design relates to the degree to which individual glycosylation sites are occupied, which could influence the efficacy of any therapy. Previous analyses of recombinant material have identified varying levels of partial occupancy at certain sites (Go et al., 2011, 2013, 2014; Pabst et al., 2012). Evidence of partial occupancy for some sites was observed here, however further investigations were not carried out for the following reasons: Firstly, the most common method used for analysing site occupancy is to treat glycopeptides with PNGase F and compare the abundance of deglycosylated peptides (which carry a +1 Da modification due to deamidation of Asn to Asp) with the abundance of non-glycosylated peptides. However such analysis can be unreliable since deamidation of asparagine residues can occur spontaneously (Palmisano et al., 2012). More accurate analysis requires additional, targeted experiments. For example expressing gp120 in $GnTI^{-/-}$ deficient HEK 293S cells or in the presence of kifunensine, to

produce all oligomannose-type glycans, followed by Endo H treatment, would leave a single N-acetlyglucosamine tag (+203 Da modification) at sites of glycosylation (Depetris et al., 2012), enabling more reliable assignment of occupancy; Secondly, the fractionation step separates the equivalent glycosylated and non-glycosylated peptides. For example the DGGPEDN^{463}KTEVFRPGGGDMR tryptic glycopeptides eluted in fractions 28 and 29, whereas the non-glycosylated DGGPEDNKTEVFRPGGGDMR peptides eluted several minutes later in fraction 32. Due to variable ionisation suppression effects across the different fractions, it is therefore not possible to accurately compare the observed abundances of the PNGase F-treated species in the separate fractions; Finally, the information regarding occupancy of a transiently expressed recombinant protein may be a poor representation of native virus, since occupancy has been shown to vary depending on transfection conditions and whether or not the gene sequence has been codon-optimised (Ebrahimi et al., 2014).

5 Resilience of the oligomannose population to sequence mutation

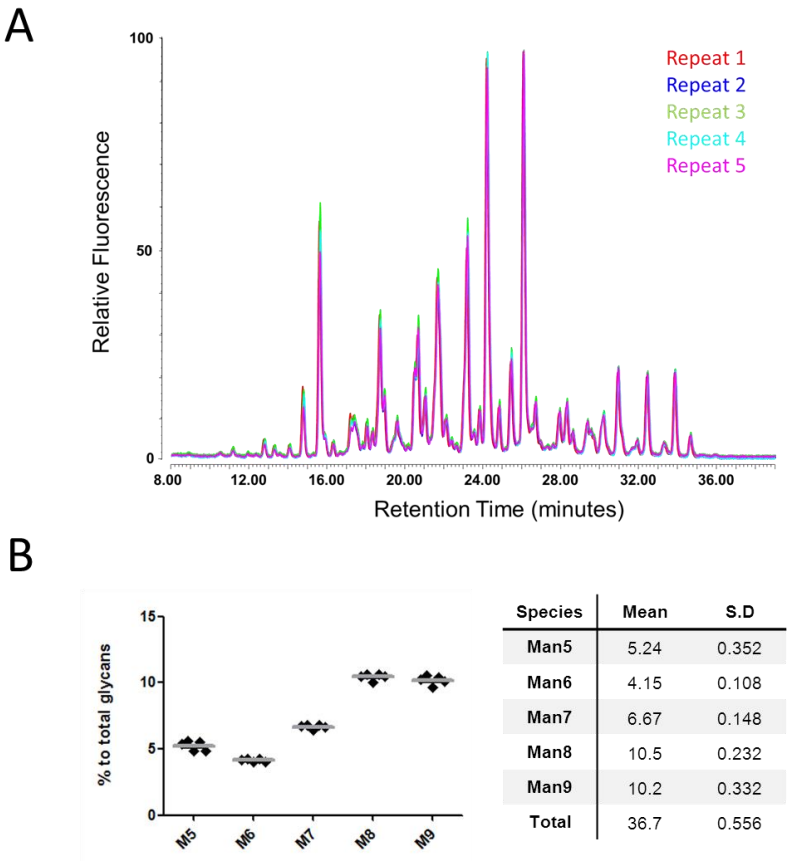
As demonstrated in chapter 3, oligomannose-type glycans are a conserved feature of gp120 and their abundance is correlated with the number of potential N-glycosylation sites (PNGSs) within the outer domain, suggesting that density is an important factor for their presence. The number of PNGSs harboured by a given strain is subject to variation over time in an infected individual, since addition or deletion of PNGSs is a common evasion strategy used by HIV-1 to escape adaptive immune responses (Moore et al., 2012; Wei et al., 2003). An important consideration for the design of glycan-targeted vaccines is thus the effect of such PNGS-deletions on the glycan population; does loss of an individual PNGS, and thus a reduction in glycan density, lead to significant changes in the abundance and type of oligomannose-type glycans? In this chapter the role of individual PNGSs in shaping the oligomannose population is investigated, at both the global and site-specific level.

5.1 Site-directed mutagenesis of the 23 PNGSs of $gp120_{BaL}$.

To investigate the factors that drive the formation of oligomannose-type glycans, all 23 PNGSs of $gp120_{BaL}$ were individually deleted by site-directed mutagenesis. This was achieved by mutating asparagine residues within the consensus sequence Asn-X-Ser/Thr (where X is any amino acid except Pro) to alanine. The panel of $gp120_{BaL}$ mutants was expressed, purified and analysed as described in Chapter 3, with the expression performed in parallel to minimise the intrinsic variation of glycosylation. Fig. 5.1 shows the glycan profiles of five batches of wild-type $gp120_{BaL}$, expressed in parallel. The high degree of reproducibility gives confidence that any observed variations seen with the PNGS-deletion mutants are genuine.

Changes in the total abundance of oligomannose-type glycans, relative to wild-type, were calculated from the HILIC-UPLC data (Fig. 5.2A; Appendix IV). Overall, the direction of change corresponded well with the predicted glycan species at each site; loss of predicted complex sites showed an increased relative abundance of oligomannose-type glycans, while loss of predicted sites of oligomannose generally resulted in lower relative abundances of oligomannose-type glycans. For a few PNGS-deletions, for example N386 and N392, larger than expected losses of oligomannose were observed. These glycans may represent sites that normally impede the access of α-mannosidases to neighbouring sites, meaning that loss of one glycan site has a knock-on effect by increasing accessibility to other sites for processing. However generally the overall abundance of oligomannose remained remarkably stable, with the largest overall effect (at site N392) displaying just a 15% decrease in overall oligomannose (Fig. 5.2A; Appendix IV). Several double PNGS-deletion mutants were also analysed (Appendix IV, Table S4.1) and no dramatic decreases in the overall abundances of oligomannose-type glycans were observed. The largest effect observed for the mutants

analysed was seen with the combined knockout of N295A and N386A, which showed an overall decrease of 27% in oligomannose-type glycans.



Species	Mean	S.D
Man5	5.24	0.352
Man6	4.15	0.108
Man7	6.67	0.148
Man8	10.5	0.232
Man9	10.2	0.332
Total	36.7	0.556

Fig. 5.1. Reproducibility of glycosylation profiles upon parallel expression. Wild-type $gp120_{BaL}$ was expressed, purified and analysed five times in parallel, starting with the same passage of 293T cells. (A) HILIC-UPLC glycan profiles of the five batches of wild-type $gp120_{BaL}$. (B) Quantitation of the individual oligomannose-type glycan species by integration of the corresponding peaks. Quantitation was achieved by Endo H-digestion of the total glycan pools, in duplicate, as described in chapter 3. Data processing was performed using Empower 3 software. The mean and standard deviation (S.D.) are given.

While the overall abundance of oligomannose-type glycans remained relatively unperturbed upon deletion of individual glycan sites, the abundances of the individual oligomannose species ($Man_{5-9}GlcNAc_2$) varied more considerably (Fig. 5.2B; Appendix IV). Similar to observations with the cross-clade panel of gp120s, the levels of $Man_9GlcNAc_2$ displayed the greatest sensitivity to removal of glycan sites. A decrease of over 25% was observed for nine sites: N130, N262, N295, N339, N356, N386, N392 and

N448. These contrasting effects on the overall abundance of oligomannose and $Man_9GlcNAc_2$ only can be reconciled by the changing proportions of the other oligomannose species (Fig. 5.2B).

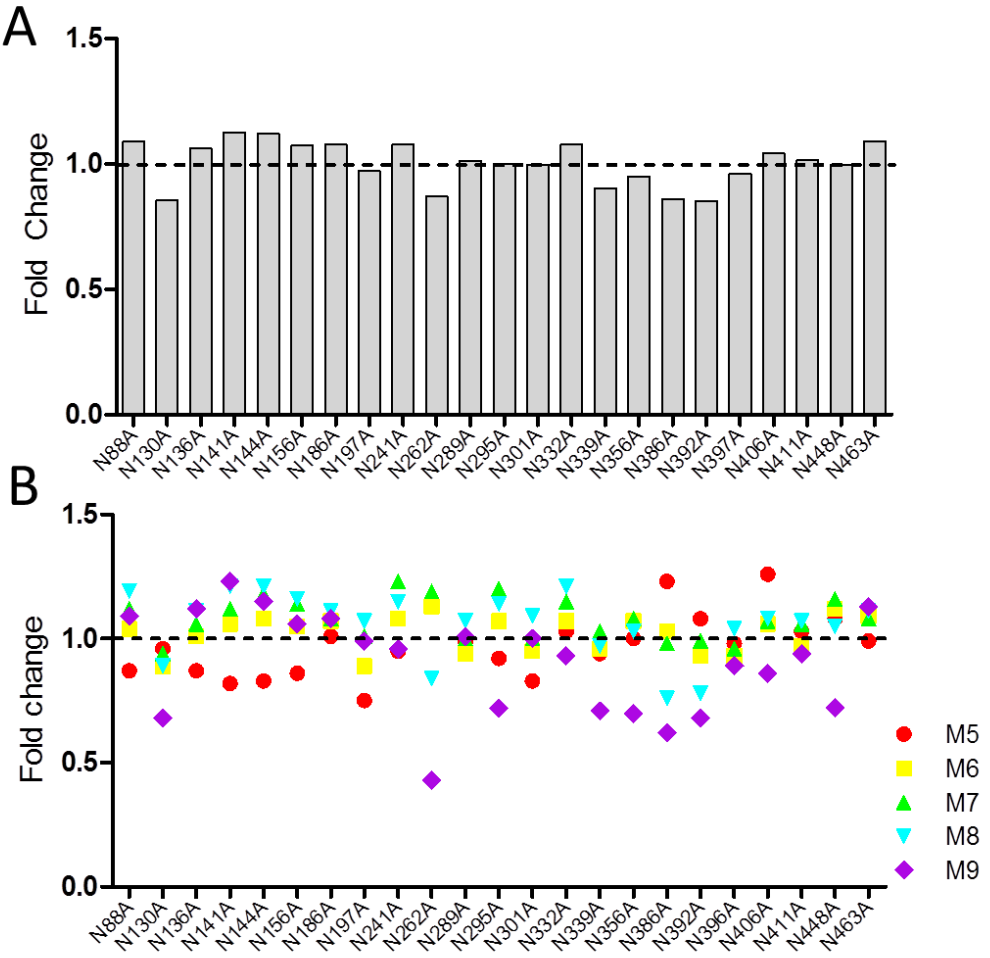


Fig. 5.2. Effect of PNGS-deletion on the oligomannose population of $gp120_{BaL}$. Values were obtained by Endo H digestion of the total glycan profiles followed by integration of corresponding HILIC-UPLC peaks, and represent the fold change, relative to wild-type, of the percentage abundances of oligomannose-type glycans. (A) Changes in overall abundance of oligomannose-type glycans upon PNGS-deletion (B) Changes in abundances of individual oligomannose-type species upon PNGS-deletion. M5, $Man_5GlcNAc_2$; M6, $Man_6GlcNAc_2$; M7, $Man_7GlcNAc_2$; M8, $Man_8GlcNAc_2$; M9, $Man_9GlcNAc_2$. Values represent analysis obtained from PNGS deletion-mutants expressed once, in parallel. Endo H digestion and quantitation was performed in duplicate.

While it was only possible to fully express the whole panel of PNGS-deletion mutants once, several mutants were chosen (WT, N130A, N136A, N262A and N386A) to repeat and verify the observed results. As expected there were minor differences in the

profile of the wild-type compared to the first expression (data not shown). However the observed changes in oligomannose-content, relative to wild-type, for the mutants were highly similar, and thus it is reasonable to have confidence in the observed effects of the PNGS-deletions.

5.2 The role of glycan-glycan interactions in maintenance of the oligomannose population.

The apparent re-distribution of glycan structures upon PNGS-deletion, from larger oligomannose-type glycans to smaller species, indicates that individual PNGSs have a role in influencing the processing of glycans at neighbouring sites. To understand this phenomenon at the structural level, a model of a fully glycosylated $gp120_{BaL}$ was constructed based on the recently solved crystal structure of the BG505 SOSIP.664 gp140 trimer (Julien et al., 2013a), with $Man_8GlcNAc_2$ glycans being incorporated at sites of predicted oligomannose-type glycans and $Man_5GlcNAc_2$ glycans at remaining sites (Fig. 5.3A). $Man_5GlcNAc_2$ were used at sites of predicted 'complex-type' sites as it is commonly found alongside complex-type glycans (chapter 4) and thus was chosen as a representative 'average' glycan that does not over-represent the density present.

The glycosylated model demonstrates the uniquely high density of glycans on the gp120 surface, which represents a significant barrier to processing by glycosidases and glycosyltransferases. Closer inspection of individual PNGSs within the outer domain reveals sites of glycan clusters with two or more glycans positioned in very close proximity. One such cluster involves the N386 and N392 glycans, in which the Asn Cα atoms are located 12 Å apart (Fig. 5.3B). Deletion of the N392 glycan led to an overall decrease in oligomannose abundance of 15%, primarily driven by losses of $Man_9GlcNAc_2$ and $Man_8GlcNAc_2$ (32% and 22% relative to wild-type, respectively; Fig. 5.2). Another cluster was identified based

around the N295 glycan, which occupies a central location between the N332 and N448 sites (Fig. 5.3C). While loss of the N295 glycan had no overall effect on the total abundance of oligomannose, it resulted in a 28% decrease in $Man_9GlcNAc_2$ (Fig. 5.2).

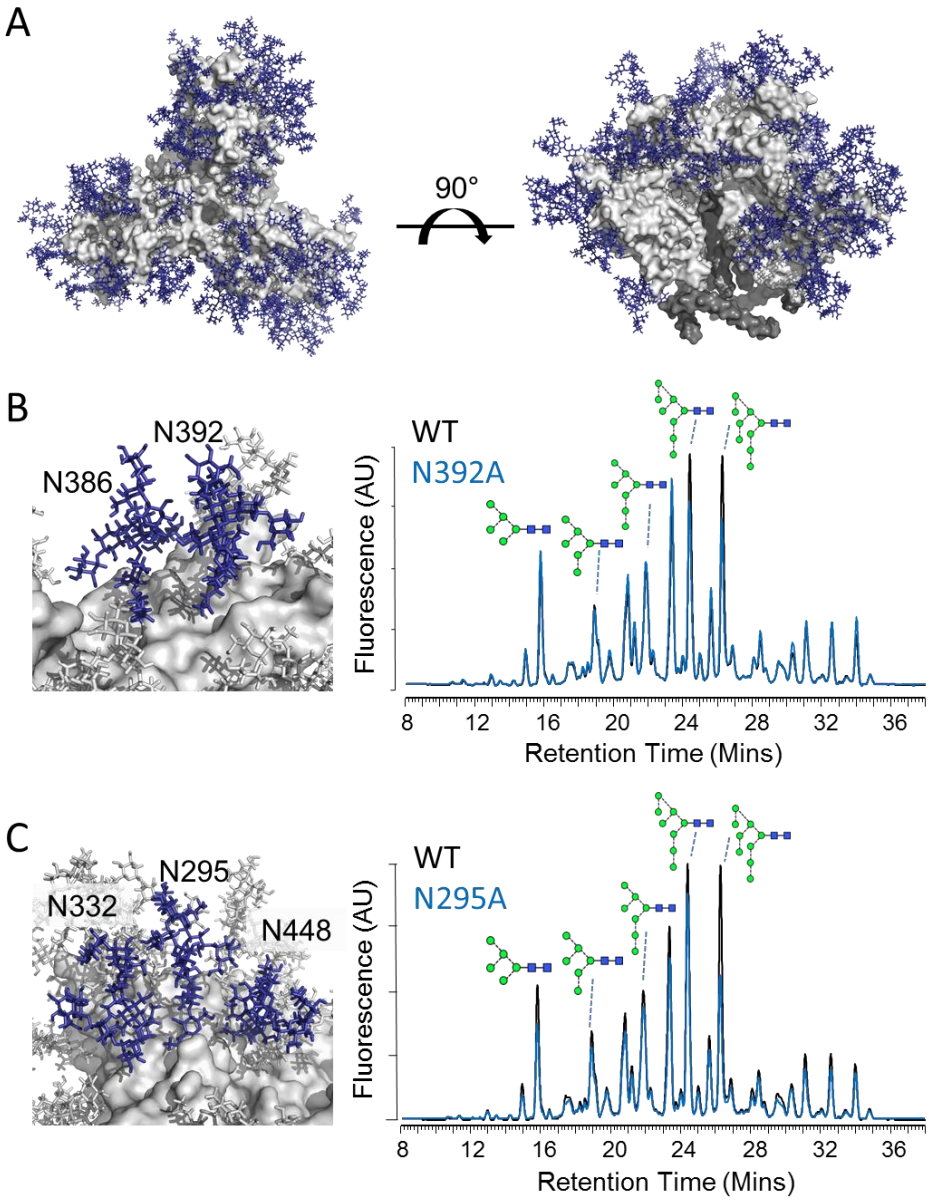


Fig. 5.3. Role of clustered glycan-glycan interactions in maintaining the oligomannose population. (A) A homology model of a BaL gp140 trimer (in grey) was constructed based on the recent gp140 trimer crystal structure (Julien et al., 2013a) (PDB code: 4NCO). Glycans are represented by blue sticks. Views show the trimer looking down on to the trimer apex (left) and from the side (right). Homology modelling and glycan building were performed by Dan Kulp and Sergey Menis (William Schief group, Scripps). (B) Image of glycan cluster comprising N386 and N392 glycans with HILICP-UPLC glycan profile of N392A mutant. (C) Image of glycan cluster comprising N295, N332 and N448 glycans with HILICP-UPLC glycan profile of N295A mutant. Wild-type profiles are in black and are overlaid with the mutant profiles in blue. Peaks corresponding to oligomannose-type glycans are indicated. AU – arbitrary units.

Steric constraints within these tight glycan clusters not only explain the limited processing of these glycans (chapter 4), but could rationalise the more widespread changes to glycan processing observed upon removal of a single glycan. It is likely that removal of a glycan within such a cluster increases the accessibility of its neighbouring glycan(s) to glycosidases, thereby increasing the level of processing that can occur at these nearby sites.

5.3 Changes in protein folding affect glycan processing

While many of the PNGS-mutants displayed very modest changes in glycosylation, a few displayed larger than expected differences, which couldn't be rationalised by involvement in glycan clusters. One explanation is that either the Asn→Ala mutation, or loss of the glycan, induces protein misfolding leading to altered glycan processing. To test this possibility ELISAs were performed using b12, a CD4-binding site conformation-dependent antibody (Fig. 5.4A).

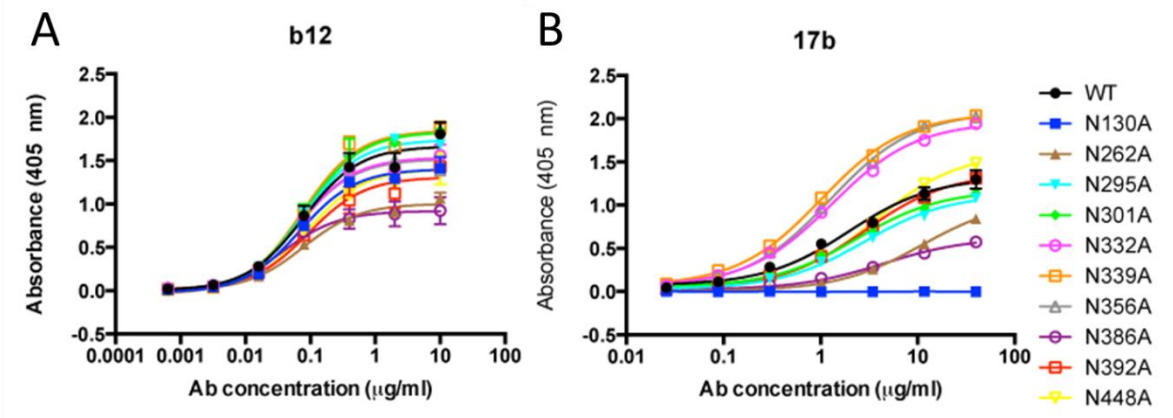


Fig. 5.4. Recognition of PNGS-deletion mutants by conformation-dependent antibodies. Binding of the conformation-dependent monoclonal antibodies b12 (A) and 17b (B) was tested against the panel of $gp120_{BaL}$ PNGS-deletion mutants by capture ELISA. ELISAs were performed by Dr Katie Doores.

The binding of the majority of the PNGS-mutants was unchanged relative to wild-type, with the exception of N262A and N386A which displayed significantly reduced binding plateaus. Reduced binding was also seen for these mutants with a second

conformation-dependent antibody, 17b, which additionally displayed an almost complete loss of binding to the N130A mutant (Fig. 5.4B). These three mutants displayed greater levels of aggregation by Western blotting relative to wild-type (data not shown), supporting a degree of misfolding. Since the samples used for ELISA contained the higher molecular weight aggregates, the reduced binding plateaus observed for the N130A, N262A and N386A mutants may reflect the presence of a higher concentration of aggregates that cannot be bound by b12 or 17b. These results suggest that misfolding of the N130A, N262A and N386A mutants (as indicated by the higher levels of aggregation) could explain the disproportionately large changes observed in their glycan processing (Fig. 5.2C). To determine if the significant effect of N262-deletion was common to other gp120 sequences, the N262A mutation was made across a small cross-clade panel (Fig. 5.5). A similar magnitude of perturbation was observed across the other isolates tested, supporting an integral role for the N262 glycan in protein folding.

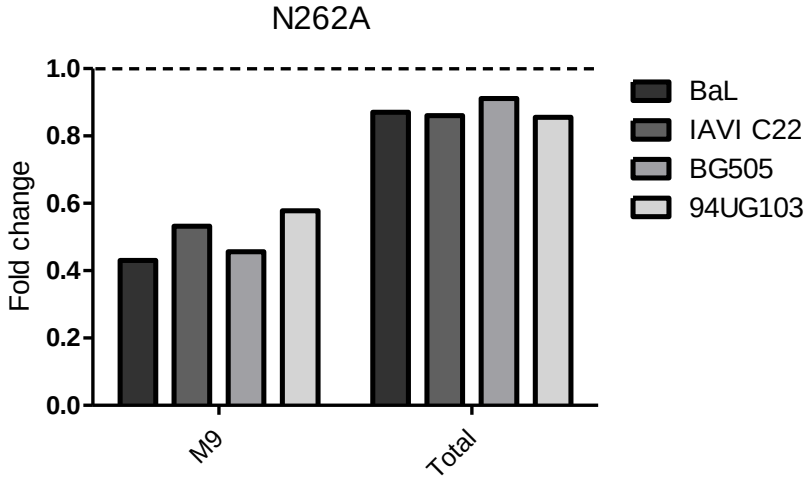


Fig. 5.5. Loss of the N262 glycan leads to reduced oligomannose abundance. The N262A mutation was made in the gp120 sequence from three additional isolates: IAVI C22 (clade C), BG505 (clade A) and 94UG103 (clade A), and compared to BaL (clade B). Mutants were expressed, purified and analysed for their oligomannose content as described above. Graphs represent the fold change, relative to wild-type, of the percentage abundances of oligomannose-type glycans M9; $Man_9GlcNAc_2$, Total, sum of $Man_{5\text{-}9}GlcNAc$.

5.4 Effect of PNGS-mutation on processing at the N332 site of vulnerability.

Analysis of the global glycan profiles of the PNGS-deletion mutants (Fig. 5.2) indicated that removal of a single glycan site may lead to bystander effects on the processing of nearby glycans. To directly address this possibility, the effect of removing neighbouring PNGSs on the processing of the N332 'site of vulnerability' (section 1.4.3) was assessed (Fig. 5.6). A tryptic glycopeptide containing the N332 site (sequence QAHCN^{332}LSR) was isolated by RP-HPLC fractionation and analysed by MALDI-MS to determine the glycoforms present. The method used was similar to that in chapter 4, with the exception that the tryptic digest was performed in-gel on the monomeric form of gp120 (in chapter 4 monomeric gp120 was purified by size exclusion chromatography). Consistent with its structurally constrained location, the glycans at the N332 site of the wild-type (WT) $gp120_{BaL}$ were found to be exclusively $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$, with approximately 35% and 65% abundances, respectively (Fig. 5.6B).

To determine if nearby sites played a role in determining the glycans present at N332, the same glycopeptide was purified and analysed from the N301A, N295A, N339A, N392A and N448A mutants. The MALDI mass spectra of these profiles revealed an altered distribution of glycans, with $Man_9GlcNAc_2$ being lost in favour of $Man_8GlcNAc_2$ (Fig. 5.6B,C). Consistent with its close proximity to N332, loss of the N295 glycan was found to have the largest impact on processing at N332, with an almost complete loss of $Man_9GlcNAc_2$. It is possible that the influence of the N448 glycan, which is not immediately adjacent to the N332 glycan, is mediated through N295, with which it is clustered. In this scenario, loss of the N448 glycan could increase the processing of the glycan at N295, presumed to be a large structure such as $Man_8GlcNAc_2$ or $Man_9GlcNAc_2$, to smaller structures like $Man_5GlcNAc_2$. This would in turn increase the accessibility of the

neighbouring N332 site to glycosidases. Loss of the N301 glycan, which reaches over the gp120 surface towards N332, also showed a distinct loss of $Man_9GlcNAc_2$ at the N332 site, and smaller yet significant changes were also observed for the N339A and N392A mutants.

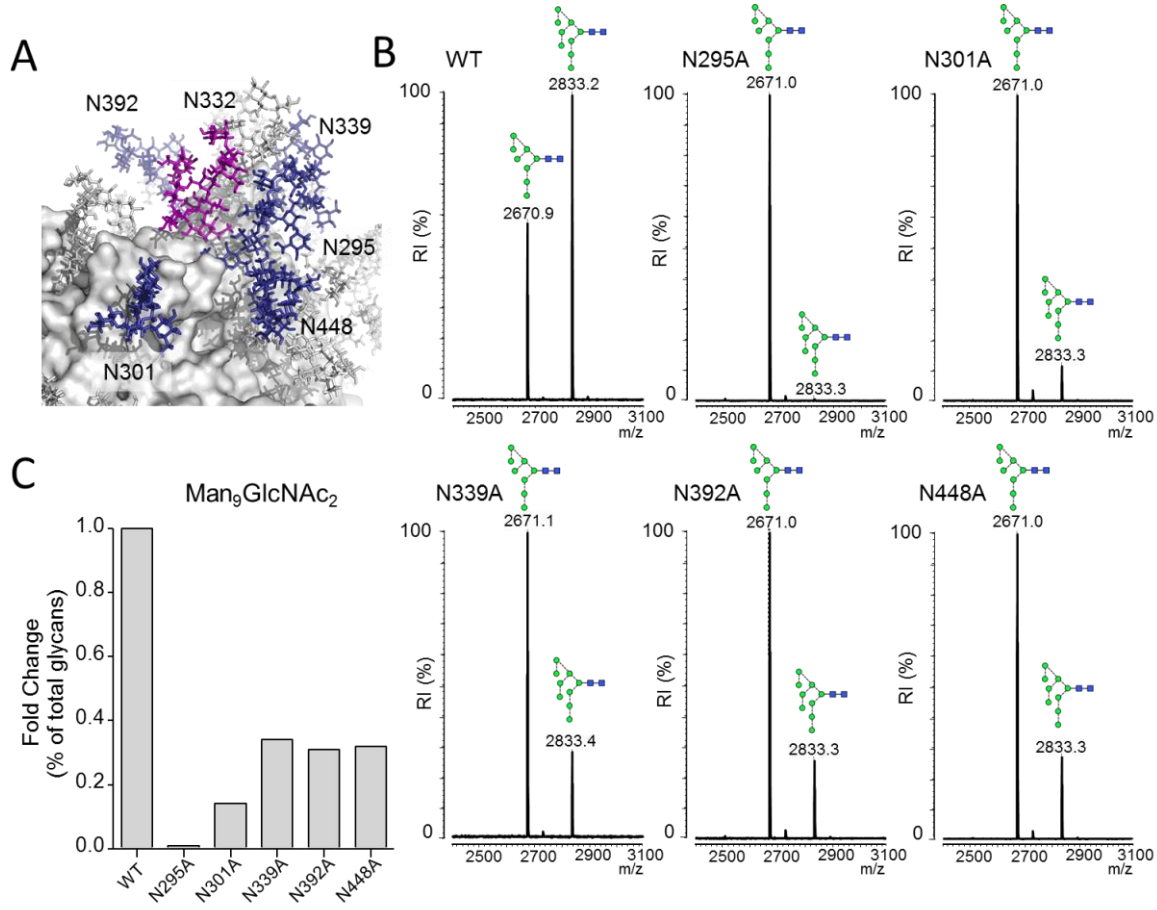


Fig. 5.6. Location of the N332 site of vulnerability. (A) Glycan modelling reveals the central location of the N332 glycan (purple) on the outer domain, surrounded by a dense pack of other glycans. Five neighbouring PNGS-deletion mutants (dark blue) were analysed for their glycan composition at the N332 site. (B) MALDI-MS spectra of glycopeptides (sequence QAHCN^{332}LSR) carrying the N332 site, isolated from wild-type (WT) $gp120_{BaL}$ and five PNGS-deletion mutants. (C) Changes in the glycan composition of the N332 site within the PNGS-deletion mutants relative to WT.

To rule out any effects deriving from natural variation in the processing of the N332 site between samples, three further batches of wild-type $gp120_{BaL}$ that had been expressed in parallel were analysed (Fig. 5.7). This revealed excellent reproducibility of the method, and

thus the data represent convincing evidence that removal of a single glycan site can modulate glycan processing at a neighbouring site.

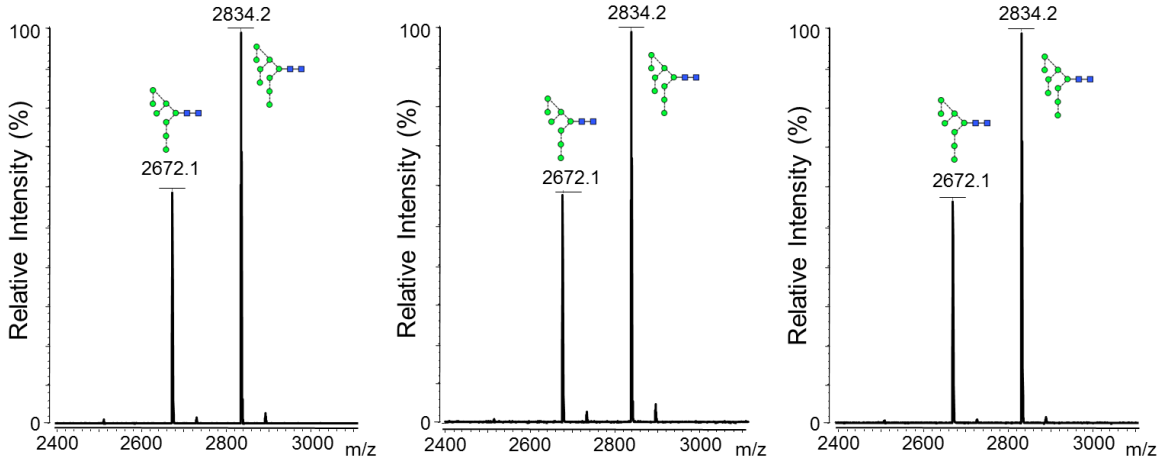


Fig. 5.7. Reproducibility of the method for site-specific glycosylation analysis of the N332 site. N332-glycopeptides (sequence QAHCN^{332}LSR) from three wild-type gp120$_{BaL}$ batches, expressed in parallel, were analysed by MALDI-MS to determine any intrinsic variation of the glycoforms present.

5.5 Conclusions

Given the shifting nature of the HIV-1 glycan shield it is important to determine any widespread effects that loss of individual glycan sites may have on glycan processing. Data reported here suggest that deletion of a single glycan site of gp120 has a limited effect on the overall abundance of oligomannose-type glycans, with the overall density of glycosylation being sufficient to withstand loss of at least one glycan site. However re-distribution of larger oligomannose structures to smaller species was found to be a common phenomenon, indicating that certain mutations can lead to altered processing of nearby sites. This sensitivity of individual oligomannose species to the number of surrounding sites suggests that across different clades, which have different distributions of PNGSs, there may be a considerable degree of heterogeneity in terms of the oligomannose glycans present at a given site. Such microheterogeneity may have consequences for bnAb recognition, and a

robust response will depend upon a tolerance to these subtle variations. These ideas are explored further in the following chapter.

6 (In)tolerance of N332-dependent broadly-neutralising antibodies to glycan microheterogeneity

Data presented in the previous chapters highlight how different factors can influence the glycan profile of HIV-1 gp120. The steric properties of gp120 and the nature of the glycosylation processing machinery mean there is an intrinsic degree of microheterogeneity at individual glycan sites, although this is variable across the molecule. The fine control of glycan processing at each site is a product of the number and distribution of glycan sites, with the density of a given region determining the level of processing that can occur. In order to provide sufficient protection, any broadly neutralising response must be able to accommodate a degree of glycan microheterogeneity at their target epitopes. This chapter will explore the glycan tolerance of a group of potent N332-dependent antibodies, so named because they share overlapping epitopes that include the N332 glycan (Kong et al., 2013; Sok et al., 2014; Walker et al., 2011). These antibodies are highly potent and represent promising vaccine leads, and these experiments sought to probe their precise glycan specificity.

6.1 Effect of PNGS-deletion on N332-dependent bnAbs

In the previous chapter it was found that removal of a single glycan site could alter the glycans present at the N332 site; conversion of $Man_9GlcNAc_2$ to $Man_8GlcNAc_2$ was commonly observed. To determine if such glycoform changes could affect antibody recognition, the panel of $gp120_{BaL}$ PNGS-mutants was screened for binding to 2G12, PGT121, PGT128 and PGT135 (Fig. 6.1).

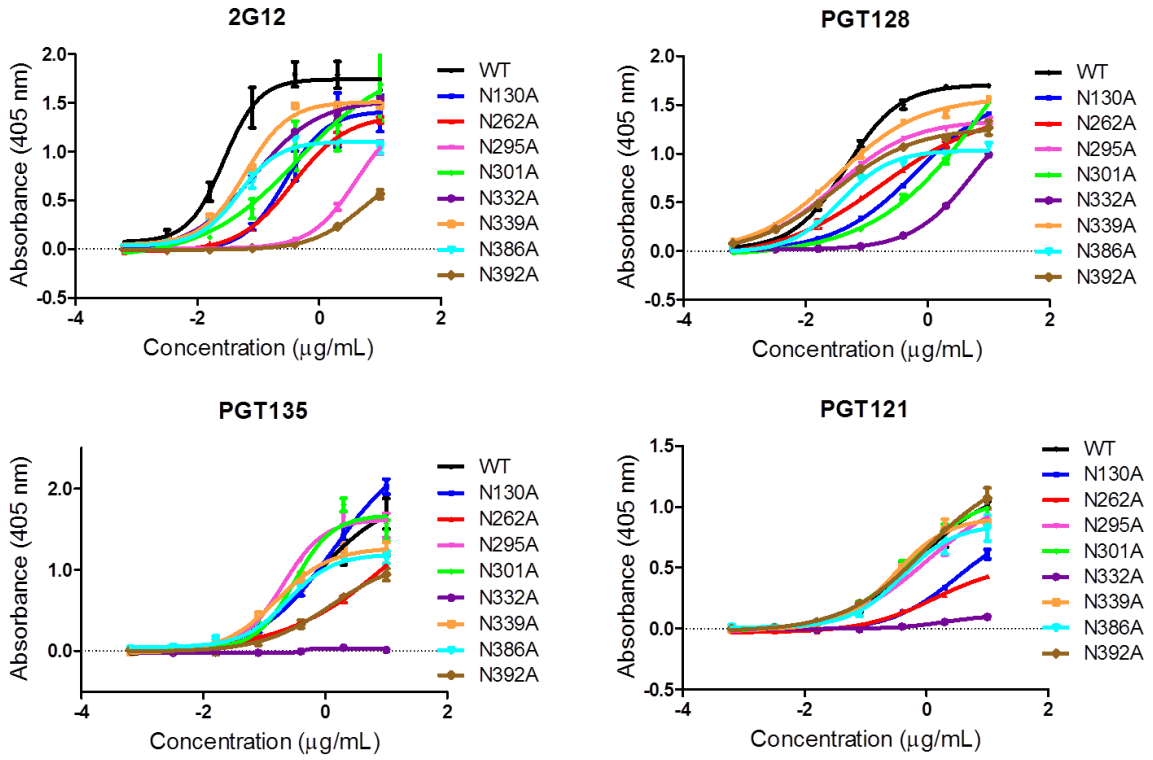


Fig. 6.1. Recognition of PNGS-deletion mutants by N332-dependent bnAbs. Binding of a panel of N332-specific bnAbs, including 2G12, PGT128, PGT135 and PGT121, was tested against the panel of $gp120_{BaL}$ PNGS-deletion mutants (Chapter 5) by capture ELISA. ELISAs were performed in collaboration with Dr Katie Doores (King's College, London).

Binding to the N130A and N262A mutants was reduced for all antibodies and, with the exception of PGT121, binding was also reduced for N386A. However given the indicated alterations in protein conformation for these mutants (chapter 5) it is not possible to interpret the role of their glycan sites in bnAb binding any further. Consistent with previous reports (Calarese et al., 2003; Sanders et al., 2002b; Scanlan et al., 2002), 2G12 binding was reduced in the N332A, N295A, N339A and N392A mutants. The loss of N295 and N392 displayed the largest decrease in binding, consistent with a recent single-particle EM structure of 2G12 Fab in complex with trimeric Env, where modelling of the glycan epitope revealed primary contacts with the N295 and N392 glycans (Murin et al., 2014). The model indicated a weaker secondary interaction with the N332 glycan, mediated through the

unique V_H/V_H' binding site formed by the domain-exchanged configuration of 2G12. A minor role for the N332 glycan in 2G12-binding was supported by the ELISA data, which showed that 2G12 retained significant binding to the N332A mutant. The N339 glycan was also suggested to interact with the V_H/V_H' binding site (Murin et al., 2014), consistent with the reduced binding of the N339A mutant. Binding to the N301A mutant was also reduced, however no role for this glycan was observed in the EM model (Murin et al., 2014). Loss of the N301 glycan was shown to lead to a reduction in $Man_9GlcNAc_2$ structures at the N332 site (Fig. 5.6), and thus the reduced binding may be a result of a reduction in the number of Manα1→2Man-containing structures at the N332 site, or possibly at another of the critical recognition sites (N295 and N392).

The loss of N332 was tolerated less well by the PGT antibodies, with deletion of known additional binding sites also causing reduced binding (N301 for PGT128, N392 for PGT135). Binding by PGT121, which has been observed to recognise both oligomannose- and complex-type glycans (Julien et al., 2013b; Mouquet et al., 2012), was robust across the remaining panel of mutants, consistent with its ability to utilise alternative sites with varying glycan specificities (Sok et al., 2014). In contrast, PGT128 binding was modestly reduced upon loss of the N295 and N392 sites, which were not observed to be contacted by PGT128 in the crystal structure of the antibody-eODmV3 complex (Pejchal et al., 2011). Glycan array data, and the extensive contacts made by PGT128 with the terminal mannose residues of the D1 and D3 arms, support a preferred specificity for $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ glycans, and thus it is possible that the observed decrease in binding is a consequence of altered glycan processing at the N332 site. Binding of PGT135 was significantly reduced in the N332A and N392A mutants, consistent with their observed roles in antibody binding (Kong et al., 2013). Loss of the N295 site, which has been found to be important in a strain-dependent manner (Kong et al., 2013), was tolerated in $gp120_{BaL}$, and in fact resulted in

modestly enhanced binding. A similar result was observed for the N301A mutant. Glycan array data for PGT135, and neutralisation of pseudovirus with altered glycosylation profiles, suggests a preference for $Man_7GlcNAc_2$ and $Man_8GlcNAc_2$ glycans at the N392 site (Kong et al., 2013), and it is possible that the observed increases in binding for some PNGS-deletion mutants arises from a conversion of unfavourable $Man_9GlcNAc_2$ structures at the N392 site to smaller, more favourable structures.

6.2 Effect of PNGS-deletion on virus neutralisation by N332-dependent bnAbs.

To assess the impact of glycan site removal on glycan processing in the context of the HIV-1 envelope trimer, BaL pseudoviruses lacking the N130, N262, N295, N301 or N332 glycans were made and the impact on neutralisation by the panel of N332-dependent bnAbs and b12 was measured (Fig. 6.2). Neutralisation by b12 was unaffected for the N130A, N295A and N332A mutants. However the N301A mutant displayed increased sensitivity, consistent with previous observations that it exposes the CD4 binding site (Binley et al., 2010). An increase in sensitivity was also observed for the N262A mutant, which could similarly arise from local changes in protein conformation increasing accessibility to the CD4 binding site. The N262A mutant was also more sensitive to several of the N332-dependent bnAbs, indicating increased accessibility of these epitopes. In contrast to binding of the gp120 monomer by ELISA, the N130A mutation had no effect on neutralisation by b12, either reflecting differences in the assays or a trimer-specific phenomenon. However reduced plateaus were observed for PGT128 and PGT130, which may indicate altered accessibility to their epitopes, either through changes in protein conformation or neighbouring glycan structures that mediate access.

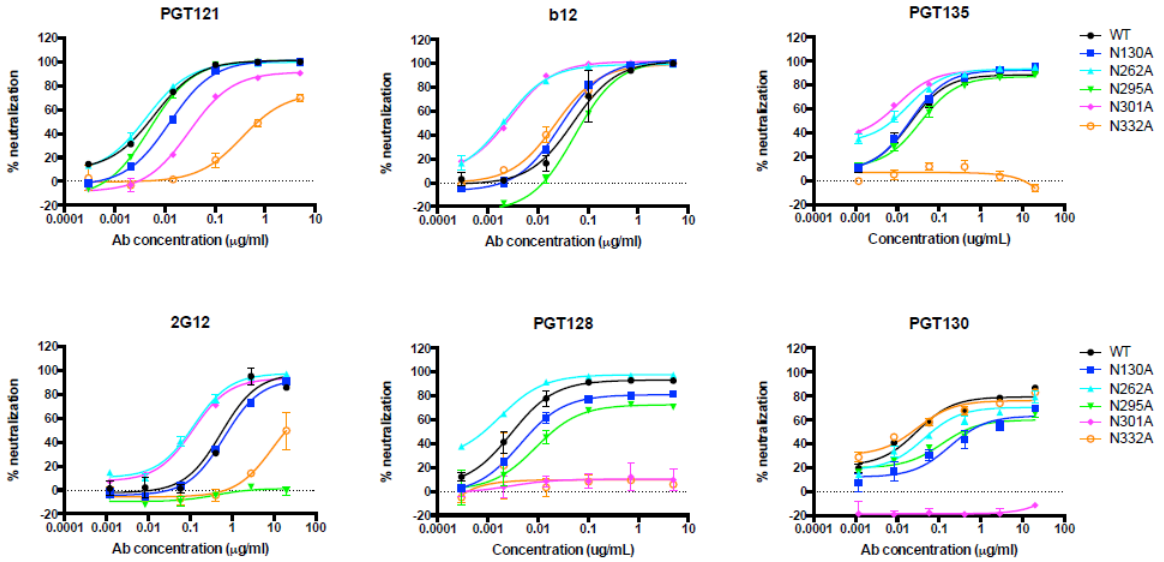


Fig. 6.2. Neutralisation of BaL virus, carrying PNGS-knockout mutations, by N332-dependent bnAbs. Neutralisation potency of N332-specific bnAbs, including PGT121, PGT135, 2G12, PGT128 and PGT130. Neutralisation by conformation-dependent monoclonal antibody b12 was also assessed. Neutralisation assays were performed by Dr Katie Doores.

Loss of glycan sites known to form bnAb epitopes resulted in reduced neutralisation (e.g. N332A for PGT128 and 2G12, N301A for PGT128 and PGT130, and N295A for 2G12). The N332A mutation had a more limited impact on neutralisation by PGT130 and PGT121, presumably arising from the ability of these bnAbs to bind promiscuously to glycans within the high-mannose patch (Sok et al., 2014). Interestingly bnAbs PGT128, PGT130 and PGT135 were unable to neutralise the wild-type virus to 100%, which could indicate the presence of glycoforms that they are unable to bind. In contrast the less selective PGT121 was able to neutralize 100% of the wild-type virus. The N295A mutation was tolerated to varying degrees by the bnAbs, and broadly correlated with the results of the gp120 monomer; neutralisation by PGT121 and PGT135 was maintained, while reduced plateaus were observed for PGT128 and PGT130.

6.3 Microheterogeneity of the PGT135 epitope

The observation that some bnAbs could not neutralise 100% of wild-type virus raises the possibility that, for some of the N332-dependent bnAbs, a population of wild-type pseudoviruses exist that carry unfavourable glycoforms which cannot be effectively targeted. To investigate this hypothesis, site-specific glycan analysis data (chapter 4) for the three sites targeted by PGT135, as revealed by a high-resolution atomic crystal structure (Fig. 6.3) (Kong et al., 2013), were analysed.

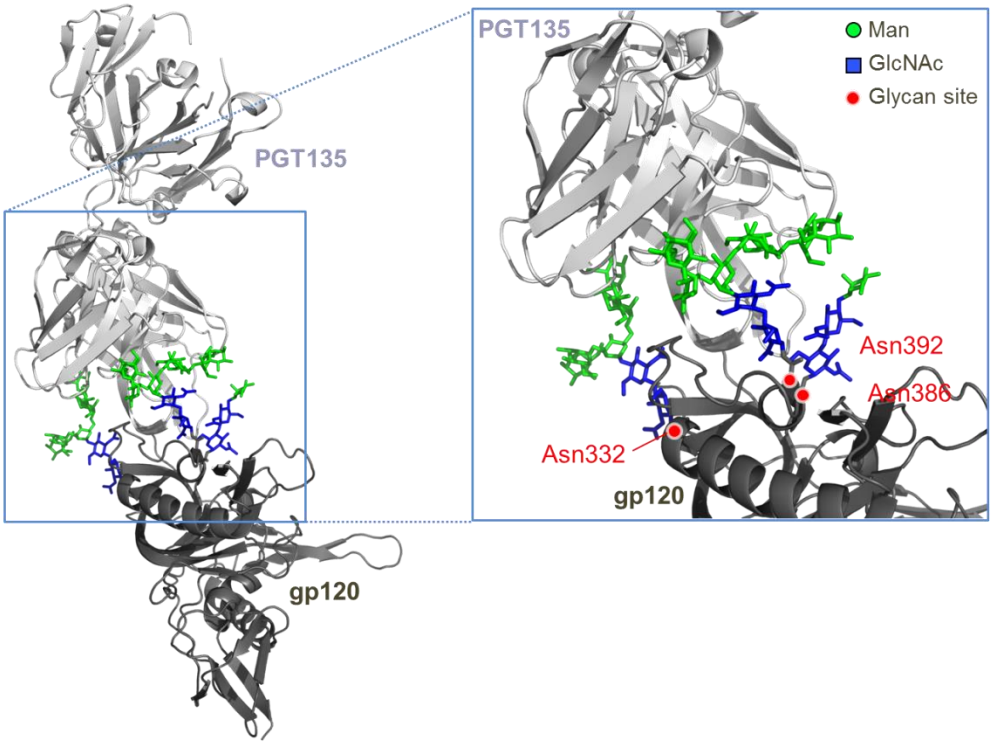


Fig. 6.3. The glycan epitope of PGT135 encompasses the N332, N392 and N386 sites. A previously reported crystal structure reveals the interaction of a PGT135 Fab domain with the N332 ($Man_6GlcNAc_2$), N392 ($Man_8GlcNAc_2$) and N386 ($Man_1GlcNAc_2$) glycans from a $gp120_{JR\text{-}FL}$ core (Kong et al., 2013). Mannose (Man) residues are coloured in green and N-acetylglucosamine residues (GlcNAc) are coloured in blue. Image was made in PyMol using PDB ID: 4JM2.

Glycopeptides containing the target glycan sites were isolated from soluble recombinant $gp120_{BaL}$ by RP-HPLC fractionation as described in chapter 4. N332-containing glycopeptides (sequence $QAHCN^{332}LSR$) were isolated in a fraction from

a tryptic digest, while N386 (YCN^{386}STQLF) and N392 (N^{392}STW) glycopeptides were isolated from the chymotryptic digest. Fractions of glycopeptides were analysed by MALDI-MS (Fig. 6.4) and the relative abundances of different oligomannose-type glycans quantitated (Table 6.1). This revealed the glycans at the N332 site to be overwhelmingly dominated by $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ glycans, with very low levels of $Man_5GlcNAc_2$, $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$. The N392 site was primarily populated by $Man_8GlcNAc_2$ (62%), with lower levels of $Man_5GlcNAc_2$ and $Man_9GlcNAc_2$, and trace amounts of $Man_4GlcNAc_2$, $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$. Meanwhile the N386 site displayed a more even distribution of oligomannose-type glycans (Fig 6.4; Table 6.1).

The oligomannose-type nature of the glycans at the N332, N392 and N386 sites is consistent with their location on the highly glycosylated outer domain. However despite the more restricted microheterogeneity at these sites, there remains a distribution of the individual oligomannose-type glycans, which may affect the capacity for PGT135 binding given its glycan specificity. High density glycan microarray data showed that PGT135 bound $Man_7GlcNAc_2$, $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$, while results from a neoglycolipid microarray demonstrated binding to $Man_7GlcNAc_2$ and $Man_8GlcNAc_2$ only (Kong et al., 2013). This preference for high oligomannose-type glycans, at least at the N392 site, was rationalised by the extensive interactions of PGT135 along the full length of the D1 arm of a $Man_8GlcNAc_2$ glycan present in the crystal structure (Kong et al., 2013).

While the population of smaller oligomannose glycans at the N392 site might not capitalise on all possible binding interactions with PGT135, their presence does not necessarily preclude binding. However the plateau in neutralisation observed here (Fig. 6.2) and elsewhere (Kong et al., 2013), at around 85-90% for BaL pseudovirus, suggests there is a population of virus that cannot be bound effectively. Modelling of a $Man_9GlcNAc_2$ glycan at the N392 site revealed that the presence of a terminal D2 mannose residue would induce a

steric clash with the CDR H3 loop of PGT135 (Kong et al., 2013). The results presented here reveal that 16% of the glycans at the N392 site of recombinant $gp120_{BaL}$ were $Man_9GlcNAc_2$, which corresponds well with the observed neutralisation incapacity against BaL pseudovirus. This constraint - excluding $Man_9GlcNAc_2$ at the N392 site - is supported by the significant decrease in neutralisation observed upon kifunensine treatment (Kong et al., 2013), which produces exclusively $Man_9GlcNAc_2$ glycans (Elbein et al., 1990). In contrast modelling of a $Man_9GlcNAc_2$ at the N332 site suggested that a D2 terminal mannose could be accommodated with a minor conformational change, and thus the microheterogeneity observed at this site is likely to be better tolerated. The interaction between PGT135 and the glycan at N386 was observed to make a much smaller contribution (Kong et al., 2013), suggesting that an imperfect glycan match at this site would not be as detrimental to binding.

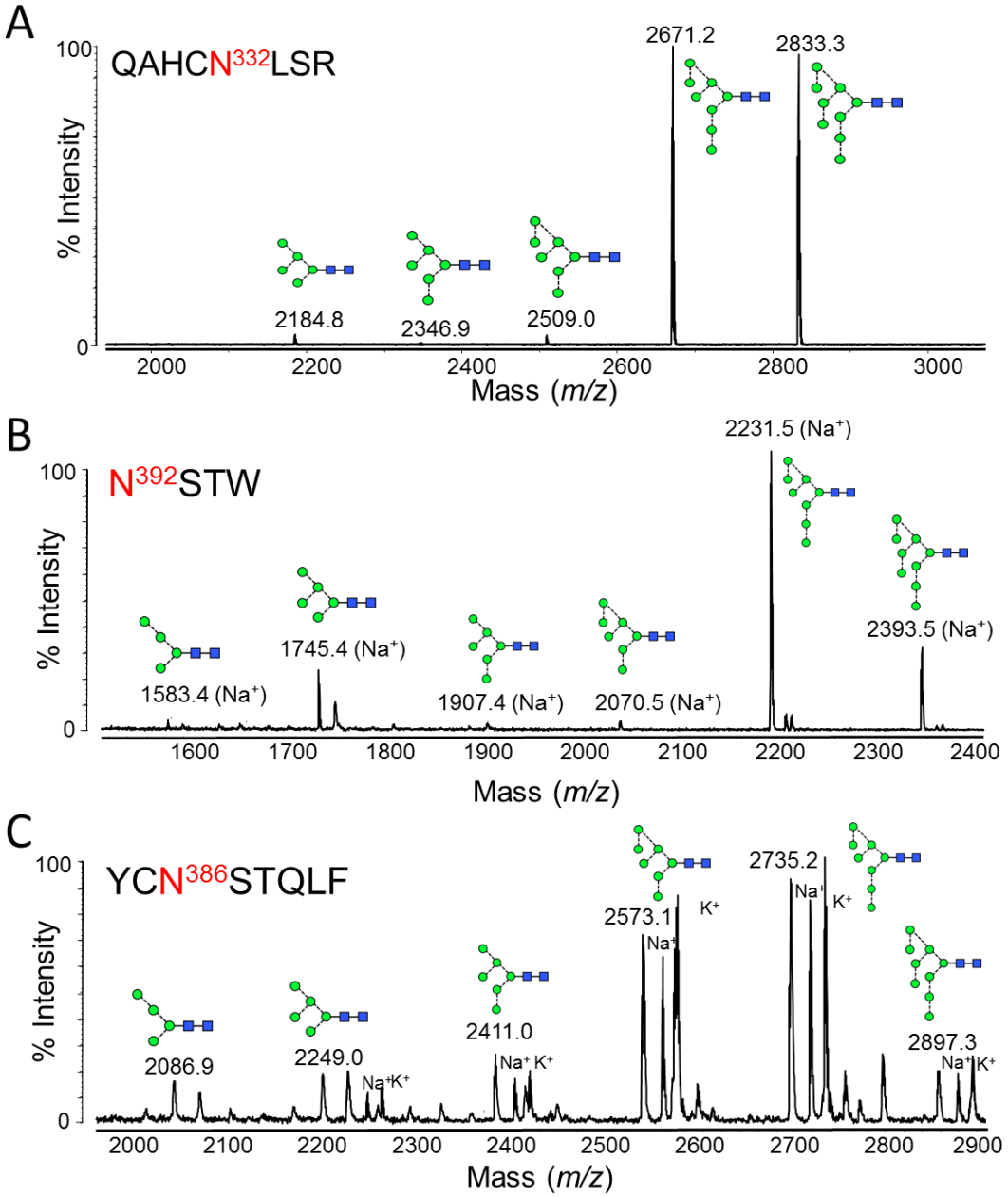


Fig. 6.4. MALDI-MS analysis of the PGT135 glycan epitope. Gp120 was digested with trypsin or chymotrypsin before RP-HPLC and MALDI-MS analysis. (A) MALDI-MS of pooled fractions containing the N332-containing glycopeptide, QAHCNLSR. The glutamine carried a pyro-Glu modification (-17), and the cysteine was modified due to treatment with iodoacetamide (carbamidomethyl, +57). Glycan structures corresponding to the observed glycopeptide masses are indicated. (B) MALDI-MS of N392 glycopeptide, NSTW. Sodium, $[M+23]^+$, and potassium, $[M+39]^+$, adducts were observed: both peaks were used for measuring abundances. (B) MALDI-MS of N386 glycopeptide, YCNSTQLF. The cysteine was modified due to treatment with iodoacetamide (carbamidomethyl, +57). Protonated glycopeptides, as well as sodium and potassium adducts, were detected: all were used for calculating abundances.

Table 6.1. Summary of glycan structures and their abundances found at the epitope of PGT135 by MALDI-MS.

Site	Peptide	Peptide mass $[M+H]^+$	m/z		Glycan	% of total
			Obsv.	Calc.		
N332	QAHCNLSR[§¶]	968.4	2184.8[a]	2184.9	M5	1.7
			2346.9[a]	2346.9	M6	0.4
			2509.0[a]	2509.0	M7	1.3
			2671.2[a]	2671.0	M8	49.4
			2833.3[a]	2833.1	M9	47.2
N392	NSTW	507.2	1583.4[b]	1583.6	M4	1.1
			1745.4[b]	1745.6	M5	18.8
			1907.4[b]	1907.7	M6	0.6
			2069.5[b]	2069.7	M7	1.9
			2231.5[b]	2231.8	M8	61.8
			2247.4[c]	2247.9		
			2393.5[b]	2393.8	M9	15.9
			2409.4[c]	2409.9		
N386	YCNSTQLF[¶]	1032.4	2086.9[a]	2086.8	M4	2.0
			2249.0[a]	2248.8	M5	6.5
			2270.9[b]	2270.8		
			2286.9[c]	2286.9		
			2411.0[a]	2410.9	M6	6.8
			2433.0[b]	2432.9		
			2449.0[c]	2449.0		
			2573.1[a]	2573.0	M7	33.4
			2595.1[b]	2594.9		
			2611.1[c]	2611.0		
			2735.2[a]	2735.0	M8	41.9
			2757.1[b]	2757.0		
			2773.1[c]	2773.1		
			2897.3[a]	2897.1	M9	9.4
			2919.2[b]	2919.0		
			2935.2[c]	2935.1		

[a] $[M+H]^+$
[b] $[M+Na]^+$
[c] $[M+K]^+$
[§] Pyro-Glu modification of glutamine (-17)
[¶] Carbamidomethyl modification of cysteine (+57)

6.4 Conclusions

Elucidating the glycan specificity of known bnAbs is an important step towards vaccine design, and data presented here provide insights into the glycan-reactivity of a group of N332-dependent bnAbs. ELISA data from the $gp120_{BaL}$ PNGS-deletion mutants revealed varying levels of tolerance to sequence variation by different bnAbs. Antibodies with known specificities for the larger oligomannose-type glycans, for example 2G12 and PGT128, were found to be sensitive to loss of certain glycan sites, including those that had been demonstrated to influence the processing of the N332 site in chapter 5. In comparison PGT121 appears to possess a broader glycan specificity, since its binding was unaffected by mutations that altered glycan processing at the N332 site.

The efficacy of a given bnAb is therefore a product of more than simply the prevalence of its target glycan sites; the presence of susceptible glycans at the necessary sites is a further requirement. A recent study highlighted the promiscuity of N332-dependent antibodies in recognising alternative sites when the N332 site was absent, which was suggested to derive from an ability to alter binding conformation (Sok et al., 2014). An outstanding question is whether this flexibility may be extended, in some circumstances, to accommodate non-ideal glycoforms. The crystal structure of PGT135 in complex with its gp120 glycan epitope indicates that the presence on $Man_9GlcNAc_2$ precludes binding (Kong et al., 2013), however other bnAbs, such as PGT121, may be more tolerant in their glycan specificity. Indeed PGT135 displayed a much lower breadth of neutralisation (Walker et al., 2011) which, while attributed to the lower prevalence of its target glycan sites (Kong et al., 2013), may additionally reflect a higher selectivity in glycan recognition. Nonetheless the tolerance of bnAbs to glycan microheterogeneity remains an outstanding question and warrants further investigation.

7 Characterisation of candidate recombinant immunogens

It is widely believed that an effective prophylactic vaccine against HIV-1 will include a recombinant component based on the surface envelope spike (Env). Vaccination trials to date have typically used recombinant, monomeric gp120 or soluble gp140 constructs, which have had little success (Barouch, 2013). It is hypothesised that this may in part be due to the presence of epitopes that are targeted by non-neutralising antibodies, as well as the absence of broadly neutralising epitopes, on the monomer (Burton et al., 2012). However efforts to produce recombinant soluble trimers as immunogens have been hampered by the intrinsic instability of Env constructs bearing a deleted membrane-spanning domain (Ringe et al., 2013). Efforts to overcome these problems have led to the development of BG505 SOSIP.664 gp140 trimer constructs, which are currently best-in-class and produce stable and homogeneous, cleaved trimers (Chung et al., 2014; Ringe et al., 2013; Sanders et al., 2013; Yasmeen et al., 2014). These trimers appear to resemble virus envelope spikes by negative stain electron microscopy (Sanders et al., 2013), and antigenicity studies have demonstrated that they bear a number of important bnAb epitopes, while the majority of known non-neutralising epitopes appear to be absent or occluded (Sanders et al., 2013).

The favourable thermodynamic and antigenic properties of BG505 SOSIP.664 trimers have supported their advancement to immunogenicity trials, which are currently underway in animal models. Stable CHO and 293T cell lines have recently been described, which are capable of producing sufficiently high yields of homogenous BG505 SOSIP.664 trimers on a scale required for current and future clinical studies (Chung et al., 2014). Given the role of glycans in many bnAb epitopes, the glycosylation of these potential immunogens

is an important consideration. In this chapter HILIC-UPLC was used to characterise the N-linked glycans present on BG505 SOSIP.664 gp140 trimers and related constructs, which revealed they are promising mimics of the native HIV-1 glycan shield.

7.1 Glycosylation of BG505 SOSIP.664 gp140 trimers produced in 293T and CHO cells.

BG505 SOSIP.664 trimers contain 28 potential N-linked glycosylation sites (PNGSs): 24 distributed across gp120 and a further 4 within the $gp41_{ECTO}$ domain (Fig. 7.1A). The N332 glycan site was introduced since it is not present in the native BG505 virus. When resolved by SDS-PAGE, gp140s expressed in CHO and 293T cells migrated at similar molecular weights (Fig. 7.1B). Upon treatment with dithiothreitol (DTT), the gp120 and $gp41_{ECTO}$ domains were resolved, confirming that the trimers were completely cleaved. While the gp120 domains appeared to migrate together, the $gp41_{ECTO}$ domain from the CHO-derived trimer migrated slightly faster than the 293T-derived material, potentially reflecting differing glycan compositions. Following SDS-PAGE, N-linked glycans were released directly from gel bands using PNGase F. After fluorescent labelling with 2-AA the glycans were analysed by HILIC-UPLC.

The profiles of the gp140 trimers from CHO and 293T cells were remarkably similar and were both dominated by oligomannose-type glycans (Fig. 7.2). Separation of the gp140s into their corresponding gp120 and $gp41_{ECTO}$ subunits revealed that glycans from gp120 accounted for the majority of the gp140 profiles, consistent with their greater number of glycan sites. Integration of the peaks revealed that the oligomannose-type glycans accounted for 65% and 63% of the total gp120 glycan population for CHO and 293T-derived trimers, respectively (Table 7.1). $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ were particularly dominant, together representing 43% and 49% of CHO and 293T-derived glycans, respectively. This

contrasted with the smaller populations of oligomannose-type glycans on the $gp41_{ECTO}$ domains, which were dominated by complex-type glycans. There were also notable differences between the structures of the complex-type glycans, with the CHO material containing a higher proportion of glycans bearing one or more terminal sialic acid residues. Digestion with neuraminidase and integration of the HILIC-UPLC profiles revealed that while approximately 30% of CHO-derived gp140 glycans contained sialic acid, only 7% of the glycans on 293T gp140 contained any sialic acid (Fig. 7.3). These differences could explain the differing mobilities of the $gp41_{ECTO}$ domains by SDS-PAGE (Fig. 7.1); the $gp41_{ECTO}$ domains are dominated by complex-type glycans and it has previously been shown that a greater level of sialylation on a glycoprotein correlates with increased electrophoretic mobility (Durocher et al., 1975). However the differing mobilities could also reflect differences in glycan site-occupancy between the two cell lines.

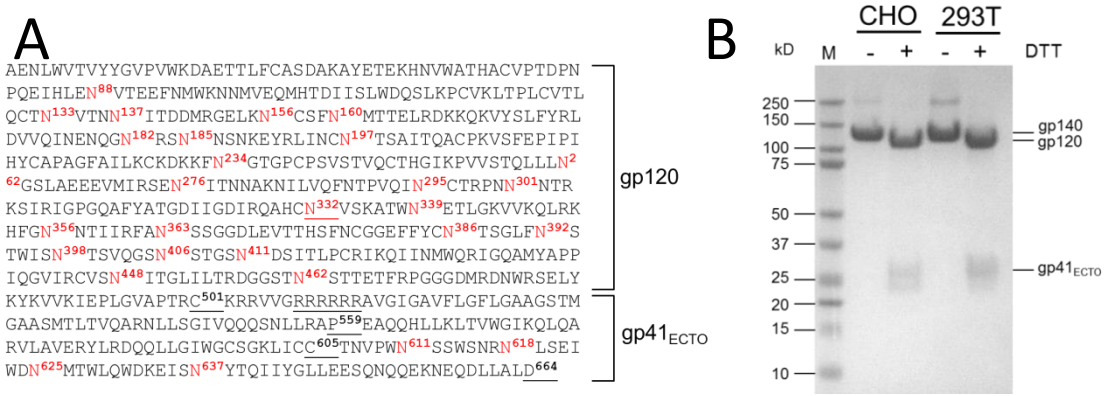


Fig. 7.1. Overview of the BG505 SOSIP.664 gp140 trimers. (A) Protein sequence of the BG505 SOSIP.664 gp140 trimer construct. The gp120 and $gp41_{ECTO}$ domains are indicated. Potential N-linked glycosylation sites (PNGSs) are highlighted in red (HXB2 numbering). Key features are underlined; A T332N mutation introduces a PNGS at the 332 site; The cysteine residues at 501 and 605 form a disulphide bridge to secure the gp120 and $gp41_{ECTO}$ domains together; Replacement of the gp120 furin cleavage site (RXXR) with a hexaarginine (R6) sequence promotes furin cleavage; An I559P mutation improves stability between gp41 subunits; The MPER region is deleted, with the sequence terminated at residue 664, to improve solubility of the trimers. (B) SDS-PAGE analysis of the BG505 SOSIP.664 gp140s produced from stable CHO and 293T cell lines, under non-reducing and reducing conditions. Trimers were purified by 2G12-affinity chromatography followed by size exclusion chromatography (SEC), as previously described (Chung et al., 2014). DTT, dithiothreitol; M, markers. Trimers were provided by Dr Albert Cupo/Prof. John Moore (Weill Cornell Medical University).

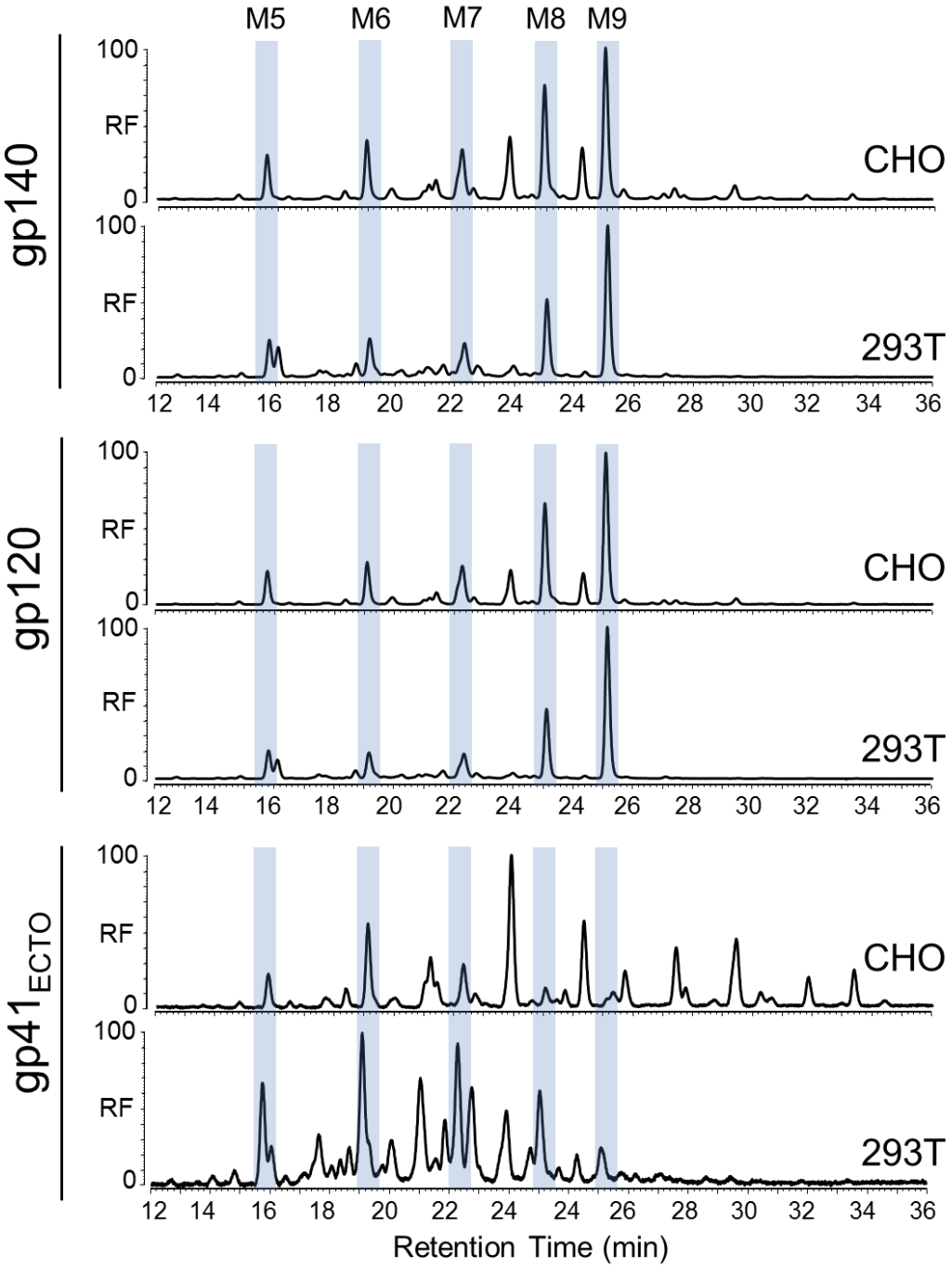


Fig. 7.2. HILIC-UPLC glycan profiles of the BG505 SOSIP.664 gp140 trimers from CHO and 293T cells. Following SDS-PAGE, N-linked glycans were released from CHO- and 293T-produced gp140 trimers by in-gel PNGase F digestion. Trimers were analysed from gels run under non-reducing (gp140) and reducing (gp120 and $gp41_{ECTO}$) conditions. Released glycans were labelled with 2-AA and resolved by HILIC-UPLC. Peaks corresponding to oligomannose-type glycans are highlighted in blue. M5-9, $Man_{5-9}GlcNAC_2$; RF, relative fluorescence.

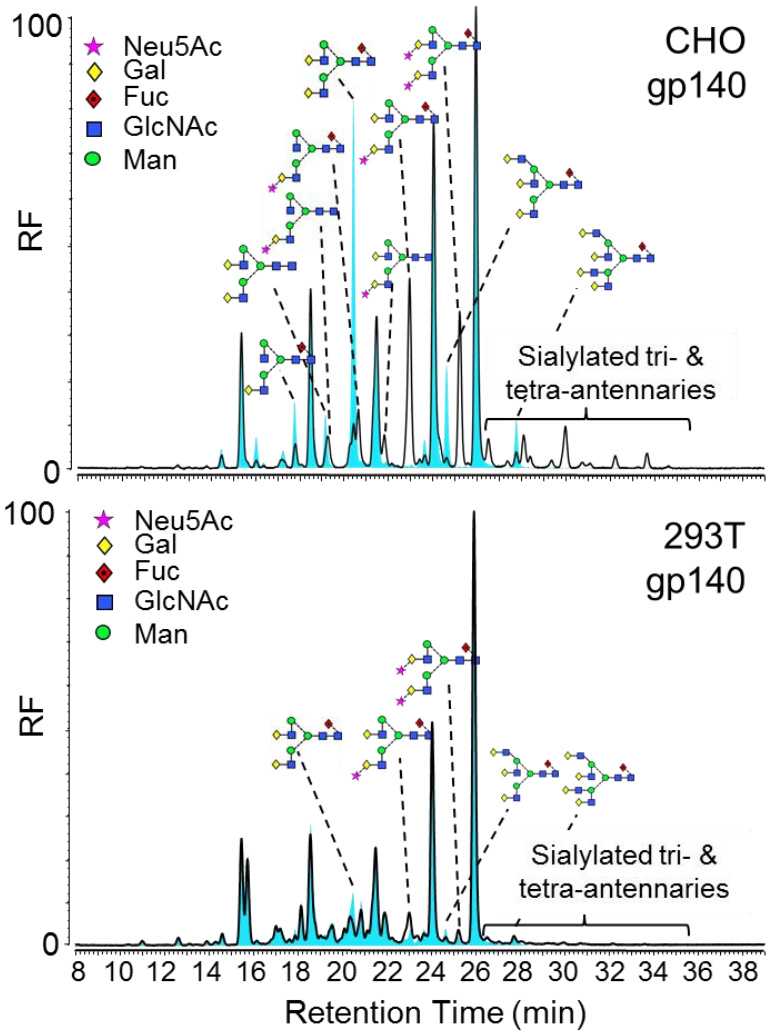


Fig. 7.3. Neuraminidase digestion of gp140 from BG505.664 trimers produced in CHO and 293T cells. Glycans from CHO and 293T-derived gp140s were digested with α2-3,6,8 neuraminidase from *Clostridium perfringens*. The black line represents the undigested glycan profile, and is overlaid with the neuraminidase-digested profile, coloured in cyan.

7.2 BG505 SOSIP.664 glycan profiles are independent of the glycan specificity of the bnAb used for purification.

The BG505 SOSIP.664 trimers characterised above and in previous reports were purified using 2G12 affinity chromatography prior to size exclusion chromatography (Chung et al., 2014; Sanders et al., 2013). Given the specificity of 2G12 for terminal Manα1→2Man residues (Scanlan et al., 2002), it is possible that the observed glycan profiles for the trimers, which are dominated by terminal Manα1→2Man-containing structures, are an artefact of glycan-biasing during the purification procedure. To investigate this BG505 SOSIP.664

trimers, produced in CHO cells, were purified by PGT145 or PGT151, followed by SEC, and their glycan compositions analysed (Fig. 7.4). The glycan profiles were almost identical to that seen following 2G12 purification, with each gp120 comprising approximately two-thirds oligomannose-type glycans (Table 7.1). This convergence of the glycan profiles, using bnAbs with diverse glycan specificities (Blattner et al., 2014; Falkowska et al., 2014; Sanders et al., 2002b; Scanlan et al., 2002; Walker et al., 2011; Table 1.2), indicates that these glycan profiles represent the true, unbiased, glycosylation pattern of soluble, cleaved BG505 SOSIP.664 trimers. It also suggests that there is a considerable degree of homogeneity in terms of the glycoforms present, since it might be expected that 2G12 affinity purification would highlight the presence of any glycoforms that have a higher proportion of Manα1→2Man-containing structures.

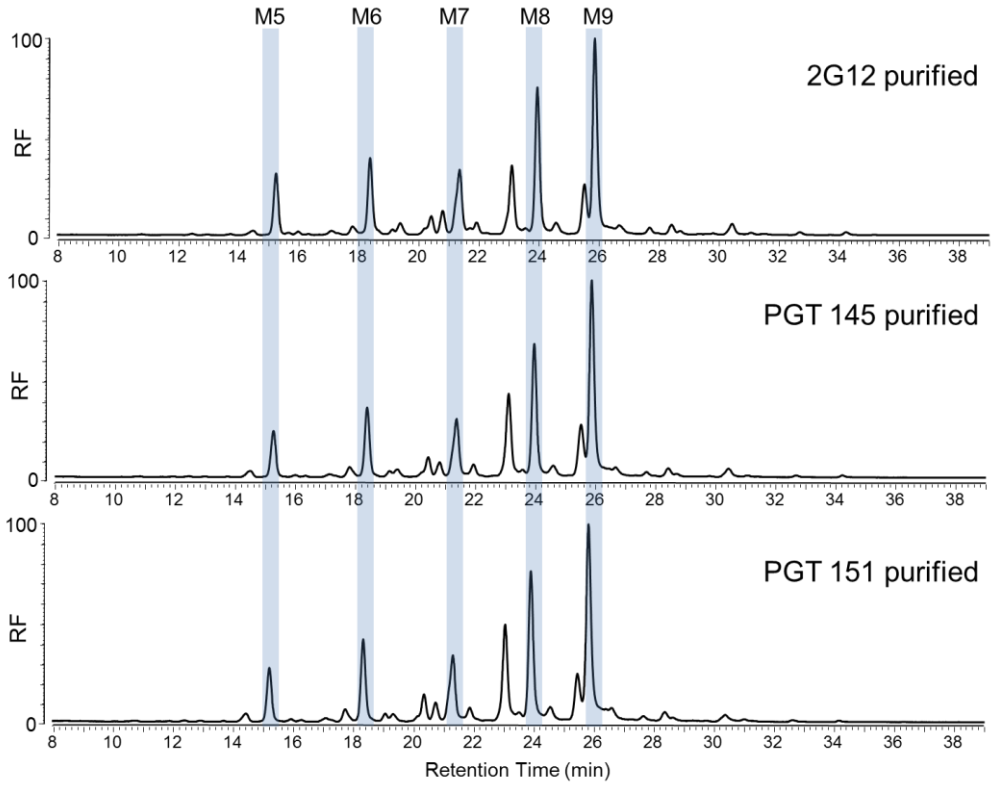


Fig. 7.4. Convergence of gp140 trimer glycan profiles following affinity purification. Glycan HILIC-UPLC chromatograms of BG505 SOSIP.664 gp140 trimers following affinity purification using bnAbs 2G12, PGT145 or PGT151. Trimers were additionally purified by SEC. Purified trimer samples were obtained from Dr Albert Cupo/Prof. John Moore, and were expressed using stable CHO cells as previously described (Chung et al., 2014).

7.3 Steric factors influence the degree of glycan processing

Comparison of the BG505 SOSIP.664 trimer gp120 glycan profile with the profile obtained for recombinant monomeric BG505 gp120 (chapter 3; Fig. 7.5) highlights the effect of oligomerisation on the processing of gp120 glycans. The monomer had approximately a third less total oligomannose than the trimer, and approximately 60% less $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ combined (Table 7.1). This suggests that trimerisation has a protective effect against glycan processing, and occurs rapidly enough to prevent significant trimming by α-mannosidases.

A number of potential immunogens lack the furin cleavage site as an alternative to disulphide stabilisation, in order to maintain the structural integrity of the trimer. Negative-stain electron microscopy experiments have revealed that in contrast to cleaved trimers, uncleaved gp140 trimers appear as highly heterogeneous mixtures bearing irregular shapes (Chung et al., 2014; Ringe et al., 2013). To assess whether these observed structural differences are reflected in altered glycosylation patterns, glycan analysis of uncleaved BG505 WT.SEKS gp140 trimers was performed. These trimers are based on the wild-type BG505 sequence, but incorporate the T332N mutation and lack the furin-cleavage site due to a REKR→SEKS mutation (Ringe et al., 2013). The trimers were found to undergo a greater degree of glycan processing (Fig. 7.5), containing approximately a quarter less oligomannose-type glycans than the cleaved SOSIP.664 equivalents (Table 7.1). This is consistent with negative stain-EM images which show that the gp120 subunits of the uncleaved trimers appear to splay out from the $gp41_{ECTO}$ core (Ringe et al., 2013). This open structure presumably allows greater access by processing enzymes, leading to a glycosylation profile that more closely resembles the recombinant monomer. A similarly highly-processed glycan profile was also observed for another candidate immunogen, CZA97.012, which comprises uncleaved, gp140 trimers that contain a fold-on domain to

promote oligomerisation (Fig. 7.5) (Go et al., 2014; Kovacs et al., 2012). By negative-stain EM these trimers display highly irregular, splayed structures (Dr Andrew Ward, unpublished data), and the increased level of processing observed in their glycosylation profile is consistent with a more open structure that is accessible to glycosidases.

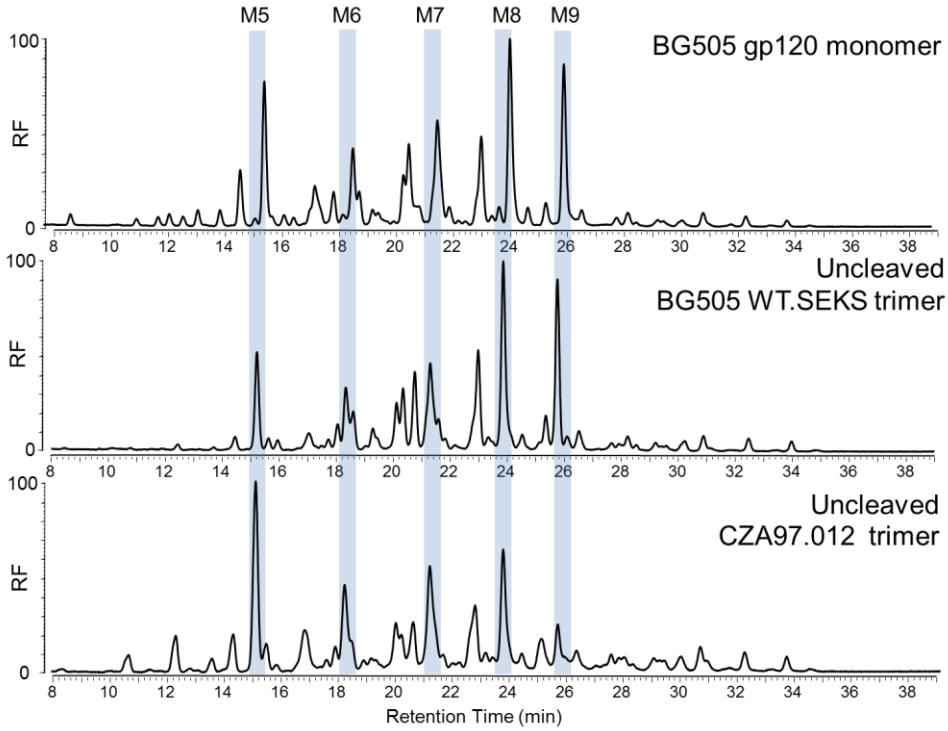


Fig. 7.5. Resemblance of uncleaved trimers to monomeric gp120 glycosylation. Glycans were released from SDS-PAGE bands, labelled and analysed as described above. Recombinant monomeric BG505 was produced and Ni^{2+}-NTA-purified as described in chapter 3. BG505 WT.SEKS trimers were 2G12/SEC purified, and CZA97.012 trimers were Ni^{2+}-NTA/SEC purified. Both were obtained from Dr Al Cupo/Prof. John Moore. All three constructs were produced in 293T cells.

7.4 The effect of membrane-tethering on glycan processing

While recombinant immunogens to date have predominantly been soluble constructs, transmembrane constructs have also been investigated for their use in eliciting neutralising antibodies. To investigate how glycan processing of membrane-tethered HIV-1 Env trimers may differ to soluble forms, glycan analysis was performed on membrane-tethered BG505.705 trimers, expressed on the surface of 293F cells, detergent-extracted and

PGT151-affinity purified (Fig. 7.6) (Blattner et al., 2014). These constructs are based on the wild-type BG505 sequence, but contain a truncated C-terminal domain (terminating in residue 705). The trimers displayed a high abundance of oligomannose (approximately 50%) with $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ together accounting for 33% of the total glycans (Table 7.1). Comparison with the soluble BG505 SOSIP.664 trimers indicates that the membrane-tethered trimers undergo a slightly greater degree of processing; the level of oligomannose was slightly lower while the abundance of highly branched and sialylated structures was greater. This is consistent with the closer proximity of membrane-tethered trimers to glycosyltransferases, which are typically type II transmembrane proteins (Breton et al., 2006). A greater degree of accessibility to processing enzymes is also reflected in the processing of the gp41 glycans; these were found to mostly be large, heavily-processed sialylated tri- and tetra-antennaries, which were present in lower abundances in the soluble SOSIP.664 constructs.

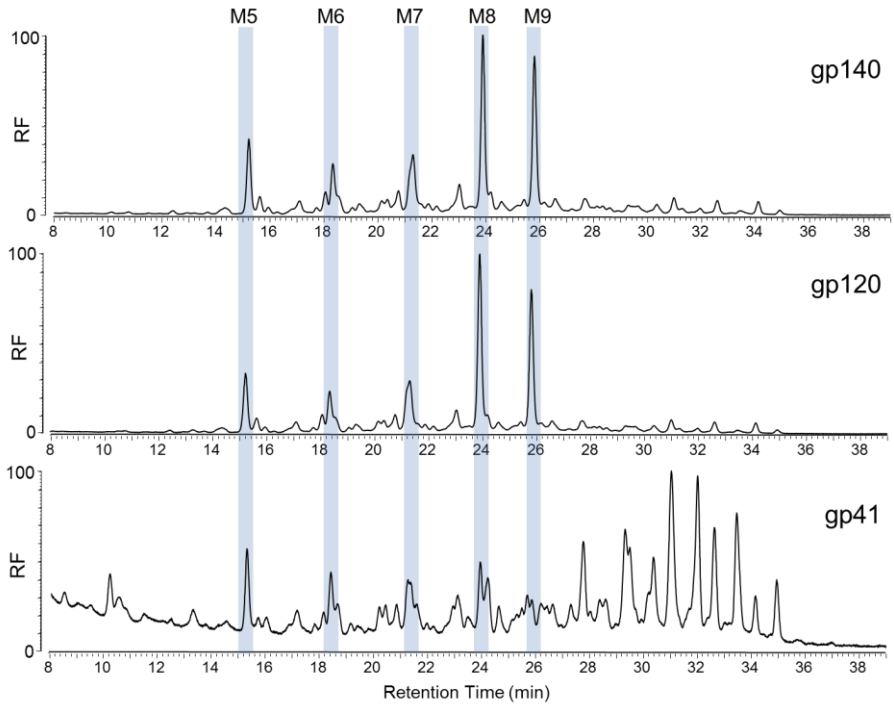


Fig. 7.6. Glycosylation of membrane-tethered BG505.705 gp140 trimers. HILIC-UPLC chromatograms of glycans released from membrane-tethered BG505.705 trimers. Trimers were produced in 293F cells, affinity-purified using PGT151 followed by SEC, and were obtained from Dr Andrew Ward (The Scripps Research Institute).

Table 7.1. Summary of the abundances of oligomannose-type glycans across different recombinant trimer constructs

Env protein	Producer cell	Affinity purification scheme	Abundance (% of total glycans)					
			M5[‡]	M6	M7	M8	M9	**Total**
$gp120_{BG505}$ from recombinant monomer								
$gp120_{BG505}$	293T	Ni^{2+}-NTA	6	6	8	10	9	**39**
$gp120_{BG505}$ from cleaved trimers								
SOSIP.664	CHO	PGT151	5	7	8	18	27	**66**
SOSIP.664	CHO	PGT145	5	7	8	17	27	**65**
SOSIP.664	CHO	2G12	6	7	9	18	25	**65**
SOSIP.664	293T	2G12	4	4	6	15	34	**63**
BG505.705	293F	PGT151	6	5	8	18	15	**51**
$gp41ECTO_{BG505}$ from cleaved trimers								
SOSIP.664	CHO	2G12	3	8	5	1	1	**17**
SOSIP.664	293T	2G12	6	11	12	5	1	**36**
Cleaved & uncleaved trimers[§#]								
SOSIP.664	CHO	2G12	6	8	9	15	20	**57**
SOSIP.664	293T	2G12	5	4	8	14	27	**59**
WT.SEKS	293T	2G12	6	6	9	13	11	**45**
CZA97.012	293T	Ni^{2+}-NTA	10	4	7	7	3	**30**

[‡]Mx refers to $Man_xGlcNAc_2$ glycans.
[§]Glycan profiles for gp140 constructs include contributions from both gp120 and $gp41_{ECTO}$ domains.
[#]All the Env proteins are based on the BG505 sequence except for the CZA97.012 gp140. Glycan profiles of gp140 include contributions from both gp120 and $gp41_{ECTO}$ subunits.

7.5 Conclusions

It has been argued that the success of a recombinant HIV-1 immunogen will require a faithful replication of the antigenic surface of native viral Env (Burton et al., 2012). In light of the number of bnAbs which are reactive to Env glycans, an important consideration remains the resemblance of the immunogen glycan profile to that of the native virus. Previous glycan characterisation of virus grown in PBMCs revealed a high abundance of oligomannose-type glycans (Bonomelli et al., 2011; Doores et al., 2010a), which is

replicated in the glycan profiles of cleaved BG505 SOSIP.664 trimers independently of the method of bnAb-affinity purification. In contrast monomeric gp120, and irregularly-shaped, uncleaved trimers, displayed much greater levels of processing, consistent with the notion that trimerisation induces additional steric constraints beyond those intrinsic to the monomer. This phenomenon has been referred to as the 'trimer-induced mannose patch' (Bonomelli et al., 2011; Doores et al., 2010a), and recent studies have revealed some of the sites where glycan processing is altered in the context of the trimer (Guttman et al., 2014; Yang et al., 2014). Plotting of these onto the trimer surface reveals that the glycans that appear to be affected (N156, N160, N197 and N301) lie at the trimer apex where the monomers form an interface (Fig. 7.7).

It is interesting to note that the BG505 SOSIP.664 trimers are not only dominated by oligomannose-type glycans, but $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ in particular. These structures represent a very limited degree of glycan processing, indicating that oligomerisation must happen rapidly in the ER in order to protect against ER α-mannosidase processing. Since Env is thought to spend several hours within the ER (Land et al., 2003), this also suggests that the trimers have a high degree of stability once trimerised, thereby preventing access to α-mannosidases. Differences in folding times may help explain the slightly higher degree of processing observed for the membrane-tethered BG505 construct, since it has been shown that gp160 takes longer to achieve its native conformation compared to gp120 (Land et al., 2003). A longer folding time would provide glycosidases with greater opportunity to access their substrates. In contrast the uncleaved trimers that were examined displayed more substantial processing, and thus either fail to oligomerise quickly enough to confer resistance to processing, or oligomerise rapidly but fail to fully fold at the apex.

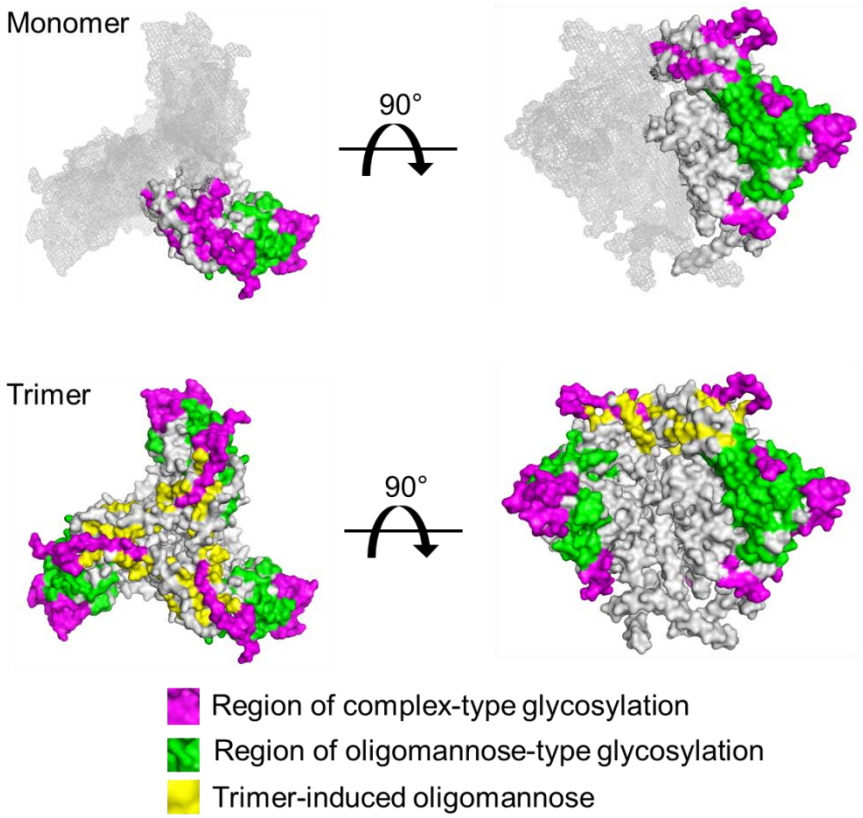


Fig. 7.7. Steric constraints from trimerisation lead to elevated levels of oligomannose. The Env protein surface is shown in grey. Sites containing complex-type glycans are shown in pink and sites containing oligomannose-type glycans are shown in green. Sites are allocated based on site-specific glycosylation analysis of recombinant monomeric gp120. Sites thought to contain oligomannose-type glycans within the context of the trimer are indicated in yellow (Guttman et al., 2014; Yang et al., 2014). Image was made in PyMol using PDB: 4NCO (Julien et al., 2013a).

In comparison to the BG505 SOSIP.664 and BG505.705 trimers, a dominance of $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ glycans was not evident from the characterisation of PBMC-derived Env, which showed a more even distribution of the different oligomannose-type structures (Bonomelli et al., 2011; Doores et al., 2010a). This could reflect cell-specific differences in glycan processing, or properties of the full length gp160. However in these studies a cocktail of non-quaternary-dependent nAbs was used for purification, which could result in the inclusion of misfolded or non-trimeric gp120, which undergo greater processing, in the analysis. Alternatively the observed distribution could arise from the use of MALDI-MS for analysis, since larger glycans ionise less efficiently than smaller glycans. Preliminary analysis of gp120 derived from JRCSF pseudovirus grown in 293T cells, and

purified using PGT151-afffinity chromatography, revealed a dominance of the larger oligomannose structures, in particular $Man_9GlcNAc_2$ (Appendix V). A similar result was observed for $gp120_{JRCSF}$ isolated from cell-surface trimers. The similarity of these profiles with those obtained from the SOSIP.664 and BG505.705 gp140 trimers, suggests that this dominance of the largest $Man_8GlcNAc_2$ and $Man_9GlcNAc_2$ glycans may be a hallmark of native, properly folded HIV-1 Env trimers.

BG505 SOSIP.664 trimers are currently being developed for use in clinical trials (Chung et al., 2014), and here the glycan compositions of BG505 SOSIP.664 gp140 trimers produced from stable CHO and 293T cell lines were assessed. Both cell lines are options for immunogen production under good manufacturing practice (GMP), and thus information on their glycosylation patterns could help inform future manufacturing decisions. Consistent with their matched antigenic properties, the glycan profiles from the CHO and 293T cells were found to be very similar with each dominated by a population of oligomannose-type glycans. While this analysis cannot exclude differences in glycosylation at the site-specific level, a recent study comparing CHO- and 293T-produced monomeric gp120 identified very few variations between the glycan compositions at individual sites (Go et al., 2013). However differences were observed here in the structures of the complex-type glycans, specifically the CHO-derived trimers contained a greater degree of sialylation than the 293T-derived trimers. While this difference does not appear to affect the antigenic properties of the trimers (Chung et al., 2014), the degree to which glycoproteins are sialylated can influence their half-life in the body. Asialo-glycoprotein receptors, expressed on hepatocytes, bind terminal galactose residues on circulating glycoproteins and trigger their clearance from serum. The presence of sialic acid caps prevents this clearance, and prolongs the half-life of therapeutic glycoproteins (Morell et al., 1971). This is therefore an important factor to consider when selecting a cell type for the production of vaccine candidates.

8 Discussion

8.1 Insights into the architecture of the HIV-1 glycan shield

The observed cross-clade conservation of the oligomannose population (chapter 3) is consistent with its essential role. Glycosylation sites that map to the neutralising face on the gp120 outer domain display a relatively high degree of conservation, indicating that there are constraints on the number and positions of glycans that are required for effective immune evasion and/or receptor interaction. HIV-1 subtypes are associated with different arrangements of glycosylation sites at certain positions, however these did not translate into differences observed in the overall sizes of the oligomannose populations. This suggests that individual glycan sites are mostly redundant in terms of maintaining the oligomannose population, and this was observed upon alanine scanning of recombinant $gp120_{BaL}$ (chapter 5) – the oligomannose population was found to be largely resilient to individual deletions. Such resilience was not observed at the level of the individual oligomannose species, however, and the levels of $Man_9GlcNAc_2$ in particular were found to be sensitive to sequence variation. This suggests that removal of the first mannose residue, typically from the D2 arm, has a lower kinetic barrier compared with the removal of further mannose residues.

Temporal and spatial constraints on glycan processing explain the phenomenon of glycan microheterogeneity, whereby individual glycosylation sites display a distribution of different glycan structures (Rudd and Dwek, 1997). Glycosylation sites in relatively exposed regions of gp120 were found to possess considerable microheterogeneity, presumably as a consequence of the variable proximity and availability of processing enzymes at a given moment during the passage of Env through the secretory pathway. However the level of

microheterogeneity was found to be more restricted at sites within glycan clusters, indicating that further spatial constraints must exist which are governed by the close proximity of neighbouring glycans. This role for glycans in influencing the processing of their neighbours was demonstrated in chapter five.

Data presented here are based on analysis of recombinant material, produced in cell lines that are non-permissive for HIV-1 infection. Glycosylation is typically cell type-dependent and is a product of the varying expression levels of glycosidases and glycosyltransferases (Kornfeld and Kornfeld, 1985). Indeed the glycosylation of monomeric gp120 has been found to differ depending on the producer cell-type used (Raska et al., 2010). However while reliable quantitation of the relative proportions of oligomannose- and complex-type glycans was not possible, all cell lines tested in the study possessed a population of oligomannose-type glycans. Such protein-directed glycosylation has been observed previously in cases where steric factors restrict typical processing reactions (Crispin et al., 2014, 2004). While the exposed complex-type glycans of Env likely reflect the characteristic glycosylation of the producer cell-type, the sterically protected sites would be expected to remain as oligomannose-type, regardless of producer cell. This is supported by analysis of the BG505 SOSIP.664 trimers produced in 293T and CHO cells, which contained equivalent abundances of oligomannose-type glycans but displayed variation in their complex-type glycans. Given the additional steric constraints of the trimer relative to the monomer, the influence of the producer cell on glycosylation would be expected to be weaker.

Data presented in chapter 7 support the hypothesis of a ‘trimer-induced oligomannose patch’, which describes the elevated abundance of under-processed oligomannose-type glycans on trimeric Env beyond what is referred to as the ‘intrinsic’ oligomannose patch of monomeric gp120 (Bonomelli et al., 2011; Doores et al., 2010a). The

higher abundance of oligomannose-type glycans observed on BG505.664 gp140 trimers, which appear structurally compact and uniform by EM, (Chung et al., 2014; Ringe et al., 2013; Sanders et al., 2013), compared to BG505 WT.SEKS, which appear disordered and heterogeneous, is consistent with a role for the quaternary structure of Env in imposing significant constraints on glycan processing.

8.2 The glycan shield as a drug and vaccine target

One of the measures of utility for conventional HIV-1 drug strategies is the level of their genetic barrier to resistance. Many mutations have been identified which confer resistance to licensed anti-retrovirals (Johnson et al., 2010). In extreme cases a single mutation is sufficient to confer complete resistance, as is the case for the M185V mutation that ablates the activity of the reverse transcriptase inhibitor, lamivudine (Boucher et al., 1993). Mannose-specific CBAs and bnAbs vary in their precise specificities of glycan recognition, however oligomannose glycans on Env – comprising both the intrinsic mannose patch and the trimer-induced oligomannose patch – provide an extended surface to which these agents can target. In particular many bnAbs which target the N332 glycan have been demonstrated to maintain neutralisation capability in its absence (Sok et al., 2014). Such promiscuity in recognition was also indicated by the retention of binding and neutralisation by some bnAbs to PNGS-deletion mutants in chapter six. While deletion of multiple glycosylation sites *in vitro* can lead to resistance against CBAs, the functional roles of glycans suggest this may be difficult *in vivo*; however this has yet to be fully investigated.

From a vaccine perspective the goal is to provide sterilising immunity against transmitted viruses. An important consideration to achieve this is therefore the prevalence of the target epitope(s) within the initial viral swarm, however knowledge of the glycosylation of early transmitter/founder viruses is currently lacking. The priorities of the virus evolve

throughout infection, with the immediate pressure being to establish an infection and disseminate throughout the body before effective immune responses can be mobilised. Evidence is increasing that early transmitted viruses, at least for some clades, possess shorter V1/V2 regions in particular and lack the associated glycosylation sites (Chohan et al., 2005; Derdeyn et al., 2004; Li et al., 2006). Shortened V1/V2 domains have been associated with increased incorporation of Env proteins into virions and enhanced fusion (Cavrois et al., 2014), thus conferring a fitness advantage in early infection. Once infection is established the objective then shifts to survival and evasion of neutralising antibody responses. One study followed the evolution of clade C viruses which initially lacked the N332 glycan, but evolved it within 6 months following immune selection pressure (Moore et al., 2012), illustrating the formative role of the immune response in shaping the glycan shield.

Characterising the nature of the glycan epitopes on early transmitted viruses will therefore be crucial for effective vaccine design. This includes an understanding of the glycan microheterogeneity at different sites, since data presented in chapter six suggest this could represent a mechanism for neutralisation resistance. Site-specific glycosylation analysis of native, early virus would be particularly informative in understanding the degree of natural microheterogeneity and the degree of site occupancy. If combined with more detailed analyses of the glycan specificities of known bnAbs, then immunogens could be designed to elicit the broadest neutralising responses possible.

8.3 Guided vaccination to elicit glycan-reactive bnAbs

The glycosylation of the HIV-1 envelope spike is unusual in two regards. Firstly, the unusually high number of glycosylation sites creates extended clusters of glycan density on the envelope surface. Secondly, as a consequence of this high density, many of the glycans

are incompletely processed and remain of the oligomannose type. These two features represent a divergence from the normal glycosylation of mammalian glycoproteins, and may explain the apparent immunogenicity of the glycan shield.

How broadly neutralising responses against Env glycans are generated is not well understood, however studies mapping the evolution of bnAbs are providing insights (Doria-Rose et al., 2014; Garces et al., 2014; Sok et al., 2013). The bnAbs isolated to date share the common feature of being heavily mutated from their germline precursors (Klein et al., 2013; Scheid et al., 2011), which may partially explain the delay in their elicitation. Many of the bnAbs possess extended CDR H3 loops, which are exploited to penetrate deep into the glycan shield to make contacts with both glycan surfaces and the protein beneath (Doria-Rose et al., 2014; Garces et al., 2014; Kong et al., 2013; McLellan et al., 2011; Mouquet et al., 2012; Pancera et al., 2013; Pejchal et al., 2011; Walker et al., 2011). 2G12 is a rare exception which can achieve sufficient affinity through glycan contacts alone, using its domain-exchanged configuration to extend the surface of interaction (Calarese et al., 2003; Doores et al., 2010b). Recent studies on the maturation process of the PGT121-124 lineage of antibodies revealed that their germline precursor possessed some capacity to bind glycans, which increased over time through somatic hypermutation and enabled additional contacts with alternative glycans (PGT121-123) or with the protein backbone (PGT124) (Garces et al., 2014; Sok et al., 2013). Since the appearance of such potent bnAbs typically takes around two years during natural infection, eliciting the same quality of antibody response may not be possible through vaccination, however polyclonal responses that display intermediate potencies may still be sufficient for protection (Sok et al., 2013).

Given the differing glycan specificities of germline and mature bnAbs, an outstanding question remains how to elicit these types of antibodies through vaccination. Initial stimulation of appropriate germline precursors is critical, and a recent study

demonstrated the rational design of an engineered outer gp120 domain which was successful in stimulating B cells expressing the germline precursor of a CD4 binding site-specific antibody, VRC01 (Jardine et al., 2013). Since breadth and potency of immune responses appears to follow extensive diversification of the virus (Doria-Rose et al., 2014), a vaccination strategy will therefore likely require a boost with multiple Env mimics. Inducing adaptability to glycan microheterogeneity could be attempted at this point, by immunisation with different glycoforms. These could be generated by synthesising SOSIP.664 (or similar) constructs from different clades, mutation of specific glycan sites, or use of glycosylation inhibitors, to fine tune glycosylation at critical sites. The uniformity of glycosylation profiles observed upon purification by different glycan-reactive bnAbs suggests that the BG505 SOSIP.664 trimers display a limited degree of microheterogeneity, suggesting that such modifications would be needed to expand the breadth of a glycan-dependent immune response.

Due to its extreme antigenic diversity HIV-1 is an extremely challenging pathogen to counter. Future methods for battling the HIV/AIDS pandemic will require new options for treatment, in order to overcome existing drug resistance, as well as the development of more robust preventative strategies such as vaccination. The glycan shield of HIV-1 is functionally conserved, offering a window of vulnerability for targeting the virus. Data presented here provide insights into the architecture of the HIV-1 glycan shield and can help guide future drug and vaccine design.

Abbreviations

2-AA	2-aminobenzoic acid
2-AB	2-aminobenzamide
ADCC	Antibody-dependent cell-mediated cytotoxicity
AIDS	Acquired immunodeficiency syndrome
Ala	Alanine
Asn	Asparagine
bnAb	Broadly neutralising antibody
BSA	Bovine serum albumin
C1-4	Constant regions of HIV gp120
CA	HIV-1 capsid protein
CBA	Carbohydrate binding agent
CCR5	Chemokine (C-C motif) receptor 5
CHO	Chinese hamster ovary
CD4bs	CD4 binding site
CDR	Complementarity determining region
CID	Collision induced dissociation
CLR	Calreticulin
CNX	Calnexin
CRF	Circulating recombinant form
CTL	Cytotoxic T-lymphocyte
CXCR4	Chemokine (CXC motif) receptor 4
DC	Dendritic cell
DC-SIGN	Dendritic cell-specific intercellular adhesion molecule (ICAM)-3-grabbing non-integrin
DHB	2,5-dihydroxybenzoic acid
DMEM	Dulbecco's modified Eagles medium
DTT	Dithiothreitol
ELISA	Enzyme-linked immunosorbent assay
Endo H	Endoglycosidase H
Env	Envelope glycoprotein of HIV-1

ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ESI-MS/MS	Electrospray ionisadtion tandem mass spectrometry
ETD	Electron transfer dissociation
Fab	Fragment antigen-binding
Fc	Fragment crystallisable
FCS	Fetal calf serum
Fuc	Fucose
Gag	Polypeptide precursor containing structural HIV-1 proteins MA, CA, NC and p6.
Gal	Galactose
GALT	Gut-associated lymphoid tissue
GlcNAc	N-acetylglucosamine
GnT I	N-acetylglucosaminyltransferase I
HAART	Highly active antiretroviral therapy
HEK	Human embryonic kidney
HILIC	Hydrophilic interaction chromatography
HIV-1	Human immunodeficiency virus type-1
HPLC	High performance liquid chromatography
HRP	Horseradish peroxidase
Ig	Immunoglobulin
IN	Integrase
Kif	Kifunensine
LTR	Long terminal repeat
MA	HIV-1 matrix protein
Man	Mannose
MALDI-MS	Matrix-assisted laser desorption/ionisation mass spectrometry
MHC	Major histocompatibility class
MPER	Membrane proximal external region
nAb	Neutralising antibody
Neu5Ac	N-acetylneuraminic acid
NC	HIV-1 nucleocapsid protein
PBMC	Peripheral blood mononuclear cells

PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PIC	Pre-integration complex
PNGS	Potential N-linked glycosylation site
PNGase F	Protein *N*-glycosidase F
Pol	Polypeptide precursors containing HIV-1 enzymes RT, IN and PR
PR	HIV-1 aspartyl protease
R6	Hexa-arginine furin cleavage site
RF	Relative fluorescence
RP-HPLC	Reverse phase-high performance liquid chromatography
RT	HIV-1 reverse transcriptase
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SIV	Simian immunodeficiency virus
SHIV	A chimera based on the SIV backbone with HIV Env, Tat and Rev genes
SOSIP	Gp140 trimer construct stabilised by a disulphide bond (SOS) and an isoleucine to proline (IP) mutation.
UGGT	UDP-glucose-glucosyltransferase
UPLC	Ultra-performance liquid chromatography
V_H	Variable domain within an antibody heavy chain
V_L	Variable domain within an antibody light chain
V1-V5	Variable loops of HIV-1 gp120
WT	Wild-type

Appendix I

Primers for cloning and mutagenesis

Primer name	Gp120 construct	Sequence (5'-3')
pHLsec_fwd2	All[§]	GCTGGTTGTTGTGCTGTCTCATC
pHLsec_rev	All[§]	CAGGGCATTGGCCACACCAGC
Gp120_Age1_fwd_A	BG505	CGACACCGGTTTATGGGTCACTGTCTACTATG
Gp120_Age1_fwd_B	92RW020	CGACACCGGTTTGTGGGTTACTGTCTACTATG
Gp120_Age1_fwd_C	92BR020 JRCSF 1012_11_TC21_3257 ZM214M_SGA_A3 X1254_C3 P0402_c2_11 JRFL X2088-c9	CGACACCGGTTTGTGGGTCACAGTCTATTATG
Gp120_Age1_fwd_D	98ZADu156	CGACACCGGTTCGTGGGTCACAGTCTATTATG
Gp120_Age1_fwd_E	TRJO4551.58	CGACACCGGTTTGTGGGTCACAGTTTATTATG
Gp120_Age1_fwd_F	1056_10_TA11_1826	CGACACCGGTGCGTGGGTCACAGTTTATTATG
Gp120_Age1_fwd_G	92TH021 BJOXO15000.11.5 C3347.c11	CGACACCGGTTTGTGGGTTACAGTTTATTATG
Gp120_Age1_fwd_I	IAVI C22 94UG103 92TH021	CGACACCGGTCTGTGGGTCACCGTGTACTAC
Gp120_Age1_fwd_J	HXB2	CGACACCGGTATGAGAGTGAAGGAG
Gp120_Age1_fwd_K	Q23-17	CGACACCGGTTTGTGGGTTACTGTCTATTATG
Gp120_Age1_fwd_L	191084.B7.19	CGACACCGGTTTGTGGGTGACTGTCTACTATG
Gp120_Age1_fwd_M	0260.v5.c36	CGACACCGGTTTGTGGGTTACTGTCTACTACG
Gp120_Age1_fwd_R	704809221_1B3	CGACACCGGTTTGTGGGTAACAGTCTATTATG
Gp120_Xma1_fwd_a	Q842.d12	CATCGACCCGGGTTGTGGGTTACTGTCTACTATG
Gp120_BsiWI_rev_A	BG505 94UG103 1056_10_TA11_1826 1012_11_TC21_3257 X2088-c9	CAGCGTACGTGCCCTGGTGGGTGCTAC
Gp120_BsiWI_rev_B	92RW020	CAGCGTACGTGCCCTGGAGGGTGCTAC
Gp120_BsiWI_rev_C	92BR020 BJOXO15000.11.5 X1254_C3 C3347.c11	CAGCGTACGTGCCCTGGTGGGTGCTATTC
Gp120_BsiWI_rev_D	JRCSF TRJO4551.58 JRFL HXB2	CAGCGTACGTGCCTTGGTGGGTGCTAC
Gp120_BsiWI_rev_E	98ZADu156	CAGCGTACGTGCCCCAGTGGGTGCTAC

Gp120_BsiWI_rev_F	ZM214M_SGA_A3	CAGCGTACGTGCTTTAGTGGGTGCTATTC
Gp120_BsiWI_rev_J	Q23-17	CAGCGTACGTGCCCTGGTAGGTGCTACTC
Gp120_BsiWI_rev_K	Q842.d12 191084.B7.19 P0402_c2_11	CAGCGTACGTGCCCTGGTGGGTGCTACTC
Gp120_BsiWI_rev_L	0260.v5.c36	CAGCGTACGTGCCTTGGTGGGTGCTACTC
Gp120_BsiW1_rev_R	704809221_1B3	CAGCGTACGTGCCCCAGTTGGAGCTAC
C22_BsiW1_rev	IAVI C22	CAGCGTACGGGCCTTGGTGGGGGCGATTCC
92TH021_BsiW1_rev	92TH021	CAGCGTACGGGCCCGGGTGGGGGCGATTCC
94UG103_BsiW1_rev	94UG103	CAGCGTACGGGCCCGGGTGGGGGCCACGCC
BaL_fwd_Age1	All BaL mutants[¶]	CGACACCGGTATGGACGCCATGAAG
BaL_rev_Kpn1	All BaL mutants[¶]	CATGGTACCCACCACGCTGCTGATG
BaL_N88A_fwd	BaL N88A	GAAGTAGAATTGGAAGCTGTGACAGAAAATTTTAAC
BaL_N88A_rev	BaL N88A	GTTAAAATTTTCTGTCACAGCTTCCAATTCTACTTC
BaL_N130A_fwd	BaL N130A	CTGTGCGTGACCCTGGCTTGCACTGATTTGAGG
BaL_N130A_rev		CCTCAAATCAGTGCAAGCCAGGGTCACGCACAG
BaL_N136A_fwd	BaL N136A	GCACTGATTTGAGGGCTGCTACTAATGGGAAC
BaL_N136A_fwd		GTTCCCATTAGTAGCAGCCCTCAAATCAGTGC
BaL_N141A_fwd	BaL N141A	GCTACTAATGGGGCCGACACCAACACCAC
BaL_N141A_rev		GTGGTGTTGGTGTCGGCCCCATTAGTAGC
BaL_N144A_fwd	BaL N144A	GGAACGACACCGCCACCACCAGCAGCAG
BaL_N144A_rev		CTGCTGCTGGTGGTGGCGGTGTCGTTCC
BaL_N156A_fwd	BaL N156A	GGCGAGATGAAGGCCTGCAGCTTCAAGATC
BaL_N156A_rev		GATCTTGAAGCTGCAGGCCTTCATCTCGCC
BaL_N186A_fwd	BaL N186A	GTGCCCATCGACGCCAACAGCAACAACC
BaL_N186A_rev		GGTTGTTGCTGTTGGCGTCGATGGGCAC
BaL_N197A_fwd	BaL N197A	CGCCTGATCAGCTGTGCCACCTCAGTCATTAC
BaL_N197A_rev		GTAATGACTGAGGTGGCACAGCTGATCAGGCG
BaL_N241A_fwd	BaL N241A	GAAAAGGACCATGTTCAGCTGTCAGCACAGTACAATG
BaL_N241A_rev		CATTGTACTGTGCTGACAGCTGAACATGGTCCTTTTC
BaL_N262A_fwd	BaL N262A	CAGCTGCTGCTGGCCGGCAGCCTGGC
BaL_N262A_rev		GCCAGGCTGCCGGCCAGCAGCAGCTG
BaL_N289A_fwd	BaL N289A	CATAATAGTACAGCTGGCTGAATCTGTAGAAATTAATTG
BaL_N289A_rev		CAATTAATTTCTACAGATTCAGCCAGCTGTACTATTATG
BaL_N301A_fwd	BaL N301A	GTACAAGACCCAACGCCAATACACGCAAGAGCATC
BaL_N301A_rev		GATGCTCTTGCGTGTATTGGCGTTGGGTCTTGTAC
BaL_N356_fwd	BaL N356A	GCGAGCAGTTCGGCGCCAAGACCATCGTC
BaL_N356_rev		GACGATGGTCTTGGCGCCGAACTGCTCGC
BaL_N386A_fwd	BaL N386A	GGAGGGGAATTTTTCTACTGTGCTTCAACACAACTGTTTAATAG
BaL_N386A_rev		CTATTAAACAGTTGTGTTGAAGCACAGTAGAAAAATTCCCCTCC
BaL_N392A_fwd	BaL N392A	CAACACAACTGTTTGCTAGTACTTGGAATG
BaL_N392A_rev		CATTCCAAGTACTAGCAAACAGTTGTGTTG
BaL_N396A_fwd	BaL N396A	GTTTAATAGTACTTGGGCTGTTACTGAAGAGTC
BaL_N396A_rev		GACTCTTCAGTAACAGCCCAAGTACTATTAAAC
BaL_N406A_fwd	BaL N406A	GAATGTTACTGAAGAGTCAGCTAACACTGTAG
BaL_N406A_rev		CTACAGTGTTAGCTGACTCTTCAGTAACATTC
BaL_N411A_fwd	BaL N411A	CAAATAACACTGTAGAAGCTAACACAATCACAC
BaL_N411A_rev		GTGTGATTGTGTTAGCTTCTACAGTGTTATTTG

BaL_N448A_fwd	BaL N448A	CAAATTCGCTGCAGCAGCGCCATCACCGGCC
BaL_N448A_rev		GGCCGGTGATGGCGCTGCTGCAGCGAATTTG
BaL_N463A_fwd	BaL N463A	CGGCCCAGAGGACGCCAAGACCGAGGTCTTC
BaL_N463A_rev		GAAGACCTCGGTCTTGGCGTCCTCTGGGCCG

[§]Primers used for sequencing

[¶]Flanking primers used for PCR mutagenesis

Appendix II

Cross-clade gp120 panel supplementary data

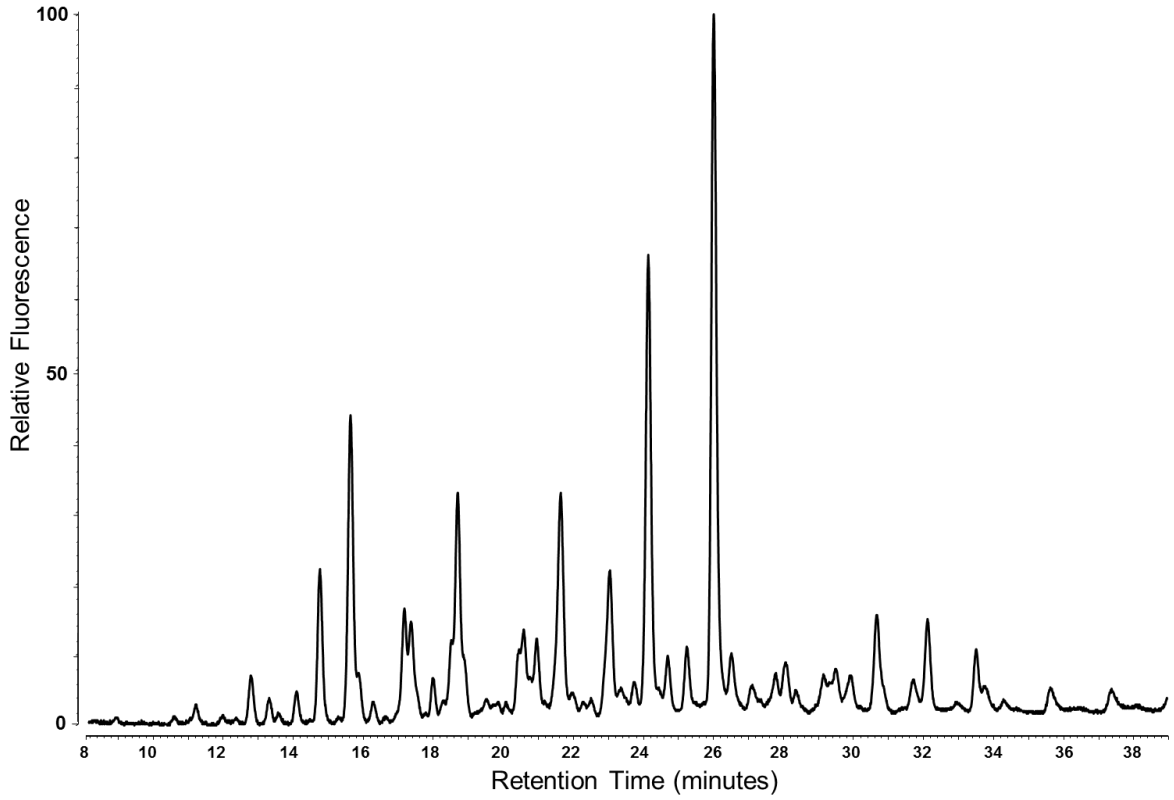
Table 1 – Panel of gp120s

Isolate	Clade
0260.V5.C36	A
191084 B7.19	A
Q23-17	A
BG505	A
94UG103	A
92RW020	A
Q842.d12	A
BaL	B
JRCSF	B
HXB2	B
SF162	B
TRJO_A2	B
1056_TA11_1826	B
1012_11_TC21	B
JRFL	B
96ZM651	C
98ZADu156	C
ZM214M_SGA_A3	C
IAVI C22	C
704809221_1B3	C
92TH021	CRF01_AE
BJOX015000.11.5	CRF01_AE
C3347.C11	CRF01_AE
P0402_c2_11	G
X1254_3	G
X2088-c9	G

Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

0260.V5.C36 (A)

LWVTVYYGVPVWKDAETTLFCASDAKAYDKEVHNVWATHACVPTDPNPQEVFLENVTEE
FNMWKNNMVEQMHTDIISLWDQSLQPCVQLTPLCVTLNCTDYKGNYTNYNYTKNNGTGH
NITSDMEGEIKDCSFNVTTELRDKRQKVHSLFYRLDIVQINSSQTNSSEYRLINCNTSAITQAC
PKISFNPIPIHYCAPAGFAILKCNDKKFHGTGPCKNVSSVQCTHGIKPIVSTQLLLNGSLAEEK
VAIRSENITNNAKNILVQLKSPVNITCVRPNNNTRTSVRIGPGQTFYATGAITGDIRQAHCNV
SKAKWNETLRQVAVQLKTYFNKTIIFNSSAGGDLEITTHSFNCGGEFFYCNTSGLFNSTWND
TSNHNTSGTESNANITLQCRIKQIINMWQRAGQAIYAPPIQGIIKCESNITGLILTRDGGDNNN
TTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKARTKHHHHHH

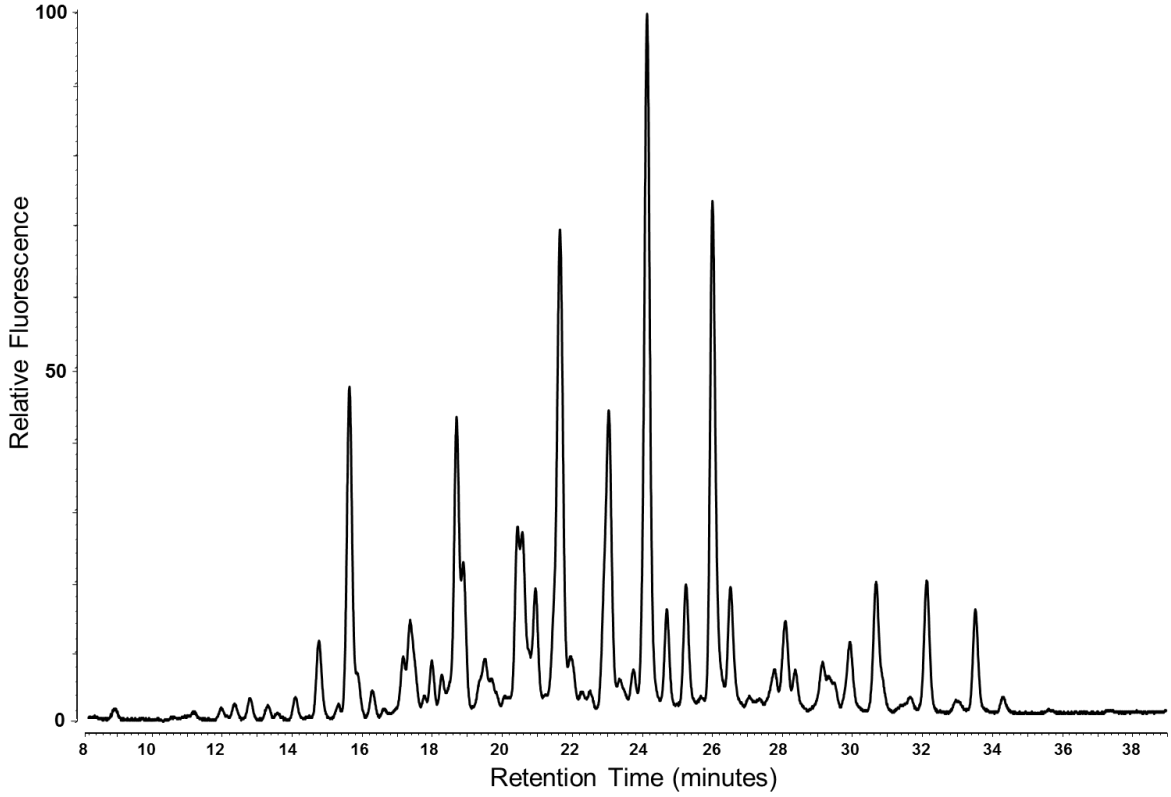


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

191084 B7.19 (A)

LWVTVYYGVPVWRDAETTLFCASDAKAYDTEMHNVWATHACVPTDPNPQEIDLENVTEK
FNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCTAITNDTRGNETGINRTVETTE
MTNCSFNMTTELRDRKKKVNALFYKLDIVQIGENSSSQYRLINCNTSVITQACPKVTFEPIPI
HYCAPAGFAILKCKDKEFNGTGTCRNVSSVQCTHGIKPVVSTQLLLNGSLAEGQVKIRSENI
SDNAKTIIVQLNESVPINCTRPNNNTRRGIHLGPGQTFFATDIIGDIRQAHCNVSESKWNKAL
QEVVKQLRQHWNKTIIFKSSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWNIAGNRTNDTK
SNETITLPCRIKQIVNVWQRVGQAIYAPPIAGVIRCNSNITGLLLVRDGGATNNTDETFRPGG
GNMRDNWRSELYKYKVVKIEPLGVAPTRARTKHHHHHH

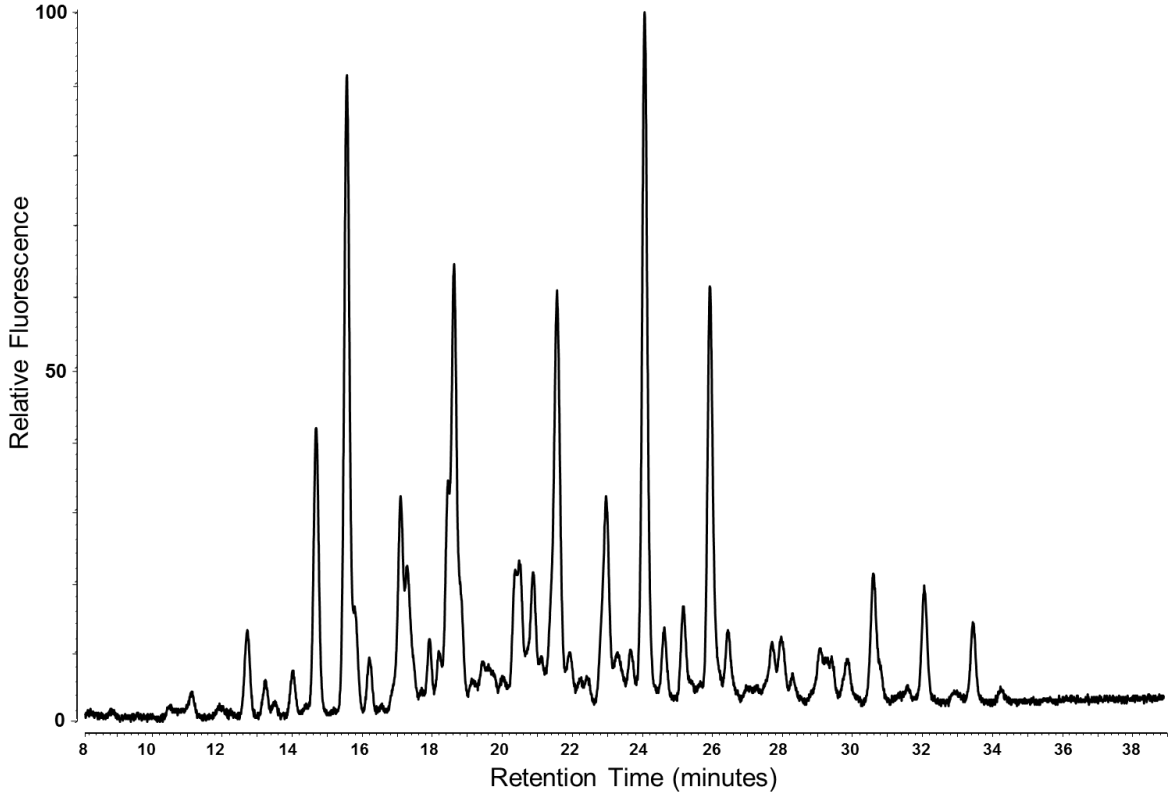


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

Q23-17 (A)

LWVTVYYGVPVWRDADTTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLDNVTEK
FNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLHCTNVTSVNTTGDREGLKNCSFN
MTTELRDKRQKVYSLFYRLDIVPINENQGSEYRLINCNTSAITQTCPKVSFEPIPIHYCTPAGF
AILKCKDEGFNGTGLCKNVSTVQCTHGIKPVVSTQLLLNGSLAEKNITIRSENITNNAKIIIVQ
LVQPVTIKCIRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVTRSRWNKTLQEVAEKLRT
YFGNKTIIFANSSGGDLEITTHSFNCGGEFFYCNTSGLFNSTWYVNSTWNDTDSTQESNDTIT
LPCRIKQIINMWQRAGQAMYAPPIPGVIKCESNITGLLLTRDGGKDNNVNETFRPGGGDMR
DNWRSELYKYKVVEIEPLGVAPTRARTKHHHHHH

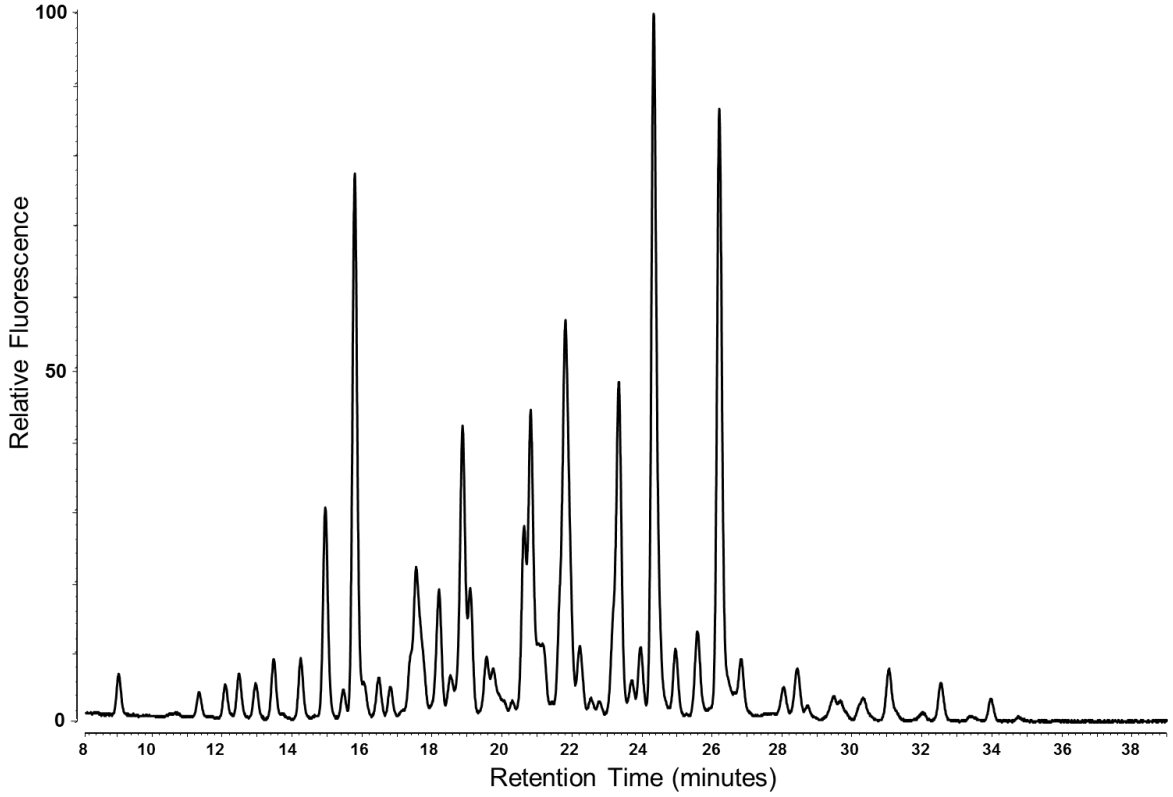


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

BG505 (A)

LWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEEF
NMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELKNCSFN
MTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPKVSFEPIPI
HYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEEVMIRSENIT
NNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNVSKATWNETL
GKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNSTWISNTSVQGS
NSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDGGSTNSTTETFRP
GGGDMRDNWRSELYKYKVVKIEPLGVAPTRGTKHHHHHH

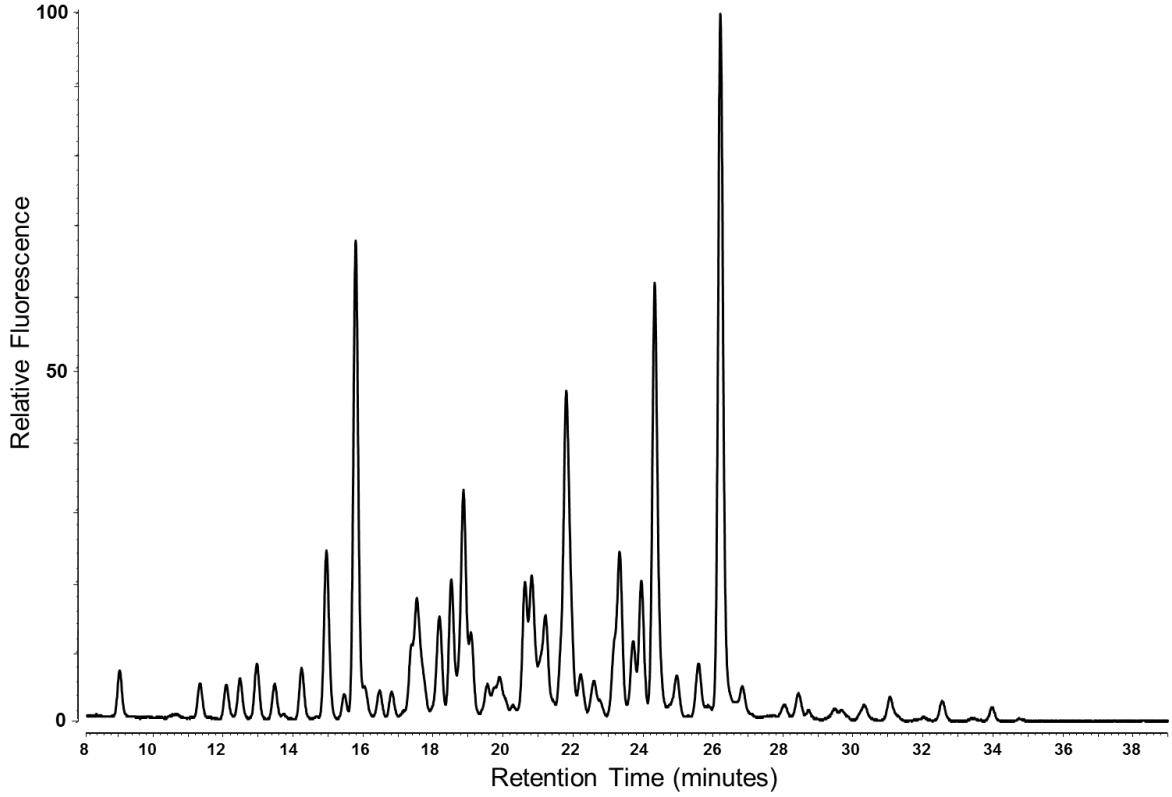


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

94UG103 (A)

LWVTVYYGVPVWRDAETTLFCASDAKAYKTEVHNVWATHACVPTDPNPQEIHLENVTED
FNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLNCTNKINAKGNSSIVPTIDPITNMS
MEGEIKNCSYNMTTELRDKKQKVYSLFYKLDVVQMDNSSSNSSGEYRLINCNTSAITQACP
KVSFEPIPIHYCAPAGFAILKCKDDEFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAERE
VRIRSENISDNAKTLIVQLSSPVTINCTRPNNNTRKAIRTGPGQGQALYAIGAIIGDIRQAHCN
VSKSDWNKTLHQVVKQLRKHWNNKTIIFTKPSGGDLEITTHSFICGGEFFYCNTSRLFNSTW
NSTWENGTEPNGTEVNGTITLPCRIKQVINMWQRVGQAMYAPPIRGVIRCQSNITGLLLTRD
GGINNGTNETFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPTRARTKHHHHHH

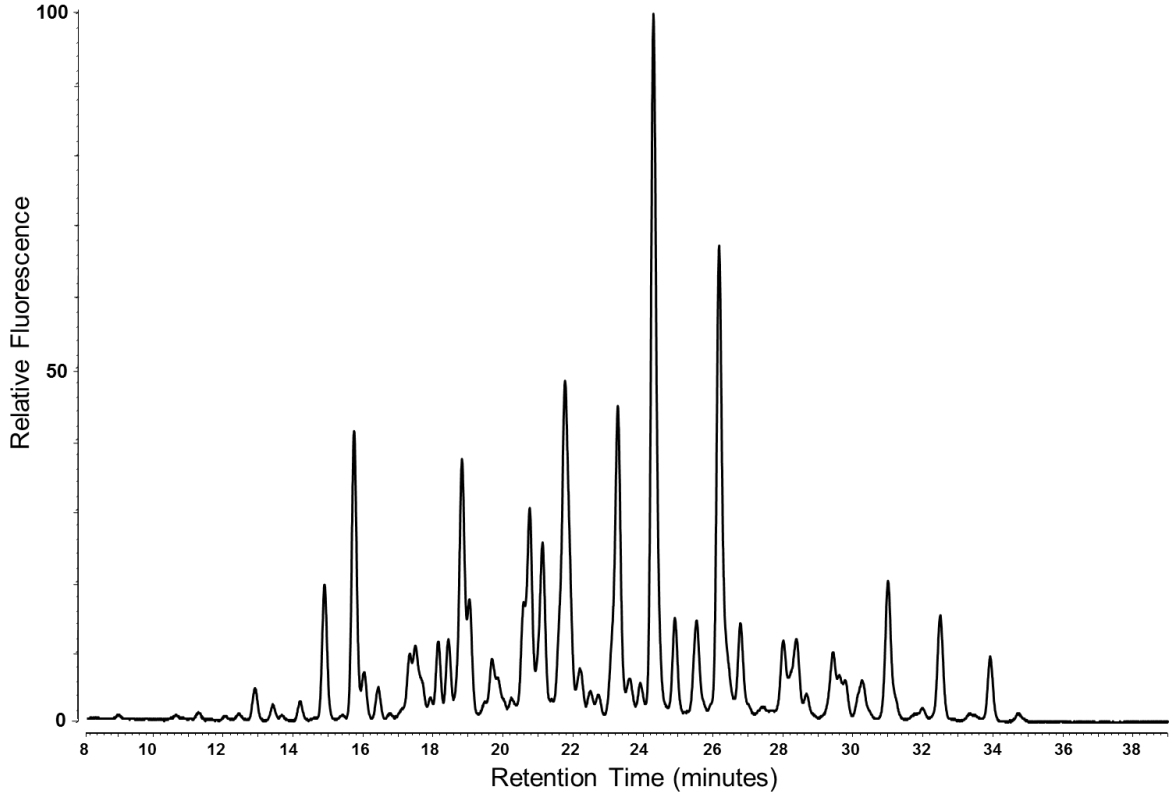


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

92RW020 (A)

LWVTVYYGVPVWKDAETTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIHLENVTED
FNMWRNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLDCNATASNVTNEMRNCSFNITTEL
KDKKQQVYSLFYKLDVVQINEKNETDKYRLINCNTSAITQACPKVSFEPIPIHYCAPAGFAIL
KCKDTEFNGTGPCKNVSTVQCTHGIRPVISTQLLLNGSLAEEGIQIRSENITNNAKTIIVQLDK
AVKINCTRPNNNTRKGVRIGPGQAFYATGGIIGDIRQAHCNVSRAKWNDTLRGVAKKLREH
FKNKTIIFEKSSGGDIEITTHSFNCGGEFFYCNTSGLFNSTWESNSTESNNTTSNDTITLTCRIK
QIINMWQKVGQAMYAPPIQGVIRCESNITGLLLTRDGGNNSTNEIFRPGGGNMRDNWRSEL
YKYKVVKIEPLGVAPSRARTKHHHHHH

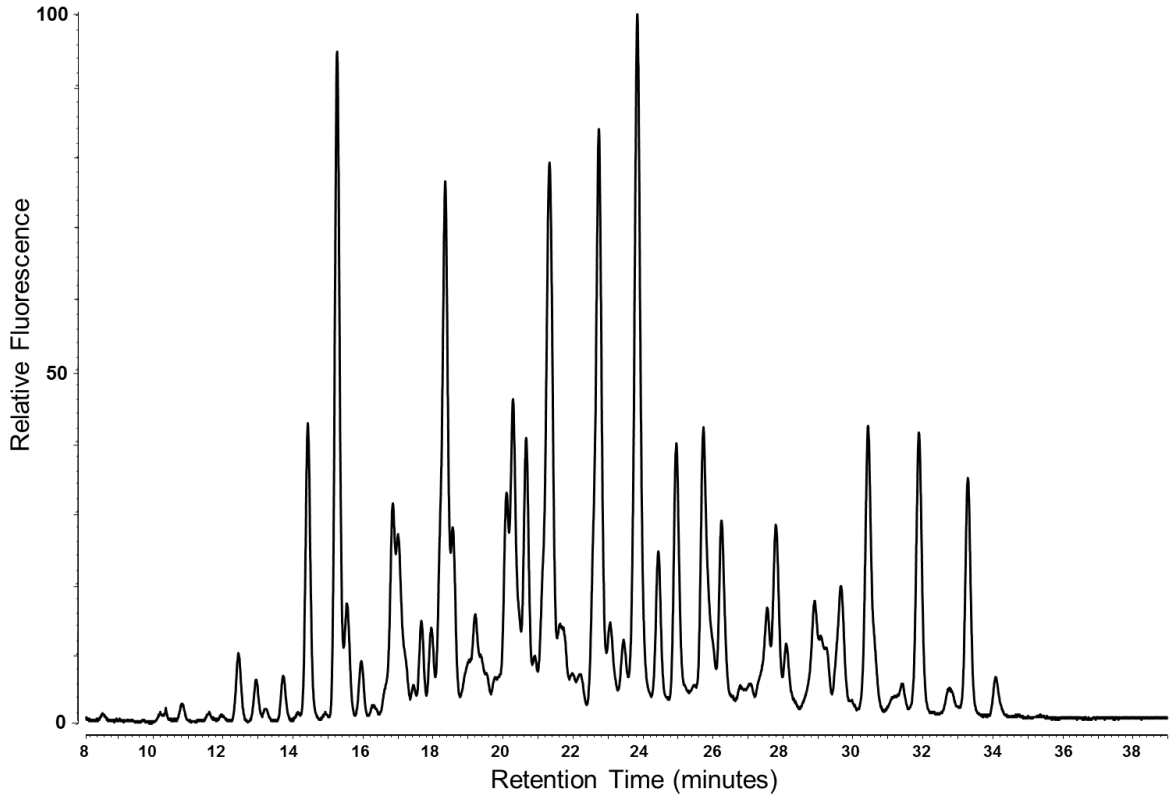


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

Q842.d12 (A)

LWVTVYYGVPVWKEAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLENVTEEF
NMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLDCNNVTNNGTSDMREEIKNCSFNM
TTELRDKRQKVYSLFYKLDIVQINEDQGNSSNNKYRLITCNTSAITQACPKVTFEPIPIHYCAP
AGFAILKCKDEEFNGIGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEKEVKIRCENITNNAKT
IIVQLVNPVKINCTRPNNNTRKSIHIGPGQAFYATGDIIGDIRQAHCNVNRTEWNNTLHQVVE
QLRKHFNKTINFANSTGGDLEITTHSFNCGGEFFYCNTTNLFNSTWNHTASMNSTESNDTIIL
PCRIKQIINMWQRVGQAMYAPPIRGVIRCESNITGLILTRDGGNTNSTRETFRPGGGDMRDN
WRSELYKYKVVKIEPLGVAPTRARTKHHHHHH

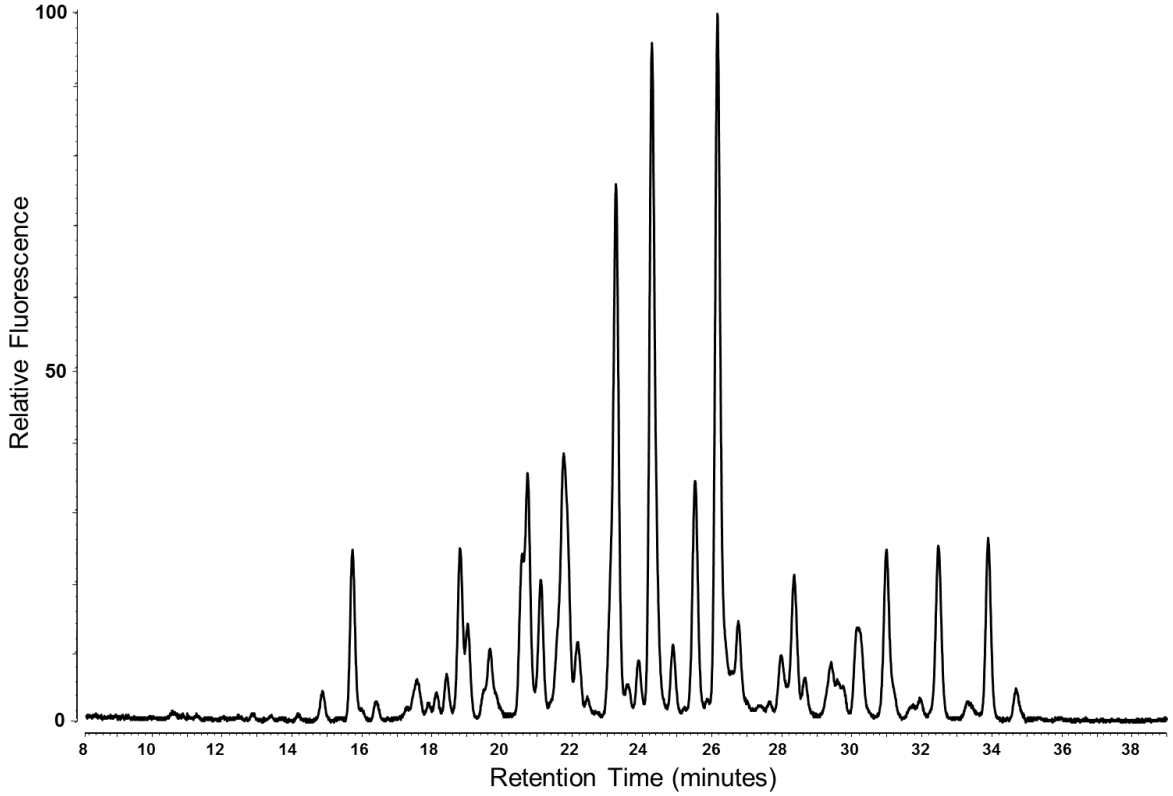


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

BaL (B)

LWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVELENVTEN
FNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLRNATNGNDTNTTSSSREM
MGGGEMKNCSFKITTNIRGKVQKEYALFYELDIVPIDNNSNNRYRLISCNTSVITQACPKISF
EPIPIHYCAPAGFAILKCKDKKFNGKGPCSNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIR
SENFADNAKTIIVQLNESVEINCTRPNNNTRKSIHIGPGRALYTTGEIIGDIRQAHCNLSRAKW
NDTLNKIVIKLREQFGNKTIVFKHSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWNVTEESN
NTVENNTITLPCRIKQIINMWQKVGRAMYAPPIRGQIRCSSNITGLLLTRDGGPEDNKTEVFR
PGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAISSVVGTKHHHHHH

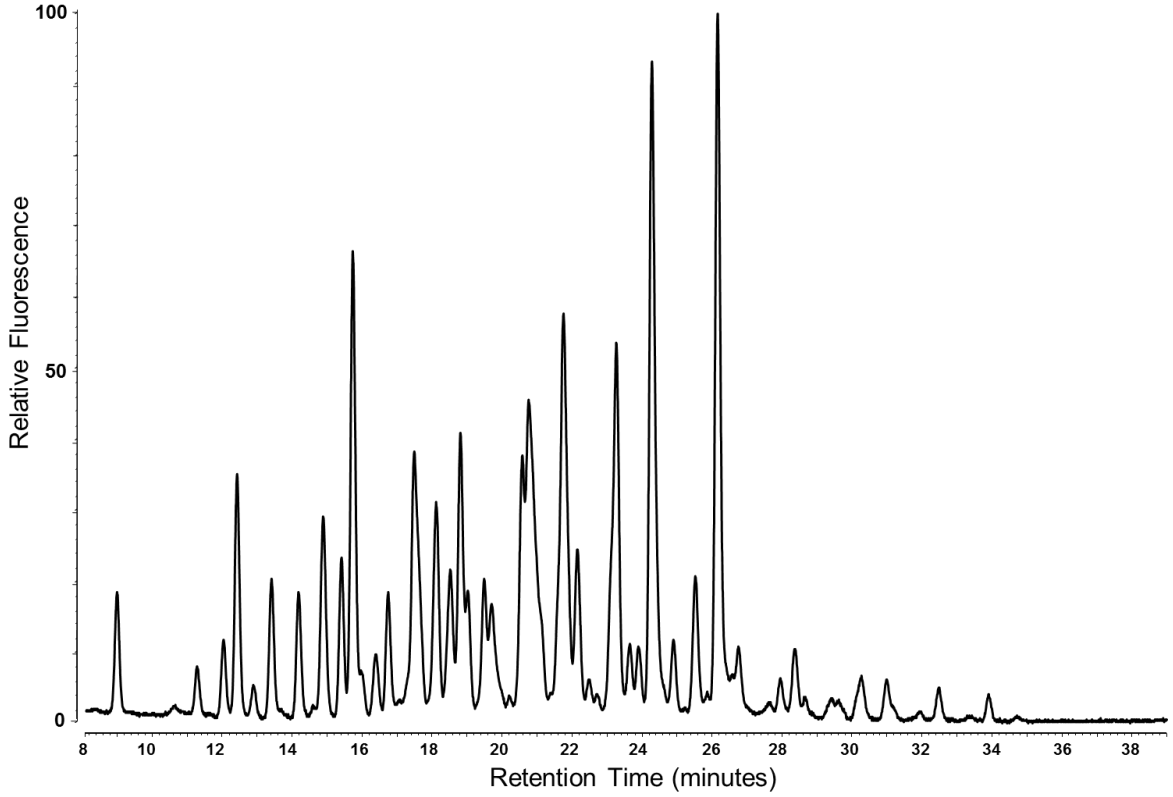


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

JRCSF (B)

VWKETTTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLENVTEDFNMWKNNMV
EQMQEDVINLWDQSLKPCVKLTPLCVTLNCKDVNATNTTSSSEGMMERGEIKNCSFNITKSI
RDKVQKEYALFYKLDVVPIDNKNNTKYRLISCNTSVITQACPKVSFEPIPIHYCAPAGFAILK
CNNKTFNGKGQCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEKVVIRSDNFTDNAKTIIVQLN
ESVKINCTRPSNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAQWNNTLKQIVEKLREQFN
NKTIVFTHSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNDTEKSSGTEGNDTIILPCRIKQI
INMWQEVGKAMYAPPIKGQIRCSSNITGLLLTRDGGKNESEIEIFRPGGGDMRDNWRSELYK
YKVVKIEGTKHHHHHH

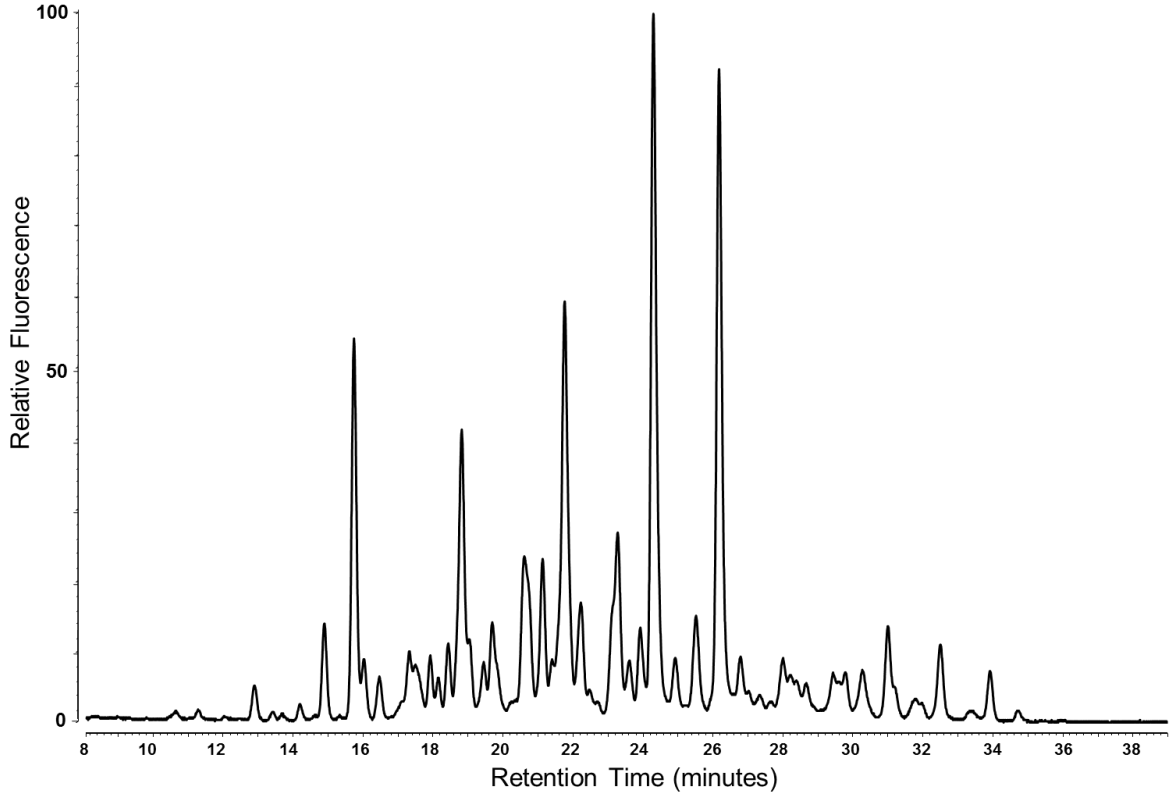


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

HXB2 (B)

LWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTEN
FNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIMEKG
EIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSCNTSVITQACPKVSFEPIPIHYC
APAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSVNFTDN
AKTIIVQLNTSVEINCTRPNNNTRKRIRIQRGPGRAFVTIGKIGNMRQAHCNISRAKWNNTLK
QIASKLREQFGNNKTIIFKQSSGGDPEIVTHSFNCGGEFFYCNSTQLFNSTWFNSTWSTEGSN
NTEGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGNSNNESEIFRP
GGGDTRDNWRSELYKYKVVKIEPLGVAPTKARTKHHHHHH

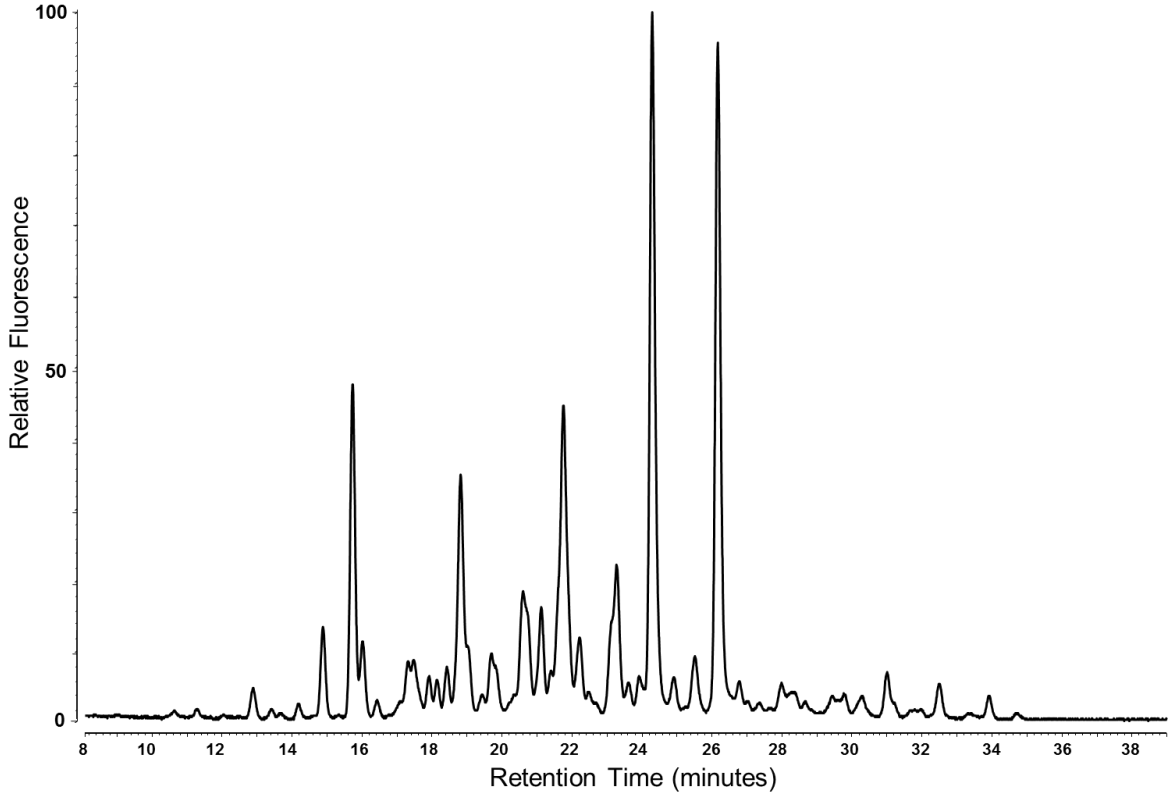


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

SF162 (B)

LWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEIVLENVTENF
NMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLHCTNLKNATNTKSSNWKEMDRGE
IKNCSFKVTTSIRNKMQKEYALFYKLDVVPIDNDNTSYKLINCNTSVITQACPKVSFEPIPIHY
CAPAGFAILKCNDKKFNGSGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEGVVIRSENFTD
NAKTIIVQLKESVEINCTRPNNNTRKSITIGPGRAFYATGDIIGDIRQAHCNISGEKWNNTLKQ
IVTKLQAQFGNKTIVFKQSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNNTIGPNNTNGTI
TLPCRIKQIINRWQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGKEISNTTEIFRPGGGDMRD
NWRSELYKYKVVKIEPLGVAPTKARTKHHHHHH

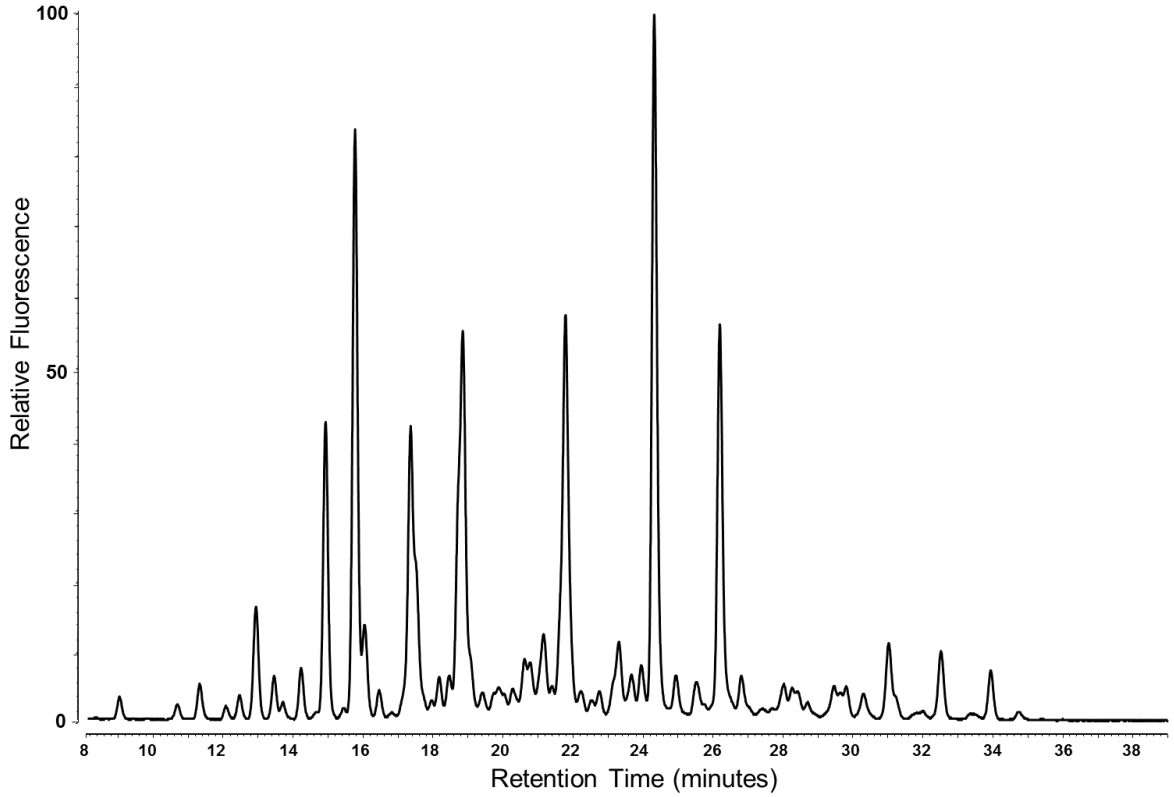


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

TRJO_A2 (B)

LWVTVYYGVPVWKEATTTLFCASDAKAYETEVHNVWATHACVPTDPNPQELVLENVTEY
FDMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDWTNGTDWNTTNSNNTTISK
EETIEGGEMKNCSFNITTATGDKKKERAFFYKLDVAPIDNSNTSYRLISCNTSVITQACPKISF
EPIPIHYCAPAGFAILKCNDKKFNGTGSCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRS
KNFSDNAKIIIVQLNESVPINCTRPHNNTRKSIHIGPGRAWYATGDIIGDIRKAYCNISEAKWN
NTLKQITEKLKEQFNKTIIVFNQSSGGDPEVTMHSFNCGGEFFYCNTSKLFNGTWNSTKRAN
NTEGIIILQCRIKQIINRWQEVGKAMYAPPIEGQIKCSSNITGLLLTRDGGKTTNNTTEFFRPG
GGNMKDNWRSELYKYKVVRIEPLGVAPTKARTKHHHHHH

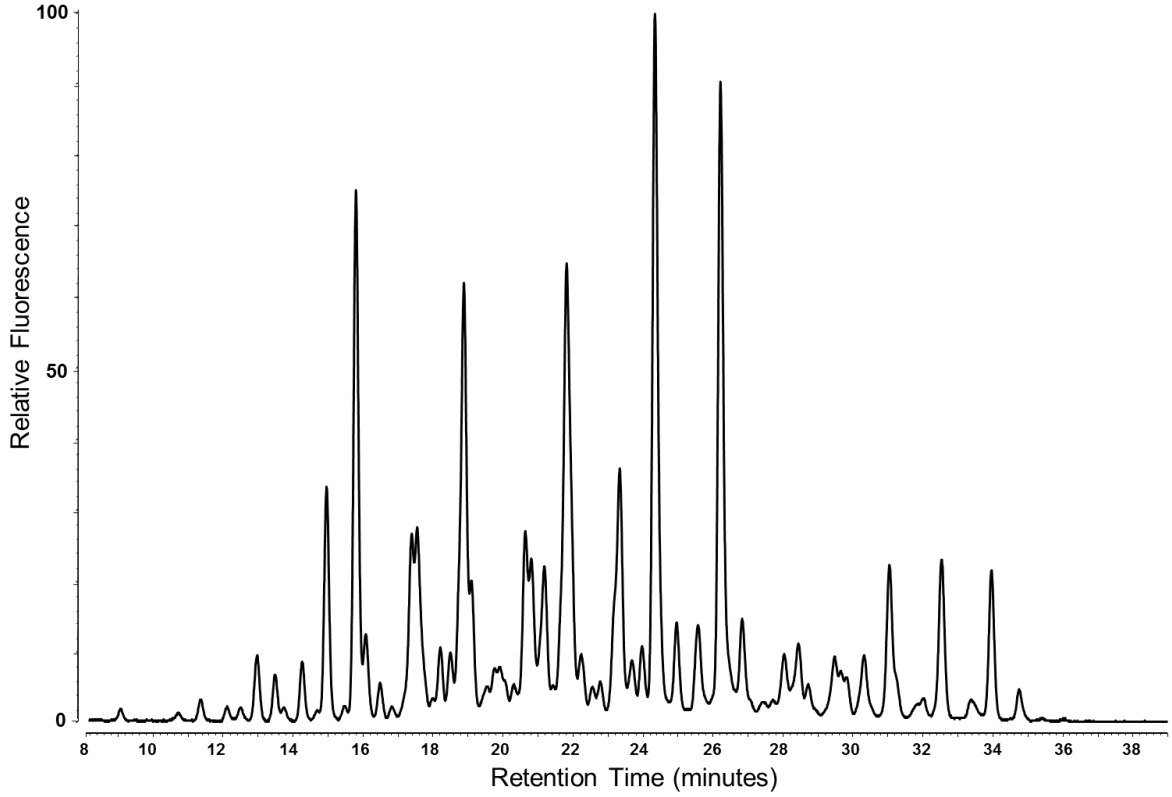


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

1056_TA11_1826 (B)

AWVTVYYGVPVWKEAVTTLFCASDAKAYDTEVHNVWTTHACVPTDPDPQEVHMENVTE
DFNMWKNNMADQMHEDIISLWDQSLKPCVKLTPLCVTLNCADWKNNTDTNTNSSVRIME
KGEIKNCSFNITTNIRDKYQKAYALFYKLDVVPIDDDNATGNNDTRNYRLISCNTSVITQACP
KVSFEPIPIHYCAPAGFAILRCNNKTFSGKGQCTNVSTVQCTHGIKPVVSTQLLLNGSLAEEE
VIIRSDNFSDNAKTIIVHLNSSVDINCTRPGNNTRKSITIGPGRAFYATGDIIGDIRQAHCNISGE
KWNNTLKQVVKKLREQFGNKTIVFNQSSGGDPEITMHTFNCGGEFFYCNTAQLFNSTWEA
NSTWENDNERVGHSNKTIILQCRIKQIINMWQEVGKAMYAPPISGQIRCSSNITGLLLTRDGG
NGNETNRTEVFRPGGGNMKDNWRSELYKYKVVKIEPLGVAPTRARTKHHHHHH

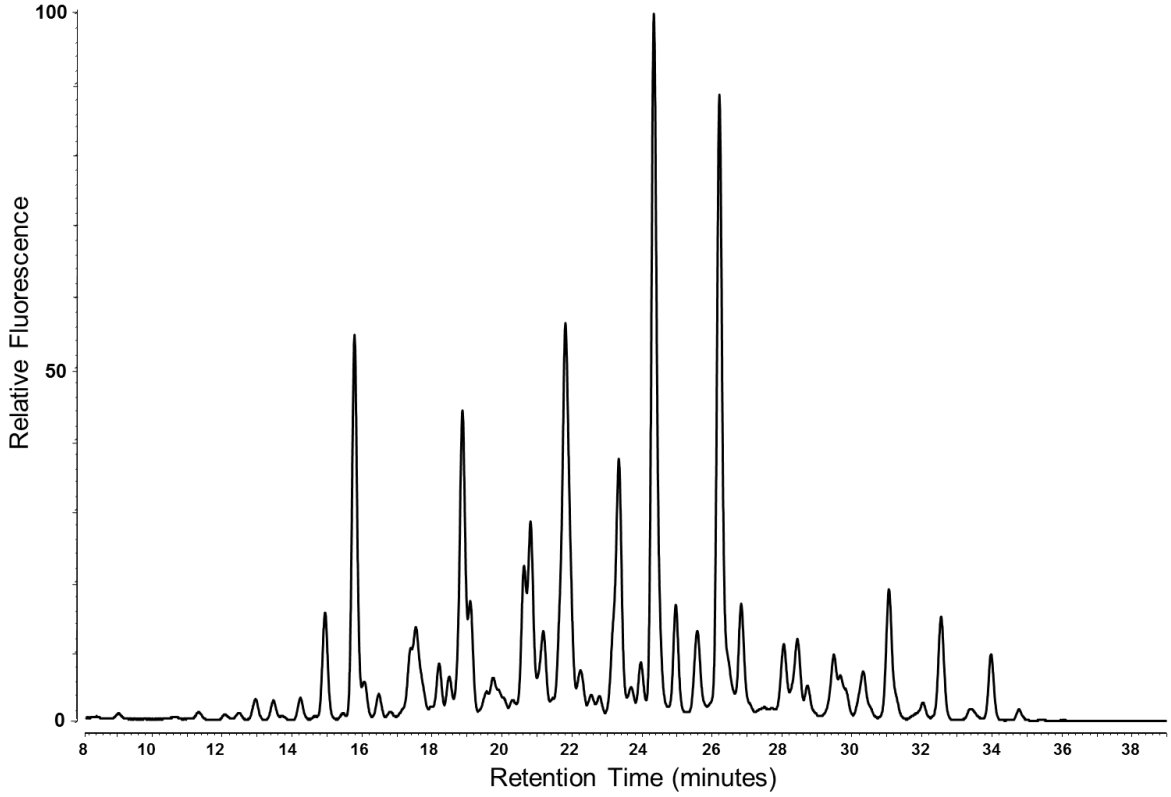


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

1012_11_TC21 (B)

LWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQELVLANVTEN
FNMWNNTMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDVTNATNINATNINNSSGGVE
SGEIKNCSFNITTSVRDKVQKEYALFYKLDIVPITNESSKYRLISCNTSVLTQACPKVSFEPIPI
HYCAPAGFAILKCNNETFNGKGPCINVSTVQCTHGIRPVVSTQLLLNGSLAEKEVIIRSDNFS
DNAKNIIVQLKEYVKINCTRPNNNTRKSIHIGPGRAFYATGEIIGNIRQAHCNISRSKWNDTL
KQIAAKLGEQFRNKTIVFNPSSGGDLEIVTHSFNCGGEFFYCNTTKLFNSTWIREGNNGTWN
GTIGLNDTAGNDTIILPCKIKQIINMWQEVGKAMYAPPIRGQIRCSSNITGLILTRDGGKDDSN
GSEILEIFRPGGGDMRDNWRSELYKYKVVRIEPLGVAPTRARTKHHHHHH

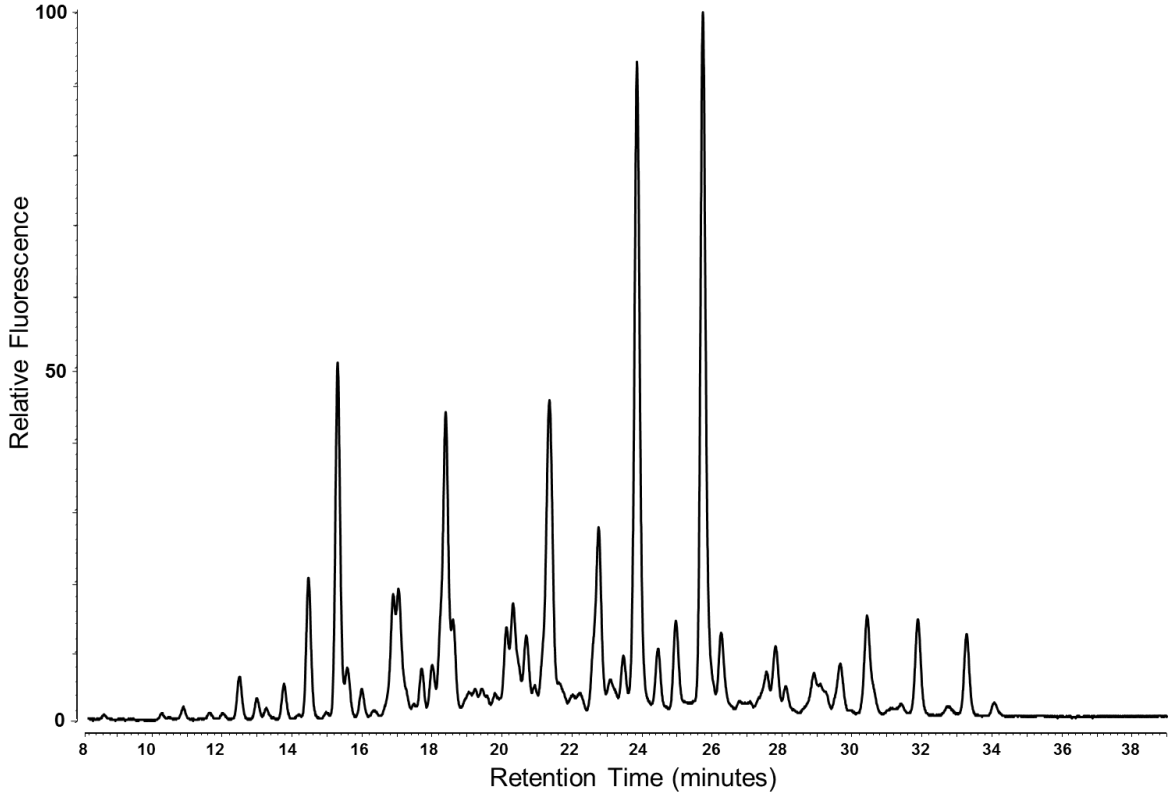


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

JRFL (B)

LWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLENVTEH
FNMWKNNMVEQMQEDIISLWDQSLKPCVKLTPLCVTLNCKDVNATNTTNDSEGTMERGEI
KNCSFNITTSIRDEVQKEYALFYKLDVVPIDNNNTSYRLISCDTSVITQACPKISFEPIPIHYCA
PAGFAILKCNDKTFNGKGPCKNVSTVQCTHGIRPVVSTQLLLNGSLAEEEVVIRSDNFTNNA
KTIIVQLKESVEINCTRPNNNTRKSIHIGPGRAFYTTGEIIGDIRQAHCNISRAKWNDTLKQIVI
KLREQFENKTIVFNHSSGGDPEIVMHSFNCGGEFFYCNSTQLFNSTWNNNTEGSNNTEGNTI
TLPCRIKQIINMWQEVGKAMYAPPIRGQIRCSSNITGLLLTRDGGINENGTEIFRPGGGDMRD
NWRSELYKYKVVKIEPLGVAPTKARTKHHHHHH

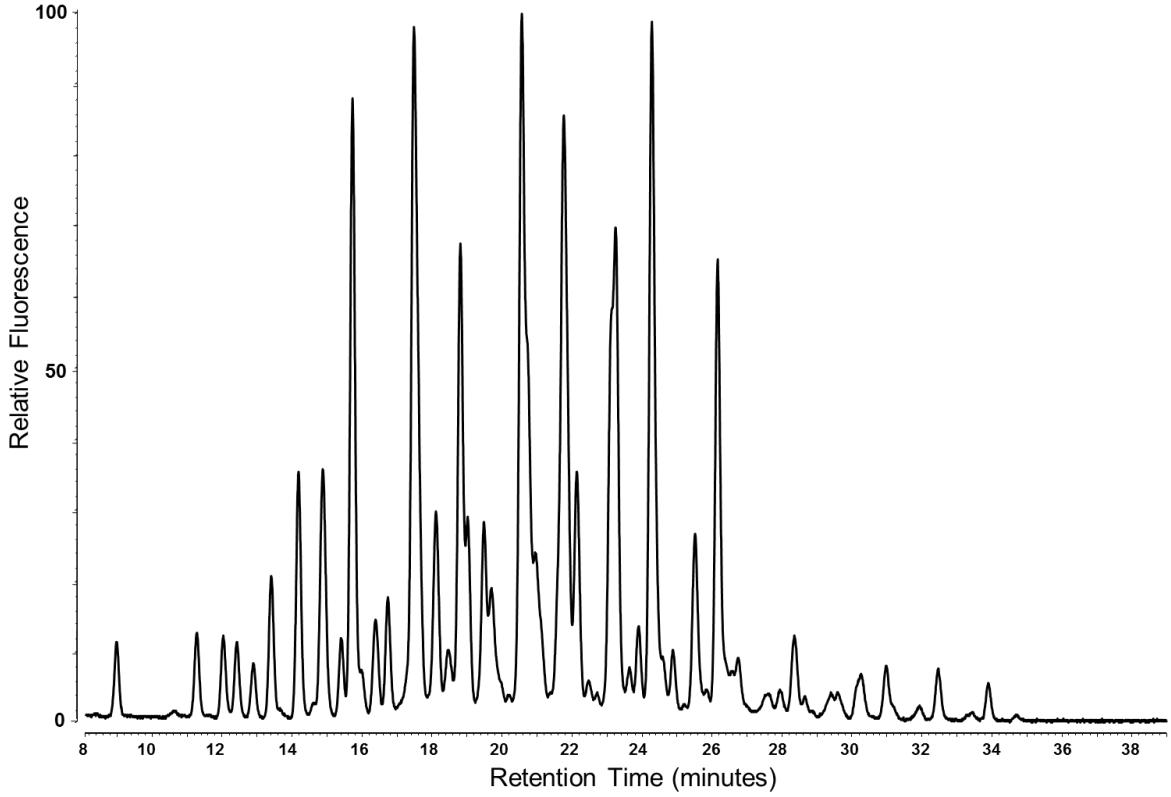


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

96ZM651 (C)

LWVTVYYGVPVWKEAKTTLFCASDAKSYEKEVHNVWATHACVPTDPNPQEIVLGNVTEN
FNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCTEVNVTRNVNNSVVNNTTNV
NNSMNGDMKNCSFNITTELKDKKKNVYALFYKLDIVSLNETDDSETGNSSKYYRLINCNTS
ALTQACPKVSFDPIPIHYCAPAGYAILKCNNKTFNGTGPCHNVSTVQCTHGIKPVVSTQLLL
NGSLAEEGIIIRSENLTNNVKTIIVHLNRSIEIVCVRPNNNTRQSIRIGPGQTFYATGDIIGDIRQ
AHCNISRTNWTKTLREVRNKLREHFPNKNITFKPSSGGDLEITTHSFNCRGEFFYCNTSGLFSI
NYTENNTDGTPITLPCRIRQIINMWQEVGRAMYAPPIEGNIACKSDITGLLLVRDGGSTNDST
NNNTEIFRPAGGDMRDNWHSELYKYKVVEIKPLGIAPTEAKRRVVEREKRGTKHHHHHH

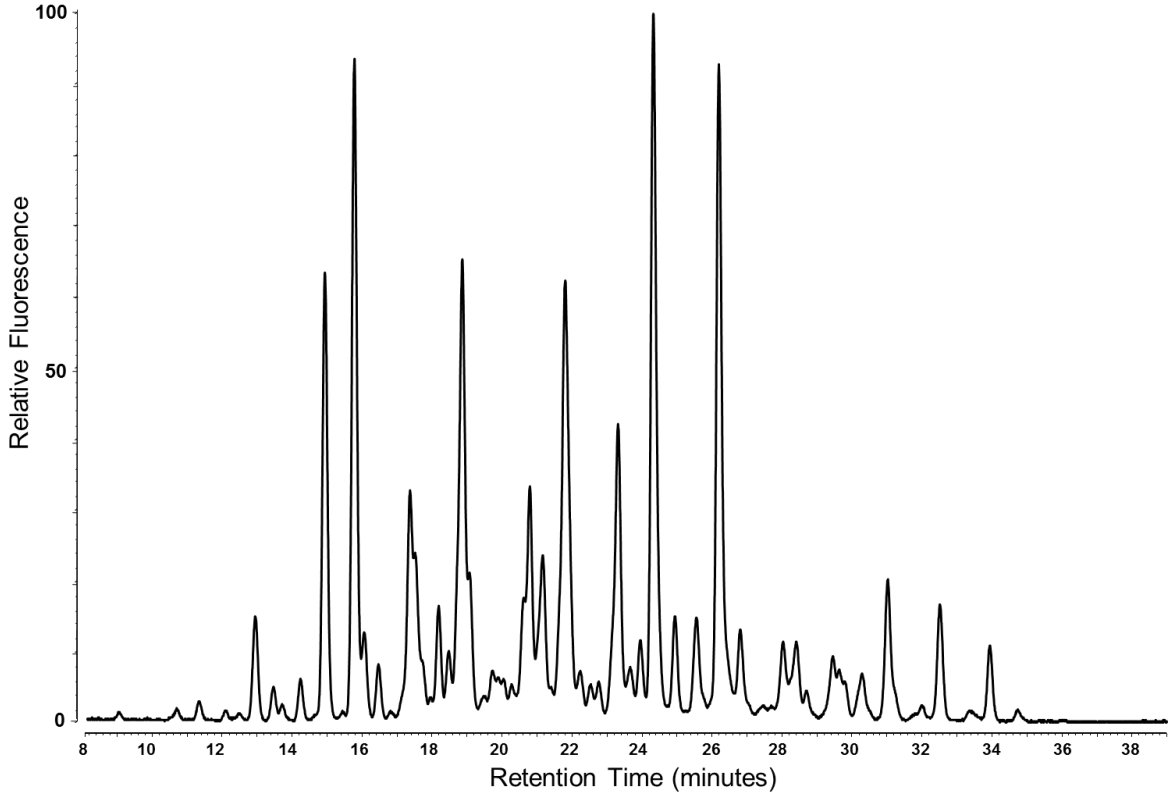


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

98ZADu156 (C)

SWVTVYYGVPVWTEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEIFLKNVTENF
NMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLNCVTYNNSMNSSATYNNSMNGEI
KNCSFNTTTELRDKKQKVYALFYRTDVVPLNNNNNNSEYILINCNTSTITQACPKVSFDPIPI
HYCAPAGYAILKCTDKKFNGTGSCNNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIKSENLT
DNIKTIIVQLNQSIGINCTRPNNNTRKSVRIGPGQTFYATGDIIGDIRQAHCNISRNQWNETLE
QVKKKLGEHFHNQTKIKFEPPSGGDLEITTHSFNCRGEFFYCNTADLFTNATKLVNDTENKA
VITIPCRIKQIINMWQGVGRAMYAPPIEGNITCNSNITGLLLTRDGGGNVTEINRTEIFRPGGG
NMKDNWRNELYKYKVVEIKPLGVAPTGARTKHHHHHH

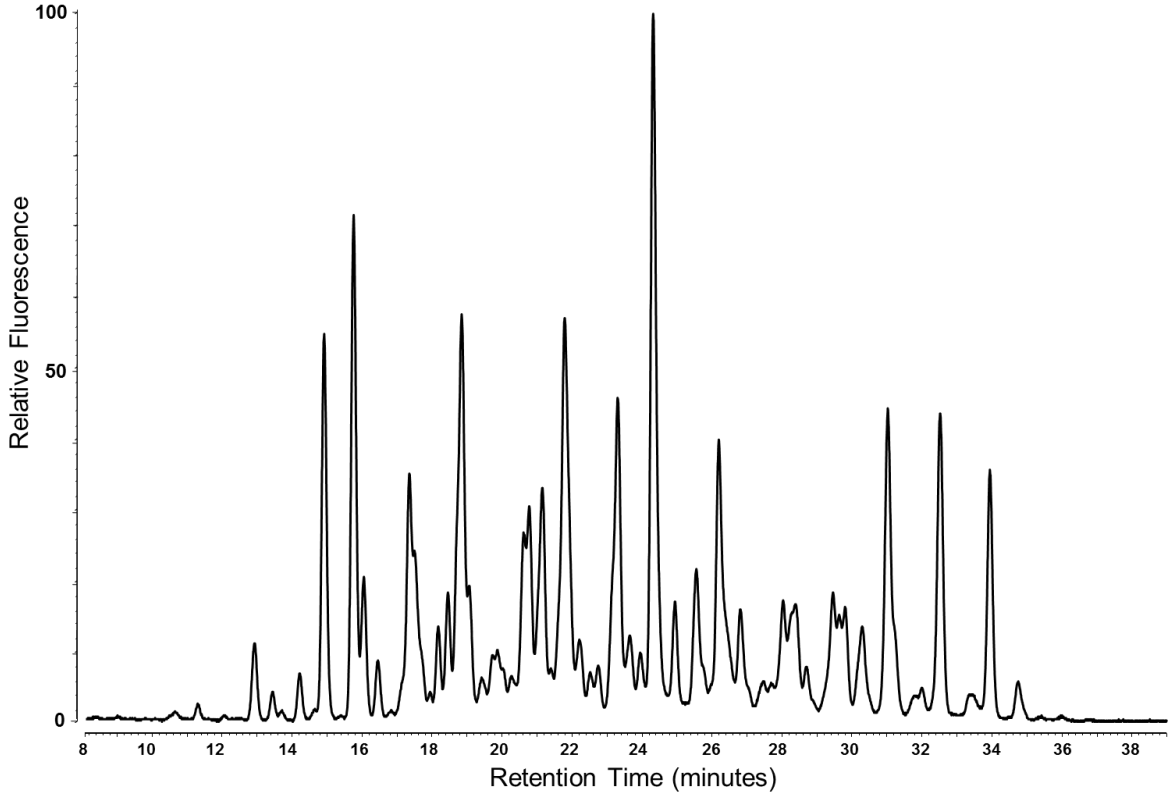


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

ZM214M_SGA_A3 (C)

LWVTVYYGVPVWREAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQELVLENVTEN
FNMWKNDMVNQMHEDIISLWDQSLKPCVKLTPLCVTLNCSNVNINETSIDFNVTSNISMKE
EMKNCSFKVNSELRDKNRREHALFYKLDIVQLNDEGNDSYSYRLINCNTSTIKQACPKVSFE
PIPIHYCAPAGYAILKCNNETFNGSGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEKEIMIRSE
NLTNNAKTIIVQLTEAVNITCMRPGNNTRRSVRIGPGQTFYATGEIIGDIRQAHCNISKDKWN
QILQNVRAKLGEHFHDKTIKFEPSSGGDLEITTHSFNCGGEFFYCNTTNLFSRTYTNGSNSNV
NITSATITLPCRIKQIINMWQEVGRAMYAPPIAGNITCISNITGLLLTRDGGNGNDTNDTETFR
PAGGDMRDNWRSELYKYKVVEIKPLGIAPTKARTKHHHHHH

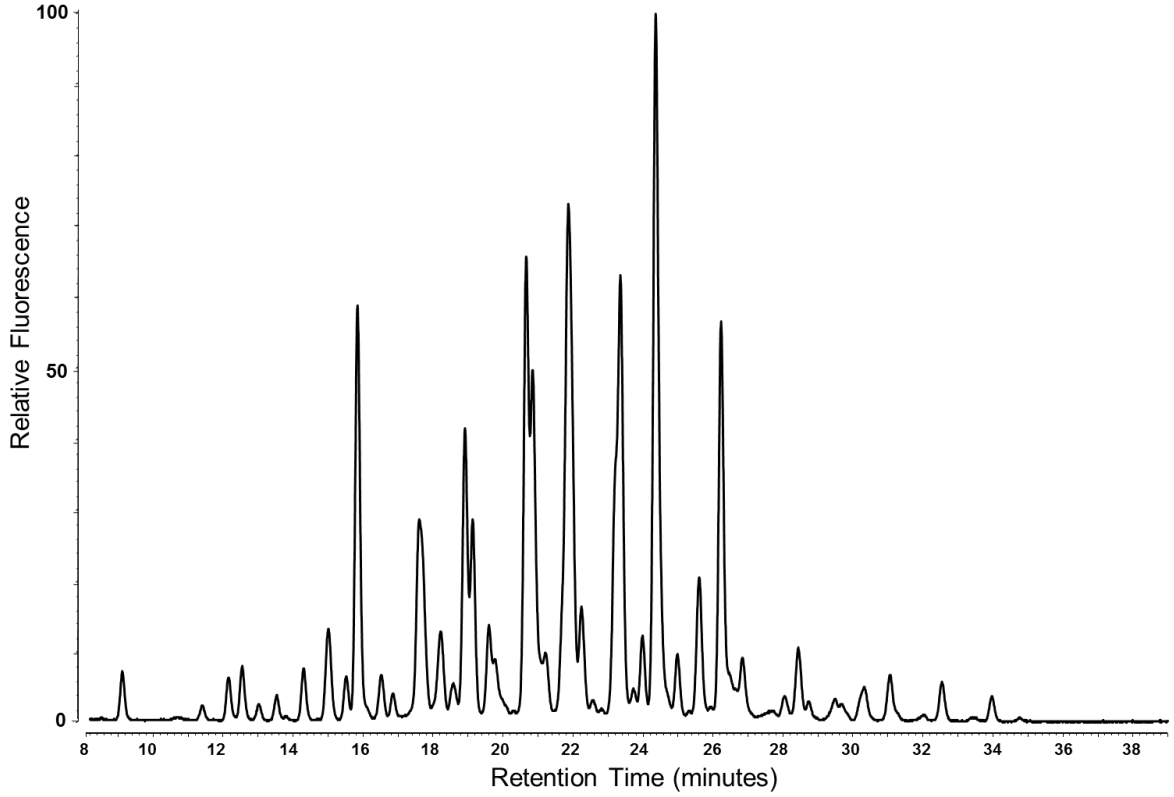


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

IAVI C22 (C)

LWVTVYYGVPVWREAKTTLFCASDAKAYEREVHNVWATHACVPTDPNPQEMVLGNVTE
NFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLECVNVNASSEMRNCSFNATTEL
KDKKRKEQALFYKLDIVPLNGDNSSYRLINCNTSTVAQACPKVTFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCHNVSTVQCTHGIRPVVSTQLLLNGSIAEGEIIIRSENLTDNAKTIIVHLNESL
SIVCTRPNNNTRKSIRIGPGQTFYATNGIIGDIRQAHCNISEGKWNRTLQMVRGKLKEHFPNK
TIKFQPSSGGDLEVTTHSFNCRGEFFYCNTSALYNISNLSNGTESNDSTITLPCKIKQIINMWQ
EVGRAMYAPPIEGNITCNSSITGLLLVRDGGYNGSNQEEIFRPGGGNMRDNWRSELYKYKV
VEIKPLGIAPTKARTKHHHHHH

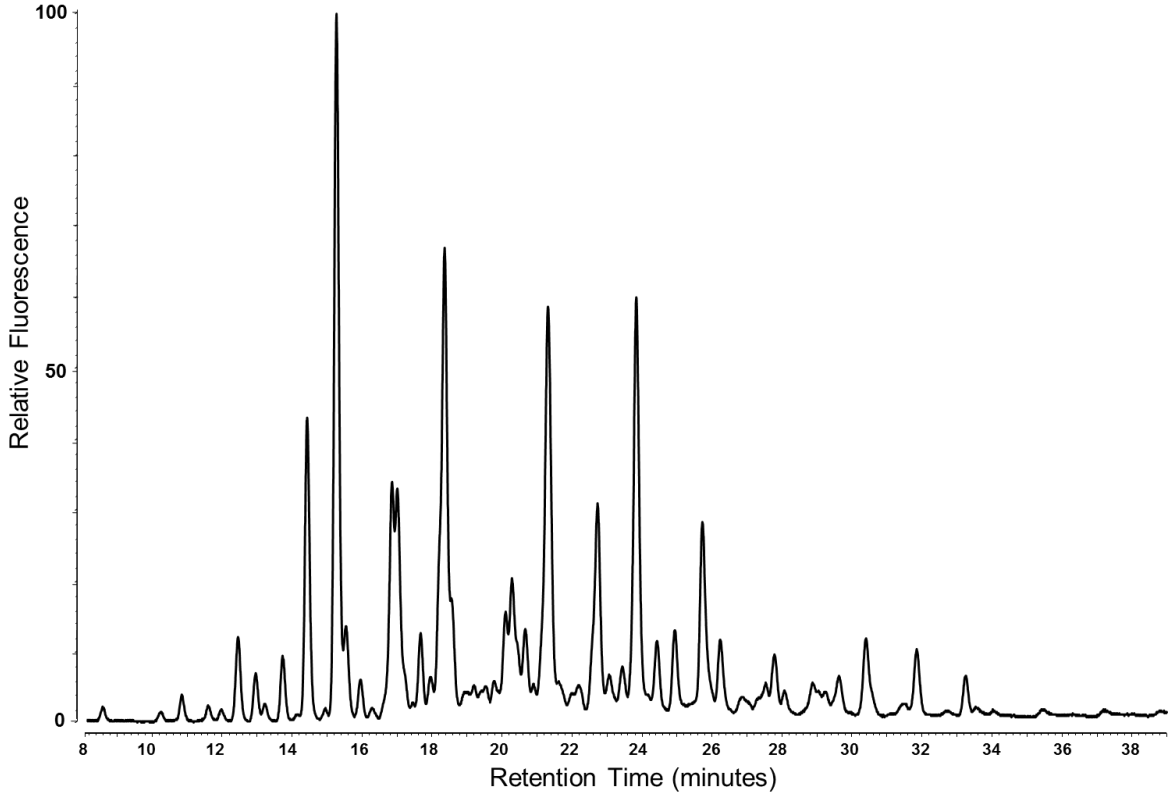


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

704809221_1B3 (C)

LWVTVYYGVPVWRDANTTLFCASDAKAYDREVHNVWATHACVPTDPSPQEMVLRNVTE
NFNMWKNDMVDQMHEDIISLWDQSLKPCVKLTPLCVTLSCTNANITNTNANSTNSTSTNA
NKTSINNDMQEIKNCSFNMTTELRDKQKKEYALFYRLDIVPLEKGNNASNYSDYRLINCNTS
AIKQACPKVSFEPIPIHYCAPAGYAILKCNSKTFNGTGPCLNVSTVQCTHGIKPVVSTQLLLN
GSLAEEEIVIRSENITNNAKTIIVHLNESVGIVCTRPGNNTRKSIRIGPGQAFYATGDIIGDIRQA
HCNISEEAWNRTLLRVAKKLREYFPNKTIAFDSPSGGDLEIVTHTFNCGGEFFYCNTSDLFNR
VYNTTGTYNSTERRNSTITIQCRIKQIINMWQRVGQAMYAPPIAGNITCKSNITGLLLTRDGG
QNIKNETNKETFRPGGGDMRDNWRSELYKYKVVEIKPLGVAPTGARTKHHHHHH

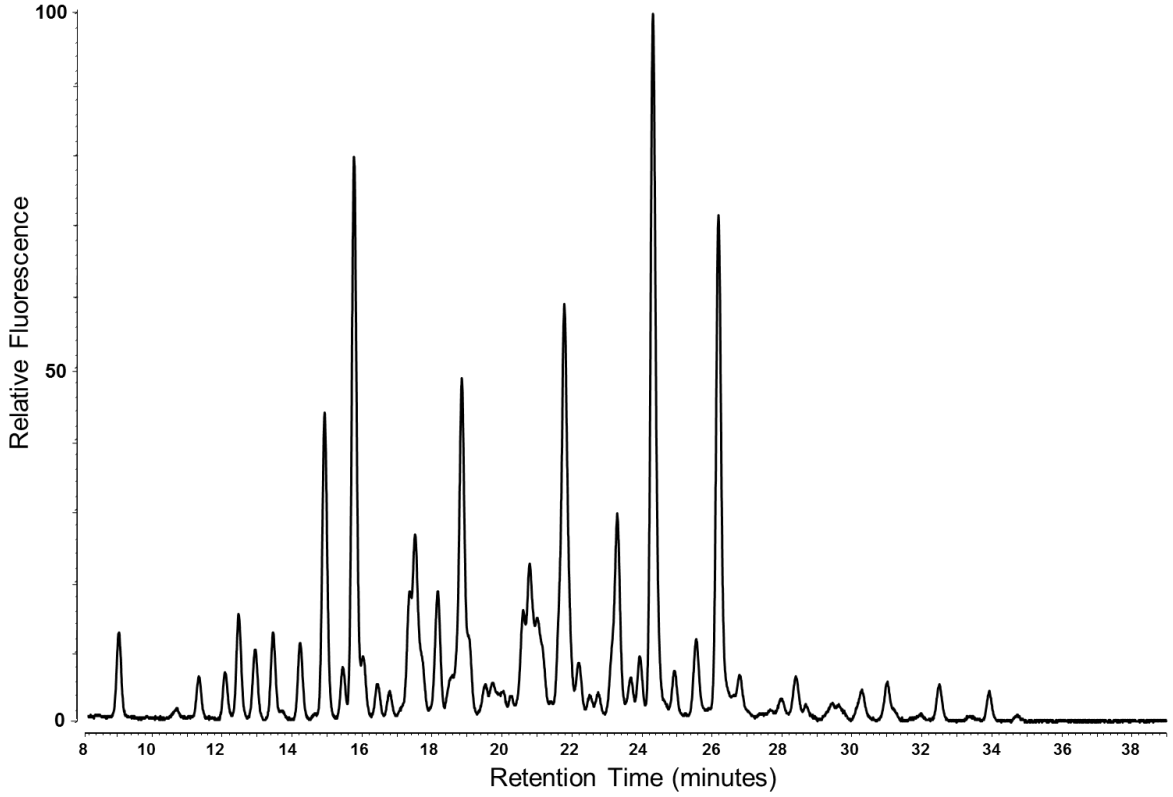


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

92TH021 (CRF01_AE)

LWVTVYYGVPVWRDADTTLFCASDAKAHETEMHNVWATHACVPTDPNPQEIHLENVTEN
FNMWKNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLKCTNANLANVNNRTNDSNIIGNI
TDEIRNCSFNMTTEIRDRKQKVHALFYKLDIVQIEDDKNSSEEYRLINCNTSVIKQACPKISFD
PIPIHYCTPAGYAILKCNDKNFNGTGPCKNVSSVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSE
NLTNNAKTIIVHLNKSVEINCTRPSNNTRTSITIGPGQVFYRTGDIIGDIRKAYCEINGTKWNE
ALKQVAEKLKEHFNNKTIIFQPPSGGDLEITTHHFNCRGEFFYCNTTQLFNSTCIGNETMEGC
NGTIILPCKIKQIMNMWQGVGQAMYAPPISGRINCVSNITGILLTRDGGANTTNNETFRPGGG
NIKDNWRSELYKYKVVQIEPLGIAPTRARTKHHHHHH

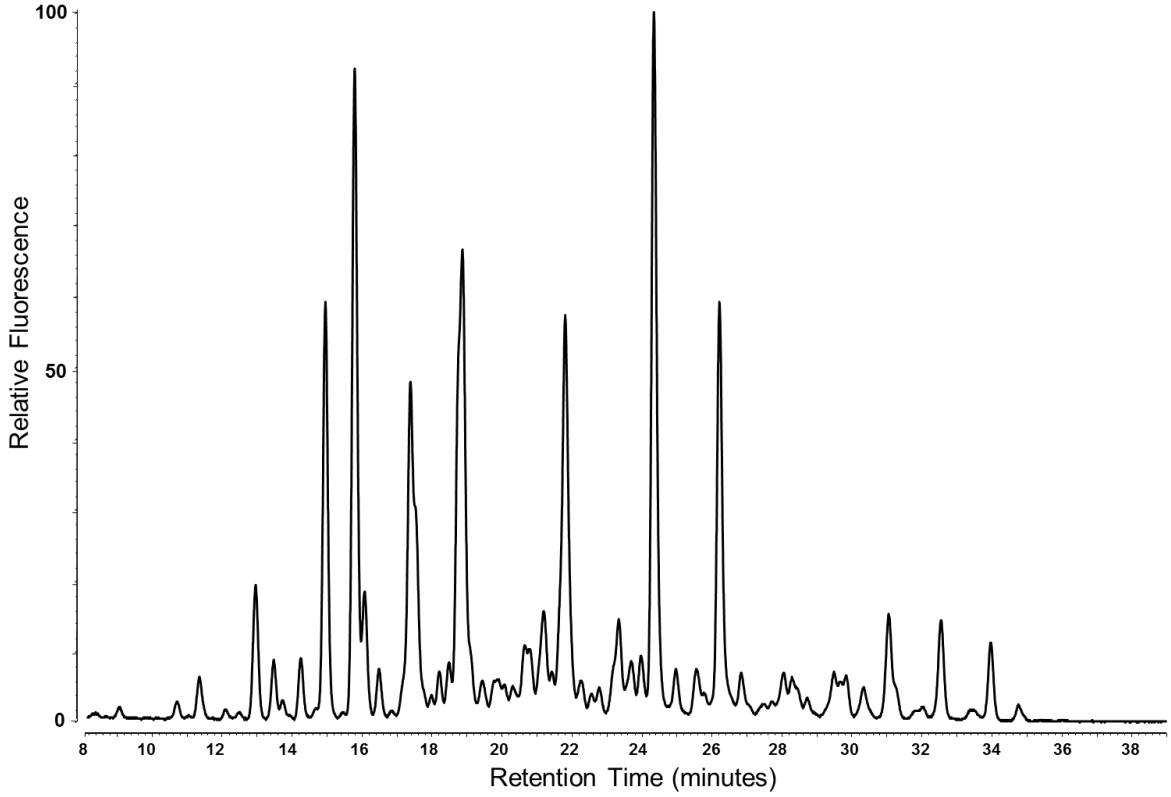


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

BJOX015000.11.5 (CRF01_AE)

LWVTVYYGVPVWRDADTTLFCASDAKAHMTEVHNVWATHACVPTDPNPQEIPLENVTEN
FNMWKNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTNANLNHNLTHNITYGNNNIG
NITDEVKNCTFNMTTEIRGKQQKVHALFYALDIVHMGENGSEYRLISCNTSVIKQACPKISFD
PIPIHYCAPAGYAILKCNDKKFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSE
NLTNNVKTIIVQFNKSVSINCTRPSNNTRTSVRIGPGQMFYRTGDIIGDIRKAYCEINGTEWNE
TLNQVTEKLKEHFKNKTIVFQPPSGGDLETTMHHFNCRGEFFYCNTTKLFNSTENGTMEGR
NTTIILPCKIKQIVNMWQGVGQTMYAPPISGIISCTSNITGILLTRDGGSGDNATETFRPGGGNI
KDNWRSELYKYKVVQIEPLGIAPTRARTKHHHHHH

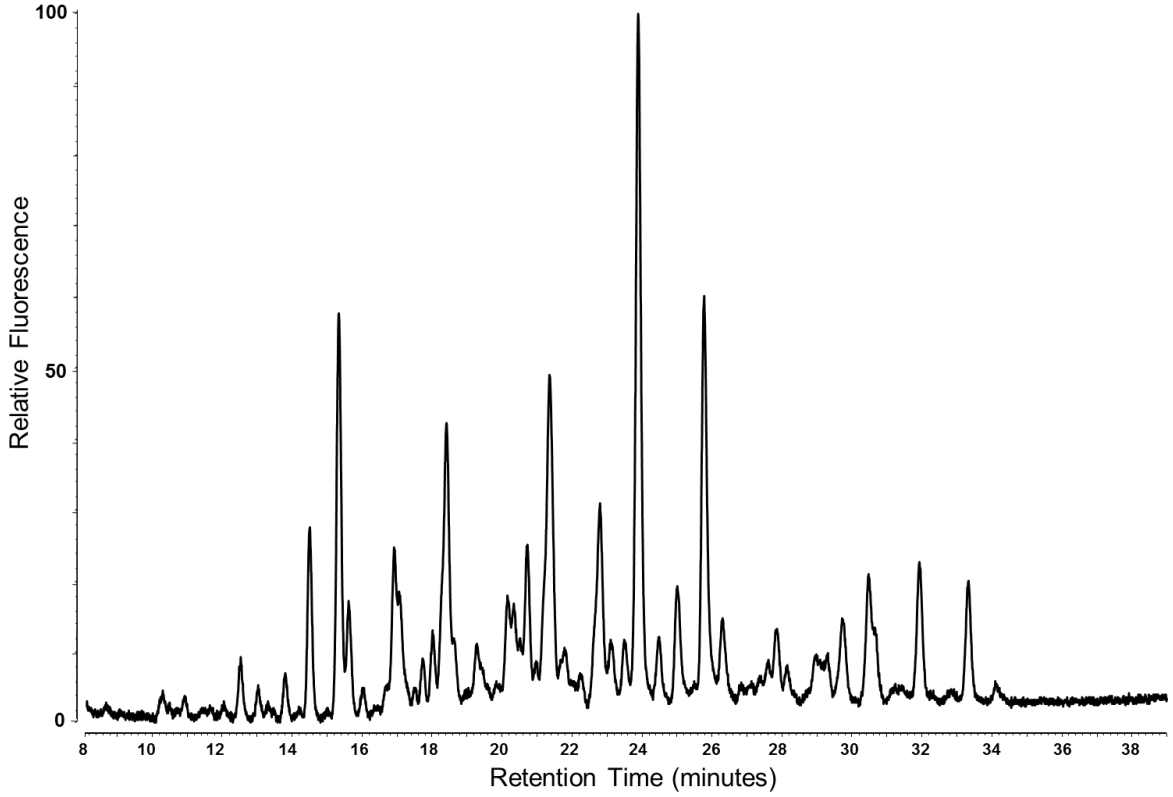


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

C3347.C11 (CRF01_AE)

LWVTVYYGVPVWRDAKTTLFCASDAKAHEREVHNVWATHACVPTDPNPQEILMKNVTEN
FNMWTNNMVEQMQEDVISLWDQSLKPCVKLTPLCVTLNCTKANLTSDTTNRTTGNRIDEV
GNMTDEVKNCTFNMTTELKDKKQKVHALFYKLDIVPIKGNENSSGEYRLINCNTSVIKQAC
PKISFDPIPIHYCTPAGYAILKCNDKNFNGTGPCKNVSSVQCTHGIKPVVSTQLLLNGSLAEEE
IIIRAENLENNAKTIIVHLNKSVEINCTRPSNNTRTSIPIGPGQVFYRTGDIIGNIRKAYCEINGT
KWHKVLKQVAEKLREHFINKTIKFRPPLGGDLEITMHHFNCRGEFFYCNTTKLFNSTYRENE
TREGNDSTIIIPCKIKQIINMWQGTGQAMYAPPISGNINCVSNITGILLTRDGGNNSADNETFR
PGGGNIKDNWRSELYKYKVVEIEPLGIAPTRARTKHHHHHH

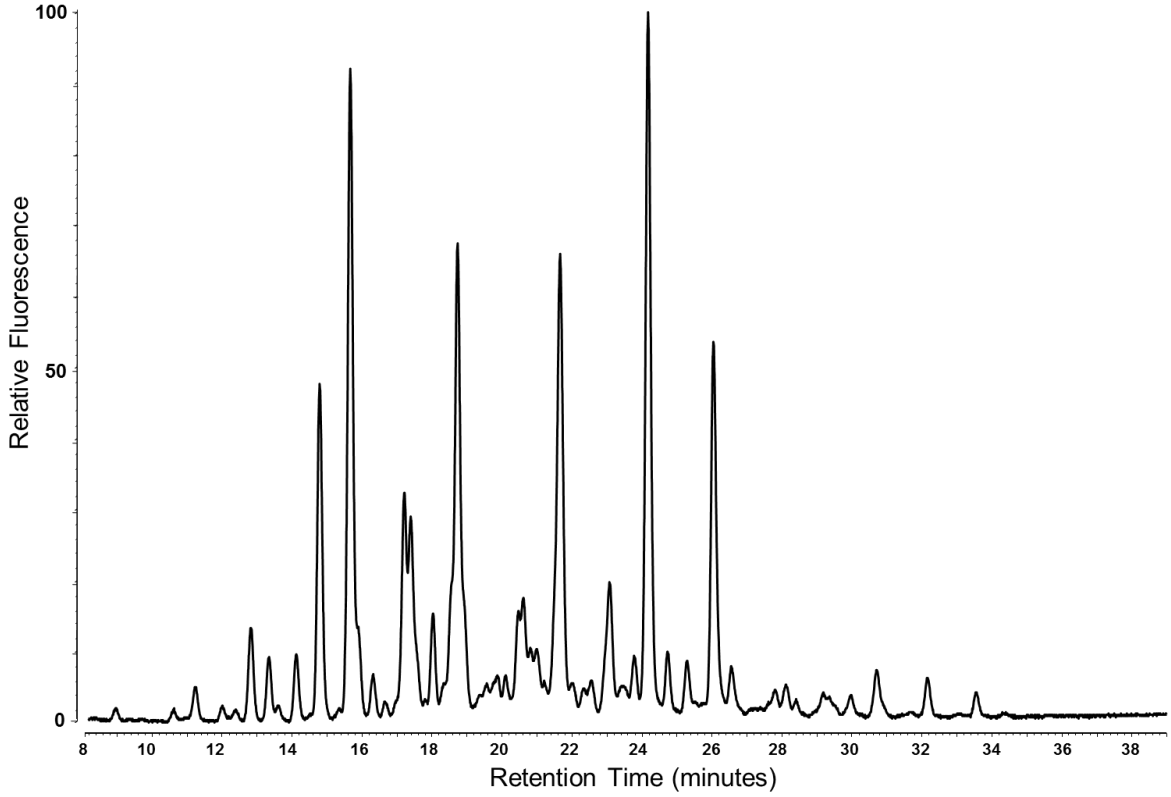


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

P0402_c2_11 (G)

LWVTVYYGVPVWEDADTPLFCASDAKAYSTESHNVWATHACVPTDPSPQEISLDNVTENF
NMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLNCTNVNNSSATNNSMVDDREGLK
NCSFNITTELRDKKKQEHALFYRLDIVPINGNSNSNSSVGDYRLINCNVSTIKQACPKMSFDPI
PIHYCAPAGFAILKCRDKKFNGTGSCKNVSTVQCTHGIKPVISTQLLLNGSVAEEEIMIRSENF
TNNAKNIIVQFNKTIDIMCTRPNNNTRKSISLGPGQAIYATGDIIGNIRQAHCNISGADWGNMI
RNVSEKLKEIFNKTTITFKASAGGDLEITTHSFNCRGEFFYCDTSDLFNSSRFNNSSNDTNDTI
TLPCKIKQIVRMWQRVGQAMYAPPIAGNITCRSNITGLLLTRDGGGNNTNETETFRPAGGD
MRDNWRSELYKYKVVKINPLGVAPTRARTKHHHHHH

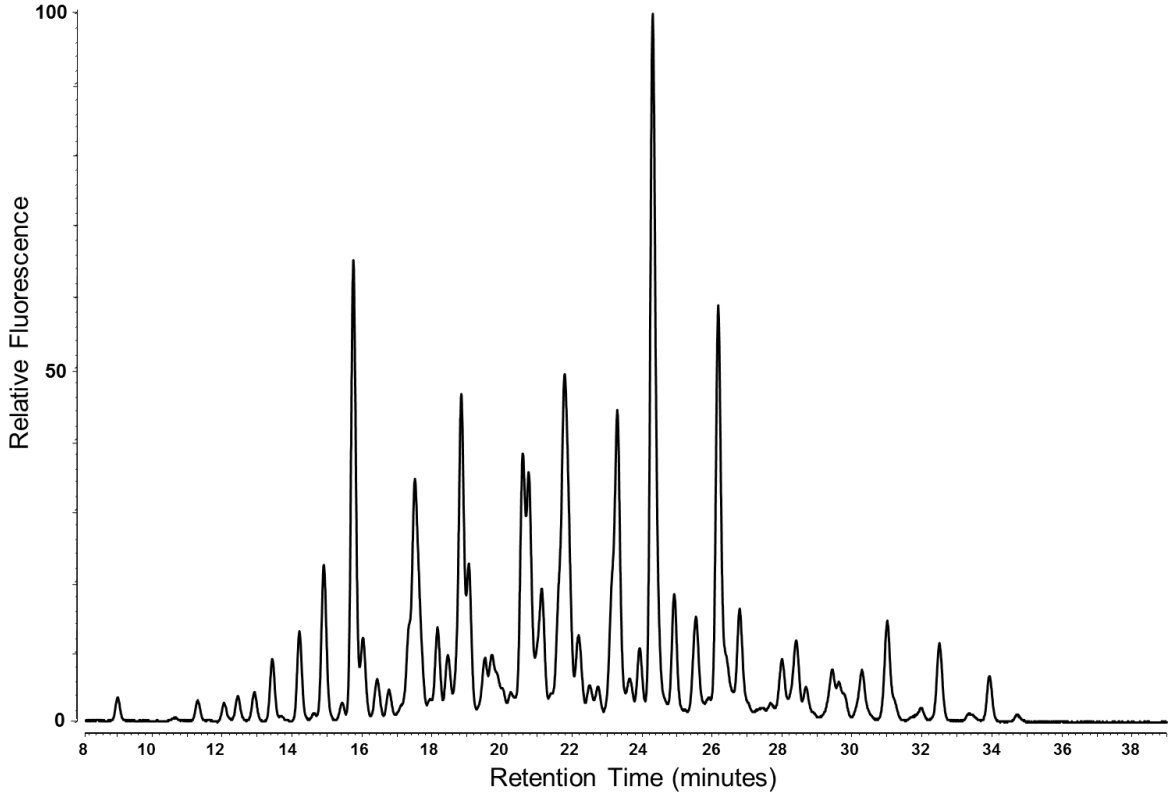


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

X1254_3 (G)

LWVTVYYGVPVWEDANTTLFCASDAKAYSTESHNVWATHACVPTDPNPQEIPLENVTENF
NMWKNNMVEQMHEDIISLWDESLKPCVKLTPLCVTLTCTNVTNNNNATNSSTVESNREEL
RNCSFNTTTAIQDKKKQEYALFYKLDIVPITGSGNSSAGNYRLINCNVSTIKQACPKVTFDPIP
IHYCAPAGFAILKCRDKKFNGTGPCKNVSTVQCTHGIKPVISTQLLLNGSLAEEEIIIRSENITD
NTKNIIVQLKEAINIICTRPSNNTRKSISLGPGQAFFATGDIIGDIRQAHCNISGTDWNNMIQNV
SEKLKGIFNKSSITFNPPSGGDLEIITHSFNCRGEFFYCNTSKLFNSSLFNSSTNSNNETITIPCKI
KQIVRMWQRVGQAMYAPPIAGNITCSSNITGLILVRDGGTNNNTNETVRPAGGDMRDNWR
SELYKYKIVKIKPLGIAPTRARTKHHHHHH

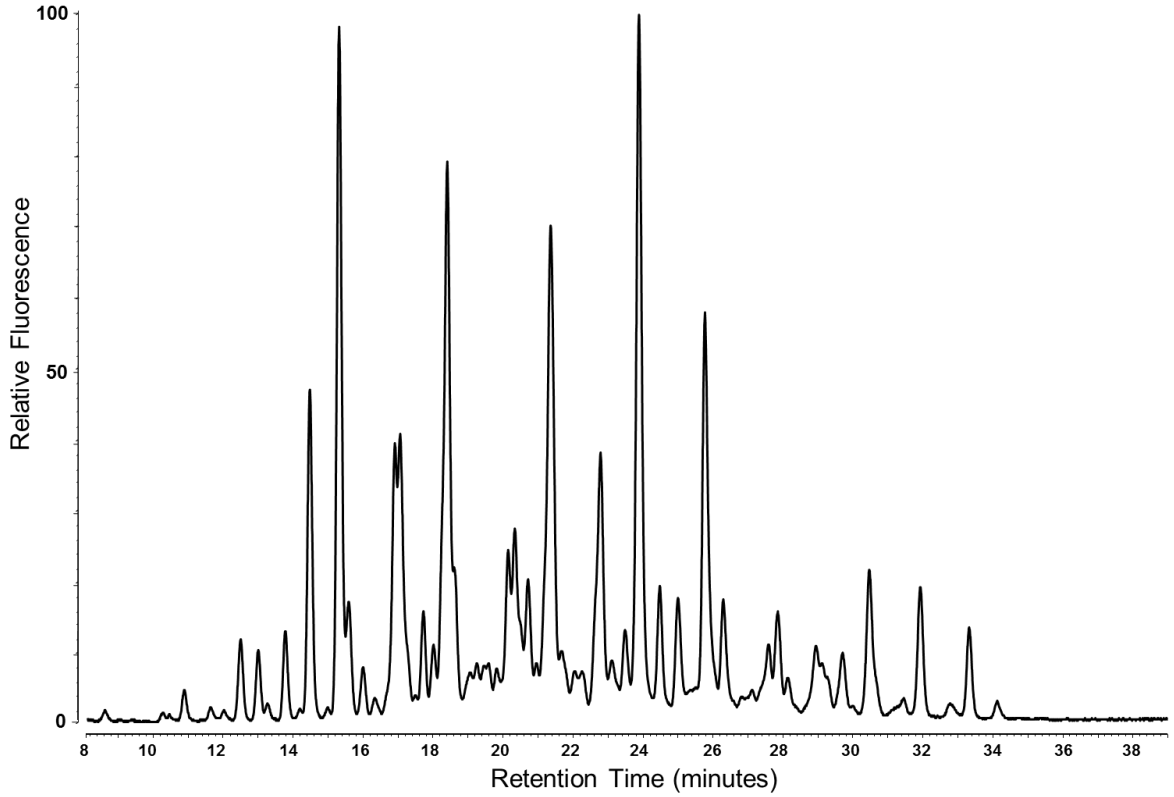


Dataset A – Protein sequences and HILIC-UPLC glycan profiles for the gp120 panel

(continued)

X2088-c9 (G)

LWVTVYYGVPAWEDADTTLFCASDAKAYSTEKHNVWATHACVPTDPDPQEIPLENVTENF
NMWKNNMVEQMHEDIISLWDESLKPCVMLTPLCVTLNCSAVNGTMNEQMKNCTFNITTEI
RDKKKEEYALFYKLDIEQISKTANSTTANSNYSSYRLINCNVSTIKQACPKVTFEPIPIHYCAP
AGFAILKCRDEDFNGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAKGDIVIRSENLTNNAK
VIIVQLNEPVQIVCIRPNNNTRRSIHLGPGRAFYATDIIGDIRQAHCNVTRGKWVNITKNVKE
QLWKIFNKTTNITFNNTIFNSPAGGDLEITTHSFNCGGEFFYCNTSDLFNETNLSANHTDTNE
NITLQCRIKQIVRMWQRVGQAMYAPPIAGNITCISNITGLLLTRDGVNDTHDKENETFRPTG
GDMRDNWRSELYKYKVIKLKPLGVAPTKARTKHHHHHH



Dataset B – Abundances of oligomannose-type glycans

Isolate	Abundance (% of total glycans)					
	M5	M6	M7	M8	M9	**Total**[†]
0260.V5.C36	7	6	6	10	14	**42**
191084 B7.19	5	7	10	12	8	**42**
Q23-17	9	7	7	9	5	**38**
BG505	5	6	8	10	9	**39**
94UG103	8	7	8	9	14	**47**
92RW020	4	5	7	11	8	**35**
Q842.d12	6	6	7	6	2	**27**
BaL	3	3	6	11	12	**34**
JRCSF	3	4	7	8	9	**29**
HXB2	6	5	8	12	10	**41**
SF162	6	4	8	14	13	**44**
TRJO_A2	8	6	8	12	7	**41**
1056_TA11_1826	5	6	7	9	8	**35**
1012_11_TC21	5	6	8	12	10	**41**
JRFL	5	6	7	11	12	**42**
96ZM651	5	5	8	7	5	**29**
98ZADu156	7	6	7	9	8	**37**
ZM214M_SGA_A3	4	4	5	7	3	**24**
IAVI C22	6	6	10	11	6	**39**
704809221_1B3	9	8	8	7	3	**35**
92TH021	7	6	9	12	9	**42**
BJOX015000.11.5	8	6	7	10	6	**37**
C3347.C11	5	5	7	10	6	**33**
P0402_c2_11	11	9	9	12	6	**46**
X1254_3	5	5	7	10	6	**33**
X2088-c9	7	7	7	8	4	**34**
Average	**6**	**6**	**8**	**10**	**8**	**37**

[†]Represents the sum of the $Man_{5\text{-}9}GlcNAc_2$ abundances.

Appendix III

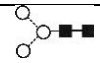
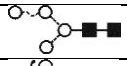
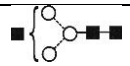
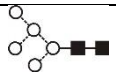
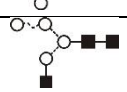
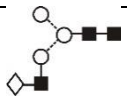
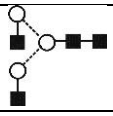
Site-specific glycosylation analysis of gp120$_{BaL}$ – supplementary data.

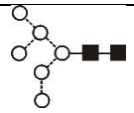
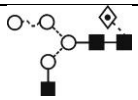
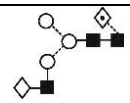
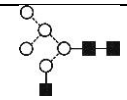
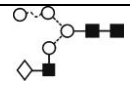
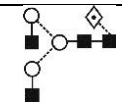
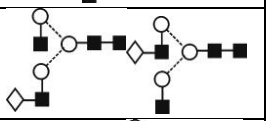
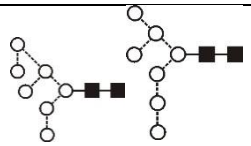
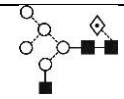
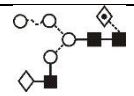
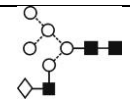
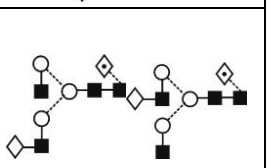
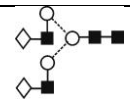
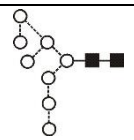
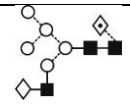
Supplementary methods

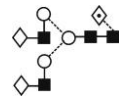
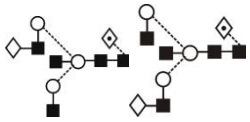
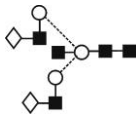
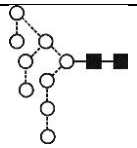
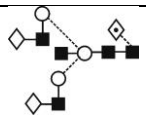
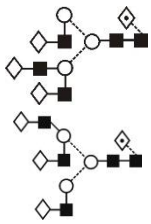
Ion-mobility ESI-MS/MS analysis of gp120$_{BaL}$ glycans

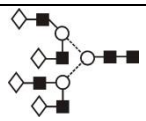
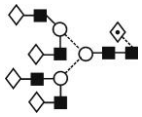
Gp120BaL was expressed and purified as described the main text. Glycans were released using PNGase F and dried down. The sample was dissolved in 5 µL water. 1.0 µL was cleaned with a Nafion membrane (approx 2 h), diluted with water (4 µL) and methanol (4 µL), and 1 µL on a 1 M solution of ammonium phosphate was added. The sample was spun at 1200 rpm for 1 min and infused into a Waters Synapt G2 ion mobility Q-TOF mass spectrometer with Waters thin-wall nanospray capillaries. MS and MS/MS spectra were acquired with ion mobility on (wave velocity 450 m/sec and wave height 40 V). The ion source temperature was 120^{o}C. Ion mobility data were processed with Waters DriftScope software (version 2.1) and instrument control and MS data processing employed Waters MassLynx software. Spectra were recorded in negative ion mode.

Table S3.1 - Compositions and proposed structures for the glycans from gp120$_{BaL}$

m/z[1]		Composition				Charge	Frag.[2]	Structure
Found	Calc.	Hex	HexNAc	dHex	Neu5Ac			
1007.3	1007.3	3	2	0	0	1	-	
1169.4	1169.3	4	2	0	0	1	-	
1210.4	1210.4	3	3	0	0	1	-	
1331.4	1331.4	5	2	0	0	1	+	
1372.4	1372.4	4	3	0	0	1	+	
1372.4	1372.4	4	3	0	0	1	+	
1565.5	1565.5				1	1	-	
1413.5	1413.5	3	4	0	0	1	-	

1493.5	1493.5	6	2	0	0	1	+	
1518.5	1518.5	4	3	1	0	1	+	
1518.5	1518.5	4	3	1	0	1	+	
1711.6	1711.6				1	1	+	
1534.5	1534.5	5	3	0	0	1	+	
1534.5	1534.5	5	3	0	0	1	+	
1727.6	1727.6				1	1	+	
1559.5	1559.5	3	4	1	0	1	+	
1575.5	1575.5	4	4	0	0	1	+	
1655.5	1655.5	7	2	0	0	1	+	
1680.5	1680.5	5	3	1	0	1	+	
1680.5	1680.5	5	3	1	0	1	+	
1873.7	1873.7				1	1	+	
1696.5	1696.5	6	3	0	0	1	+	
1889.7	1889.7				1	1	+	
1721.6	1721.6	4	4	1	0	1	+	
1914.7	1914.7				1	1	-	
1762.6	1762.6	3	5	1	0	1	-	-
1737.6	1737.6	5	4	0	0	1	+	
1930.7	1930.7				1	1	+	
1778.6	1778.6	4	5	0	0	1	-	-
1817.6	1817.6	8	2	0	0	1	+	
957.3	957.3				0 ($(H_2PO_4)_2$)	2	-	
1842.7	1842.6	6	3	1	0	1	+	
2035.8	2035.7				1	1	+	
1856.6	1856.6	7	3	0	0	1	-	-

1883.7	1883.6	5	4	1	0	1	+	
990.3	990.3				0 ($(H_2PO_4)_2$)	2	-	
2076.8	2076.7				1	1	+	
1086.9	1086.9				1(Na, H_2PO_4)	2	+	
2367.8	2367.8				2	1	-	
2389.8	2389.8				2(Na)	1	-	
1183.4	1183.4				2	2	+	
1899.6	1899.6	6	4	0	0	1	-	-
1924.6	1924.6	4	5	1	0	1	+	
2117.8	2117.7				1	1	-	
1940.7	1940.6	5	5	0	0	1	-	
1979.6	1979.6	9	2	0	0	1	+	
1038.3	1038.3				0 ($(H_2PO_4)_2$	2	-	
2045.7	2045.7	6	4	1	0	1	-	-
2086.7	2086.7	5	5	1	0	1	+	
2279.8	2279.8				1	1	-	
2102.8	2102.7	6	5	0	0	1	-	Triantennary
2295.8	2295.8				1	1	-	
1293.0	1292.9				2	2	-	
958.7	958.7				3	3	-	
2127.8	2127.7	4	6	1	0	1	-	-
2248.8	2248.8	6	5	1	0	1	-	
2441.9	2441.9				1	1	-	
1269.4	1269.2				1(Na, H_2PO_4)	2	+	
1366.0	1366.0				2	2	+	
1522.5	1522.5				3(Na)	2	-	
1007.4	1007.3				3	3	-	
2289.8	2289.8	5	6	1	0	1	-	
2482.9	2482.9				1	1	-	

2305.9	2305.8	6	6	0	0	1	-	Tetra-antennary lacking galactose?
2644.9	2644.9				1	1	-	
2451.9	2451.8	6	6	1	0	1	-	Fucosylated tetra-antennary lacking galactose?
2660.9	2660.9	7	6	0	1	1	-	
1475.5	1475.5				2	2	-	
1084.4	1084.4				3	3	-	
2613.9	2613.9	7	6	1	0	1	-	
2807.1	2807.0				1	1	-	
1608.5	1608.5				1(Na, H_2PO_4)	2	-	
3098.1	3098.1				2	1	-	
3120.2	3120.2				2(Na)	1	-	
1548.6	1548.5				2	2	+	
1694.1	1694.1				3	2	+	
1705.1	1705.1				3(Na)	2	-	
1129.1	1129.1				3	3	+	
1839.7	1839.6				4	2	-	
1850.6	1850.6				4(Na)	2	-	
1226.1	1226.1				4	3	-	
1233.4	1233.4				4(Na)	3	+	

1) Charge type as in the tables above
2) Frag. = A + sign indicates that fragmentation spectra were obtained.

Table S3.2 – Glycopeptides identified from tryptic digestion of $gp120_{BaL}$ by MALDI MS.

m/z		Peptide	Mods	Peptide mass [M+1]	Site	Fraction	Composition				Predicted structure	Intensity[§]
Found	Calc.						Hex	HexNAc	dHex	Neu5 Ac		
1615.3	1614.7	EQFGNK		722.3	356	T10	3	2	0	0	M3	295
1939.5	1938.8	EQFGNK		722.3	356	T10	5	2	0	0	M5	195
2101.6	2100.8	EQFGNK		722.3	356	T10	6	2	0	0	M6	232
2263.9	2262.9	EQFGNK		722.3	356	T10	7	2	0	0	M7	5857
2426.0	2424.9	EQFGNK		722.3	356	T10	8	2	0	0	M8	1172
2588.1	2587.0	EQFGNK		722.3	356	T10	9	8	0	0	M9	181
2101.3	2100.8	EQFGNK		722.3	356	T11	6	2	0	0	M6	3526
2263.3	2262.9	EQFGNK		722.3	356	T11	7	2	0	0	M7	416
1777.3	1776.7	EQFGNK		722.3	356	T12	4	2	0	0	M4	492
1939.4	1938.8	EQFGNK		722.3	356	T12	5	2	0	0	M5	4630
1980.4	1979.8	EQFGNK		722.3	356	T12	4	3	0	0	M4A1	633
2141.4	2141.8	EQFGNK		722.3	356	T12	5	3	0	0	M5A1	455
2304.6	2303.9	EQFGNK		722.3	356	T12	6	3	0	0	M5A1G1	295
1709.9	1709.6	NCSFK	C	655.3	156	T13	4	2	0	0	M4	669
1872.0	1871.7	NCSFK	C	655.3	156	T13	5	2	0	0	M5	6584
1913.0	1912.7	NCSFK	C	655.3	156	T13	4	3	0	0	M4A1	604
2059.1	2058.8	NCSFK	C	655.3	156	T13	4	3	1	0	FM4A1	283
2075.1	2074.8	NCSFK	C	655.3	156	T13	5	3	0	0	M4A1G1 or M5A1	959
2237.0	2236.8	NCSFK	C	655.3	156	T13	6	3	0	0	M5A1G1 or M6A1	839
2277.3	2277.8	NCSFK	C	655.3	156	T13	5	4	0	0	A2G2	189
2528.3	2527.9	NCSFK	C	655.3	156	T13	6	3	0	1	M5A1G1S1	410

2673.7	2674.0	NCSFK	C	655.3	156	T13	6	3	1	1	FM5A1G1S1	258
2526.5	2526.0	QAHCNLSR	C	985.4	332	T13	7	2	0	0	M7	138
2688.6	2688.0	QAHCNLSR	C	985.4	332	T13	8	2	0	0	M8	2679
2850.8	2850.1	QAHCNLSR	C	985.4	332	T13	9	2	0	0	M9	2530
1963.4	1963.8	EQFGNK		722.3	356	T14	3	3	1	0	FA1	309
2126.2	2125.9	EQFGNK		722.3	356	T14	4	3	1	0	FM4A1	249
2141.5	2141.8	EQFGNK		722.3	356	T14	5	3	0	0	M4A1G1 or M5A1	299
2166.3	2166.9	EQFGNK		722.3	356	T14	3	4	1	0	FA2	1264
2329.3	2328.9	EQFGNK		722.3	356	T14	4	4	1	0	FA2G1	376
2491.2	2491.0	EQFGNK		722.3	356	T14	5	4	1	0	FA2G2	335
2782.1	2782.1	EQFGNK		722.3	356	T14	5	4	1	1	FA2G2S1	394
2099.2	2099.8	NCSFK	C	655.3	156	T14	3	4	1	0	FA2	243
2115.6	2115.8	NCSFK	C	655.3	156	T14	4	4	0	0	A2G1	286
2262.5	2261.9	NCSFK	C	655.3	156	T14	4	4	1	0	FA2G1	468
2022.8	2022.8	QAHCNLSR	C, P-Q	968.4	332	T17	4	2	0	0	M4	103
2184.9	2184.9	QAHCNLSR	C, P-Q	968.4	332	T17	5	2	0	0	M5	276
2346.9	2346.9	QAHCNLSR	C, P-Q	968.4	332	T17	6	2	0	0	M6	309
2509.0	2509.0	QAHCNLSR	C, P-Q	968.4	332	T17	7	2	0	0	M7	632
2671.0	2671.0	QAHCNLSR	C, P-Q	968.4	332	T17	8	2	0	0	M8	5281
2833.1	2833.1	QAHCNLSR	C, P-Q	968.4	332	T17	9	2	0	0	M9	4925
2022.8	2022.8	QAHCNLSR	C, P-Q	968.4	332	T18	4	2	0	0	M4	303
2185.0	2184.9	QAHCNLSR	C, P-Q	968.4	332	T18	5	2	0	0	M5	1655
2210.0	2209.9	QAHCNLSR	C, P-Q	968.4	332	T18	3	3	1	0	FA1	233
2372.0	2371.9	QAHCNLSR	C, P-Q	968.4	332	T18	4	3	1	0	FM4A1	187
2388.1	2387.9	QAHCNLSR	C, P-Q	968.4	332	T18	5	3	0	0	M4A1G1 or M5A1	85

2413.1	2413.0	QAHCNLSR	C, P-Q	968.4	332	T18	3	4	1	0	FA2	476
2509.1	2509.0	QAHCNLSR	C, P-Q	968.4	332	T18	7	2	0	0	M7	99.7
2534.1	2534.0	QAHCNLSR	C, P-Q	968.4	332	T18	5	3	1	0	FM4A1G1 or FM5A1	274
2575.1	2575.0	QAHCNLSR	C, P-Q	968.4	332	T18	4	4	1	0	FA2G1	176
2616.2	2616.1	QAHCNLSR	C, P-Q	968.4	332	T18	3	5	1	0	FA3	240
2697.1	2696.1	QAHCNLSR	C, P-Q	968.4	332	T18	6	3	1	0	FM5A1G1	123
2737.2	2737.1	QAHCNLSR	C, P-Q	968.4	332	T18	5	4	1	0	FA2G2	272
2778.2	2778.1	QAHCNLSR	C, P-Q	968.4	332	T18	4	5	1	0	FA3G1	205
2940.3	2940.2	QAHCNLSR	C, P-Q	968.4	332	T18	5	5	1	0	FA3G2	153
3028.4	3028.2	QAHCNLSR	C, P-Q	968.4	332	T18	5	4	1	1	FA2G2S1	120
2532.2	2532.1	LREQFGNK		991.5	356	T19	7	2	0	0	M7	4306
2694.2	2694.1	LREQFGNK		991.5	356	T19	8	2	0	0	M8	3575
2856.3	2856.2	LREQFGNK		991.5	356	T19	9	2	0	0	M9	922
3125.5	3125.3	LREQFGNK		991.5	356	T19	6	5	1	0	FA3G3	301
3416.5	3416.4	LREQFGNK		991.5	356	T19	6	5	1	1	FA3G3S1	243
3490.7	3490.4	LREQFGNK		991.5	356	T19	7	6	1	0	FA4G4	221
3781.8	3781.5	LREQFGNK		991.5	356	T19	7	6	1	1	FA4G4S1	92
4072.9	4072.6	LREQFGNK		991.5	356	T19	7	6	1	2	FA4G4S2	79
2046.0	2045.9	LREQFGNK		991.5	356	T20	4	2	0	0	M4	485
2087.0	2086.9	LREQFGNK		991.5	356	T20	3	3	0	0	A1	336
2192.0	2192.0	LREQFGNK		991.5	356	T20	4	2	1	0	FM4	287
2208.0	2208.0	LREQFGNK		991.5	356	T20	5	2	0	0	M5	7671
2233.1	2233.0	LREQFGNK		991.5	356	T20	3	3	1	0	FA1	864
2249.1	2249.0	LREQFGNK		991.5	356	T20	4	3	0	0	M4A1	666
2290.1	2290.0	LREQFGNK		991.5	356	T20	3	4	0	0	A2	307
2370.1	2370.0	LREQFGNK		991.5	356	T20	6	2	0	0	M6	1100

2395.1	2395.0	LREQFGNK	991.5	356	T20	4	3	1	0	FM4A1	3609
2411.1	2411.0	LREQFGNK	991.5	356	T20	5	3	0	0	M4A1G1 or M5A1	1423
2436.1	2436.1	LREQFGNK	991.5	356	T20	3	4	1	0	FA2	4607
2532.1	2532.1	LREQFGNK	991.5	356	T20	7	2	0	0	M7	1488
2557.2	2557.1	LREQFGNK	991.5	356	T20	5	3	1	0	FM5A1	3285
2573.2	2573.1	LREQFGNK	991.5	356	T20	6	3	0	0	M5A1G1 or M6A1	591
2598.2	2598.1	LREQFGNK	991.5	356	T20	4	4	1	0	FA2G1	1357
2639.2	2639.1	LREQFGNK	991.5	356	T20	3	5	1	0	FA3	1693
2719.2	2719.1	LREQFGNK	991.5	356	T20	6	3	1	0	FM5A1G1 or FM6A1	740
2760.3	2760.2	LREQFGNK	991.5	356	T20	5	4	1	0	FA2G2	1485
2801.3	2801.2	LREQFGNK	991.5	356	T20	4	5	1	0	FA3G1	431
2842.3	2842.2	LREQFGNK	991.5	356	T20	3	6	1	0	FA4	934
2848.2	2848.2	LREQFGNK	991.5	356	T20	5	3	1	1	FM4A1G1S1	493
2963.3	2963.3	LREQFGNK	991.5	356	T20	5	5	1	0	FA3G2	383
3004.4	3004.3	LREQFGNK	991.5	356	T20	4	6	1	0	FA4G1	245
3010.3	3010.2	LREQFGNK	991.5	356	T20	6	3	1	1	FM5A1G1S1	189
3051.3	3051.3	LREQFGNK	991.5	356	T20	5	4	1	1	FA2G2S1	789
3125.5	3125.3	LREQFGNK	991.5	356	T20	6	5	1	0	FA3G3	176
3207.5	3166.3	LREQFGNK	991.5	356	T20	5	6	1	0	FA4G2	100
1945.0	1944.8	WNDTLNK	890.4	339	T21	4	2	0	0	M4	462
2107.0	2106.9	WNDTLNK	890.4	339	T21	5	2	0	0	M5	229
2269.1	2268.9	WNDTLNK	890.4	339	T21	6	2	0	0	M6	184
2431.2	2431.0	WNDTLNK	890.4	339	T21	7	2	0	0	M7	332
2593.2	2593.0	WNDTLNK	890.4	339	T21	8	2	0	0	M8	1511
2755.3	2755.1	WNDTLNK	890.4	339	T21	9	2	0	0	M9	4979

1782.9	1782.8	WNDTLNK	890.4	339	T22	3	2	0	0	M3	264
1944.9	1944.8	WNDTLNK	890.4	339	T22	4	2	0	0	M4	1046
1985.9	1985.8	WNDTLNK	890.4	339	T22	3	3	0	0	A1	328
2107.0	2106.9	WNDTLNK	890.4	339	T22	5	2	0	0	M5	2880
2269.0	2268.9	WNDTLNK	890.4	339	T22	6	2	0	0	M6	1236
2310.1	2309.9	WNDTLNK	890.4	339	T22	5	3	0	0	M4A1G1 or M5A1	326
2431.1	2431.0	WNDTLNK	890.4	339	T22	7	2	0	0	M7	1663
2456.1	2456.0	WNDTLNK	890.4	339	T22	5	3	1	0	FM4A1G1 or FM5A1	183
2472.1	2472.0	WNDTLNK	890.4	339	T22	6	3	0	0	M5A1G1 or M6A1	257
2593.1	2593.0	WNDTLNK	890.4	339	T22	8	2	0	0	M8	1222
2755.2	2755.1	WNDTLNK	890.4	339	T22	9	2	0	0	M9	523
2305.3	2306.0	AKWNDTLNK	1089.6	339	T24	5	2	0	0	M5	86
2630.4	2630.1	AKWNDTLNK	1089.6	339	T24	7	2	0	0	M7	120
2792.4	2792.2	AKWNDTLNK	1089.6	339	T24	8	2	0	0	M8	259
2954.5	2954.2	AKWNDTLNK	1089.6	339	T24	9	2	0	0	M9	3737
2144.1	2143.9	AKWNDTLNK	1089.6	339	T25	4	2	0	0	M4	821
2306.2	2306.0	AKWNDTLNK	1089.6	339	T25	5	2	0	0	M5	2710
2331.1	2331.0	AKWNDTLNK	1089.6	339	T25	3	3	1	0	FA1	247
2347.2	2347.0	AKWNDTLNK	1089.6	339	T25	4	3	0	0	M4A1	275
2468.2	2468.0	AKWNDTLNK	1089.6	339	T25	6	2	0	0	M6	1641
2493.3	2493.1	AKWNDTLNK	1089.6	339	T25	4	3	1	0	FM4A1	390
2509.3	2509.1	AKWNDTLNK	1089.6	339	T25	5	3	0	0	M4A1G1 or M5A1	439
2534.3	2534.1	AKWNDTLNK	1089.6	339	T25	3	4	1	0	FA2	611
2630.3	2630.1	AKWNDTLNK	1089.6	339	T25	7	2	0	0	M7	3305

2655.3	2655.1	AKWNDTLNK	1089.6	339	T25	5	3	1	0	FM4A1G1 or FM5A1	387
2671.3	2671.1	AKWNDTLNK	1089.6	339	T25	6	3	0	0	M5A1G1 or M6A1	238
2696.4	2696.2	AKWNDTLNK	1089.6	339	T25	4	4	1	0	FA2G1	171
2737.4	2737.2	AKWNDTLNK	1089.6	339	T25	3	5	1	0	FA3	205
2792.4	2792.2	AKWNDTLNK	1089.6	339	T25	8	2	0	0	M8	7970
2858.4	2858.2	AKWNDTLNK	1089.6	339	T25	5	4	1	0	FA2G2	261
2954.4	2954.2	AKWNDTLNK	1089.6	339	T25	9	2	0	0	M9	7643
3149.6	3149.3	AKWNDTLNK	1089.6	339	T25	5	4	1	1	FA2G2S1	130
3172.5	3172.3	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	3	2	1	0	FM3	155
3188.5	3188.3	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	4	2	0	0	M4	197
3229.6	3229.4	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	3	3	0	0	A1	250
3350.7	3350.4	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	5	2	0	0	M5	1829
3375.7	3375.4	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	3	3	1	0	FA1	453
3391.7	3391.4	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	4	3	0	0	M4A1	154
3432.8	3432.4	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	3	4	0	0	A2	176
3512.7	3512.4	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	6	2	0	0	M6	190
3537.9	3537.5	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	4	3	1	0	FM4A1	215
3553.8	3553.5	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	5	3	0	0	M4A1G1 or M5A1	202
3578.9	3578.5	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	3	4	1	0	FA2	1756
3594.8	3594.5	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	4	4	0	0	A2G1	296
3674.9	3674.5	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	7	2	0	0	M7	99.5
3699.9	3699.5	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	5	3	1	0	FM4A1G1 or FM5A1	265
3741.0	3740.5	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	4	4	1	0	FA2G1	510
3756.8	3756.5	DGGPEDNKTEVFRPGGGDMR	2134.0	463	T29	5	4	0	0	A2G2	114

3782.0	3781.6	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	3	5	1	0	FA3	966
3798.0	3797.6	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	4	5	0	0	A3G1	288
3863.0	3836.5	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	8	2	0	0	M8	176
3903.1	3902.6	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	5	4	1	0	FA2G2	357
3944.1	3943.6	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	4	5	1	0	FA3G1	390
3985.2	3984.7	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	3	6	1	0	FA4	670
4106.3	4105.7	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	5	5	1	0	FA3G2	275
4147.3	4146.7	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	5	6	1	0	FA4G2	253
4194.3	4193.7	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	5	4	1	1	FA2G2S1	152
4268.4	4267.7	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	6	5	1	0	FA3G3	144
4309.5	4308.8	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	5	6	1	0	FA4G2	188
4471.6	4470.8	DGGPEDNKTEVFRPGGGDMR		2134.0	463	T29	6	6	1	0	FA4G3	189
2468.2	2468.0	AKWNDTLNK		1089.6	339	T31	6	2	0	0	M6	743
2630.2	2630.1	AKWNDTLNK		1089.6	339	T31	7	2	0	0	M7	310
2671.3	2671.1	AKWNDTLNK		1089.6	339	T31	7	3	0	0	M5A1G1 or M6A1	202
2792.3	2792.2	AKWNDTLNK		1089.6	339	T31	8	2	0	0	M8	389
2817.2	2817.2	AKWNDTLNK		1089.6	339	T31	6	3	1	0	FM5A1G1 or FM6A1	228
2954.3	2954.2	AKWNDTLNK		1089.6	339	T31	9	2	0	0	M9	327
2908.5	2908.2	LISCNTSVITQACPK	C	1691.8	197	T35	5	2	0	0	M5	1248
2933.5	2933.3	LISCNTSVITQACPK	C	1691.8	197	T35	3	3	1	0	FA1	394
3070.6	3070.3	LISCNTSVITQACPK	C	1691.8	197	T35	6	2	0	0	M6	302
3095.6	3095.3	LISCNTSVITQACPK	C	1691.8	197	T35	4	3	1	0	FM4A1	218
3111.6	3111.3	LISCNTSVITQACPK	C	1691.8	197	T35	5	3	0	0	M4A1G1 or M5A1	200
3136.6	3136.3	LISCNTSVITQACPK	C	1691.8	197	T35	3	4	1	0	FA2	473
3232.5	3232.3	LISCNTSVITQACPK	C	1691.8	197	T35	7	2	0	0	M7	249

3257.8	3257.4	LISCNTSVITQACPK	C	1691.8	197	T35	5	3	1	0	FM4A1G1 or FM5A1	164
3273.7	3273.4	LISCNTSVITQACPK	C	1691.8	197	T35	6	3	0	0	M5A1G1 or M6A1	157
3298.7	3298.4	LISCNTSVITQACPK	C	1691.8	197	T35	4	4	1	0	FA2G1	297
3339.7	3339.4	LISCNTSVITQACPK	C	1691.8	197	T35	3	5	1	0	FA3	292
3394.8	3394.4	LISCNTSVITQACPK	C	1691.8	197	T35	8	2	0	0	M8	257
3460.8	3460.4	LISCNTSVITQACPK	C	1691.8	197	T35	5	4	1	0	FA2G2	218
3501.8	3501.5	LISCNTSVITQACPK	C	1691.8	197	T35	4	5	1	0	FA3G1	228
3542.8	3542.5	LISCNTSVITQACPK	C	1691.8	197	T35	3	6	1	0	FA4	231
3663.8	3663.5	LISCNTSVITQACPK	C	1691.8	197	T35	5	5	1	0	FA3G2	203
3751.8	3751.5	LISCNTSVITQACPK	C	1691.8	197	T35	5	4	1	1	FA2G2S1	140
2892.6	2892.2	LTPLCVTLNCTDLR	C	1675.8	130	T47	5	2	0	0	M5	1575
2917.7	2917.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	3	3	1	0	FA1	291
2933.6	2933.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	4	3	0	0	M4A1	146
2974.8	2974.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	3	6	0	0	A4	110
3054.7	3054.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	6	2	0	0	M6	527
3079.7	3079.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	4	3	1	0	FM4A1	212
3095.7	3095.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	5	3	0	0	M4A1G1 or M5A1	263
3120.8	3120.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	3	4	1	0	FA2	274
3136.8	3136.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	4	4	0	0	A2G1	156
3178.0	3177.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	3	5	0	0	A3	123
3216.8	3216.3	LTPLCVTLNCTDLR	C	1675.8	130	T47	7	2	0	0	M7	648
3241.8	3241.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	5	3	1	0	FM4A1G1 or FM5A1	283
3257.8	3257.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	6	3	0	0	M5A1G1 or M6A1	394
3282.9	3282.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	4	4	1	0	FA2G1	358

3298.8	3298.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	5	4	0	0	A2G2	300
3323.9	3323.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	3	5	1	0	FA3	303
3378.9	3378.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	8	2	0	0	M8	508
3444.9	3444.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	5	4	1	0	FA2G2	285
3486.0	3485.5	LTPLCVTLNCTDLR	C	1675.8	130	T47	4	5	1	0	FA3G1	215
3527.1	3526.5	LTPLCVTLNCTDLR	C	1675.8	130	T47	3	6	1	0	FA4	218
3541.0	3540.4	LTPLCVTLNCTDLR	C	1675.8	130	T47	9	2	0	0	M9	169
3648.0	3647.5	LTPLCVTLNCTDLR	C	1675.8	130	T47	5	5	1	0	FA3G2	228
2568.5	2568.1	LTPLCVTLNCTDLR	C	1675.8	130	T48	3	2	0	0	M3	193
2730.5	2730.2	LTPLCVTLNCTDLR	C	1675.8	130	T48	4	2	0	0	M4	165
2771.6	2771.2	LTPLCVTLNCTDLR	C	1675.8	130	T48	3	3	0	0	A1	172
2892.6	2892.2	LTPLCVTLNCTDLR	C	1675.8	130	T48	4	2	1	0	FM4	243
2933.6	2933.3	LTPLCVTLNCTDLR	C	1675.8	130	T48	3	3	1	0	FA1	123
2372.5	2372.1	CSSNITGLLLTR	C, P-C	1317.7	448	T50	4	2	0	0	M4	157
2696.7	2696.2	CSSNITGLLLTR	C, P-C	1317.7	448	T50	6	2	0	0	M6	231
2762.7	2762.2	CSSNITGLLLTR	C, P-C	1317.7	448	T50	3	3	1	0	FA2	183
2858.7	2858.2	CSSNITGLLLTR	C, P-C	1317.7	448	T50	7	2	0	0	M7	992
3020.8	3020.3	CSSNITGLLLTR	C, P-C	1317.7	448	T50	8	2	0	0	M8	5723
3182.9	3182.3	CSSNITGLLLTR	C, P-C	1317.7	448	T50	9	2	0	0	M9	3682
2210.5	2210.0	CSSNITGLLLTR	C, P-C	1317.7	448	T51	3	2	0	0	M3	109
2372.5	2372.1	CSSNITGLLLTR	C, P-C	1317.7	448	T51	4	2	0	0	M4	359
2534.6	2534.1	CSSNITGLLLTR	C, P-C	1317.7	448	T51	5	2	0	0	M5	1414
2559.5	2559.1	CSSNITGLLLTR	C, P-C	1317.7	448	T51	3	3	1	0	FA1	107
2575.7	2575.1	CSSNITGLLLTR	C, P-C	1317.7	448	T51	4	3	0	0	M4A1	139
2616.9	2616.2	CSSNITGLLLTR	C, P-C	1317.7	448	T51	3	4	0	0	A2	118
2696.7	2696.2	CSSNITGLLLTR	C, P-C	1317.7	448	T51	6	2	0	0	M6	549

2737.7	2737.2	CSSNITGLLLTR	C, P-C	1317.7	448	T51	5	3	0	0	M4A1G1 or M5A1	145
2858.7	2858.2	CSSNITGLLLTR	C, P-C	1317.7	448	T51	7	2	0	0	M7	253
3020.8	3020.3	CSSNITGLLLTR	C, P-C	1317.7	448	T51	8	2	0	0	M8	187
3182.9	3182.3	CSSNITGLLLTR	C, P-C	1317.7	448	T51	9	2	0	0	M9	123
3020.7	3020.3	CSSNITGLLLTR	C, P-C	1317.7	448	T52	8	2	0	0	M8	151
3182.8	3182.3	CSSNITGLLLTR	C, P-C	1317.7	448	T52	9	2	0	0	M9	327

[§]Monoisotopic peak
C: carbamidomethyl (+57)
P-Q: Pyro-Glu (-17)
P-C: Pyro-Cys (-17)

Figure S3.1 – HILIC-UPLC profiles of N-linked glycans released from tryptic glycopeptide fractions.

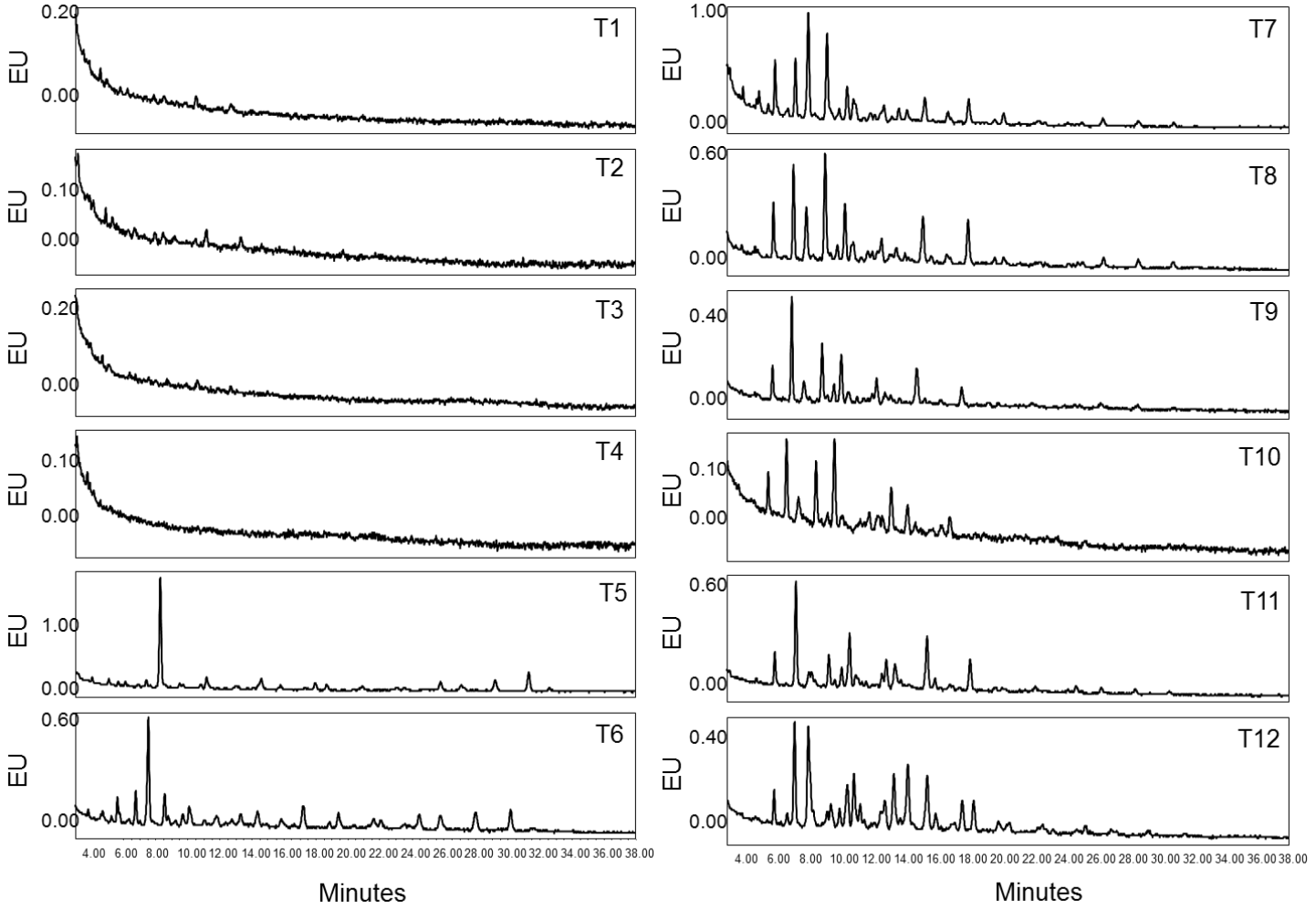


Figure S3.1 (cont.) – UPLC glycan profiles of tryptic fractions

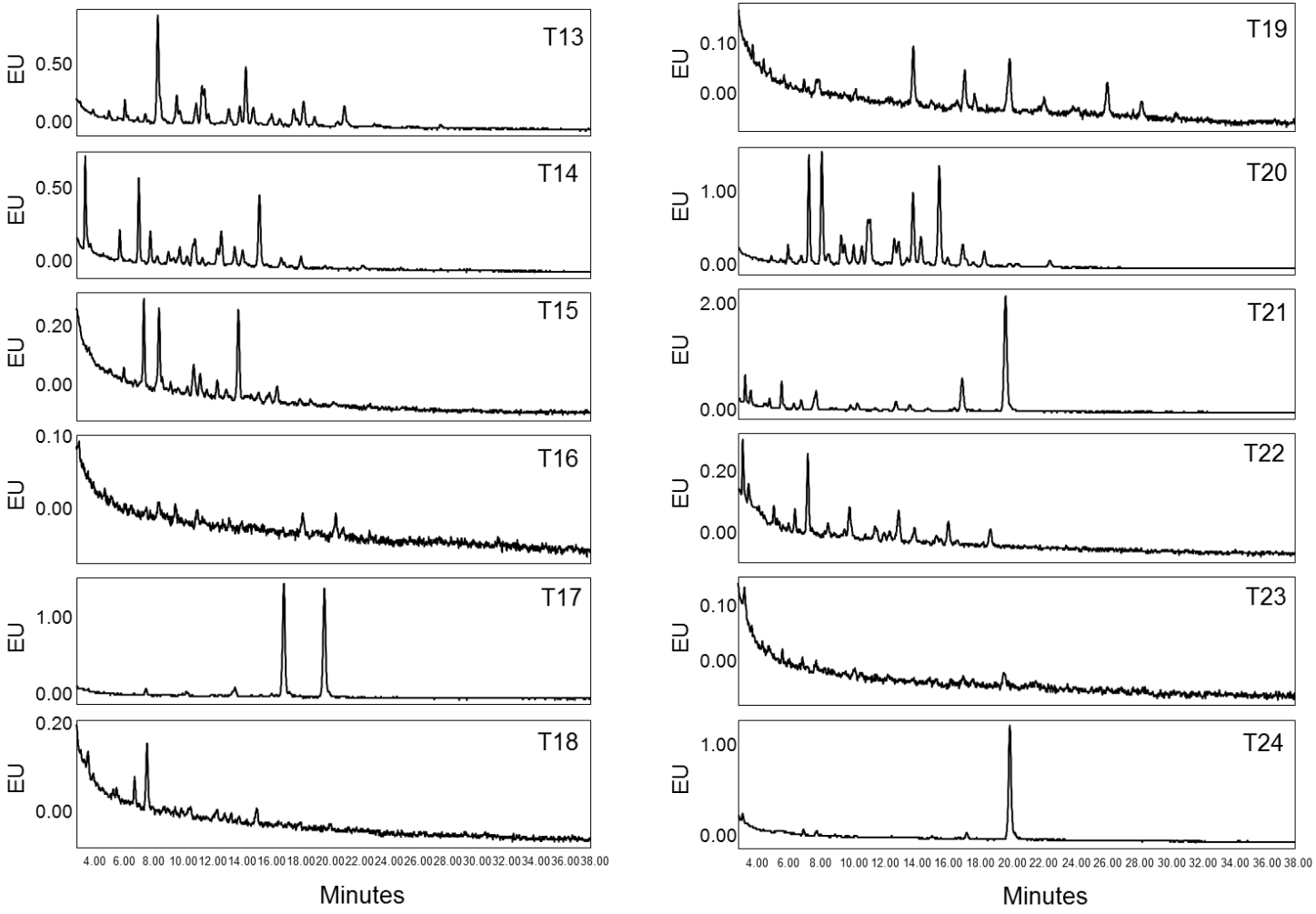


Figure S3.1 (cont.) – UPLC glycan profiles of tryptic fractions

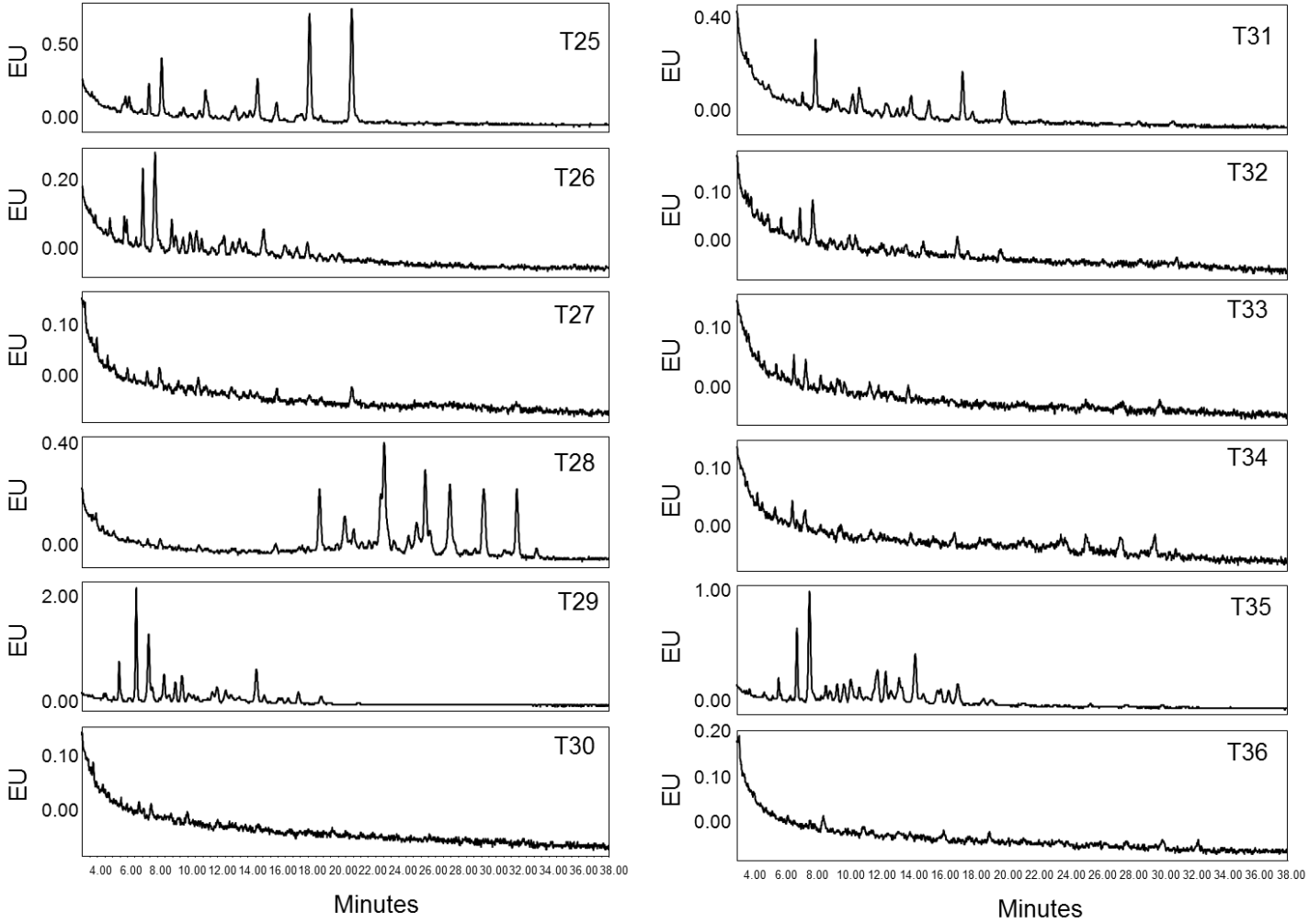


Figure S3.1 (cont.) – UPLC glycan profiles of tryptic fractions

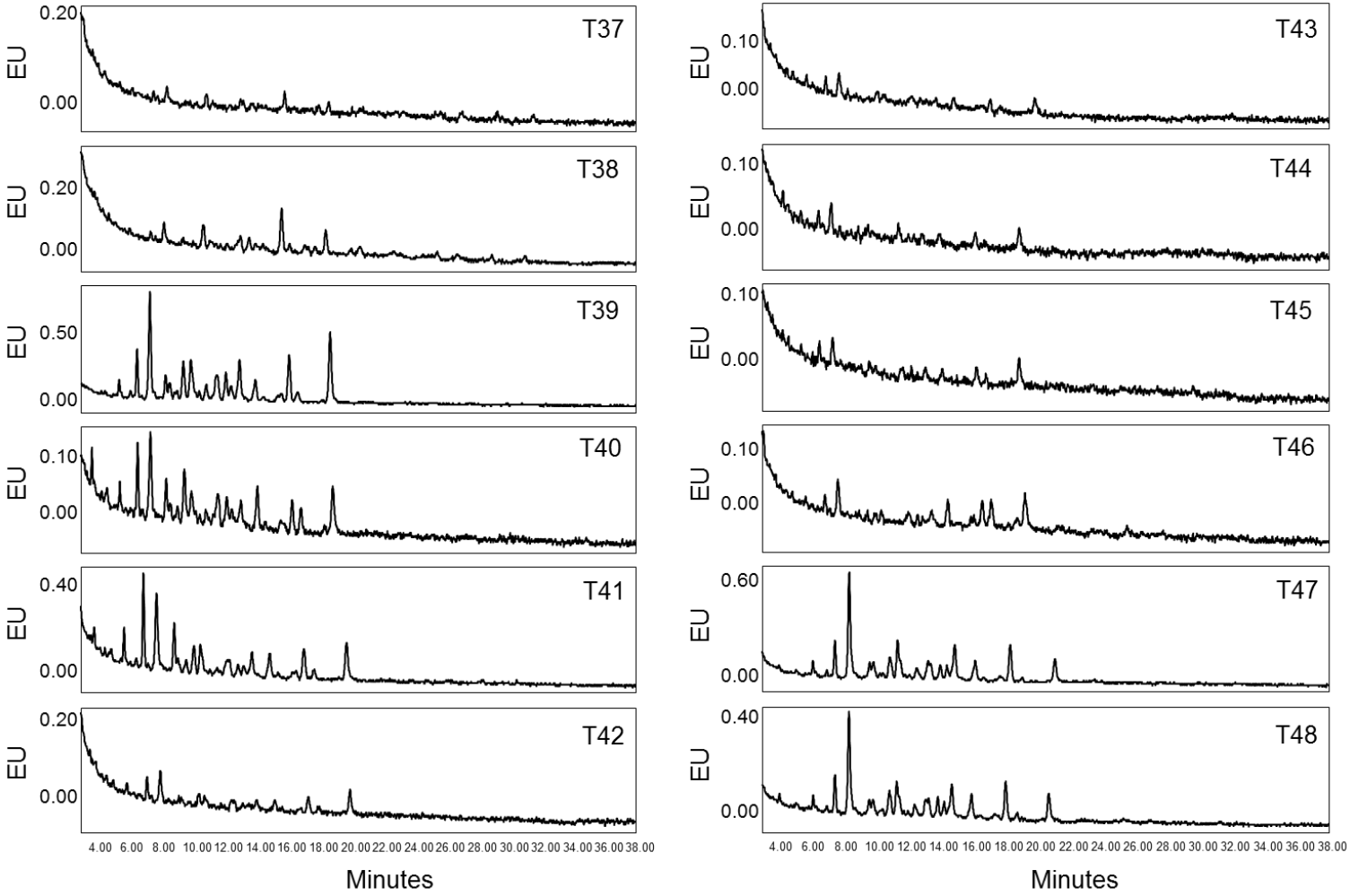


Figure S3.1 (cont.) – UPLC glycan profiles of tryptic fractions

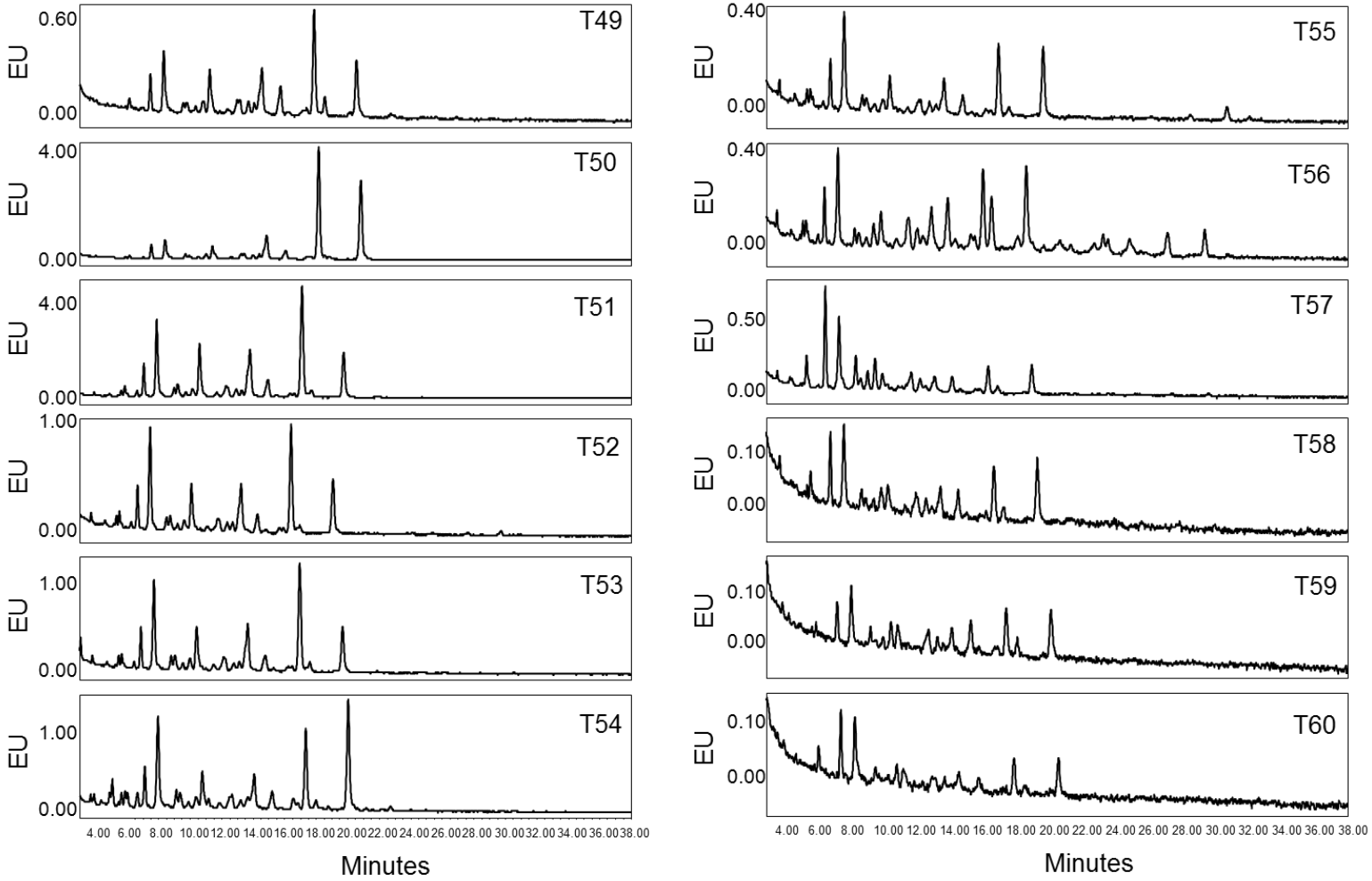


Table S3.3 – Glycopeptides identified from chymotryptic digestion of $gp120_{BaL}$ by MALDI MS.

m/z		Peptide	Mods	Peptide mass [M+1]	Site	Fraction	Composition				Predicted structure	Intensity
Found	Calc.						Hex	HexNAc	dHex	Neu5Ac		
1561.8	1561.6	NSTW		507.2	392	C13	4	2	0	0	M4	561
1885.8	1885.7	NSTW		507.2	392	C13	6	2	0	0	M6	601
1907.9	1907.7	NSTW	Na	507.2	392	C13	6	2	0	0	M6	773
1923.8	1923.7	NSTW	K	507.2	392	C13	6	2	0	0	M6	915
2047.9	2047.7	NSTW		507.2	392	C13	7	2	0	0	M7	699
2069.9	2069.7	NSTW	Na	507.2	392	C13	7	2	0	0	M7	1591
2085.9	2085.7	NSTW	K	507.2	392	C13	7	2	0	0	M7	2213
2209.9	2209.8	NSTW		507.2	392	C13	8	2	0	0	M8	2300
2232.0	2231.8	NSTW	Na	507.2	392	C13	8	2	0	0	M8	9073
2247.9	2247.8	NSTW	K	507.2	392	C13	8	2	0	0	M8	12769
2372.0	2371.9	NSTW		507.2	392	C13	9	2	0	0	M9	829
2394.0	2371.9	NSTW	Na	507.2	392	C13	9	2	0	0	M9	3605
2410.0	2409.9	NSTW	K	507.2	392	C13	9	2	0	0	M9	5394
1562.1	1561.6	NSTW		507.2	392	C14	4	2	0	0	M4	2662
1724.1	1723.6	NSTW		507.2	392	C14	5	2	0	0	M5	3089
1746.0	1745.6	NSTW	Na	507.2	392	C14	5	2	0	0	M5	12640
1762.0	1761.6	NSTW	K	507.2	392	C14	5	2	0	0	M5	17949
1884.7	1885.7	NSTW		507.2	392	C14	6	2	0	0	M6	1748
1908.1	1907.7	NSTW	Na	507.2	392	C14	6	2	0	0	M6	1428
1924.0	1923.7	NSTW	K	507.2	392	C14	6	2	0	0	M6	2089
2086.9	2086.8	YCNSTQLF		1032.4	386	C33	4	2	0	0	M4	908
2249.0	2248.8	YCNSTQLF		1032.4	386	C33	5	2	0	0	M5	1009

2270.9	2270.8	YCNSTQLF	Na	1032.4	386	C33	5	2	0	0	M5	638
2286.9	2286.8	YCNSTQLF	K	1032.4	386	C33	5	2	0	0	M5	778
2411.0	2410.9	YCNSTQLF		1032.4	386	C33	6	2	0	0	M6	1294
2433.0	2432.9	YCNSTQLF	Na	1032.4	386	C33	6	2	0	0	M6	885
2449.0	2448.9	YCNSTQLF	K	1032.4	386	C33	6	2	0	0	M6	1013
2573.1	2573.0	YCNSTQLF		1032.4	386	C33	7	2	0	0	M7	3605
2595.1	2595.0	YCNSTQLF	Na	1032.4	386	C33	7	2	0	0	M7	3091
2610.2	2611.0	YCNSTQLF	K	1032.4	386	C33	7	2	0	0	M7	4360
2735.2	2735.0	YCNSTQLF		1032.4	386	C33	8	2	0	0	M8	4087
2757.1	2757.0	YCNSTQLF	Na	1032.4	386	C33	8	2	0	0	M8	3824
2773.1	2773.0	YCNSTQLF	K	1032.4	386	C33	8	2	0	0	M8	4685
2897.3	2897.1	YCNSTQLF		1032.4	386	C33	9	2	0	0	M9	704
2919.2	2919.1	YCNSTQLF	Na	1032.4	386	C33	9	2	0	0	M9	806
2935.2	2935.1	YCNSTQLF	K	1032.4	386	C33	9	2	0	0	M9	420
1924.8	1924.7	YCNSTQLF	C	1032.4	386	C34	3	2	0	0	M3	196
1946.8	1946.7	YCNSTQLF	C, Na	1032.4	386	C34	3	2	0	0	M3	142
1962.8	1962.7	YCNSTQLF	C, K	1032.4	386	C34	3	2	0	0	M3	207
2087.0	2086.8	YCNSTQLF	C	1032.4	386	C34	4	2	0	0	M4	348
2108.9	2108.8	YCNSTQLF	C, Na	1032.4	386	C34	4	2	0	0	M4	281
2124.9	2124.8	YCNSTQLF	C, K	1032.4	386	C34	4	2	0	0	M4	429
2248.7	2248.8	YCNSTQLF	C	1032.4	386	C34	5	2	0	0	M5	158
2271.0	2270.8	YCNSTQLF	C, Na	1032.4	386	C34	5	2	0	0	M5	242
2286.9	2286.8	YCNSTQLF	C, K	1032.4	386	C34	5	2	0	0	M5	342
3007.5	3007.3	LLTRDGGPEDNKTEVF		1790.9	463	C34	5	2	0	0	M5	1575
3032.5	3032.3	LLTRDGGPEDNKTEVF		1790.9	463	C34	3	3	1	0	FA1	203
3089.7	3089.4	LLTRDGGPEDNKTEVF		1790.9	463	C34	3	4	0	0	A2	175

3169.6	3169.4	LLTRDGGPEDNKTEVF	1790.9	463	C34	6	2	0	0	M6	73.4
3194.6	3194.4	LLTRDGGPEDNKTEVF	1790.9	463	C34	4	3	1	0	FM4A1	174
3210.6	3210.4	LLTRDGGPEDNKTEVF	1790.9	463	C34	5	3	0	0	M5A1	112
3235.6	3235.4	LLTRDGGPEDNKTEVF	1790.9	463	C34	3	4	1	0	FA2	1487
3331.6	3331.4	LLTRDGGPEDNKTEVF	1790.9	463	C34	7	2	0	0	M7	93.9
3356.6	3356.5	LLTRDGGPEDNKTEVF	1790.9	463	C34	5	3	1	0	FM4A1G1 or FM5A1	167
3397.7	3397.5	LLTRDGGPEDNKTEVF	1790.9	463	C34	4	4	1	0	FA2G1	433
3438.7	3438.5	LLTRDGGPEDNKTEVF	1790.9	463	C34	3	5	1	0	FA3	848
3518.7	3518.5	LLTRDGGPEDNKTEVF	1790.9	463	C34	6	2	1	0	FM5A1G1 or FM6A1	103
3559.7	3559.5	LLTRDGGPEDNKTEVF	1790.9	463	C34	5	4	1	0	FA2G2	286
3600.8	3600.6	LLTRDGGPEDNKTEVF	1790.9	463	C34	4	5	1	0	FA3G1	324
3641.8	3641.6	LLTRDGGPEDNKTEVF	1790.9	463	C34	3	6	1	0	FA4	560
3762.9	3762.6	LLTRDGGPEDNKTEVF	1790.9	463	C34	5	5	1	0	FA3G2	208
3803.9	3803.6	LLTRDGGPEDNKTEVF	1790.9	463	C34	4	6	1	0	FA4G1	188
3851.0	3850.6	LLTRDGGPEDNKTEVF	1790.9	463	C34	5	4	1	1	FA2G2S1	162
3924.8	3924.7	LLTRDGGPEDNKTEVF	1790.9	463	C34	6	5	1	0	FA3G3	72.7
3965.9	3965.7	LLTRDGGPEDNKTEVF	1790.9	463	C34	5	6	1	0	FA4G2	122
2683.3	2683.2	LLTRDGGPEDNKTEVF	1790.9	463	C35	3	2	0	0	M3	267
2829.3	2829.3	LLTRDGGPEDNKTEVF	1790.9	463	C35	3	2	1	0	FM3	470
2886.4	2886.3	LLTRDGGPEDNKTEVF	1790.9	463	C35	3	3	0	0	A1	511
2991.5	2991.3	LLTRDGGPEDNKTEVF	1790.9	463	C35	4	2	1	0	FM4	218
3007.5	3007.3	LLTRDGGPEDNKTEVF	1790.9	463	C35	5	2	0	0	M5	135
3032.4	3032.3	LLTRDGGPEDNKTEVF	1790.9	463	C35	3	3	1	0	FA1	1924
3235.5	3235.4	LLTRDGGPEDNKTEVF	1790.9	463	C35	3	4	1	0	FA2	298
3438.6	3438.5	LLTRDGGPEDNKTEVF	1790.9	463	C35	3	5	1	0	FA3	170

3542.7	3542.5	LLTRDGGPEDNKTEVF		1790.9	463	C35	4	4	0	1	A2G1S1	291
3704.7	3704.6	LLTRDGGPEDNKTEVF		1790.9	463	C35	5	4	0	1	A2G2S1	520
2894.7	2894.3	APPIRGQIRCSSNITGL	C	1840.0	448	C36	4	2	0	0	M4	180
3056.7	3056.4	APPIRGQIRCSSNITGL	C	1840.0	448	C36	5	2	0	0	M5	392
3218.9	3218.4	APPIRGQIRCSSNITGL	C	1840.0	448	C36	6	2	0	0	M6	336
3380.9	3380.5	APPIRGQIRCSSNITGL	C	1840.0	448	C36	7	2	0	0	M7	532
3543.0	3542.5	APPIRGQIRCSSNITGL	C	1840.0	448	C36	8	2	0	0	M8	2875
3705.0	3704.6	APPIRGQIRCSSNITGL	C	1840.0	448	C36	9	2	0	0	M9	1961
2539.5	2539.1	REQFGNKTIVF		1338.7	356	C37	4	2	1	0	FM4	715
2555.5	2555.1	REQFGNKTIVF		1338.7	356	C37	5	2	0	0	M5	4045
2580.5	2580.2	REQFGNKTIVF		1338.7	356	C37	3	3	1	0	FA1	1588
2701.5	2701.2	REQFGNKTIVF		1338.7	356	C37	5	2	1	0	FM5	598
2717.5	2717.2	REQFGNKTIVF		1338.7	356	C37	6	2	0	0	M6	3533
2742.6	2742.2	REQFGNKTIVF		1338.7	356	C37	4	3	1	0	FM4A1	15012
2758.6	2758.2	REQFGNKTIVF		1338.7	356	C37	5	3	0	0	M4A1G1 or M5A1	4336
2783.6	2783.3	REQFGNKTIVF		1338.7	356	C37	3	4	1	0	FA2	22248
2879.6	2879.2	REQFGNKTIVF		1338.7	356	C37	7	2	0	0	M7	9603
2904.6	2904.3	REQFGNKTIVF		1338.7	356	C37	5	3	1	0	FM5A1	15078
2920.7	2920.3	REQFGNKTIVF		1338.7	356	C37	6	3	0	0	M5A1G1 or M6A1	2564
2945.7	2945.3	REQFGNKTIVF		1338.7	356	C37	4	4	1	0	FA2G1	6273
2986.7	2986.3	REQFGNKTIVF		1338.7	356	C37	3	5	1	0	FA3	11725
3041.7	3041.3	REQFGNKTIVF		1338.7	356	C37	8	2	0	0	M8	2556
3066.7	3066.3	REQFGNKTIVF		1338.7	356	C37	6	3	1	0	FM5A1G1 or FM6A1	3390
3107.8	3107.4	REQFGNKTIVF		1338.7	356	C37	5	4	1	0	FA2G2	8130
3148.8	3148.4	REQFGNKTIVF		1338.7	356	C37	4	5	1	0	FA3G1	4335

3189.8	3189.4	REQFGNKTIVF	1338.7	356	C37	3	6	1	0	FA4	4295
3195.8	3195.4	REQFGNKTIVF	1338.7	356	C37	5	3	1	1	FM4A1G1S1	2528
3203.8	3203.4	REQFGNKTIVF	1338.7	356	C37	9	2	0	0	M9	421
3236.8	3236.4	REQFGNKTIVF	1338.7	356	C37	4	4	1	1	FA2G1S1	662
3310.9	3310.4	REQFGNKTIVF	1338.7	356	C37	5	5	1	0	FA3G2	4199
3351.8	3351.5	REQFGNKTIVF	1338.7	356	C37	4	6	1	0	FA4G1	1069
3357.9	3357.4	REQFGNKTIVF	1338.7	356	C37	6	3	1	1	FM5A1G1S1	664
3398.9	3398.5	REQFGNKTIVF	1338.7	356	C37	5	4	1	1	FA2G2S1	4106
3472.9	3472.5	REQFGNKTIVF	1338.7	356	C37	6	5	1	0	FA3G3	682
3514.0	3513.5	REQFGNKTIVF	1338.7	356	C37	5	6	1	0	FA4G2	341
3543.1	3543.5	REQFGNKTIVF	1338.7	356	C37	5	4	0	1	A2G2S1	462
3676.0	3675.6	REQFGNKTIVF	1338.7	356	C37	6	6	1	0	FA4G3	223
3690.0	3689.5	REQFGNKTIVF	1338.7	356	C37	5	4	1	2	FA2G2S2	234
3764.1	3763.6	REQFGNKTIVF	1338.7	356	C37	6	5	1	1	FA3G3S1	313
2231.2	2231.0	REQFGNKTIVF	1338.7	356	C38	3	2	0	0	M3	633
2377.3	2377.1	REQFGNKTIVF	1338.7	356	C38	3	2	1	0	FM3	516
2393.3	2393.1	REQFGNKTIVF	1338.7	356	C38	4	2	0	0	M4	1093
2434.3	2434.1	REQFGNKTIVF	1338.7	356	C38	3	3	0	0	A1	1454
2539.4	2539.1	REQFGNKTIVF	1338.7	356	C38	4	2	1	0	FM4	500
2555.4	2555.1	REQFGNKTIVF	1338.7	356	C38	5	2	0	0	M5	19807
2580.4	2580.2	REQFGNKTIVF	1338.7	356	C38	3	3	1	0	FA1	6447
2596.4	2596.2	REQFGNKTIVF	1338.7	356	C38	4	3	0	0	M4A1	1984
2637.4	2637.2	REQFGNKTIVF	1338.7	356	C38	3	4	0	0	A2	853
2717.5	2717.2	REQFGNKTIVF	1338.7	356	C38	6	2	0	0	M6	353
2742.5	2742.2	REQFGNKTIVF	1338.7	356	C38	4	3	1	0	FM4A1	427

2758.5	2758.2	REQFGNKTIVF		1338.7	356	C38	5	3	0	0	M5A1	605
2783.5	2783.3	REQFGNKTIVF		1338.7	356	C38	3	4	1	0	FA2	233
3014.5	3014.3	TTGEIIGDIRQAHCNL	C	1797.9	332	C39	5	2	0	0	M5	199
3176.6	3176.3	TTGEIIGDIRQAHCNL	C	1797.9	332	C39	6	2	0	0	M6	378
3338.7	3338.4	TTGEIIGDIRQAHCNL	C	1797.9	332	C39	7	2	0	0	M7	867
3500.7	3500.5	TTGEIIGDIRQAHCNL	C	1797.9	332	C39	8	2	0	0	M8	8614
3662.8	3662.5	TTGEIIGDIRQAHCNL	C	1797.9	332	C39	9	2	0	0	M9	7975
2852.6	2852.2	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	4	2	0	0	M4	346
3014.7	3014.3	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	5	2	0	0	M5	2115
3039.7	3039.3	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	3	3	1	0	FA1	285
3176.8	3176.3	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	6	2	0	0	M6	382
3201.9	3201.4	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	4	3	1	0	FM4A1	243
3217.9	3217.4	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	5	3	1	0	M4A1G1 or FM5A1	186
3242.9	3242.4	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	3	4	1	0	FA2	290
3338.8	3338.4	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	7	2	0	0	M7	217
3501.0	3500.5	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	8	2	0	0	M8	823
3663.0	3662.5	TTGEIIGDIRQAHCNL	C	1797.9	332	C40	9	2	0	0	M9	717
4439.0	4438.8	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C42	6	3	1	0	FM5A1G1 or FM6A1	427
4887.1	4886.0	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C42	5	6	1	0	FA4G2	232
5064.5	5064.1	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C42	7	6	0	0	A4G4	203
4360.1	4359.8	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C42	4	3	1	0	FM4A1	277
4400.7	4400.8	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C42	3	4	1	0	FA2	242
4521.6	4521.9	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C42	5	3	1	0	FM4A1G1 or FM5A1	326
4562.4	4562.9	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C42	4	4	1	0	FA2G1	432
4579.0	4578.9	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C42	5	4	0	0	A2G2	212

4724.8	4724.9	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C42	5	4	1	0	FA2G2	336
4766.0	4766.0	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C42	4	5	1	0	FA3G1	294
4928.1	4928.0	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C42	5	5	1	0	FA3G2	207
3060.6	3060.4	DIVPIDNNSNNRYRL		1802.9	186	C42	4	3	0	0	M4A1	240
3222.2	3222.4	DIVPIDNNSNNRYRL		1802.9	186	C42	5	4	0	0	M4A1G1 or M5A1	466
3305.1	3304.5	DIVPIDNNSNNRYRL		1802.9	186	C42	3	5	0	0	A3	281
3409.2	3409.5	DIVPIDNNSNNRYRL		1802.9	186	C42	4	4	1	0	FA2G1	959
3425.4	3425.5	DIVPIDNNSNNRYRL		1802.9	186	C42	5	4	0	0	A2G2	288
3513.4	3513.5	DIVPIDNNSNNRYRL		1802.9	186	C42	5	3	0	1	M4A1G1S1	473
3653.3	3653.6	DIVPIDNNSNNRYRL		1802.9	186	C42	3	6	1	0	FA4	525
3749.3	3749.6	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	3	2	1	0	FM3	1082
3911.5	3911.7	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	4	2	1	0	FM4	915
3928.4	3927.6	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	5	2	0	0	M5	1848
3952.5	3952.7	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	3	3	1	0	FA1	1416
4009.4	4009.7	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	3	4	0	0	A2	239
4073.6	4073.7	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	5	2	1	0	FM5	338
4090.4	4089.7	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	6	2	0	0	M6	337
4114.5	4114.7	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	4	3	1	0	FM4A1	345
4277.0	4276.8	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	5	3	1	0	FM5A1	391
4561.6	4561.9	ATHACVPTDPNPQEVELENVTENF	C	2711.2	88	C43	3	6	1	0	FA4	566
4359.7	4359.8	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	4	3	1	0	FM4A1	636
3848.4	3848.6	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	3	2	0	0	M3	927
3994.6	3994.7	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	3	2	1	0	FM3	913
4010.5	4010.7	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	4	2	0	0	M4	741
4051.2	4051.7	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	3	3	0	0	A1	589
4156.6	4156.7	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	4	2	1	0	FM4	2312

4172.6	4172.7	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	5	2	0	0	M5	381
4213.6	4213.8	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	3	3	1	0	FA1	434
4254.7	4254.8	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	3	4	0	0	A2	600
4318.7	4318.8	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	5	2	1	0	FM5	804
4359.7	4359.8	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	4	3	1	0	FM4A1	636
4521.7	4521.9	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	5	3	1	0	FM4A1G1 or FM5A1	362
4562.6	4562.9	ATHACVPTDPNPQEVELENVTENFNM	C	2956.3	88	C43	4	4	1	0	FA2G1	566
2649.3	2649.2	NDTLNKIVIKL		1270.8	339	C44	6	2	0	0	M6	365
2811.9	2811.3	NDTLNKIVIKL		1270.8	339	C44	7	2	0	0	M7	318
2973.8	2973.4	NDTLNKIVIKL		1270.8	339	C44	8	2	0	0	M8	793
3135.9	3135.4	NDTLNKIVIKL		1270.8	339	C44	9	2	0	0	M9	3911
2487.5	2487.2	NDTLNKIVIKL		1270.8	339	C45	5	2	0	0	M5	298
2811.8	2811.3	NDTLNKIVIKL		1270.8	339	C45	7	2	0	0	M7	334
2973.7	2973.4	NDTLNKIVIKL		1270.8	339	C45	8	2	0	0	M8	507
3135.8	3135.4	NDTLNKIVIKL		1270.8	339	C45	9	2	0	0	M9	291
3666.9	3666.7	LLNGSLAEEEVVIRSENF		2019.0	262	C47	3	5	1	0	FA3	241
3722.4	3721.6	LLNGSLAEEEVVIRSENF		2019.0	262	C47	8	2	0	0	M8	258
3788.4	3787.7	LLNGSLAEEEVVIRSENF		2019.0	262	C47	5	4	1	0	FA2G2	349
3829.4	3828.7	LLNGSLAEEEVVIRSENF		2019.0	262	C47	4	5	1	0	FA3G1	349
3870.4	3869.7	LLNGSLAEEEVVIRSENF		2019.0	262	C47	3	6	1	0	FA4	307
3884.4	3883.7	LLNGSLAEEEVVIRSENF		2019.0	262	C47	9	2	0	0	M9	1422
3073.8	3073.4	LLNGSLAEEEVVIRSENF		2019.0	262	C48	4	2	0	0	M4	208
3236.0	3235.5	LLNGSLAEEEVVIRSENF		2019.0	262	C48	5	2	0	0	M5	452
3398.0	3397.5	LLNGSLAEEEVVIRSENF		2019.0	262	C48	6	2	0	0	M6	527
3464.0	3463.6	LLNGSLAEEEVVIRSENF		2019.0	262	C48	3	4	1	0	FA2	263
3560.2	3559.6	LLNGSLAEEEVVIRSENF		2019.0	262	C48	7	2	0	0	M7	390

3722.2	3721.6	LLNGSLAEEEVVIRSENF	2019.0	262	C48	8	2	0	0	M8	1534
3884.3	3883.7	LLNGSLAEEEVVIRSENF	2019.0	262	C48	9	2	0	0	M9	3856
4046.4	4045.7	LLNGSLAEEEVVIRSENF	2019.0	262	C48	10	2	0	0	Glc1M9	361
3073.7	3073.4	LLNGSLAEEEVVIRSENF	2019.0	262	C49	4	2	0	0	M4	371
3235.8	3235.5	LLNGSLAEEEVVIRSENF	2019.0	262	C49	5	2	0	0	M5	1138
3397.9	3397.5	LLNGSLAEEEVVIRSENF	2019.0	262	C49	6	2	0	0	M6	397
3464.2	3463.6	LLNGSLAEEEVVIRSENF	2019.0	262	C49	3	4	1	0	FA2	165
3559.9	3559.6	LLNGSLAEEEVVIRSENF	2019.0	262	C49	7	2	0	0	M7	595
3722.0	3721.6	LLNGSLAEEEVVIRSENF	2019.0	262	C49	8	2	0	0	M8	1985
3884.0	3883.7	LLNGSLAEEEVVIRSENF	2019.0	262	C49	9	2	0	0	M9	5377
4046.2	4045.7	LLNGSLAEEEVVIRSENF	2019.0	262	C49	10	2	0	0	Glc1M9	220
3235.8	3235.5	LLNGSLAEEEVVIRSENF	2019.0	262	C50	5	2	0	0	M5	419
3397.9	3397.5	LLNGSLAEEEVVIRSENF	2019.0	262	C50	6	2	0	0	M6	230
3560.4	3559.6	LLNGSLAEEEVVIRSENF	2019.0	262	C50	7	2	0	0	M7	342
3722.4	3721.6	LLNGSLAEEEVVIRSENF	2019.0	262	C50	8	2	0	0	M8	565
3884.4	3883.7	LLNGSLAEEEVVIRSENF	2019.0	262	C50	9	2	0	0	M9	1629
3115.9	3115.5	SRAKWNDTLNKIVIKL	1899.1	339	C51	5	2	0	0	M5	470
3277.9	3277.6	SRAKWNDTLNKIVIKL	1899.1	339	C51	6	2	0	0	M6	478
3344.1	3343.7	SRAKWNDTLNKIVIKL	1899.1	339	C51	3	4	1	0	FA2	287
3602.1	3601.7	SRAKWNDTLNKIVIKL	1899.1	339	C51	8	2	0	0	M8	1942
3764.2	3763.8	SRAKWNDTLNKIVIKL	1899.1	339	C51	9	2	0	0	M9	6111
3235.7	3235.5	LLNGSLAEEEVVIRSENF	2019.0	262	C51	5	2	0	0	M5	480
3397.8	3397.5	LLNGSLAEEEVVIRSENF	2019.0	262	C51	6	2	0	0	M6	230
3560.0	3559.6	LLNGSLAEEEVVIRSENF	2019.0	262	C51	7	2	0	0	M7	351
3722.1	3721.6	LLNGSLAEEEVVIRSENF	2019.0	262	C51	8	2	0	0	M8	770
3884.1	3883.7	LLNGSLAEEEVVIRSENF	2019.0	262	C51	9	2	0	0	M9	1910

[§]Monoisotopic peak
C: carbamidomethyl (+57)
Na: Na^+ adduct $[M+23]^+$
K: K^+ adduct $[M+39]^+$

Figure S3.2 – HILIC-UPLC profiles of N-linked glycans released from chymotryptic glycopeptide fractions.

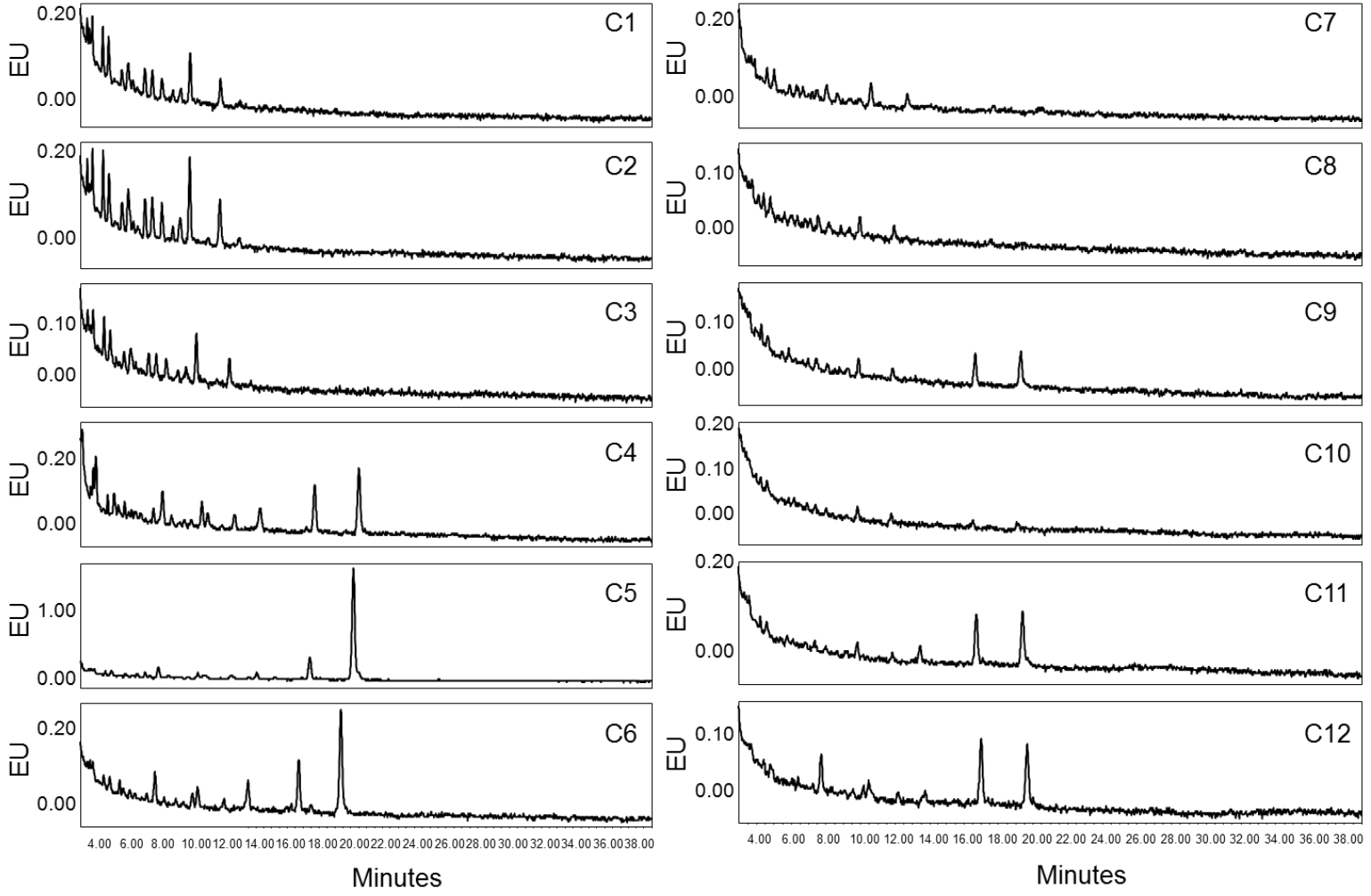


Figure S3.2 (cont.) – UPLC glycan profiles of chymotryptic fractions

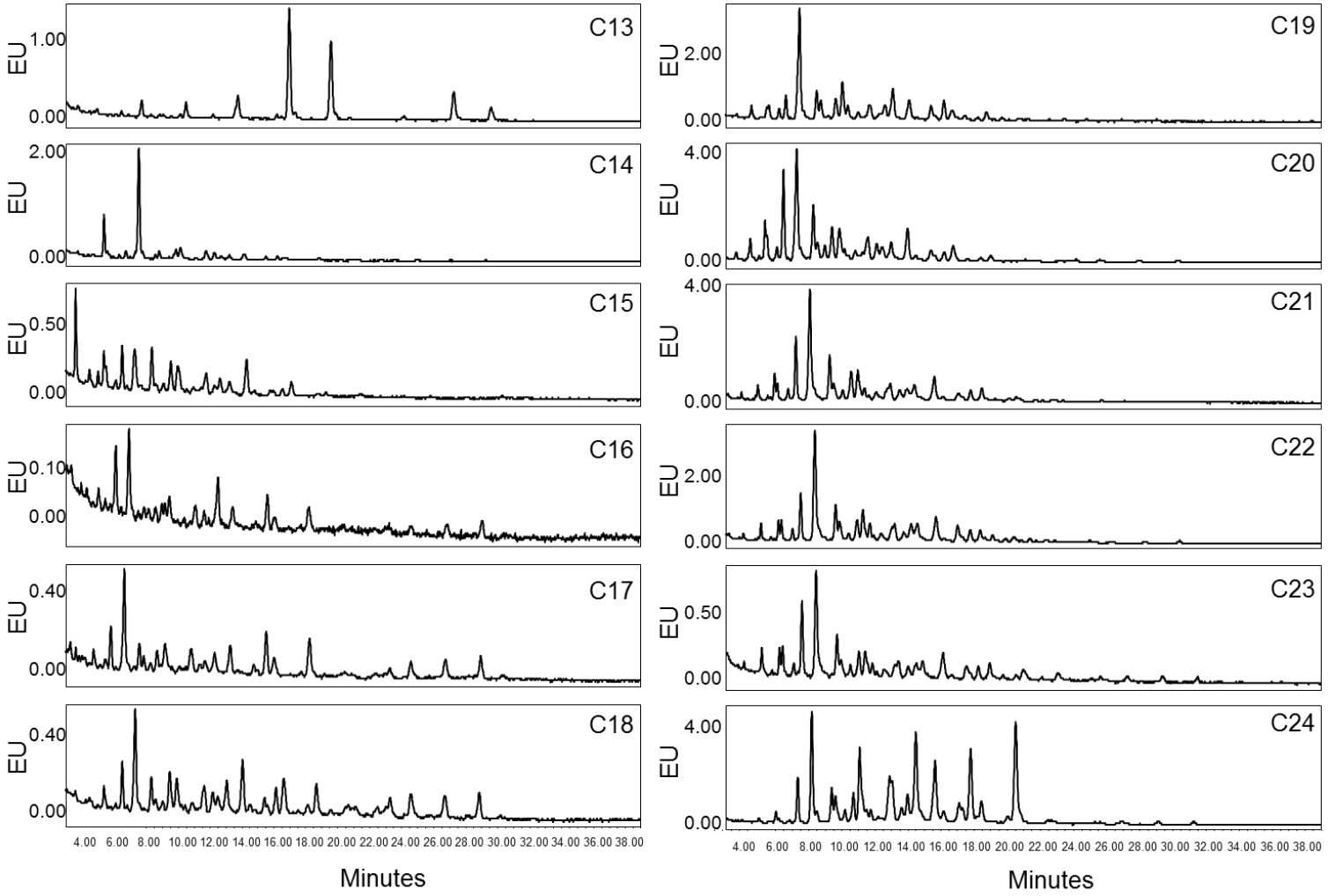


Figure S3.2 (cont.) – UPLC glycan profiles of chymotryptic fractions

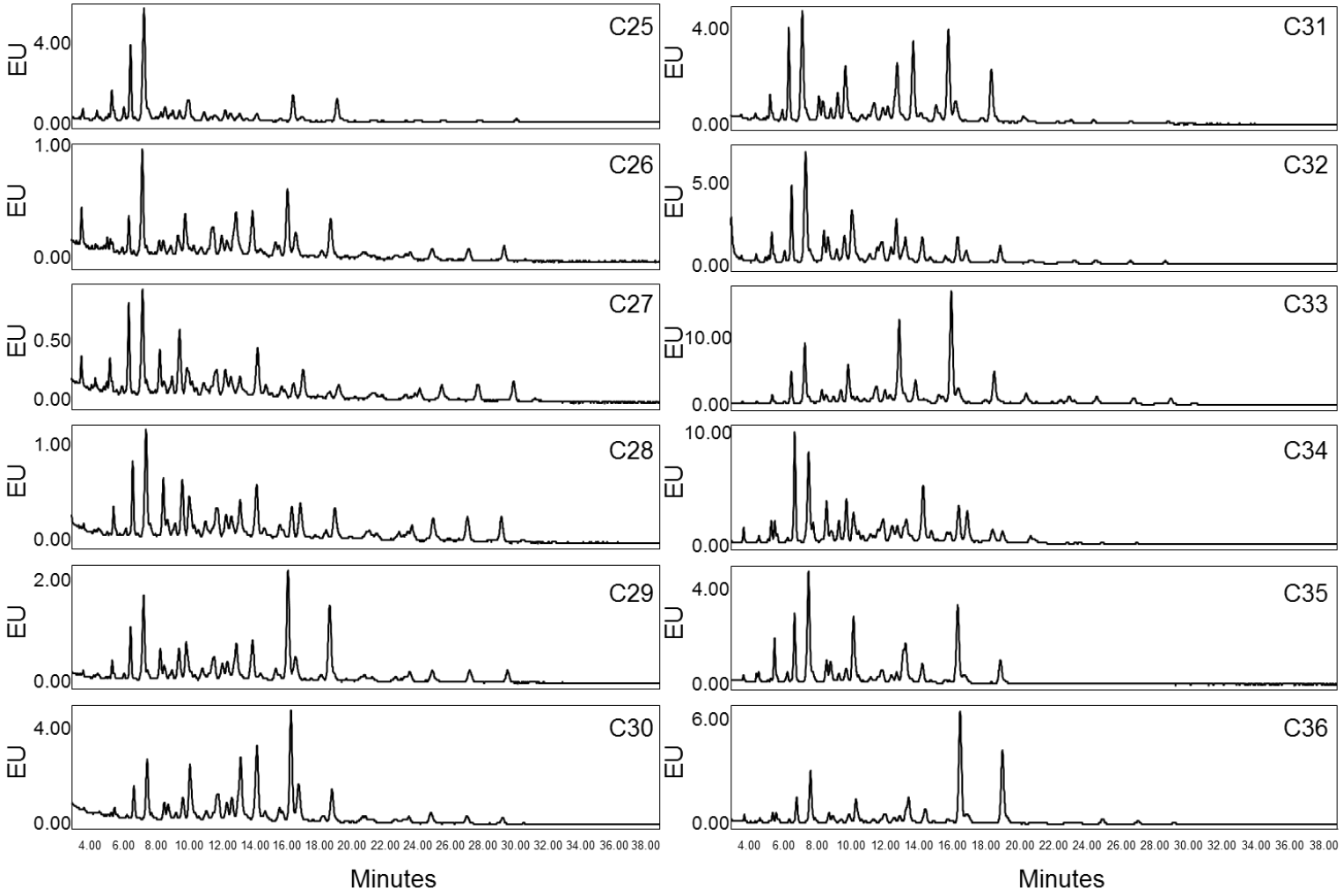


Figure S3.2 (cont.) – UPLC glycan profiles of chymotryptic fractions

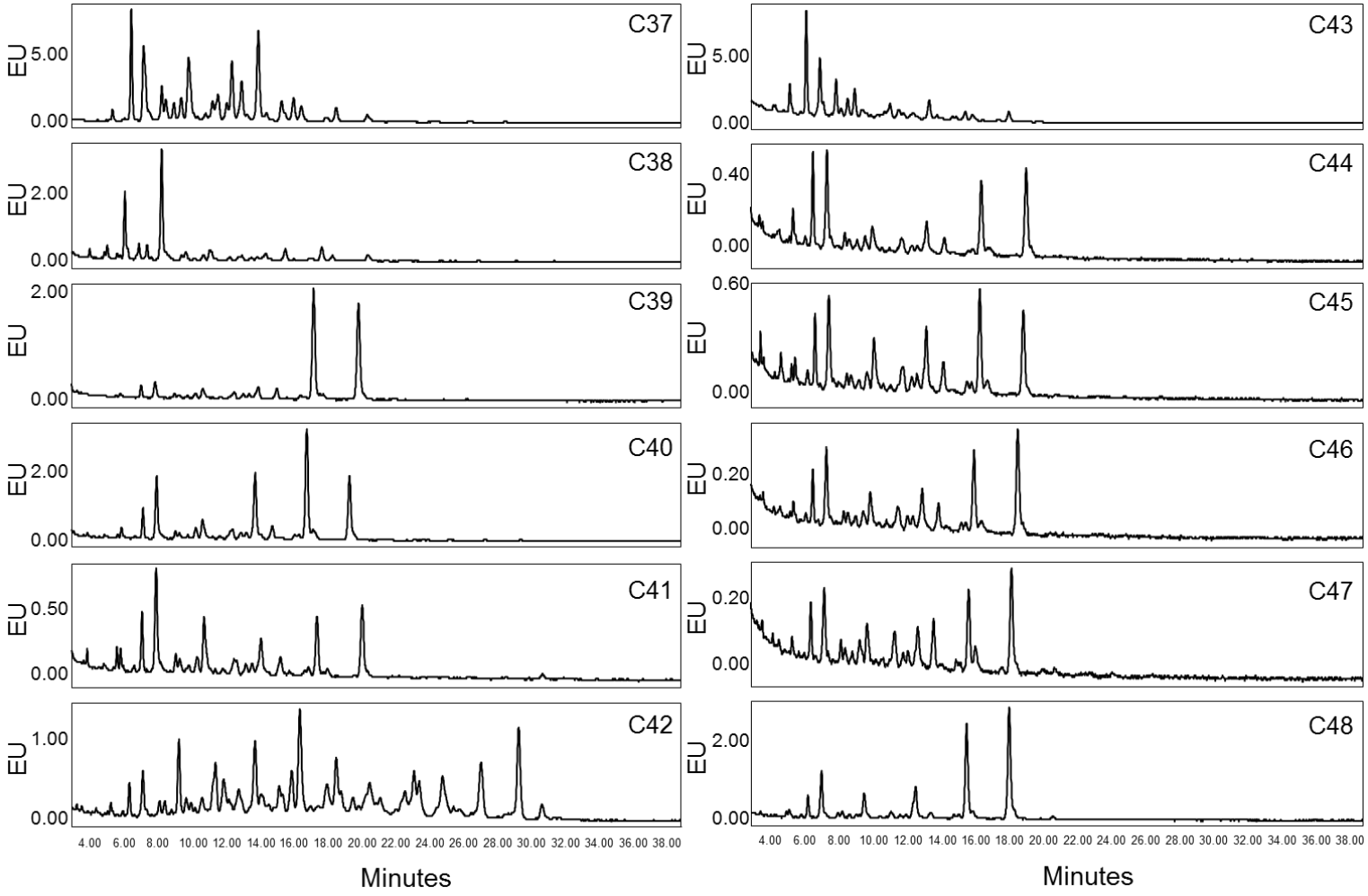
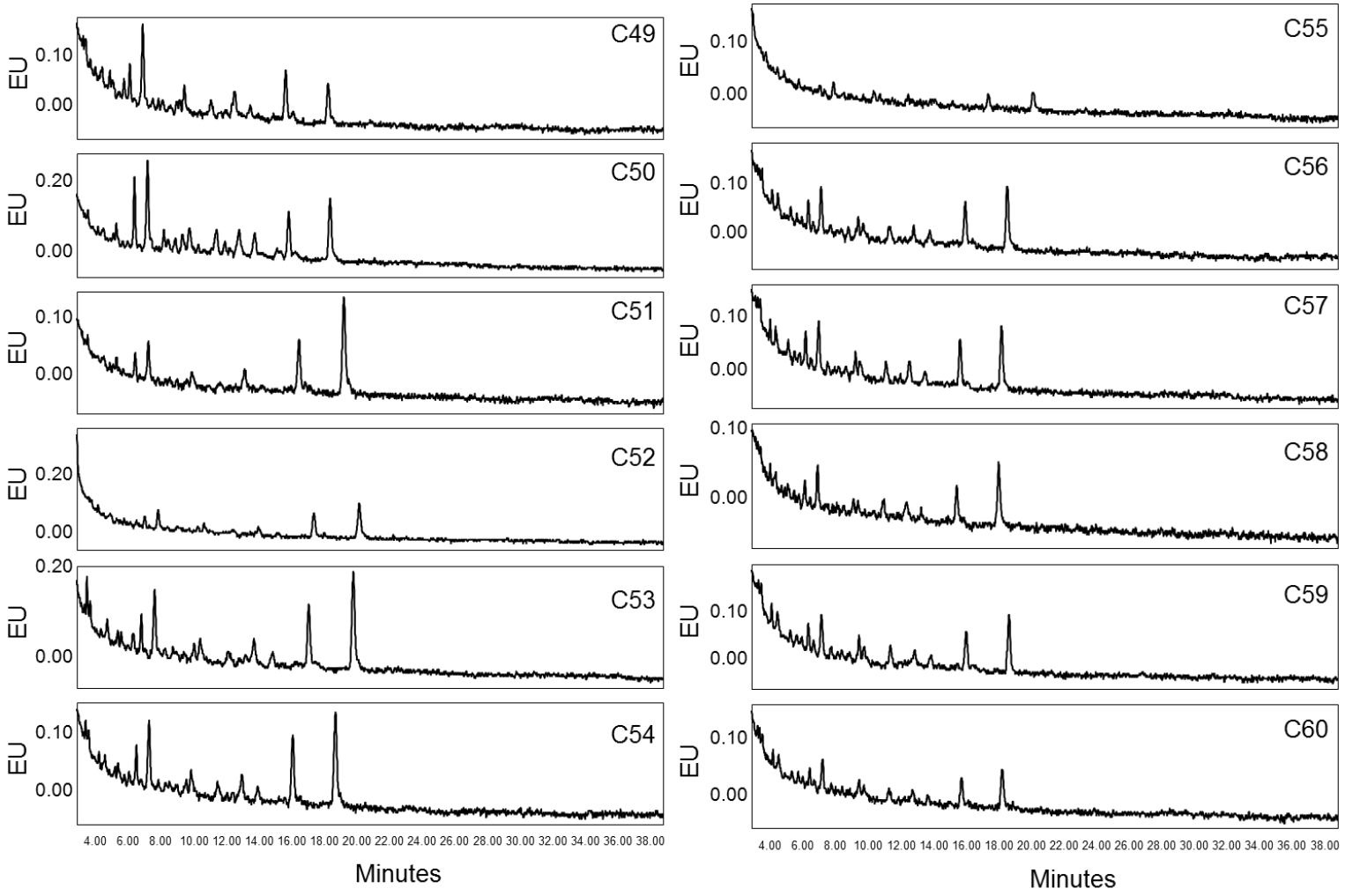


Figure S3.2 (cont.) – UPLC glycan profiles of chymotryptic fractions



Appendix IV

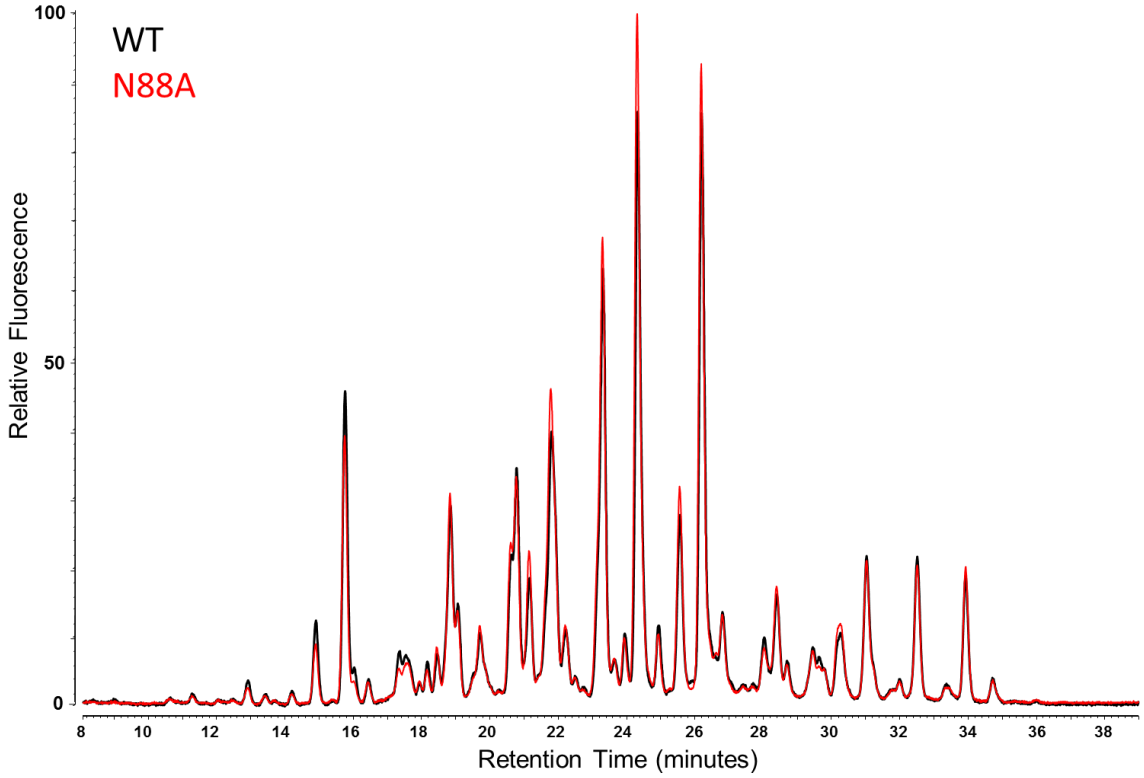
HILIC-UPLC analysis of gp120$_{BaL}$ PNGS-deletion mutants

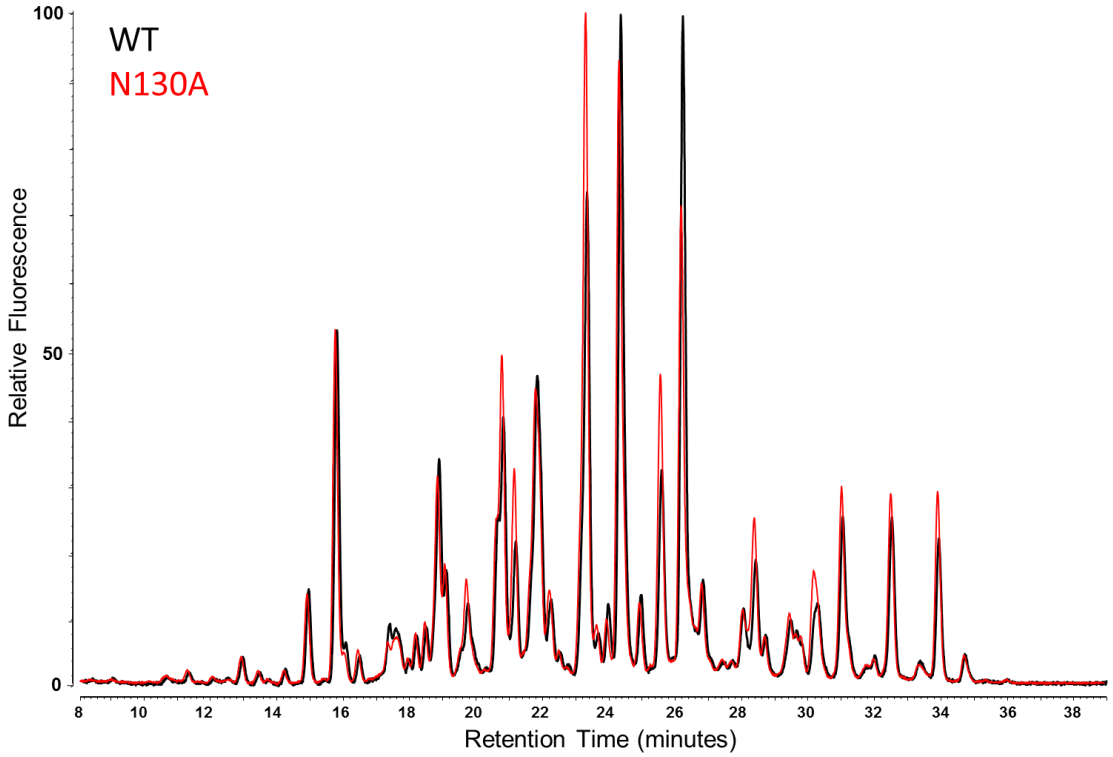
Table S4.1 – Abundances of released N-linked glycans obtained from gp120$_{BaL}$ PNGS-deletion mutants.

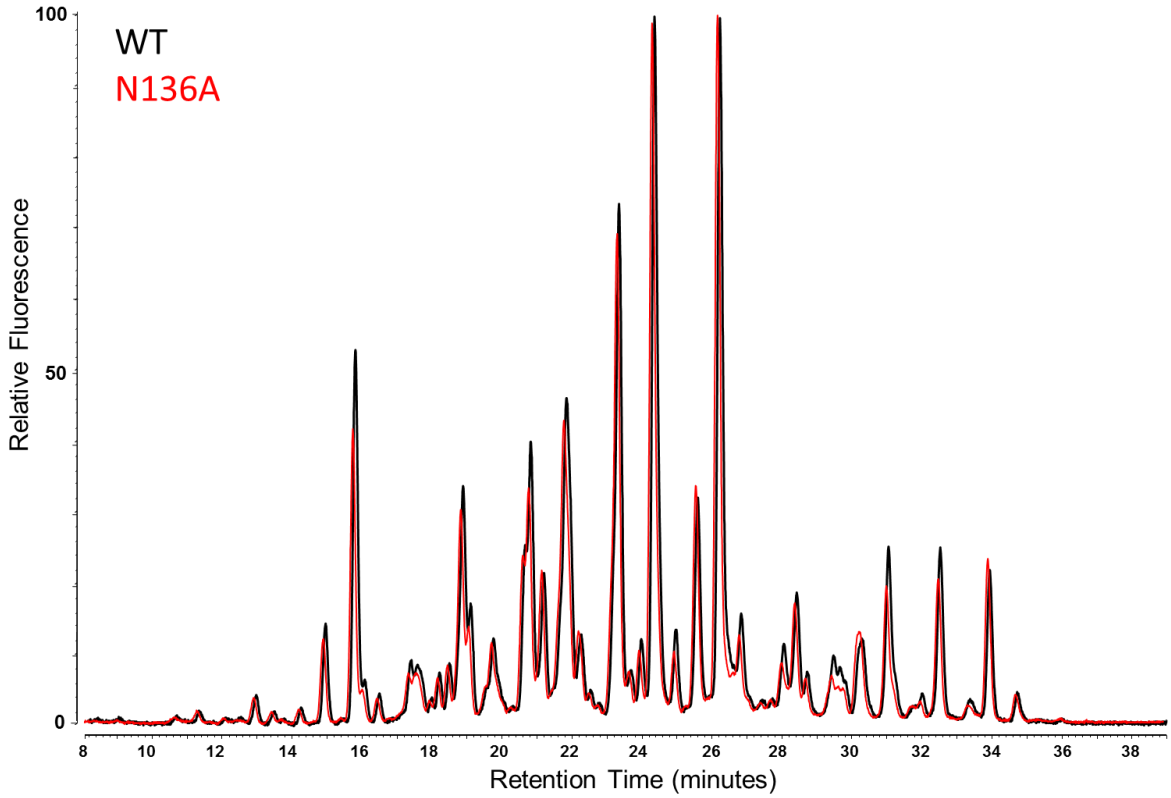
	% of total glycans (fold change)					
	M5	**M6**	**M7**	**M8**	**M9**	**Total M5-M9**
N88A	0.87	1.04	1.12	1.19	1.09	1.09
N130A	0.96	0.89	0.94	0.89	0.68	0.85
N136A	0.87	1.01	1.06	1.11	1.12	1.06
N141A	0.82	1.06	1.12	1.21	1.23	1.13
N144A	0.83	1.08	1.18	1.21	1.15	1.12
N156A	0.86	1.05	1.14	1.16	1.06	1.07
N186A	1.01	1.07	1.08	1.11	1.08	1.08
N197A	0.75	0.89	1.01	1.07	0.99	0.97
N241A	0.95	1.08	1.23	1.15	0.96	1.08
N262A	1.15	1.13	1.19	0.84	0.43	0.87
N289A	0.99	0.94	1.00	1.07	1.01	1.01
N295A	0.92	1.07	1.20	1.14	0.72	1.00
N301A	0.83	0.95	1.00	1.09	1.00	1.00
N332A	1.03	1.07	1.15	1.21	0.93	1.08
N339A	0.94	0.96	1.03	0.97	0.71	0.90
N356A	1.00	1.07	1.08	1.03	0.70	0.95
N386A	1.23	1.03	0.98	0.76	0.62	0.86
N392A	1.08	0.93	0.99	0.78	0.68	0.85
N396A	0.98	0.93	0.96	1.04	0.89	0.96
N406A	1.26	1.06	1.07	1.08	0.86	1.04
N411A	1.03	0.97	1.06	1.07	0.94	1.02
N448A	1.08	1.12	1.16	1.05	0.72	1.00
N463A	0.99	1.09	1.08	1.11	1.13	1.09
N295A + N332A	0.94	1.01	1.18	0.89	0.46	0.83
N295A + N339A	0.85	1.00	1.16	0.89	0.46	0.82
N295A + N386A	1.20	1.02	1.03	0.66	0.29	0.73
N295A + N448A	1.05	1.06	1.23	1.03	0.41	0.89

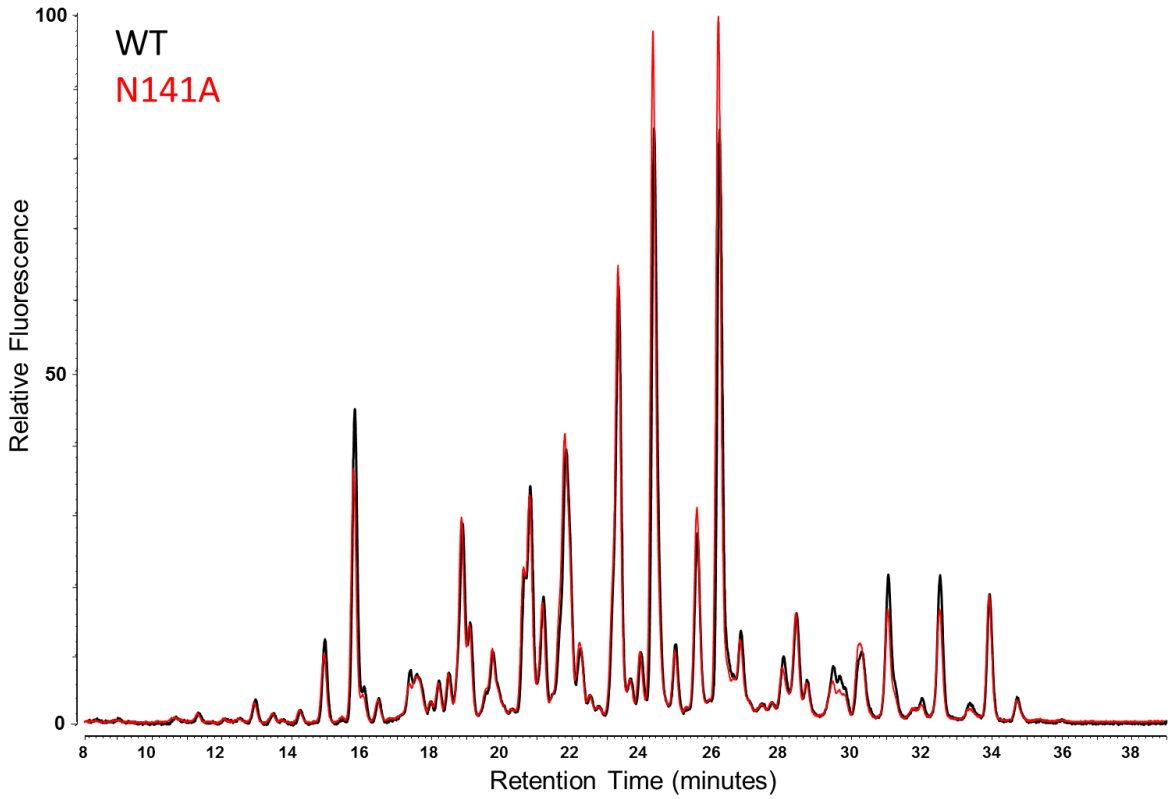
[¶]Values represent the proportion of the total glycan population that are of the oligomannose-type, and are presented as the abundance relative to wild-type. Values were obtained by integration of the corresponding HILIC-UPLC peaks using Empower 3 software.

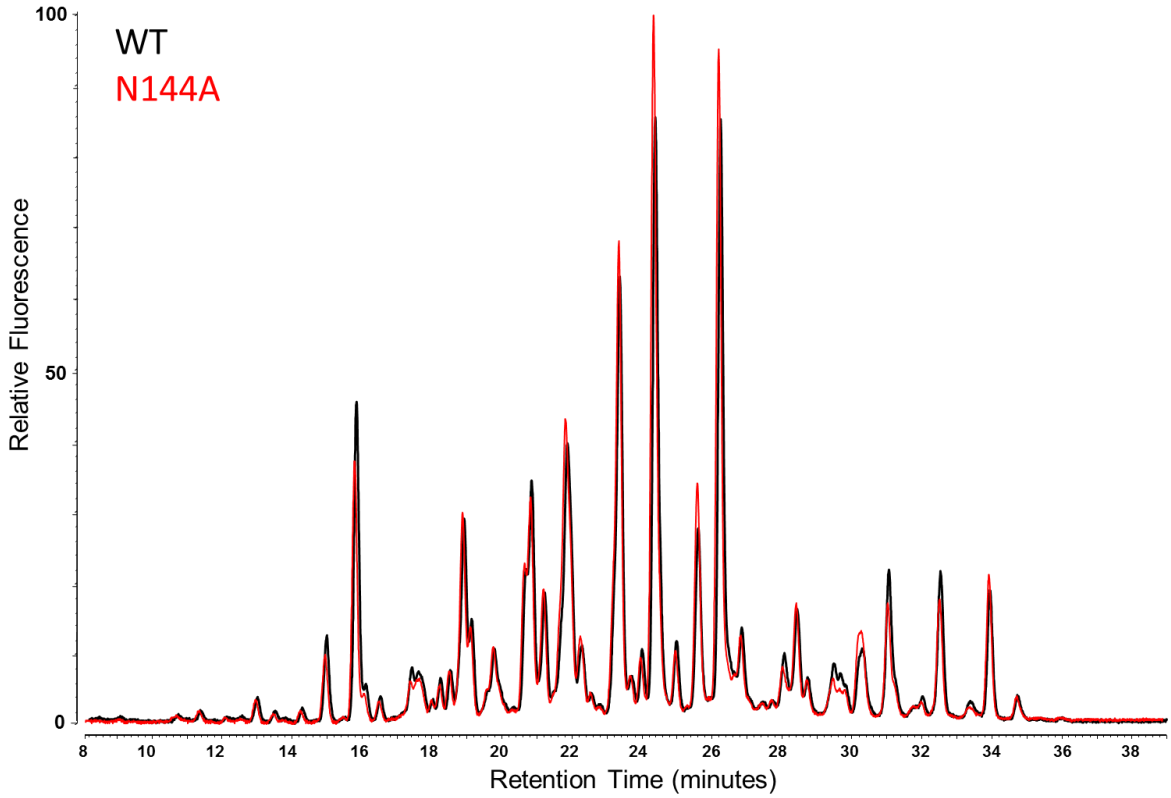
Dataset S4.A – HILIC-UPLC glycan profiles for the $gp120_{BaL}$ PNGS-deletion mutants

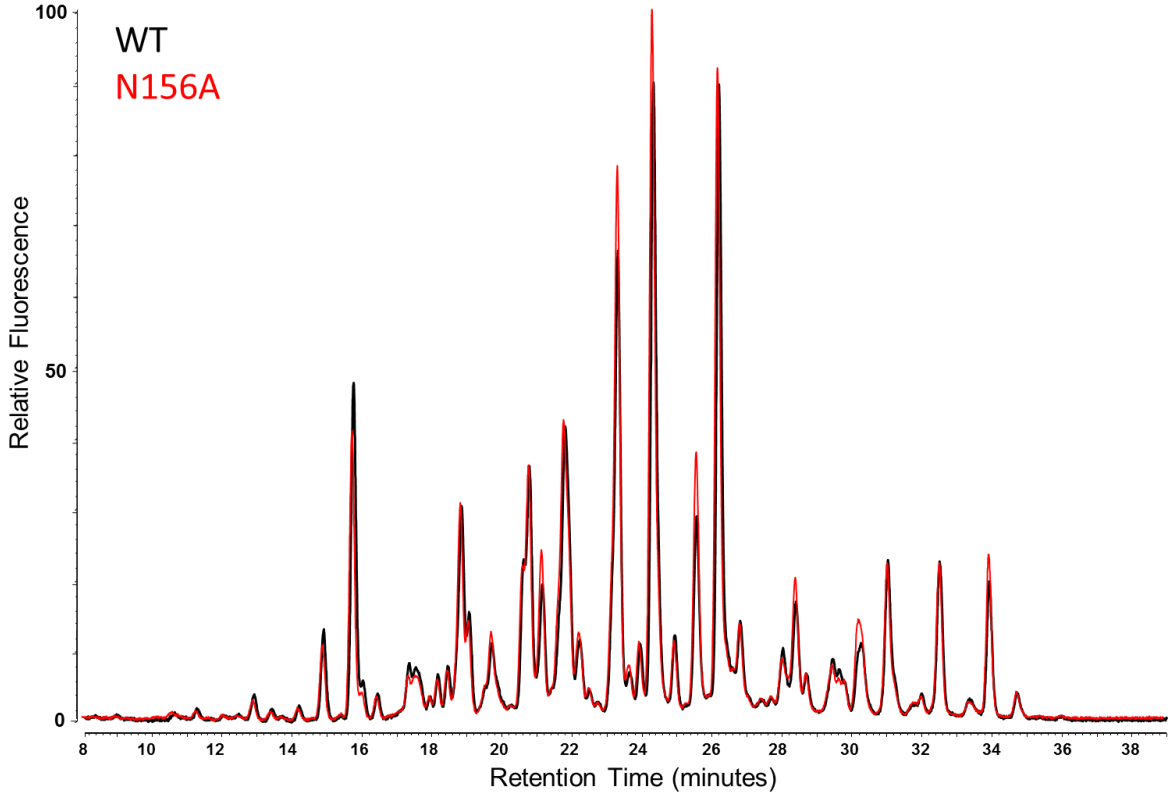


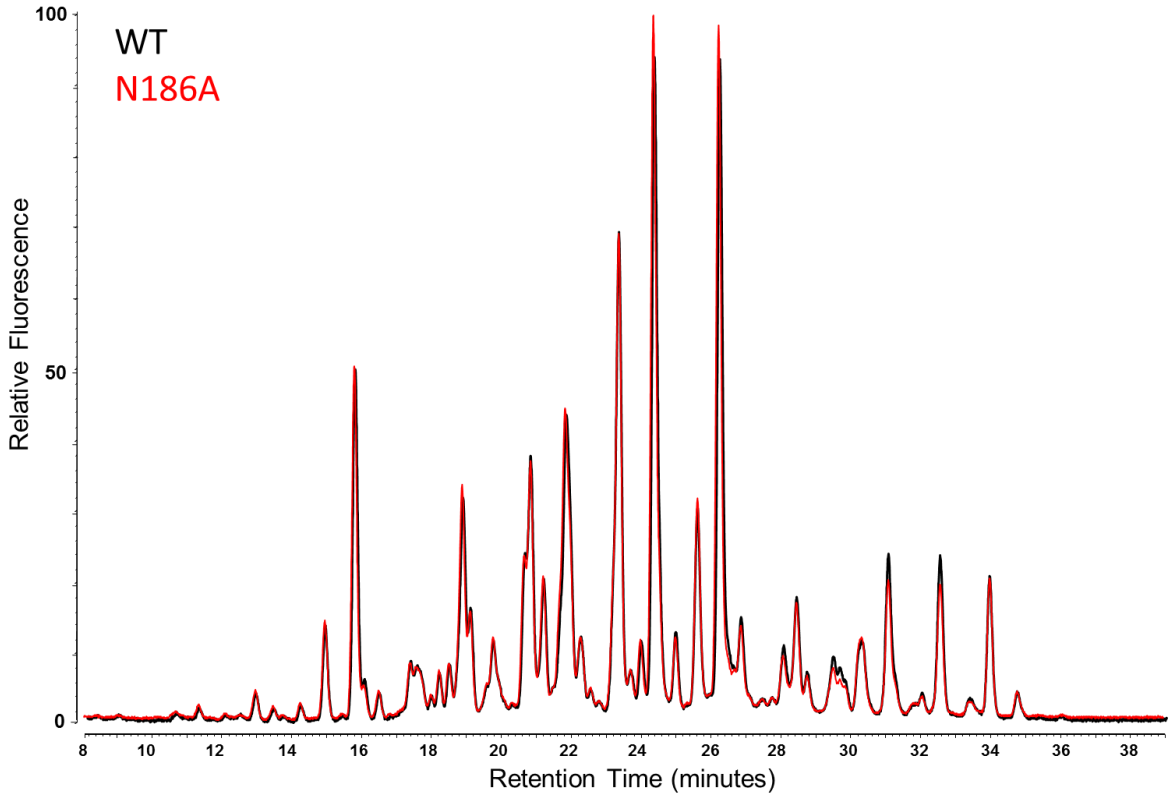


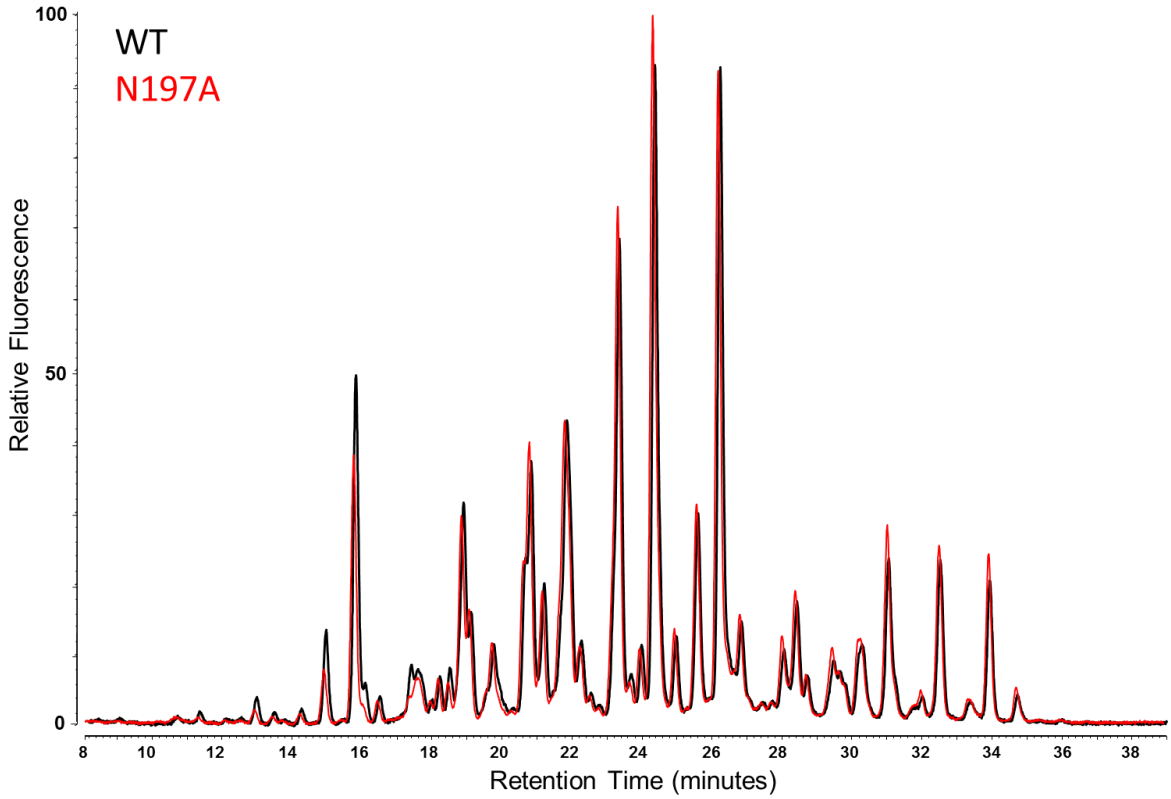

WT
N136A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

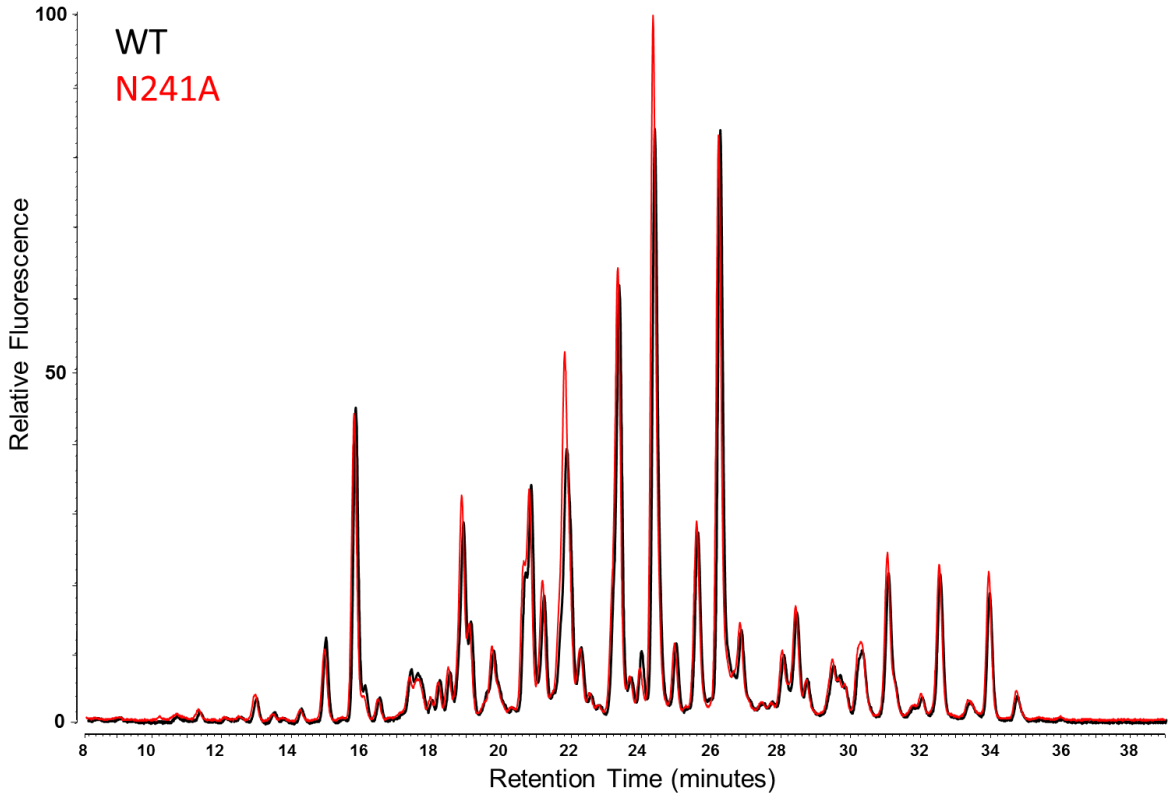

WT
N141A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

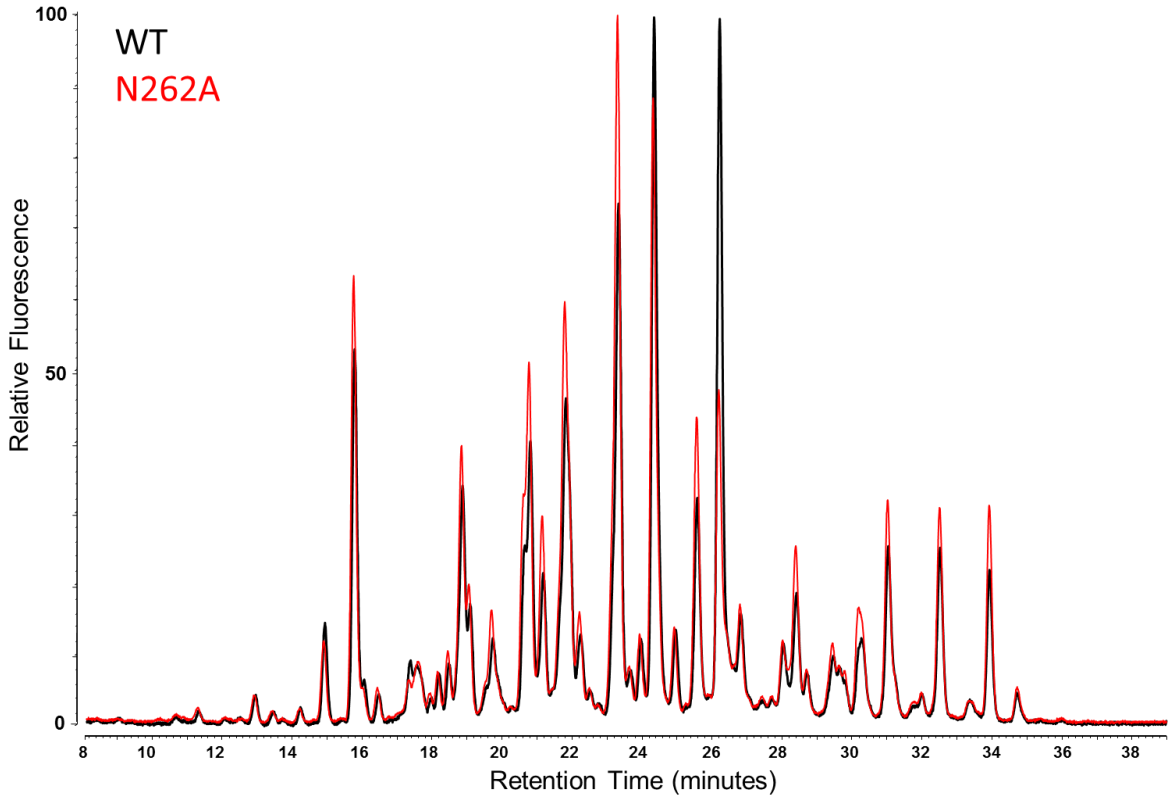

WT
N144A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

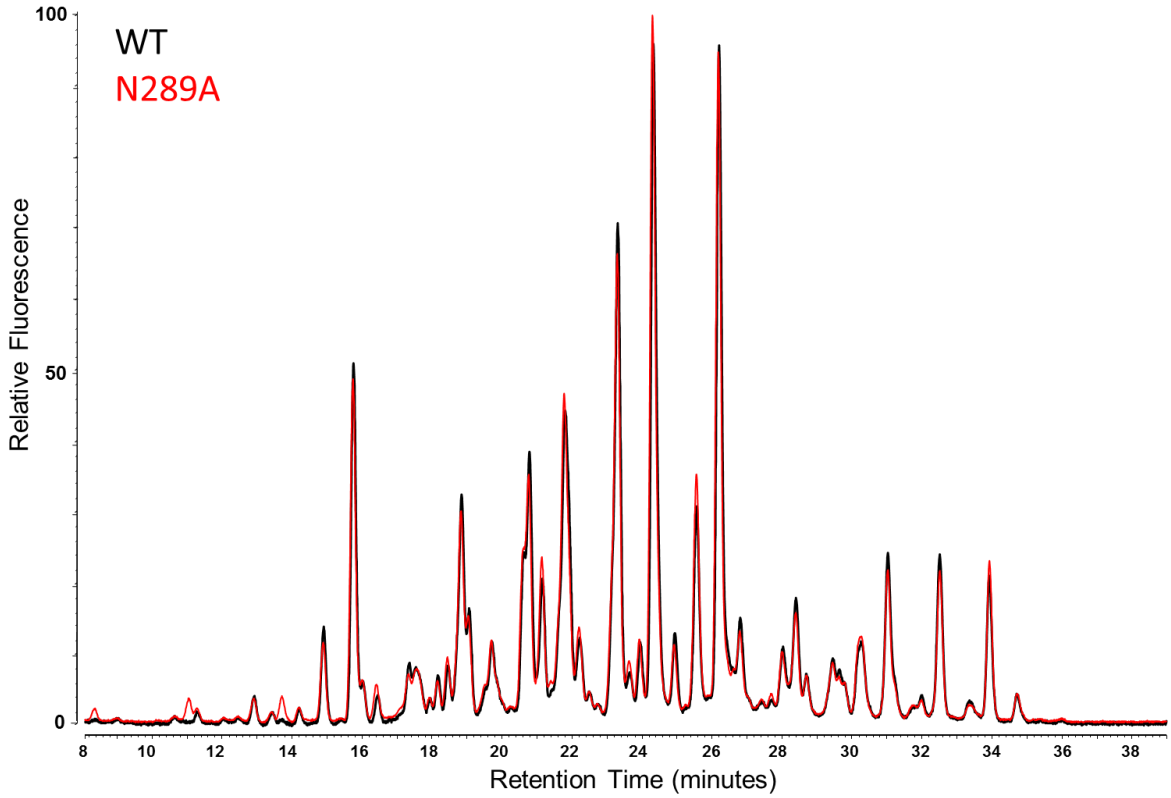

WT
N156A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

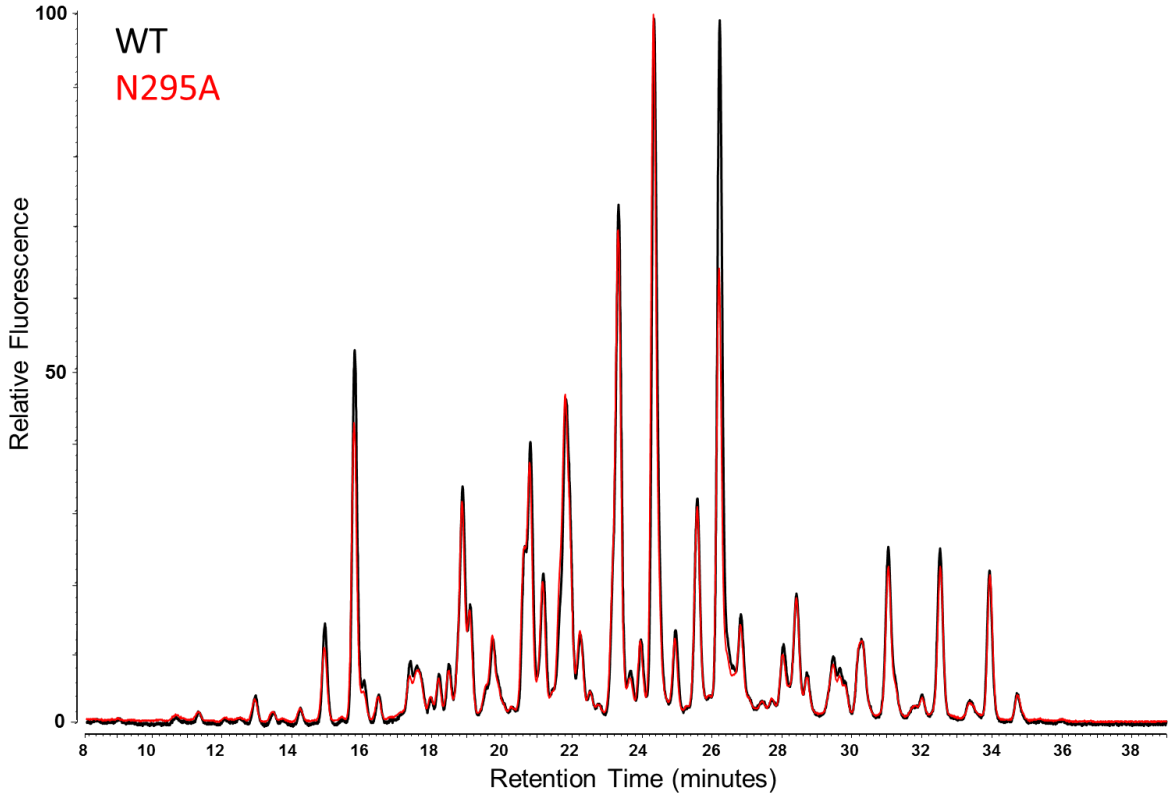

WT
N186A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

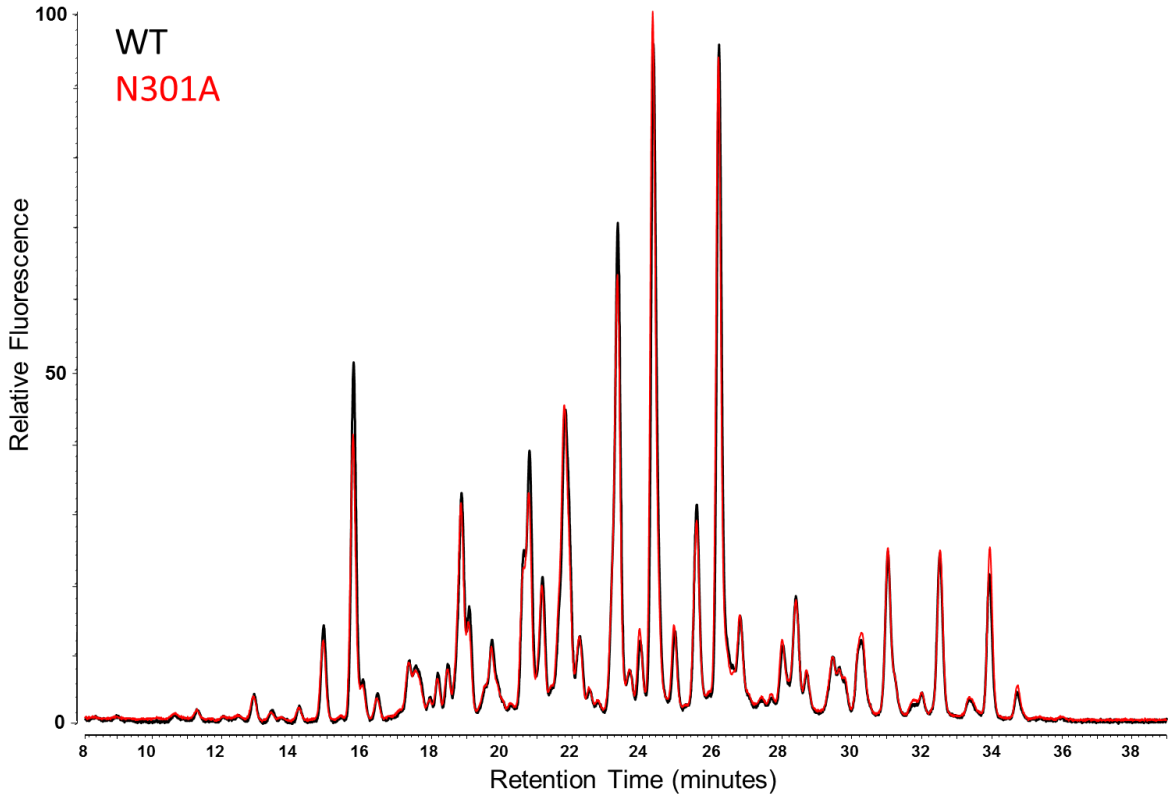

WT
N197A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

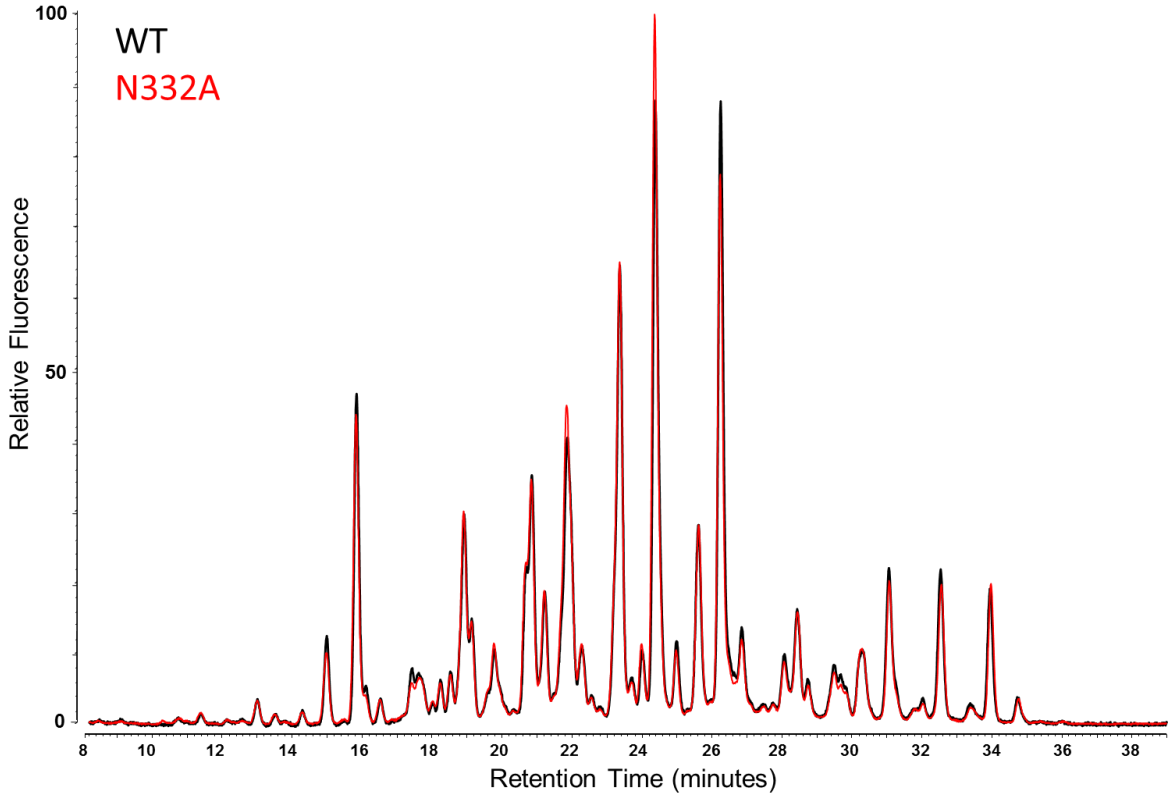

WT
N241A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

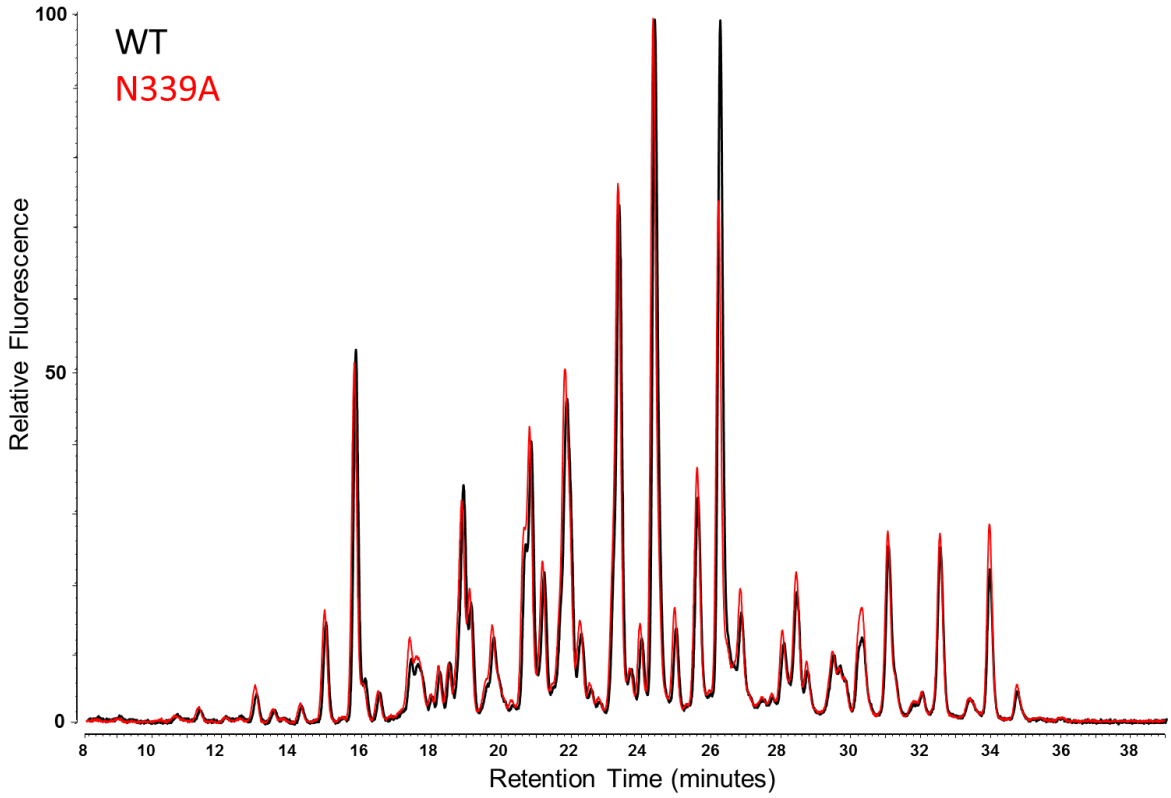

WT
N262A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

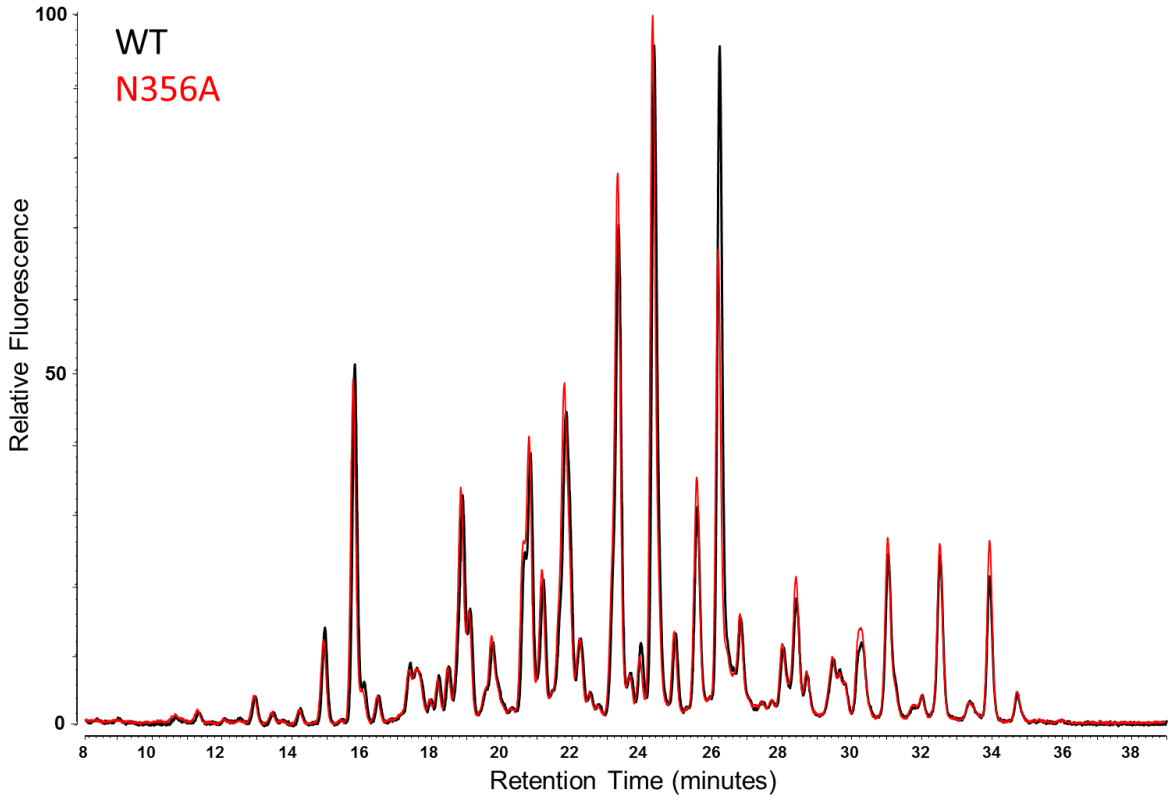

WT
N289A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

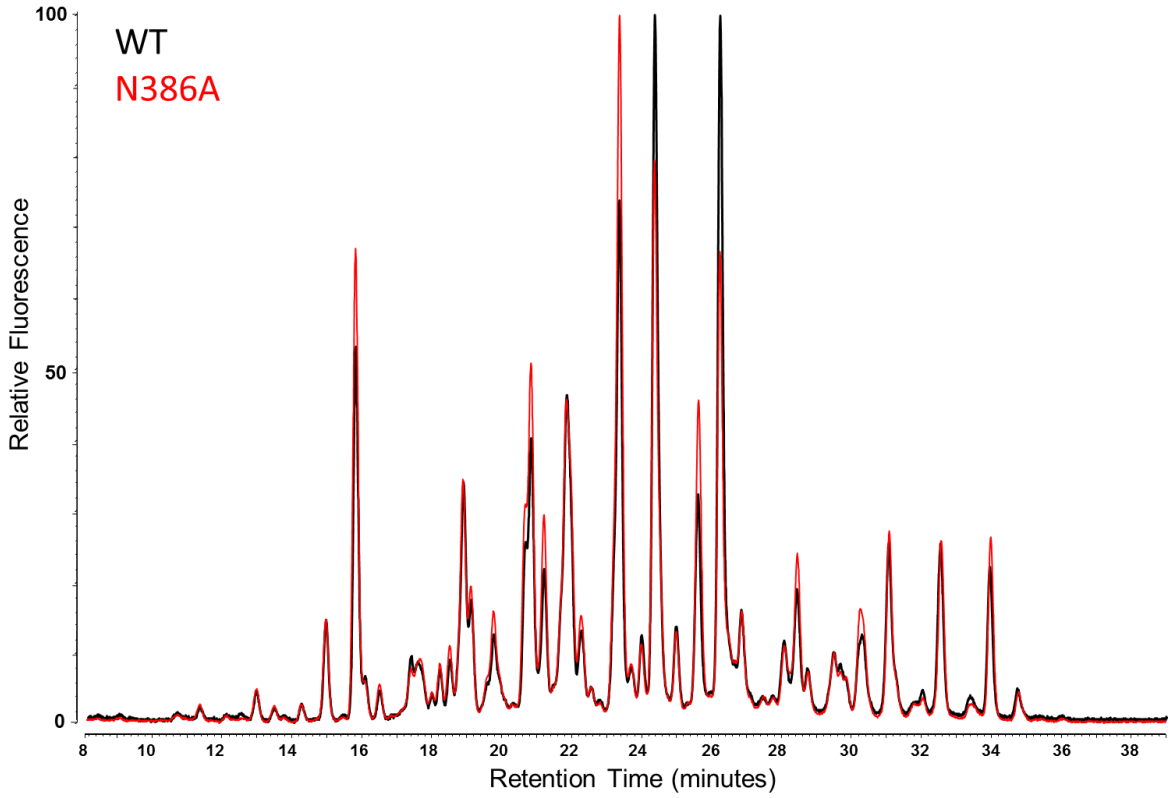

WT
N295A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

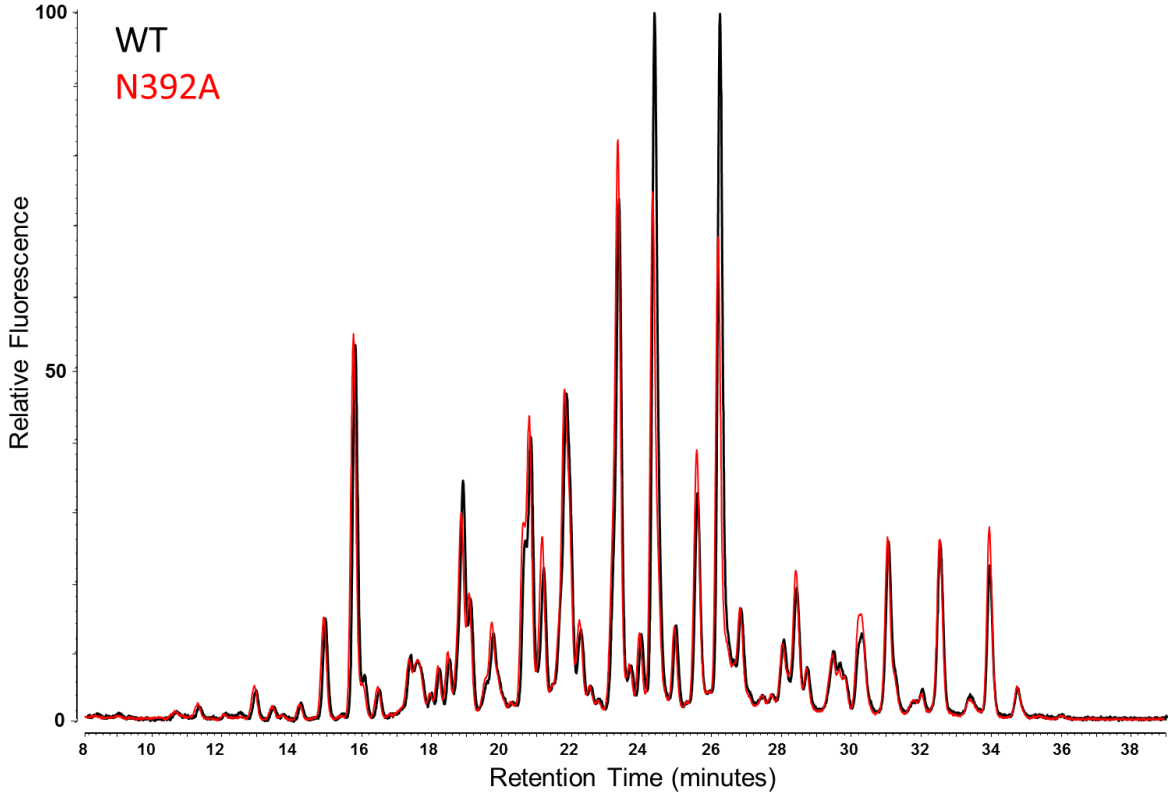

WT
N301A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

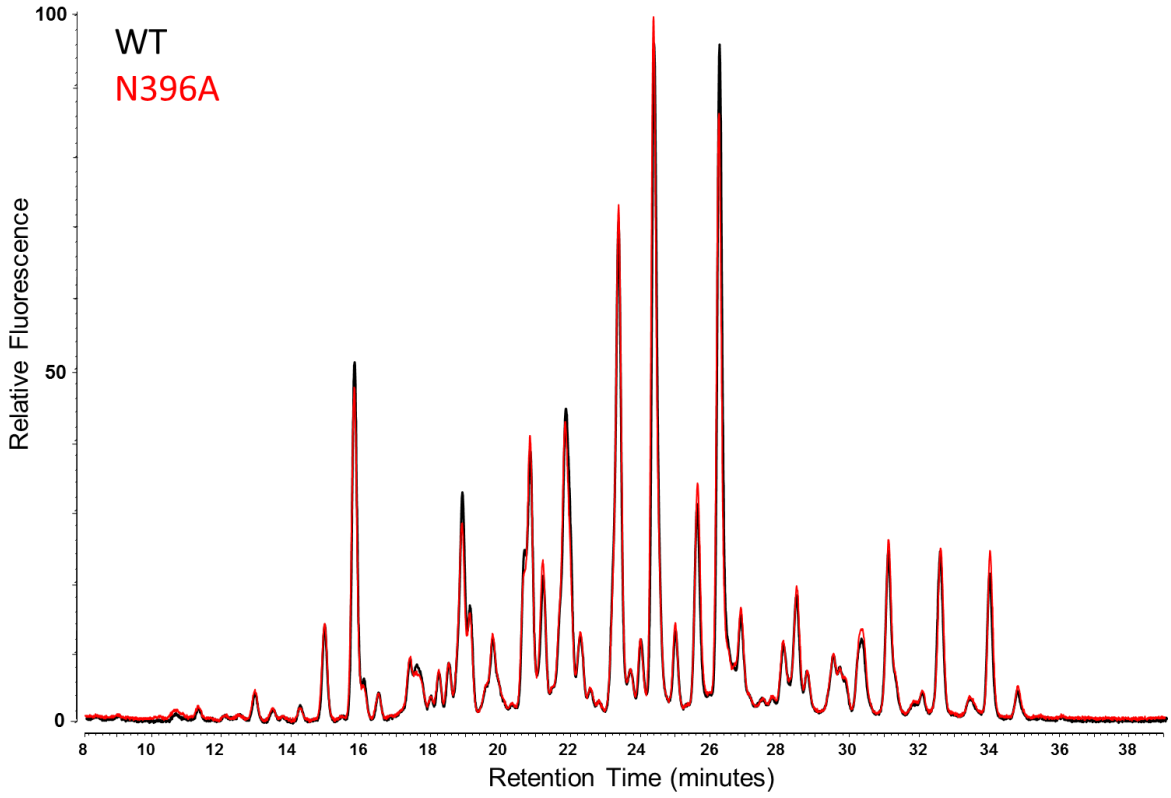

WT
N332A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

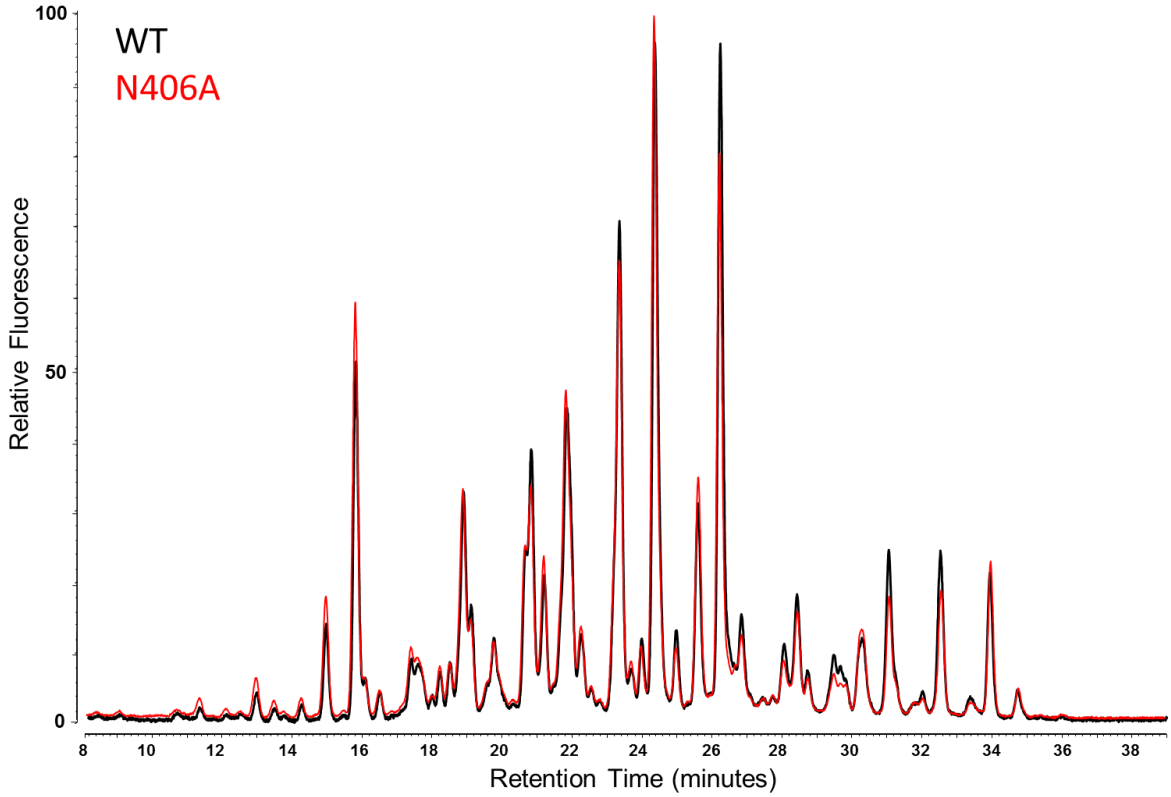

WT
N339A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

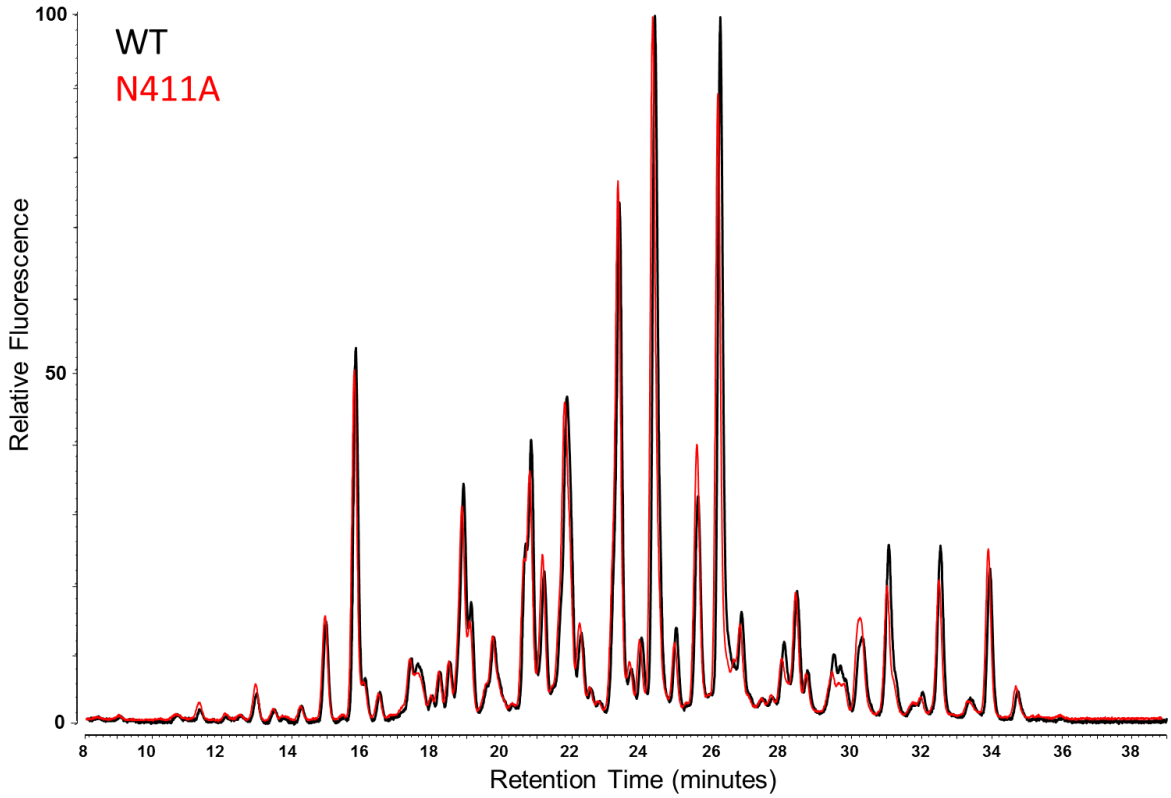

WT
N356A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

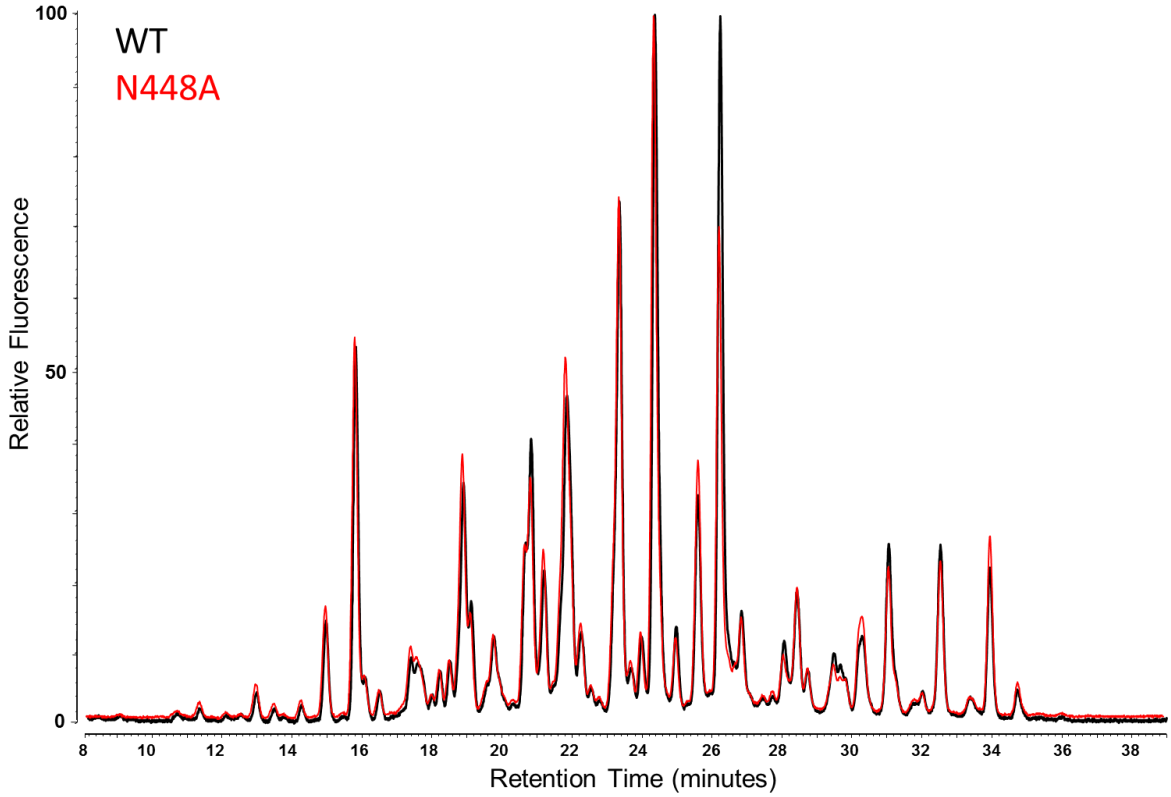

WT
N386A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

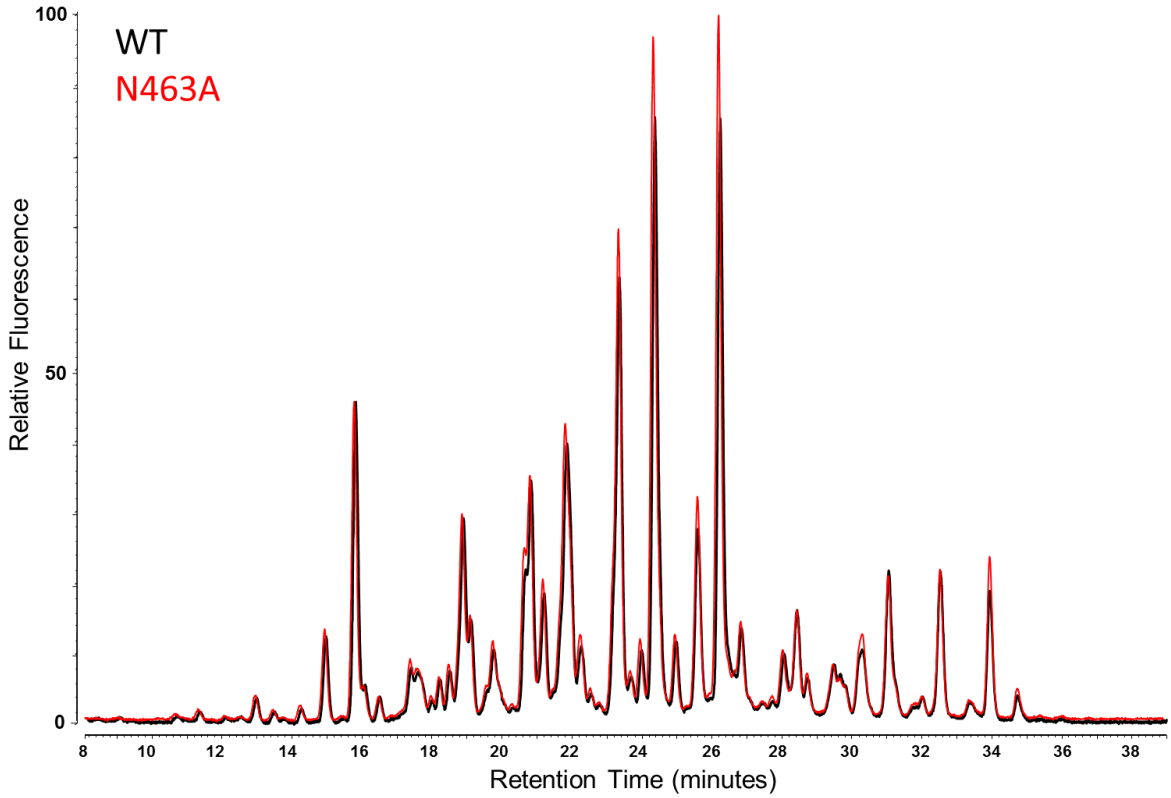

WT
N392A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)


WT
N396A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)


WT
N406A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

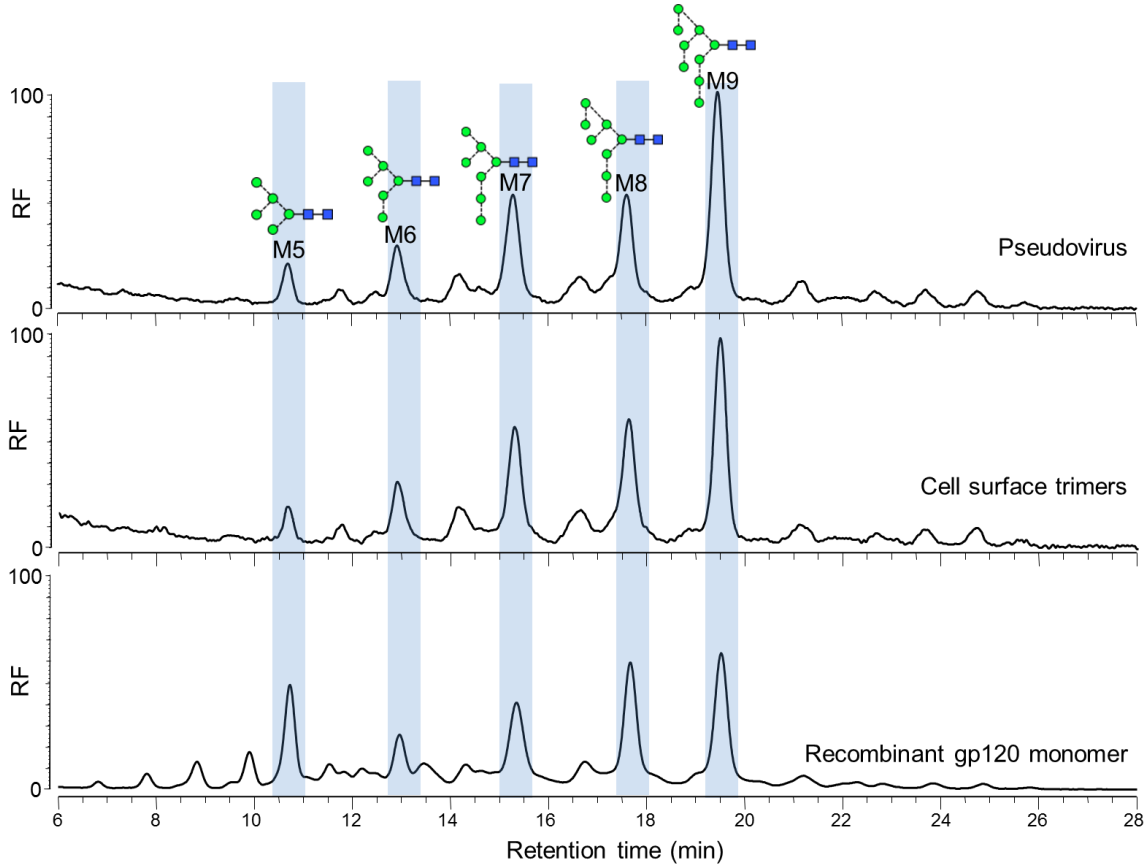

WT
N411A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)


WT
N448A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)


WT
N463A
Relative Fluorescence
100
50
0
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
38
Retention Time (minutes)

Appendix V

HILIC-HPLC glycan analysis of pseudovirus and cell-surface trimers purified using PGT151.



Dominance of large oligomannose-type glycans on gp120 derived from pseudovirus or cell surface trimers, isolated using the quaternary-dependent PGT151 bnAb. Pseudovirus was produced in 293T cells as described in section 2.12. Trimeric gp120 was isolated by immunoprecipitation with bnAb PGT151. Cell surface trimers were also isolated by immunoprecipitation with PGT151 following transfection of 293T cells with a full length Env construct. Recombinant monomeric gp120 was purified using Ni^{2+}-NTA. All samples were prepared by Dr Katie Doores (King's College, London). Glycans were released from SDS-PAGE bands using PNGase F and 2-AA labelled as described in sections 2.6.1 and 2.7. They were then analysed by HILIC-HPLC using the following gradient conditions: time = 0 min (*t*=0): 35% A, 65% B; *t*= 26: 46% A, 54% B; *t*=26.5: 80% A, 20% B; *t*=28.5: 80% A, 20% B; *t*=30: 35% A, 65% B; *t*=46: 35% A, 65% B, at a flow rate of 1 ml/min, and where solvent A was 50 mM ammonium formate, pH 4.4, and solvent B was acetonitrile. Fluorescence was measured using an excitation wavelength of 330 nm and a detection wavelength of 420 nm.

Bibliography

Abdool Karim, Q., Abdool Karim, S.S., Frohlich, J. a, Grobler, A.C., Baxter, C., Mansoor, L.E., Kharsany, A.B.M., Sibeko, S., Mlisana, K.P., Omar, Z., et al. (2010). Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science *329*, 1168–1174.

Abrahams, M.-R., Anderson, J. a, Giorgi, E.E., Seoighe, C., Mlisana, K., Ping, L.-H., Athreya, G.S., Treurnicht, F.K., Keele, B.F., Wood, N., et al. (2009). Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. J. Virol. *83*, 3556–3567.

Agrawal-Gamse, C., Luallen, R.J., Liu, B., Fu, H., Lee, F.-H., Geng, Y., and Doms, R.W. (2011). Yeast-elicited cross-reactive antibodies to HIV Env glycans efficiently neutralize virions expressing exclusively high-mannose N-linked glycans. J. Virol. *85*, 470–480.

Alexandre, K.B., Gray, E.S., Lambson, B.E., Moore, P.L., Choge, I. a, Mlisana, K., Karim, S.S.A., McMahon, J., O'Keefe, B., Chikwamba, R., et al. (2010). Mannose-rich glycosylation patterns on HIV-1 subtype C gp120 and sensitivity to the lectins, Griffithsin, Cyanovirin-N and Scytovirin. Virology *402*, 187–196.

Alexandre, K.B., Moore, P.L., Nonyane, M., Gray, E.S., Ranchobe, N., Chakauya, E., McMahon, J.B., O'Keefe, B.R., Chikwamba, R., and Morris, L. (2013). Mechanisms of HIV-1 subtype C resistance to GRFT, CV-N and SVN. Virology *446*, 66–76.

Amin, M.N., McLellan, J.S., Huang, W., Orwenyo, J., Burton, D.R., Koff, W.C., Kwong, P.D., and Wang, L.-X. (2013). Synthetic glycopeptides reveal the glycan specificity of HIV-neutralizing antibodies. Nat. Chem. Biol. *9*, 521–526.

Anumula, K.R., and Dhume, S.T. (1998). High resolution and high sensitivity methods for oligosaccharide mapping and characterization by normal phase high performance liquid chromatography following derivatization with highly fluorescent anthranilic acid. Glycobiology *8*, 685–694.

Arhel, N. (2010). Revisiting HIV-1 uncoating. Retrovirology *7*, 96.

Aricescu, a R., Lu, W., and Jones, E.Y. (2006). A time- and cost-efficient system for high-level protein production in mammalian cells. Acta Crystallogr. D. Biol. Crystallogr. *62*, 1243–1250.

Arthos, J., Cicala, C., Martinelli, E., Macleod, K., Van Ryk, D., Wei, D., Xiao, Z., Veenstra, T.D., Conrad, T.P., Lempicki, R. a, et al. (2008). HIV-1 envelope protein binds to and signals through integrin alpha4beta7, the gut mucosal homing receptor for peripheral T cells. Nat. Immunol. *9*, 301–309.

Baan, E., Ronde, A. De, Luchters, S., Vyankandondera, J., Lange, J.M., Pollakis, G., and Paxton, W.A. (2012). HIV Type 1 mother-to-child transmission facilitated by distinctive glycosylation sites in the gp120 envelope glycoprotein. AIDS Res. Hum. Retroviruses *28*, 715–724.

Baba, T.W., Liska, V., Hofmann-Lehmann, R., Vlasak, J., Xu, W., Ayehunie, S., Cavacini, L. a, Posner, M.R., Katinger, H., Stiegler, G., et al. (2000). Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection. Nat. Med. *6*, 200–206.

Back, N.K.T., Smit, L., De Jong, J.-J., Keulen, W., Schutten, M., Goudsmit, J., and Tersmette, M. (1994). An N-glycan within the Human Immunodeficiency Virus Type 1 gp120 V3 loop affects virus neutralization. Virology *199*, 431–438.

Baeten, J.M., Donnell, D., Ndase, P., Mugo, N.R., Campbell, J.D., Wangisi, J., Tappero, J.W., Bukusi, E. a, Cohen, C.R., Katabira, E., et al. (2012). Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N. Engl. J. Med. *367*, 399–410.

Ballweber, L., Robinson, B., Kreger, A., Fialkow, M., Lentz, G., McElrath, M.J., and Hladik, F. (2011). Vaginal langerhans cells nonproductively transporting HIV-1 mediate infection of T cells. J. Virol. *85*, 13443–13447.

Balzarini, J. (2005). Targeting the glycans of gp120: a novel approach aimed at the Achilles heel of HIV. Lancet Infect. Dis. *5*, 726–731.

Balzarini, J. (2006). Large-molecular-weight carbohydrate-binding agents as HIV entry inhibitors targeting glycoprotein gp120. Curr. Opin. HIV AIDS *1*, 355–360.

Balzarini, J. (2007). The alpha(1,2)-mannosidase I inhibitor 1-deoxymannojirimycin potentiates the antiviral activity of carbohydrate-binding agents against wild-type and mutant HIV-1 strains containing glycan deletions in gp120. FEBS Lett. *581*, 2060–2064.

Balzarini, J., Laethem, K. Van, Hatse, S., Vermeire, K., Clercq, E. De, Peumans, W., Damme, E. Van, Vandamme, A., Böhlmstedt, A., and Schols, D. (2004). Profile of resistance of human immunodeficiency virus mannose-specific plant lectins. J. Virol. *78*, 10617–10627.

Balzarini, J., Laethem, K. Van, Hatse, S., Froeyen, M., Damme, E. Van, Bolmstedt, A., Peumans, W., Clercq, E. De, and Schols, D. (2005). Marked depletion of glycosylation sites in HIV-1 gp120 under selection pressure by the mannose-specific plant lectins of Hippeastrum hybrid and Galanthus nivalis. Mol. Pharmacol. *67*, 1556–1565.

Balzarini, J., Van Laethem, K., Peumans, W.J., Van Damme, E.J.M., Bolmstedt, A., Gago, F., and Schols, D. (2006). Mutational pathways, resistance profile, and side effects of cyanovirin relative to human immunodeficiency virus type 1 strains with N-glycan deletions in their gp120 envelopes. J. Virol. *80*, 8411–8421.

Banerjee, K., Andjelic, S., Klasse, P.J., Kang, Y., Sanders, R.W., Michael, E., Durso, R.J., Ketas, T.J., Olson, W.C., and Moore, J.P. (2009). Enzymatic removal of mannose moieties can increase the immune response to HIV-1 gp120 in vivo. Virology *389*, 108–121.

Banerjee, K., Michael, E., Eggink, D., van Montfort, T., Lasnik, A.B., Palmer, K.E., Sanders, R.W., Moore, J.P., and Klasse, P.J. (2012). Occluding the mannose moieties on human immunodeficiency virus type 1 gp120 with griffithsin improves the antibody responses to both proteins in mice. AIDS Res. Hum. Retroviruses *28*, 206–214.

Barbouche, R., Miquelis, R., Jones, I.M., and Fenouillet, E. (2003). Protein-disulfide isomerase-mediated reduction of two disulfide bonds of HIV envelope glycoprotein 120 occurs post-CXCR4 binding and is required for fusion. J. Biol. Chem. *278*, 3131–3136.

Barouch, D.H. (2010). Novel adenovirus vector-based vaccines for HIV-1. Curr. Opin. HIV AIDS *5*, 386–390.

Barouch, D.H. (2013). The quest for an HIV-1 vaccine - moving forward. N. Engl. J. Med. *369*, 2073–2076.

Barouch, D.H., Santra, S., Scmitz, J.E., Kuroda, M.J., Fu, T.-M., Wagner, W., Bilska, M., Craiu, A., Zheng, X.X., Krivulka, G.R., et al. (2000). Control of viremia and prevention of clinical AIDS in rhesus monkeys by cytokine-augmented DNA vaccination. Science *290*, 486–492.

Barouch, D.H., Kunstman, J., Kuroda, M.J., Schmitz, J.E., Santra, S., Peyerl, F.W., Krivulka, G.R., Beaudry, K., Lifton, M. a, Gorgone, D. a, et al. (2002). Eventual AIDS vaccine failure in a rhesus monkey by viral escape from cytotoxic T lymphocytes. Nature *415*, 335–339.

Barré-Sinoussi, F., Chermann, C.J., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., et al. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science *220*, 868–871.

Ben-Dor, S., Esterman, N., Rubin, E., and Sharon, N. (2004). Biases and complex patterns in the residues flanking protein N-glycosylation sites. Glycobiology *14*, 95–101.

Bernstein, H.B., Tucker, S.P., Hunter, E., Schutzbach, J.S., and Compans, R.W. (1994). Human immunodeficiency virus type 1 envelope glycoprotein is modified by O-linked oligosaccharides. J. Virol. *68*, 463–468.

Bewley, C. a (2001). Solution structure of a cyanovirin-N:Man alpha 1-2Man alpha complex: structural basis for high-affinity carbohydrate-mediated binding to gp120. Structure *9*, 931–940.

Bewley, C. a, and Otero-Quintero, S. (2001). The potent anti-HIV protein cyanovirin-N contains two novel carbohydrate binding sites that selectively bind to Man(8) D1D3 and Man(9) with nanomolar affinity: implications for binding to the HIV envelope protein gp120. J. Am. Chem. Soc. *123*, 3892–3902.

Bieniasz, P.D. (2009). The cell biology of HIV-1 virion genesis. Cell Host Microbe *5*, 550–558.

Binley, J.M., Sanders, R.W., Clas, B., Schuelke, N., Master, A., Guo, Y., Kajumo, F., Anselma, D.J., Maddon, P.J., Olson, W.C., et al. (2000). 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. *74*, 627–643.

Binley, J.M., Wrin, T., Korber, B., Zwick, M.B., Wang, M., Chappey, C., Stiegler, G., Zolla-pazner, S., Katinger, H., Christos, J., et al. (2004). Comprehensive Cross-Clade Neutralization Analysis of a Panel of Anti-Human Immunodeficiency Virus Type 1 Monoclonal Antibodies Comprehensive Cross-Clade Neutralization Analysis of a Panel of Anti-Human Immunodeficiency Virus Type 1 Monoclonal Antibodies. J. Virol. *78*, 13232–13252.

Binley, J.M., Ban, Y.-E.A., Crooks, E.T., Eggink, D., Osawa, K., Schief, W.R., and Sanders, R.W. (2010). Role of complex carbohydrates in human immunodeficiency virus type 1 infection and resistance to antibody neutralization. J. Virol. *84*, 5637–5655.

Black, L.R., and Aiken, C. (2010). TRIM5alpha disrupts the structure of assembled HIV-1 capsid complexes in vitro. J. Virol. *84*, 6564–6569.

Blattner, C., Lee, J.H., Sliepen, K., Derking, R., Falkowska, E., de la Peña, A.T., Cupo, A., Julien, J.-P., van Gils, M., Lee, P.S., et al. (2014). Structural delineation of a quaternary, cleavage-dependent epitope at the gp41-gp120 interface on intact HIV-1 Env trimers. Immunity *40*, 669–680.

Bolmstedt, a J., O'Keefe, B.R., Shenoy, S.R., McMahon, J.B., and Boyd, M.R. (2001). Cyanovirin-N defines a new class of antiviral agent targeting N-linked, high-mannose glycans in an oligosaccharide-specific manner. Mol. Pharmacol. *59*, 949–954.

Bonomelli, C., Doores, K.J., Dunlop, D.C., Thaney, V., Dwek, R. a, Burton, D.R., Crispin, M., and Scanlan, C.N. (2011). The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade. PLoS One *6*, e23521.

Bonsignori, M., Hwang, K.-K., Chen, X., Tsao, C.-Y., Morris, L., Gray, E., Marshall, D.J., Crump, J. a, Kapiga, S.H., Sam, N.E., et al. (2011). 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. *85*, 9998–10009.

Bonsignori, M., Pollara, J., Moody, M.A., Alpert, M.D., Chen, X., Hwang, K.-K., Gilbert, P.B., Huang, Y., Gurley, T.C., Kozink, D.M., et al. (2012). Antibody-dependent cellular cytotoxicity-mediating antibodies from an HIV-1 vaccine efficacy trial target multiple epitopes and preferentially use the VH1 gene family. J. Virol. *86*, 11521–11532.

Botarelli, P., Houlden, B. a, Haigwood, N.L., Servis, C., Montagna, D., and Abrignani, S. (1991). N-glycosylation of HIV-gp120 may constrain recognition by T lymphocytes. J. Immunol. *147*, 3128–3132.

Botos, I., O'Keefe, B.R., Shenoy, S.R., Cartner, L.K., Ratner, D.M., Seeberger, P.H., Boyd, M.R., and Wlodawer, A. (2002). Structures of the complexes of a potent anti-HIV protein cyanovirin-N and high mannose oligosaccharides. J. Biol. Chem. *277*, 34336–34342.

Boucher, C.A.B., Cammack, N., Schipper, P., Schuurman, R.O.B., Rouse, P., Wainberg, M.A., and Cameron, J.M. (1993). High-level resistance to (-) enantiomeric 2'-deoxy-3'-thiacytidine in vitro is due to one amino acid substitution in the catalytic site of human immunodeficiency virus type 1 reverse transcriptase. Antimicrob. Agents Chemother. *37*, 2231–2234.

Boyd, M.R., Gustafson, K.R., Mcmahon, J.B., Shoemaker, R.H., Keefe, B.R.O., Mori, T., Gulakowski, R.J., Wu, L.I.N., Rivera, M.I., Laurencot, C.M., et al. (1997a). Discovery of cyanovirin-N , a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development . Discovery of Cyanovirin-N , a Novel Human Immunodeficiency Vir. Antimicrob. Agents Chemother. *41*, 1521–1530.

Boyd, M.R., Gustafson, K.R., Mcmahon, J.B., Shoemaker, R.H., O'Keefe, B.R., Mori, T., Gulakowski, R.J., Wu, L.I.N., Rivera, M.I., Laurencot, C.M., et al. (1997b). Discovery of cyanovirin-N , a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. Antimicrob. Agents Chemother. *41*, 1521–1530.

Brenchley, J.M., Schacker, T.W., Ruff, L.E., Price, D. a, Taylor, J.H., Beilman, G.J., Nguyen, P.L., Khoruts, A., Larson, M., Haase, A.T., et al. (2004). CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J. Exp. Med. *200*, 749–759.

Brenchley, J.M., Price, D.A., Schacker, T.W., Asher, T.E., Silvestri, G., Rao, S., Kazzaz, Z., Bornstein, E., Lambotte, O., Altmann, D., et al. (2006). Microbial translocation is a cause of systemic immune activation in chronic HIV infection. Nat. Med. *12*, 1365–1371.

Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., and Imberty, A. (2006). Structures and mechanisms of glycosyltransferases. Glycobiology *16*, 29R–37R.

Buchbinder, S.P., Mehrotra, D. V, Duerr, A., Fitzgerald, D.W., Mogg, R., Li, D., Gilbert, P.B., Lama, J.R., Marmor, M., Del Rio, C., et al. (2008). Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. Lancet *372*, 1881–1893.

Burton, D., and Weiss, R.A. (2010). A boost for HIV vaccine design. Science *329*, 770–772.

Burton, D., Pyati, J., Koduri, R., Sharp, S., Thornton, G., Parren, P., Sawyer, L., Hendry, R., Dunlop, N., Nara, P., et al. (1994). Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. Science *266*, 1024–1027.

Burton, D.R., Ahmed, R., Barouch, D.H., Butera, S.T., Crotty, S., Godzik, A., Kaufmann, D.E., McElrath, M.J., Nussenzweig, M.C., Pulendran, B., et al. (2012). A blueprint for HIV vaccine discovery. Cell Host Microbe *12*, 396–407.

Calarese, D. a, Scanlan, C.N., Zwick, M.B., Deechongkit, S., Mimura, Y., Kunert, R., Zhu, P., Wormald, M.R., Stanfield, R.L., Roux, K.H., et al. (2003). Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. Science *300*, 2065–2071.

Calarese, D. a, Lee, H.-K., Huang, C.-Y., Best, M.D., Astronomo, R.D., Stanfield, R.L., Katinger, H., Burton, D.R., Wong, C.-H., and Wilson, I. a (2005). Dissection of the carbohydrate specificity of the broadly neutralizing anti-HIV-1 antibody 2G12. Proc. Natl. Acad. Sci. U. S. A. *102*, 13372–13377.

Cavrois, M., Neidleman, J., Santiago, M.L., Derdeyn, C. a, Hunter, E., and Greene, W.C. (2014). Enhanced fusion and virion incorporation for HIV-1 subtype C envelope glycoproteins with compact V1/V2 domains. J. Virol. *88*, 2083–2094.

Chackerian, B., Rudensey, L.M., and Overbaugh, J. (1997). Specific N-linked and O-linked glycosylation modifications in the envelope V1 domain of simian immunodeficiency virus variants that evolve in the host alter recognition by neutralizing antibodies. J. Virol. *71*, 7719–7727.

Checkley, M.A., Luttge, B.G., and Freed, E.O. (2011). HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. J. Mol. Biol. *410*, 582–608.

Chen, B., Vogan, E.M., Gong, H., Skehel, J.J., Wiley, D.C., and Harrison, S.C. (2005). Structure of an unliganded simian immunodeficiency virus gp120 core. Nature *433*, 834–841.

Chen, Y., Hwang, S.-L., Chan, V.S.F., Chung, N.P.Y., Wang, S.-R., Li, Z., Ma, J., Lin, C.-W., Hsieh, Y.-J., Chang, K.-P., et al. (2013). Binding of HIV-1 gp120 to DC-SIGN promotes ASK-1-dependent activation-induced apoptosis of human dendritic cells. PLoS Pathog. *9*, e1003100.

Chertova, E., Bess, J.W., Crise, B.J., Sowder, R.C., Schaden, T.M., Hilburn, J.M., Hoxie, J.A., Benveniste, R.E., Lifson, J.D., Henderson, L.E., et al. (2002). Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), is the primary determinant of SU content of purified human immunodeficiency virus type 1 and simian immunodeficiency. J. Virol. *76*, 5315–5325.

Chohan, B., Lang, D., Sagar, M., Korber, B., Lavreys, L., Richardson, B., and Overbaugh, J. (2005). 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. *79*, 6528–6531.

Chojnacki, J., Staudt, T., Glass, B., Bingen, P., Engelhardt, J., Anders, M., Schneider, J., Müller, B., Hell, S.W., and Kräusslich, H.-G. (2012). Maturation-dependent HIV-1 surface protein redistribution revealed by fluorescence nanoscopy. Science *338*, 524–528.

Chuang, G.-Y., Acharya, P., Schmidt, S.D., Yang, Y., Louder, M.K., Zhou, T., Kwon, Y. Do, Pancera, M., Bailer, R.T., Doria-Rose, N. a, et al. (2013). Residue-level prediction of HIV-1 antibody epitopes based on neutralization of diverse viral strains. J. Virol. *87*, 10047–10058.

Chun, T.W., Stuyver, L., Mizell, S.B., Ehler, L. a, Mican, J. a, Baseler, M., Lloyd, a L., Nowak, M. a, and Fauci, a S. (1997). Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. Proc. Natl. Acad. Sci. U. S. A. *94*, 13193–13197.

Chung, N.P., Matthews, K., Kim, H.J., Ketas, T.J., Golabek, M., de Los Reyes, K., Korzun, J., Yasmeen, A., Sanders, R.W., Klasse, P.J., et al. (2014). Stable 293 T and CHO cell lines expressing cleaved, stable HIV-1 envelope glycoprotein trimers for structural and vaccine studies. Retrovirology *11*, 33.

Clavel, F., and Hance, A.J. (2004). HIV drug resistance. N. Engl. J. Med. *350*, 1023–1035.

Clevestig, P., Pramanik, L., Leitner, T., and Ehrnst, A. (2006). CCR5 use by human immunodeficiency virus type 1 is associated closely with the gp120 V3 loop N-linked glycosylation site. J. Gen. Virol. *87*, 607–612.

Coffin, J.M. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science *267*, 483–489.

Cohen, J. (2013). Is high-tech view of HIV too good to be true? Science (80-.). *341*, 443–444.

Connor, R.I., Sheridan, K.E., Ceradini, D., Choe, S., and Landau, N.R. (1997). Change in coreceptor use correlates with disease progression in HIV-1--infected individuals. J. Exp. Med. *185*, 621–628.

Crispin, M., Harvey, D.J., Bitto, D., Bonomelli, C., Edgeworth, M., Scrivens, J.H., Huiskonen, J.T., and Bowden, T. a (2014). Structural Plasticity of the Semliki Forest Virus Glycome upon Interspecies Transmission. J. Proteome Res.

Crispin, M.D.M., Ritchie, G.E., Critchley, A.J., Morgan, B.P., Wilson, I. a, Dwek, R. a, Sim, R.B., and Rudd, P.M. (2004). Monoglucosylated glycans in the secreted human complement component C3: implications for protein biosynthesis and structure. FEBS Lett. *566*, 270–274.

Cutalo, J.M., Deterding, L.J., and Tomer, K.B. (2004). Characterization if glycopeptides from HIV-1 SF2 gp120 by liquid chromatography mass spectrometry. J. Am. Soc. Mass Spectrom. *15*, 1545–1555.

Deneka, M., Pelchen-Matthews, A., Byland, R., Ruiz-Mateos, E., and Marsh, M. (2007). In macrophages, HIV-1 assembles into an intracellular plasma membrane domain containing the tetraspanins CD81, CD9, and CD53. J. Cell Biol. *177*, 329–341.

Depetris, R.S., Julien, J.-P., Khayat, R., Lee, J.H., Pejchal, R., Katpally, U., Cocco, N., Kachare, M., Massi, E., David, K.B., et al. (2012). Partial enzymatic deglycosylation preserves the structure of cleaved recombinant HIV-1 envelope glycoprotein trimers. J. Biol. Chem. *287*, 24239–24254.

Derdeyn, C. a, Decker, J.M., Bibollet-Ruche, F., Mokili, J.L., Muldoon, M., Denham, S. a, Heil, M.L., Kasolo, F., Musonda, R., Hahn, B.H., et al. (2004). Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science *303*, 2019–2022.

Dettenhofer, M., and Yu, X.F. (2001). Characterization of the biosynthesis of human immunodeficiency virus type 1 Env from infected T-cells and the effects of glucose trimming of Env on virion infectivity. J. Biol. Chem. *276*, 5985–5991.

Dey, B., Lerner, D.L., Lusso, P., Boyd, M.R., Elder, J.H., and Berger, E.A. (2000). Multiple antiviral activities of Cyanovirin-N : blocking of human immunodeficiency virus type 1 gp120 interaction with CD4 and coreceptor and inhibition of diverse enveloped viruses. J. Virol. *74*, 4562–4569.

Doitsh, G., Galloway, N.L.K., Geng, X., Yang, Z., Monroe, K.M., Zepeda, O., Hunt, P.W., Hatano, H., Sowinski, S., Muñoz-Arias, I., et al. (2014). Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. Nature *505*, 509–514.

Doms, R.W., and Moore, J.P. (2000). HIV-1 membrane fusion: targets of opportunity. J. Cell Biol. *151*, F9–14.

Doores, K.J., and Burton, D.R. (2010). Variable loop glycan dependency of the broad and potent HIV-1-neutralizing antibodies PG9 and PG16. J. Virol. *84*, 10510–10521.

Doores, K.J., Bonomelli, C., Harvey, D.J., Vasiljevic, S., Dwek, R.A., and Burton, D.R. (2010a). Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. Proc. Natl. Acad. Sci. U. S. A. *107*, 13800–13805.

Doores, K.J., Fulton, Z., Huber, M., Wilson, I. a, and Burton, D.R. (2010b). 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. *84*, 10690–10699.

Doores, K.J., Fulton, Z., Hong, V., Patel, M.K., Scanlan, C.N., Wormald, M.R., Finn, M.G., Burton, D.R., Wilson, I. a, and Davis, B.G. (2010c). A nonself sugar mimic of the HIV glycan shield shows enhanced antigenicity. Proc. Natl. Acad. Sci. U. S. A. *107*, 17107–17112.

Doria-Rose, N. a, Schramm, C. a, Gorman, J., Moore, P.L., Bhiman, J.N., DeKosky, B.J., Ernandes, M.J., Georgiev, I.S., Kim, H.J., Pancera, M., et al. (2014). Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. Nature *509*, 55–62.

Drummer, H.E., Hill, M.K., Maerz, A.L., Wood, S., Ramsland, P. a, Mak, J., and Poumbourios, P. (2013). Allosteric modulation of the HIV-1 gp120-gp41 association site by adjacent gp120 variable region 1 (V1) N-glycans linked to neutralization sensitivity. PLoS Pathog. *9*, e1003218.

Duenas-Decamp, M.J., Peters, P., Burton, D., and Clapham, P.R. (2008). Natural resistance of human immunodeficiency virus type 1 to the CD4bs antibody b12 conferred by a glycan and an arginine residue close to the CD4 binding loop. J. Virol. *82*, 5807–5814.

Dunlop, D.C., Bonomelli, C., Mansab, F., Vasiljevic, S., Doores, K.J., Wormald, M.R., Palma, A.S., Feizi, T., Harvey, D.J., Dwek, R. a, et al. (2010). Polysaccharide mimicry of the epitope of the broadly neutralizing anti-HIV antibody, 2G12, induces enhanced antibody responses to self oligomannose glycans. Glycobiology *20*, 812–823.

Durocher, J.R., Payne, R.C., and Conrad, M.E. (1975). Role of sialic acid in erythrocyte survival. Blood *45*, 11–20.

Earl, P.L., Moss, B., and Doms, R.W. (1991). Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. J. Virol. *65*, 2047–2055.

Ebrahimi, K.H., West, G.M., and Flefil, R. (2014). Mass spectrometry approach and ELISA reveal the effect of codon optimization on N-linked glycosylation of HIV-1 gp120. J. Proteome Res. *dx.doi.org*.

Eggink, D., Melchers, M., and Sanders, R.W. (2007). Antibodies to HIV-1: aiming at the right target. Trends Microbiol. *15*, 291–294.

Elbein, A.D., Tropea, J.E., Mitchell, M., and Tropea, E. (1990). Kifunensine, a Potent Inhibitor of the glycoprotein processing mannosidase I. J. Biol. Chem. *265*, 15599–15605.

Ernardin, F., Kong, D., Peddada, L., Baxter-lowe, A., Delwart, E., Bernardin, F., and Baxter-lowe, L.A. (2005). Human Immunodeficiency Virus Mutations during the First Month of Infection Are Preferentially Found in Known Cytotoxic T-Lymphocyte Epitopes. J. Virol. *79*, 11523–11528.

Esser, M.T., Mori, T., Mondor, I., Sattentau, Q.J., Dey, B., Berger, E.A., Boyd, M.R., and Lifson, J.D. (1999). Cyanovirin-N binds to gp120 to interfere with CD4-dependent human immunodeficiency virus type 1 virion binding, fusion, and infectivity but does not affect the CD4 binding site on gp120 or soluble CD4-induced conformational changes in gp120. J. Virol. *73*, 4360–4371.

Falkowska, E., Le, K.M., Ramos, A., Doores, K.J., Lee, J.H., Blattner, C., Ramirez, A., Derking, R., van Gils, M.J., Liang, C.-H., et al. (2014). Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. Immunity *40*, 657–668.

Feinberg, H., Mitchell, D. a, Drickamer, K., and Weis, W.I. (2001). Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR. Science (80-.). *294*, 2163–2166.

Fenouillet, E., Jones, I., Powell, B., Schmitt, D., Kieny, M.P., and Gluckman, J.C. (1993). Functional role of the glycan cluster of the human immunodeficiency virus type 1 transmembrane glycoprotein (gp41) ectodomain. J. Virol. *67*, 150–160.

Férir, G., Huskens, D., Noppen, S., Koharudin, L.M.I., Gronenborn, A.M., and Schols, D. (2014). Broad anti-HIV activity of the Oscillatoria agardhii agglutinin homologue lectin family. J. Antimicrob. Chemother. *69*, 2746–2758.

Figdor, C.G., Kooyk, Y. Van, and Adema, G.J. (2002). C-type lectin receptors on dendritic cells and langerhans cells. Nat. Rev. Immunol. *2*, 77–84.

Finzi, D. (1997). Identification of a Reservoir for HIV-1 in Patients on Highly Active Antiretroviral Therapy. Science *278*, 1295–1300.

Fischer, P.B., Karlsson, G.B., Butters, T.D., Dwek, R. a, and Platt, F.M. (1996). N-butyldeoxynojirimycin-mediated inhibition of human immunodeficiency virus entry correlates with changes in antibody recognition of the V1/V2 region of gp120. J. Virol. *70*, 7143–7152.

Fischl, M.A., Richman, D.D., Grieco, M.H., Gottlieb, M.S., Volverding, P.A., Laskin, O.L., Leedom, J.M., Groopman, J.E., Mildvan, D., Schooley, R.T., et al. (1987). The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex. N. Engl. J. Med. *317*, 185–191.

Fleet, G.W., Karpas, a, Dwek, R. a, Fellows, L.E., Tyms, a S., Petursson, S., Namgoong, S.K., Ramsden, N.G., Smith, P.W., and Son, J.C. (1988). Inhibition of HIV replication by amino-sugar derivatives. FEBS Lett. *237*, 128–132.

Flynn, N.M., Forthal, D.N., Harro, C.D., Judson, F.N., Mayer, K.H., and Para, M.F. (2005). Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. J. Infect. Dis. *191*, 654–665.

François, K.O., and Balzarini, J. (2011). The highly conserved glycan at asparagine 260 of HIV-1 gp120 is indispensable for viral entry. J. Biol. Chem. *286*, 42900–42910.

Frankel, A.D., and Young, J.A. (1998). HIV-1: fifteen proteins and an RNA. Annu. Rev. Biochem. *67*, 1–25.

Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B., White, G., et al. (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science *224*, 500–503.

Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Rodenburg, C.M., Michael, S.F., Cummins, L.B., Arthur, L.O., Peeters, M., Shaw, G.M., et al. (1999). Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature *397*, 436–441.

Garces, F., Sok, D., Kong, L., McBride, R., Kim, H.J., Saye-Francisco, K.F., Julien, J.-P., Hua, Y., Cupo, A., Moore, J.P., et al. (2014). Structural evolution of glycan recognition by a family of potent HIV antibodies. Cell *159*, 69–79.

Geijtenbeek, T.B., Kwon, D.S., Torensma, R., van Vliet, S.J., van Duijnhoven, G.C., Middel, J., Cornelissen, I.L., Nottet, H.S., KewalRamani, V.N., Littman, D.R., et al. (2000a). DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell *100*, 587–597.

Geijtenbeek, T.B., Torensma, R., van Vliet, S.J., van Duijnhoven, G.C., Adema, G.J., van Kooyk, Y., and Figdor, C.G. (2000b). Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. Cell *100*, 575–585.

Go, E.P., Irungu, J., Zhang, Y., Dalpathado, D.S., Liao, H., Sutherland, L.L., Alam, S.M., Haynes, B.F., and Desaire, H. (2008). 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. *7*, 1660–1674.

Go, E.P., Chang, Q., Liao, H., Sutherland, L.L., Alam, S.M., Haynes, B.F., and Desaire, H. (2009). Glycosylation site-specific analysis of clade C HIV-1 envelope proteins. J. Proteome Res. *8*, 4231–4242.

Go, E.P., Hewawasam, G., Liao, H.-X., Chen, H., Ping, L.-H., Anderson, J. a, Hua, D.C., Haynes, B.F., and Desaire, H. (2011). Characterization of glycosylation profiles of HIV-1 transmitted/founder envelopes by mass spectrometry. J. Virol. *85*, 8270–8284.

Go, E.P., Liao, H.-X., Alam, S.M., Hua, D., Haynes, B.F., and Desaire, H. (2013). Characterization of host-cell line specific glycosylation profiles of early transmitted/founder HIV-1 gp120 envelope proteins. J. Proteome Res. *12*, 1223–1234.

Go, E.P., Hua, D., and Desaire, H. (2014). Glycosylation and disulfide bond analysis of transiently and stably expressed clade C HIV-1 gp140 trimers in 293T cells identify disulfide heterogeneity present in both proteins and differences in O-linked glycosylation. J. Proteome Res.

Gorny, M.K., Pan, R., Williams, C., Wang, X.-H., Volsky, B., O'Neal, T., Spurrier, B., Sampson, J.M., Li, L., Seaman, M.S., et al. (2012). Functional and immunochemical cross-reactivity of V2-specific monoclonal antibodies from HIV-1-infected individuals. Virology *427*, 198–207.

Gougeon, M.-L. (2003). Apoptosis as an HIV strategy to escape immune attack. Nat. Rev. Immunol. *3*, 392–404.

Gray, E.S., Moore, P.L., Pantophlet, R. a, and Morris, L. (2007). N-linked glycan modifications in gp120 of human immunodeficiency virus type 1 subtype C render partial sensitivity to 2G12 antibody neutralization. J. Virol. *81*, 10769–10776.

Gray, G., Buchbinder, S., and Duerr, A. (2010). Overview of STEP and Phambili trial results: two phase IIb test-of-concept studies investigating the efficacy of MRK adenovirus type 5 gag/pol/nef subtype B HIV vaccine. Curr. Opin. HIV AIDS *5*, 357–361.

Gray, G.E., Allen, M., Moodie, Z., Churchyard, G., Bekker, L.-G., Nchabeleng, M., Mlisana, K., Metch, B., de Bruyn, G., Latka, M.H., et al. (2011). 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. Lancet Infect. Dis. *11*, 507–515.

Grove, J., and Marsh, M. (2011). The cell biology of receptor-mediated virus entry. J. Cell Biol. *195*, 1071–1082.

Gruters, R.A., Neefjes, J.J., Tersmette, M., de Goede, R.E.Y., Tulp, A., Huisman, H.G., Miedema, F., and Ploegh, H.L. (1987). Interference with HIV-indued syncytium formation and viral infectivity by inhibitors of trimming glucosidases. Nature *330*, 74–77.

Guo, Y., Feinberg, H., Conroy, E., Mitchell, D. a, Alvarez, R., Blixt, O., Taylor, M.E., Weis, W.I., and Drickamer, K. (2004). Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. Nat. Struct. Mol. Biol. *11*, 591–598.

Guttman, M., Garcia, N.K., Cupo, A., Matsui, T., Julien, J.-P., Sanders, R.W., Wilson, I. a, Moore, J.P., and Lee, K.K. (2014). CD4-induced activation in a soluble HIV-1 Env trimer. Structure *22*, 974–984.

Hahn, B.H., Shaw, G.M., De Cock, K.M., and Sharp, P.M. (2000). AIDS as a zoonosis: scientific and public health implications. Science *287*, 607–614.

Hallenberger, S., Bosch, V., Angliker, H., Shaw, E., Klenk, H.-D., and Garten, W. (1992). Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160. Nature *360*, 358–361.

Hammer, S.M., Sobieszczyk, M.E., Janes, H., Karuna, S.T., Mulligan, M.J., Grove, D., Koblin, B. a, Buchbinder, S.P., Keefer, M.C., Tomaras, G.D., et al. (2013). Efficacy trial of a DNA/rAd5 HIV-1 preventive vaccine. N. Engl. J. Med. *369*, 2083–2092.

Hansen, J.E., Nielsen, C., Arendrup, M., Olofsson, S., Mathiesen, L., Nielsen, J.O., and Clausen, H. (1991). Broadly neutralizing antibodies targeted to mucin-type carbohydrate epitopes of human immunodeficiency virus. J. Virol. *65*, 6461–6467.

Hansen, J.E., Lund, O., Tolstrup, N., Gooley, a a, Williams, K.L., and Brunak, S. (1998). NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. Glycoconj. J. *15*, 115–130.

Hansen, J.E.S., Clausen, H., Hu, S.L., Nielsen, J.O., and Olofsson, S. (1992). An O-linked carbohydrate neutralization epitope of HIV-1 gp120 is expressed by HIV-1 env gene recombinant vaccinia virus. Arch. Virol. *126*, 11–20.

Hansen, J.S., Clausen, H., Nielsen, C., Teglbjaerg, L.S., Hansen, C., Nielson, M., Dabelsteen, E., Mathiesen, L., Hakomori, S., and Nielsen, J. (1990). Inhibition of human immunodeficiency virus (HIV) infection in vitro by anticarbohydrate monoclonal antibodies: peripheral glycosylation of HIV envelope glycoprotein gp120 may be a target for virus neutralization. J. Virol. *64*, 2833–2840.

Hart, M.L., Saifuddin, M., and Spear, G.T. (2003). Glycosylation inhibitors and neuraminidase enhance human immunodeficiency virus type 1 binding and neutralization by mannose-binding lectin. J. Gen. Virol. *84*, 353–360.

Harvey, D.J., Merry, A.H., Royle, L., Campbell, M.P., Dwek, R. a, and Rudd, P.M. (2009). Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. Proteomics *9*, 3796–3801.

Harvey, D.J., Sobott, F., Crispin, M., Wrobel, A., Bonomelli, C., Vasiljevic, S., Scanlan, C.N., Scarff, C. a, Thalassinos, K., and Scrivens, J.H. (2011). Ion mobility mass spectrometry for extracting spectra of N-glycans directly from incubation mixtures following glycan release: application to glycans from engineered glycoforms of intact, folded HIV gp120. J. Am. Soc. Mass Spectrom. *22*, 568–581.

Haynes, B.F., Gilbert, P.B., McElrath, M.J., Zolla-Pazner, S., Tomaras, G.D., Alam, S.M., Evans, D.T., Montefiori, D.C., Karnasuta, C., Sutthent, R., et al. (2012). Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N. Engl. J. Med. *366*, 1275–1286.

Helenius, A., and Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. Annu. Rev. Biochem. *73*, 1019–1049.

Herrera, C., Klasse, P.J., Michael, E., Kake, S., Barnes, K., Kibler, C.W., Campbell-Gardener, L., Si, Z., Sodroski, J., Moore, J.P., et al. (2005). The impact of envelope glycoprotein cleavage on the antigenicity, infectivity, and neutralization sensitivity of Env-pseudotyped human immunodeficiency virus type 1 particles. Virology *338*, 154–172.

Hessell, A.J., Rakasz, E.G., Poignard, P., Hangartner, L., Landucci, G., Forthal, D.N., Koff, W.C., Watkins, D.I., and Burton, D.R. (2009). Broadly neutralizing human anti-HIV antibody 2G12 is effective in protection against mucosal SHIV challenge even at low serum neutralizing titers. PLoS Pathog. *5*, e1000433.

Hijazi, K., Wang, Y., Scala, C., Jeffs, S., Longstaff, C., Stieh, D., Haggarty, B., Vanham, G., Schols, D., Balzarini, J., et al. (2011). DC-SIGN increases the affinity of HIV-1 envelope glycoprotein interaction with CD4. PLoS One *6*, e28307.

Hladik, F., and McElrath, M.J. (2008). Setting the stage: host invasion by HIV. Nat. Rev. Immunol. *8*, 447–457.

Hong, P.W., Flummerfelt, K.B., de Parseval, A., Gurney, K., Elder, J.H., Lee, B., and Parseval, A. De (2002). Human immunodeficiency virus envelope (gp120) binding to DC-SIGN and primary dendritic cells is carbohydrate dependent but does not involve 2G12 or cyanovirin binding sites: implications for structural analyses of gp120-DC-SIGN binding. J. Virol. *76*, 12855–12865.

Hoorelbeke, B., Huskens, D., Férir, G., François, K.O., Takahashi, A., Van Laethem, K., Schols, D., Tanaka, H., and Balzarini, J. (2010). Actinohivin, a broadly neutralizing prokaryotic lectin, inhibits HIV-1 infection by specifically targeting high-mannose-type glycans on the gp120 envelope. Antimicrob. Agents Chemother. *54*, 3287–3301.

Hoxie, J.A. (2010). Toward an antibody-based HIV-1 vaccine. Annu. Rev. Med. *61*, 135–152.

Hu, H., Shioda, T., Moriya, C., Xin, X., Hasan, M.K., Miyake, K., Shimada, T., and Nagai, Y. (1996). Infectivities of human and other primate lentiviruses are activated by desialylation of the virion surface. J. Virol. *70*, 7462–7470.

Hu, Q., Mahmood, N., and Shattock, R.J. (2007). High-mannose-specific deglycosylation of HIV-1 gp120 induced by resistance to cyanovirin-N and the impact on antibody neutralization. Virology *368*, 145–154.

Huang, J., Kang, B.H., Pancera, M., Lee, J.H., Tong, T., Feng, Y., Georgiev, I.S., Chuang, G.-Y., Druz, A., Doria-Rose, N. a., et al. (2014). Broad and potent HIV-1 neutralization by a human antibody that binds the gp41–gp120 interface. Nature *515*, 138–142.

Huang, P.-S., Ban, Y.-E.A., Richter, F., Andre, I., Vernon, R., Schief, W.R., and Baker, D. (2011a). RosettaRemodel: a generalized framework for flexible backbone protein design. PLoS One *6*, e24109.

Huang, X., Jin, W., Griffin, G.E., Shattock, R.J., and Hu, Q. (2011b). Removal of two high-mannose N-linked glycans on gp120 renders human immunodeficiency virus 1 largely resistant to the carbohydrate-binding agent griffithsin. J. Gen. Virol. *92*, 2367–2373.

Huber, M., Le, K.M., Doores, K.J., Fulton, Z., Stanfield, R.L., Wilson, I. a, and Burton, D.R. (2010). Very few substitutions in a germ line antibody are required to initiate significant domain exchange. J. Virol. *84*, 10700–10707.

Hwang, S., Boyle, T., Lyerly, H., and Cullen, B. (1991). Identification of the envelope V3 loop as the primary determinant of cell tropism in HIV-1. Science *253*, 71–74.

Ingrosso, D., Fowler, A.V., Bleibaum, J., and Clarke, C. (1989). Specificity of endoproteinase Asp-N (pseudomonas fragi) - cleavage at glutamyl residues in two proteins. Biochem. Biophys. Res. Commun. *162*, 1528–1534.

Iyer, S.P.N., Franti, M., Krauchuk, A. a, Fisch, D.N., Ouattara, A. a, Roux, K.H., Krawiec, L., Dey, A.K., Beddows, S., Maddon, P.J., et al. (2007). Purified, proteolytically mature HIV type 1 SOSIP gp140 envelope trimers. AIDS Res. Hum. Retroviruses *23*, 817–828.

Jardine, J., Julien, J.-P., Menis, S., Ota, T., Kalyuzhniy, O., McGuire, A., Sok, D., Huang, P.-S., MacPherson, S., Jones, M., et al. (2013). Rational HIV immunogen design to target specific germline B cell receptors. Science *340*, 711–716.

Johnson, V. a, Brun-Vézinet, F., Clotet, B., Günthard, H.F., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., and Richman, D.D. (2010). Update of the drug resistance mutations in HIV-1: December 2010. Top. HIV Med. *18*, 156–163.

Johnson, W.E., Sauvron, J.M., and Desrosiers, R.C. (2001). Conserved, N-linked carbohydrates of human immunodeficiency virus type 1 gp41 are largely dispensable for viral replication. J. Virol. *75*, 11426–11436.

Julien, J.-P., Cupo, A., Sok, D., Stanfield, R.L., Lyumkis, D., Deller, M.C., Klasse, P.-J., Burton, D.R., Sanders, R.W., Moore, J.P., et al. (2013a). Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science *342*, 1477–1483.

Julien, J.-P., Sok, D., Khayat, R., Lee, J.H., Doores, K.J., Walker, L.M., Ramos, A., Diwanji, D.C., Pejchal, R., Cupo, A., et al. (2013b). Broadly neutralizing antibody PGT121 allosterically modulates CD4 binding via recognition of the HIV-1 gp120 V3 base and multiple surrounding glycans. PLoS Pathog. *9*, e1003342.

Julien, J.-P., Lee, J.H., Cupo, A., Murin, C.D., Derking, R., Hoffenberg, S., Caulfield, M.J., King, C.R., Marozsan, A.J., Klasse, P.J., et al. (2013c). Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. Proc. Natl. Acad. Sci. U. S. A. *110*, 4351–4356.

Karlsson, G.B., Butters, T.D., Dwek, R. a, and Platt, F.M. (1993). Effects of the imino sugar N-butyldeoxynojirimycin on the N-glycosylation of recombinant gp120. J. Biol. Chem. *268*, 570–576.

Karpas, A., Fleet, G.W.J., Dwek, R.A., Petursson, S., Namgoong, S.K., Ramsden, N.G., Jacob, G.S., and Rademacher, T.W. (1988). Aminosugar derivatives as potential anti-human immunodeficiency virus agents. Proc. Natl. Acad. Sci. *85*, 9229–9233.

Keeffe, J.R., Gnanapragasam, P.N.P., Gillespie, S.K., Yong, J., Bjorkman, P.J., and Mayo, S.L. (2011). Designed oligomers of cyanovirin-N show enhanced HIV neutralization. Proc. Natl. Acad. Sci. U. S. A. *108*, 14079–14084.

Keele, B.F., Giorgi, E.E., Salazar-Gonzalez, J.F., Decker, J.M., Pham, K.T., Salazar, M.G., Sun, C., Grayson, T., Wang, S., Li, H., et al. (2008). Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc. Natl. Acad. Sci. U. S. A. *105*, 7552–7557.

Kelley, B.S., Chang, L.C., and Bewley, C. a (2002). Engineering an obligate domain-swapped dimer of cyanovirin-N with enhanced anti-HIV activity. J. Am. Chem. Soc. *124*, 3210–3211.

Klasse, P.J., Depetris, R.S., Pejchal, R., Julien, J.-P., Khayat, R., Lee, J.H., Marozsan, A.J., Cupo, A., Cocco, N., Korzun, J., et al. (2013). Influences on trimerization and aggregation of soluble, cleaved HIV-1 SOSIP envelope glycoprotein. J. Virol. *87*, 9873–9885.

Klaver, B., and Berkhout, B. (1994). Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. J. Virol. *68*, 3830–3840.

Klein, J.S., and Bjorkman, P.J. (2010). Few and far between: how HIV may be evading antibody avidity. PLoS Pathog. *6*, e1000908.

Klein, F., Halper-Stromberg, A., Horwitz, J. a, Gruell, H., Scheid, J.F., Bournazos, S., Mouquet, H., Spatz, L. a, Diskin, R., Abadir, A., et al. (2012). HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. Nature *492*, 118–122.

Klein, F., Diskin, R., Scheid, J.F., Gaebler, C., Mouquet, H., Georgiev, I.S., Pancera, M., Zhou, T., Incesu, R.-B., Fu, B.Z., et al. (2013). Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. Cell *153*, 126–138.

Klein, J.S., Webster, A., Gnanapragasam, P.N.P., Rachel, P., and Bjorkman, P.J. (2010). A dimeric form of the HIV-1 antibody 2G12 elicits potent antibody-dependent cellular cytotoxicity. AIDS *24*, 1633–1640.

Koch, M., Pancera, M., Kwong, P.D., Kolchinsky, P., Grundner, C., Wang, L., Hendrickson, W. a, Sodroski, J., and Wyatt, R. (2003). Structure-based, targeted deglycosylation of HIV-1 gp120 and effects on neutralization sensitivity and antibody recognition. Virology *313*, 387–400.

Koharudin, L.M., and Gronenborn, A.M. (2014). Antiviral lectins as potential HIV microbicides. Curr. Opin. Virol. *7C*, 95–100.

Kolchinsky, P., Kiprilov, E., Bartley, P., Rubinstein, R., and Sodroski, J. (2001). Loss of a single N-linked glycan allows CD4-independent human immunodeficiency virus type 1 infection by altering the position of the gp120 V1/V2 variable loops. J. Virol. *75*, 3435–3443.

Kong, L., Sheppard, N.C., Stewart-Jones, G.B.E., Robson, C.L., Chen, H., Xu, X., Krashias, G., Bonomelli, C., Scanlan, C.N., Kwong, P.D., et al. (2010). Expression-system-dependent modulation of HIV-1 envelope glycoprotein antigenicity and immunogenicity. J. Mol. Biol. *403*, 131–147.

Kong, L., Lee, J.H., Doores, K.J., Murin, C.D., Julien, J.-P., McBride, R., Liu, Y., Marozsan, A., Cupo, A., Klasse, P.-J., et al. (2013a). Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. Nat. Struct. Mol. Biol. *20*, 796–803.

Kong, L., Lee, J.H., Doores, K.J., Murin, C.D., Julien, J., Mcbride, R., Liu, Y., Marozsan, A., Cupo, A., Klasse, P., et al. (2013b). Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. Nat. Struct. Mol. Biol. Mol. Biol. *20*, 796–803.

Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C., and Detours, V. (2001). Evolutionary and immunological implications of contemporary HIV-1 variation. Br. Med. Bull. *58*, 19–42.

Kornfeld, R., and Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. Annu. Rev. Biochem. *54*, 631–664.

Kovacs, J.M., Nkolola, J.P., Peng, H., Cheung, A., Perry, J., Miller, C. a, Seaman, M.S., Barouch, D.H., and Chen, B. (2012). HIV-1 envelope trimer elicits more potent neutralizing antibody responses than monomeric gp120. Proc. Natl. Acad. Sci. U. S. A. *109*, 12111–12116.

Kwon, D.S., Gregorio, G., Bitton, N., Hendrickson, W. a, and Littman, D.R. (2002). DC-SIGN-mediated internalization of HIV is required for trans-enhancement of T cell infection. Immunity *16*, 135–144.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., and Hendrickson, W. a (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature *393*, 648–659.

Kwong, P.D., Doyle, M.L., Casper, D.J., Cicala, C., Leavitt, S. a, Majeed, S., Steenbeke, T.D., Venturi, M., Chaiken, I., Fung, M., et al. (2002). HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. Nature *420*, 678–682.

Land, A., Zonneveld, D., and Braakman, I. (2003). Folding of HIV-1 envelope glycoprotein involves extensive isomerization of disulfide bonds and conformation-dependent leader peptide cleavage. FASEB J. *17*, 1058–1067.

Lasky, L., Groopman, J., Fennie, C., Benz, P., Capon, D., Dowbenko, D., Nakamura, G., Nunes, W., Renz, M., and Berman, P. (1986). Neutralization of the AIDS retrovirus by antibodies to a recombinant envelope glycoprotein. Science *233*, 209–212.

Lavine, C.L., Lao, S., Montefiori, D.C., Haynes, B.F., Sodroski, J.G., and Yang, X. (2012). High-mannose glycan-dependent epitopes are frequently targeted in broad neutralizing antibody responses during human immunodeficiency virus type 1 infection. J. Virol. *86*, 2153–2164.

Lawn, S.D., and Butera, S.T. (2001). Contribution of immune activation to the pathogenesis and transmission of Human Immunodeficiency Virus type 1 infection. Clin. Microbiol. Rev. *14*, 753–777.

Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N., and Gregory, T.J. (1990). 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. *265*, 10373–10382.

Li, H., Chien, P.C., Tuen, M., Visciano, M.L., Cohen, S., Blais, S., Xu, C.-F., Zhang, H.-T., and Hioe, C.E. (2008). Identification of an N-linked glycosylation in the C4 region of HIV-1 envelope gp120 that is critical for recognition of neighboring CD4 T cell epitopes. J. Immunol. *180*, 4011–4021.

Li, M., Salazar-Gonzalez, J.F., Derdeyn, C. a, Morris, L., Williamson, C., Robinson, J.E., Decker, J.M., Li, Y., Salazar, M.G., Polonis, V.R., et al. (2006). Genetic and neutralization properties of subtype C human immunodeficiency virus type 1 molecular env clones from acute and early heterosexually acquired infections in Southern Africa. J. Virol. *80*, 11776–11790.

Li, Y., Luo, L., Rasool, N., and Kang, C.Y. (1993). Glycosylation is necessary for the correct folding of human immunodeficiency virus gp120 in CD4 binding. J. Virol. *67*, 584–588.

Li, Y., Bergeron, J.J., Luo, L., Ou, W.J., Thomas, D.Y., and Kang, C.Y. (1996). 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. *93*, 9606–9611.

Lindley, H. (1956). A new synthetic substrate for trypsin and its application to the determination of the amino-acid sequence of proteins. Nature *178*, 647–648.

Liu, J., Bartesaghi, A., Borgnia, M.J., Sapiro, G., and Subramaniam, S. (2008a). Molecular architecture of native HIV-1 gp120 trimers. Nature *455*, 109–113.

Liu, W., Sohn, H.W., Tolar, P., and Pierce, S.K. (2010). It's all about change: the antigen-driven initiation of B-cell receptor signaling. Cold Spring Harb. Perspect. Biol. *2*, a002295.

Liu, Y., Curlin, M.E., Diem, K., Zhao, H., Ghosh, A.K., Zhu, H., Woodward, A.S., Maenza, J., Stevens, C.E., Stekler, J., et al. (2008b). Env length and N-linked glycosylation following transmission of human immunodeficiency virus Type 1 subtype B viruses. Virology *374*, 229–233.

Luallen, R.J., Lin, J., Fu, H., Cai, K.K., Agrawal, C., Mboudjeka, I., Lee, F.-H., Montefiori, D., Smith, D.F., Doms, R.W., et al. (2008). An engineered Saccharomyces cerevisiae strain binds the broadly neutralizing human immunodeficiency virus type 1 antibody 2G12 and elicits mannose-specific gp120-binding antibodies. J. Virol. *82*, 6447–6457.

Lyumkis, D., Julien, J.-P., de Val, N., Cupo, A., Potter, C.S., Klasse, P.-J., Burton, D.R., Sanders, R.W., Moore, J.P., Carragher, B., et al. (2013). Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. Science *342*, 1484–1490.

Mansky, L.M., and Temin, H.M. (1995). Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse These include : Lower In Vivo Mutation Rate of Human Immunodeficiency Virus Type 1 than That Predicted from the Fidelity of Purifi. J. Virol. *69*, 5087–5094.

Mao, Y., Wang, L., Gu, C., Herschhorn, A., Xiang, S.-H., Haim, H., Yang, X., and Sodroski, J. (2012). Subunit organization of the membrane-bound HIV-1 envelope glycoprotein trimer. Nat. Struct. Mol. Biol. *19*, 893–899.

Mao, Y., Wang, L., Gu, C., Herschhorn, A., Désormeaux, A., Finzi, A., Xiang, S.-H., and Sodroski, J.G. (2013). Molecular architecture of the uncleaved HIV-1 envelope glycoprotein trimer. Proc. Natl. Acad. Sci. U. S. A. *110*, 12438–12443.

Margolis, L., and Shattock, R. (2006). Selective transmission of CCR5-utilizing HIV-1: the "gatekeeper" problem resolved? Nat. Rev. Microbiol. *4*, 312–317.

Mariño, K., Bones, J., Kattla, J.J., and Rudd, P.M. (2010). A systematic approach to protein glycosylation analysis: a path through the maze. Nat. Chem. Biol. *6*, 713–723.

Mascola, J.R., Stiegler, G., VanCott, T.C., Katinger, H., Carpenter, C.B., Hanson, C.E., Beary, H., Hayes, D., Frankel, S.S., Birx, D.L., et al. (2000). Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. Nat. Med. *6*, 207–210.

McCune, J.M., Rabin, L.B., Feinberg, M.B., Lieberman, M., Kosek, J.C., Reyes, G.R., and Weissman, I.L. (1988). Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. Cell *53*, 55–67.

McElrath, M.J., and Haynes, B.F. (2010). Induction of immunity to human immunodeficiency virus type-1 by vaccination. Immunity *33*, 542–554.

McLellan, J.S., Pancera, M., Carrico, C., Gorman, J., Julien, J.-P., Khayat, R., Louder, R., Pejchal, R., Sastry, M., Dai, K., et al. (2011). Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature *480*, 336–343.

McMichael, A., Picker, L.J., Moore, J.P., and Burton, D.R. (2013). Another HIV vaccine failure: where to next? Nat. Med. *19*, 1576–1577.

McMichael, A.J., Borrow, P., Tomaras, G.D., Goonetilleke, N., and Haynes, B.F. (2010). The immune response during acute HIV-1 infection: clues for vaccine development. Nat. Rev. Immunol. *10*, 11–23.

Mehandru, S., Vcelar, B., Wrin, T., Stiegler, G., Joos, B., Mohri, H., Boden, D., Galovich, J., Tenner-Racz, K., Racz, P., et al. (2007). Adjunctive passive immunotherapy in human immunodeficiency virus type 1-infected individuals treated with antiviral therapy during acute and early infection. J. Virol. *81*, 11016–11031.

Miller, J.L., Lachica, R., Sayce, A.C., Williams, J.P., Bapat, M., Dwek, R., Beatty, P.R., Harris, E., and Zitzmann, N. (2012). Liposome-mediated delivery of iminosugars enhances efficacy against dengue virus in vivo. Antimicrob. Agents Chemother. *56*, 6379–6386.

Mitchell, D. a, Fadden, a J., and Drickamer, K. (2001). A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands. J. Biol. Chem. *276*, 28939–28945.

Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V., and Melikyan, G.B. (2009). HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. Cell *137*, 433–444.

Moldt, B., Rakasz, E.G., Schultz, N., Chan-Hui, P.-Y., Swiderek, K., Weisgrau, K.L., Piaskowski, S.M., Bergman, Z., Watkins, D.I., Poignard, P., et al. (2012). Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. Proc. Natl. Acad. Sci. U. S. A. *109*, 18921–18925.

Montefiori, D.C. (2005). Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. Curr. Protoc. Immunol. *Chapter 12*, Unit 12.11.

Van Montfort, T., Eggink, D., Boot, M., Tuen, M., Hioe, C.E., Berkhout, B., and Sanders, R.W. (2011). HIV-1 N-glycan composition governs a balance between dendritic cell-mediated viral transmission and antigen presentation. J. Immunol. *187*, 4676–4685.

Moore, P.L., Crooks, E.T., Porter, L., Zhu, P., Cayanan, C.S., Grise, H., Corcoran, P., Zwick, M.B., Franti, M., Morris, L., et al. (2006). Nature of nonfunctional envelope proteins on the surface of Human Immunodeficiency Virus Type 1. J. Virol. *80*, 2515–2528.

Moore, P.L., Gray, E.S., and Morris, L. (2010). Specificity of the autologous neutralizing antibody repsonse. Curr. Opin. HIV AIDS *4*, 358–363.

Moore, P.L., Gray, E.S., Wibmer, C.K., Bhiman, J.N., Nonyane, M., Sheward, D.J., Hermanus, T., Bajimaya, S., Tumba, N.L., Abrahams, M.-R., et al. (2012). Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. Nat. Med. *18*, 1688–1692.

Morell, A.G., Gregoriadis, G., Scheinberg, H., Hickman, J., and Ashwell, G. (1971). The role of sialic acid in determining the survival of glycoproteins in the circulation. J. Biol. Chem. *246*, 1461–1467.

Mori, T., and Boyd, M.R. (2001). Cyanovirin-N , a potent human immunodeficiency virus-inactivating protein, blocks both CD4-dependent and CD4-independent binding of soluble gp120 (sgp120) to target cells, inhibits sCD4-induced binding of sgp120 to cell-associated CXCR4, and dissociates b. Antimicrob. Agents Chemother. *45*, 664–672.

Mori, T., O'Keefe, B.R., Sowder, R.C., Bringans, S., Gardella, R., Berg, S., Cochran, P., Turpin, J. a, Buckheit, R.W., McMahon, J.B., et al. (2005). Isolation and characterization of griffithsin, a novel HIV-inactivating protein, from the red alga Griffithsia sp. J. Biol. Chem. *280*, 9345–9353.

Mouquet, H. (2014). Antibody B cell responses in HIV-1 infection. Trends Immunol. *35*, 549–561.

Mouquet, H., Scharf, L., Euler, Z., Liu, Y., Eden, C., Scheid, J.F., Halper-Stromberg, A., Gnanapragasam, P.N.P., Spencer, D.I.R., Seaman, M.S., et al. (2012). Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. Proc. Natl. Acad. Sci. U. S. A. *109*, E3268–77.

Murin, C.D., Julien, J.-P., Sok, D., Stanfield, R.L., Khayat, R., Cupo, A., Moore, J.P., Burton, D.R., Wilson, I.A., and Ward, A.B. (2014). Structure of 2G12 Fab2 in complex with soluble and fully glycosylated HIV-1 Env by negative-stain single particle electron microscopy. J. Virol. *88*, 10177–10188.

Myszka, D.G., Sweet, R.W., Hensley, P., Brigham-Burke, M., Kwong, P.D., Hendrickson, W. a., Wyatt, R., Sodroski, J., and Doyle, M.L. (2000). Energetics of the HIV gp120-CD4 binding reaction. Proc. Natl. Acad. Sci. *97*, 9026–9031.

Nahmias, A.J., Weiss, J., Yao, X., Lee, F., Kodsi, R., Schanfield, M., Matthews, T., Bolognesi, D., Durack, D., Motulsky, A., et al. (1986). Evidence for human infection with an HTLV III/LAV-like virus in central Africa, 1959. Lancet 1279–1280.

Neil, S.J.D., Zang, T., and Bieniasz, P.D. (2008). Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature *451*, 425–430.

Ni, J., Song, H., Wang, Y., Stamatos, N.M., and Wang, L.-X. (2006). Toward a carbohydrate-based HIV-1 vaccine: synthesis and immunological studies of oligomannose-containing glycoconjugates. Bioconjug. Chem. *17*, 493–500.

O'Keefe, B.R., Vojdani, F., Buffa, V., Shattock, R.J., Montefiori, D.C., Bakke, J., Mirsalis, J., d'Andrea, A.-L., Hume, S.D., Bratcher, B., et al. (2009). Scaleable manufacture of HIV-1 entry inhibitor griffithsin and validation of its safety and efficacy as a topical microbicide component. Proc. Natl. Acad. Sci. U. S. A. *106*, 6099–6104.

Otteken, A., and Moss, B. (1996). Calreticulin interacts with newly synthesized human immunodeficiency virus type 1 envelope glycoprotein, suggesting a chaperone function similar to that of calnexin. J. Biol. Chem. *271*, 97–103.

Otteken, A., Earl, P.L., and Moss, B. (1996). Folding, assembly, and intracellular trafficking of the human immunodeficiency virus type 1 envelope glycoprotein analyzed with monoclonal antibodies recognizing maturational intermediates. J. Virol. *70*, 3407–3415.

Ouellet, M., Mercier, S., Pelletier, I., Bounou, S., Roy, J., Hirabayashi, J., Sato, S., and Tremblay, M.J. (2005). Galectin-1 acts as a soluble host factorthat promotes HIV-1 infectivity through stabilization of virus attachment to host cells. J. Immunol. *174*, 4120–4126.

Overbaugh, J., and Morris, L. (2012). The Antibody Response against HIV-1. Cold Spring Harb. Perspect. Med. *2*, a007039.

Pabst, M., Chang, M., Stadlmann, J., and Altmann, F. (2012). Glycan profiles of the 27 N-glycosylation sites of the HIV envelope protein CN54gp140. Biol. Chem. *393*, 719–730.

Palmisano, G., Melo-Braga, M.N., Engholm-Keller, K., Parker, B.L., and Larsen, M.R. (2012). Chemical deamidation: a common pitfall in large-scale N-linked glycoproteomic mass spectrometry-based analyses. J. Proteome Res. *11*, 1949–1957.

Pancera, M., McLellan, J.S., Wu, X., Zhu, J., Changela, A., Schmidt, S.D., Yang, Y., Zhou, T., Phogat, S., Mascola, J.R., et al. (2010a). Crystal structure of PG16 and chimeric dissection with somatically related PG9: structure-function analysis of two quaternary-specific antibodies that effectively neutralize HIV-1. J. Virol. *84*, 8098–8110.

Pancera, M., Majeed, S., Ban, Y.-E.A., Chen, L., Huang, C., Kong, L., Kwon, Y. Do, Stuckey, J., Zhou, T., Robinson, J.E., et al. (2010b). 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. *107*, 1166–1171.

Pancera, M., Shahzad-Ul-Hussan, S., Doria-Rose, N. a, McLellan, J.S., Bailer, R.T., Dai, K., Loesgen, S., Louder, M.K., Staupe, R.P., Yang, Y., et al. (2013). Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. Nat. Struct. Mol. Biol. *20*, 804–813.

Pancera, M., Zhou, T., Druz, A., Georgiev, I.S., Soto, C., Gorman, J., Huang, J., Acharya, P., Chuang, G.-Y., Ofek, G., et al. (2014). Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature *514*, 455–461.

Papandréou, M.-J., Barbouche, R., Guieu, R., Kieny, M.P., and Fenouillet, E. (2002). The alpha-glucosidase inhibitor 1-deoxynojirimycin blocks human immunodeficiency virus envelope glycoprotein-mediated membrane fusion at the CXCR4 binding step. Mol. Pharmacol. *61*, 186–193.

Parks, C.L., Picker, L.J., and King, C.R. (2013). Development of replication-competent viral vectors for HIV vaccine delivery. Curr. Opin. HIV AIDS *8*, 402–411.

Parren, P.W.H.I., Marx, P.A., Hessell, A.J., Luckay, A., Harouse, J., Cheng-mayer, C., Moore, J.P., and Burton, D.R. (2001). Antibody protects macaques against vaginal challenge with a pathogenic R5 simian/human immunodeficiency virus at serum levels giving complete neutralization in vitro. J. Virol. *75*, 8340–8347.

Pejchal, R., Doores, K.J., Walker, L.M., Khayat, R., Huang, P.-S., Wang, S.-K., Stanfield, R.L., Julien, J.-P., Ramos, A., Crispin, M., et al. (2011). A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. Science *334*, 1097–1103.

Penn, M.L., Grivel, J.C., Schramm, B., Goldsmith, M. a, and Margolis, L. (1999). CXCR4 utilization is sufficient to trigger CD4+ T cell depletion in HIV-1-infected human lymphoid tissue. Proc. Natl. Acad. Sci. U. S. A. *96*, 663–668.

Perez, L.G., Chen, H., Liao, H.-X., and Montefiori, D.C. (2014). Envelope Glycoprotein Binding to the Integrin α4β7 Is Not a General Property of Most HIV-1 Strains. J. Virol. *88*, 10767–10777.

Picker, L.J., and Watkins, D.I. (2005). HIV pathogenesis: the first cut is the deepest. Nat. Immunol. *6*, 430–432.

Pikora, C., Wittish, C., and Desrosiers, R.C. (2005). Identification of two N-linked glycosylation sites within the core of the simian immunodeficiency virus glycoprotein whose removal enhances sensitivity to soluble CD4. J. Virol. *79*, 12575–12583.

Pitisuttithum, P., Gilbert, P., Gurwith, M., Heyward, W., Martin, M., Griensven, F. Van, Hu, D., Tappero, J.W., Choopanya, K., Vaccine, B., et al. (2006). Efficacy Trial of a Bivalent Recombinant Glycoprotein 120 HIV-1 Vaccine among Injection Drug Users in Bangkok , Thailand. J. Infect. Dis. *194*, 1671–71.

Plantier, J.-C., Leoz, M., Dickerson, J.E., De Oliveira, F., Cordonnier, F., Lemée, V., Damond, F., Robertson, D.L., and Simon, F. (2009). A new human immunodeficiency virus derived from gorillas. Nat. Med. *15*, 871–872.

Poles, M.A., Elliott, J., Taing, P., Anton, P.A., and Chen, I.S.Y. (2001). A preponderance of CCR5+ CXCR4+ mononuclear cells enhances gastrointestinal mucosal susceptibility to human immunodeficiency virus type 1 infection. J. Virol. *75*, 8390–8399.

Pollakis, G., Kang, S., Kliphuis, a, Chalaby, M.I., Goudsmit, J., and Paxton, W. a (2001). N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization. J. Biol. Chem. *276*, 13433–13441.

Pollock, S., Dwek, R. a, Burton, D.R., and Zitzmann, N. (2008). N-Butyldeoxynojirimycin is a broadly effective anti-HIV therapy significantly enhanced by targeted liposome delivery. AIDS *22*, 1961–1969.

Preston, B., Poiesz, B., and Loeb, L. (1988). Fidelity of HIV-1 reverse transcriptase. Science *242*, 1168–1171.

Raska, M., Takahashi, K., Czernekova, L., Zachova, K., Hall, S., Moldoveanu, Z., Elliott, M.C., Wilson, L., Brown, R., Jancova, D., et al. (2010). Glycosylation patterns of HIV-1 gp120 depend on the type of expressing cells and affect antibody recognition. J. Biol. Chem. *285*, 20860–20869.

Ratner, L. (1992). Glucosidase inhibitors for treatment of HIV-1 infection. AIDS Res. Hum. Retroviruses *8*, 165–173.

Reitter, J.N., Means, R.E., and Desrosiers, R.C. (1998). A role for carbohydrates in immune evasion in AIDS. Nat. Med. *4*, 679–684.

Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., Premsi, N., Namwat, C., Souza, M. De, Adams, E., et al. (2009). Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N. Engl. J. Med. *361*, 2209–2220.

Ringe, R.P., Sanders, R.W., Yasmeen, A., Kim, H.J., Hyun, J., Cupo, A., and Korzun, J. (2013). Cleavage strongly influences whether soluble HIV-1 envelope glycoprotein trimers adopt a native-like conformation. Proc. Natl. Acad. Sci. U. S. A. *110*, 18256–18261.

Roberts, J., Bebenek, K., and Kunkel, T. (1988). The accuracy of reverse transcriptase from HIV-1. Science *242*, 1171–1173.

Rudd, P.M., and Dwek, R.A. (1997). Glycosylation: Heterogeneity and the 3D Structure of Proteins. Crit. Rev. Biochem. Mol. Biol. *32*, 1–100.

Salazar-Gonzalez, J.F., Salazar, M.G., Keele, B.F., Learn, G.H., Giorgi, E.E., Li, H., Decker, J.M., Wang, S., Baalwa, J., Kraus, M.H., et al. (2009). Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. J. Exp. Med. *206*, 1273–1289.

Sanders, R.W., Vesanen, M., Schuelke, N., Master, A., Schiffner, L., Kalyanaraman, R., Paluch, M., Berkhout, B., Maddon, P.J., Olson, W.C., et al. (2002a). Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. J. Virol. *76*, 8875–8889.

Sanders, R.W., Venturi, M., Schiffner, L., Kalyanaraman, R., Katinger, H., Lloyd, K.O., Kwong, P.D., and Moore, J.P. (2002b). The mannose-dependent epitope for neutralizing antibody 2G12 on human immunodeficiency virus type 1 glycoprotein gp120. J. Virol. *76*, 7293–7305.

Sanders, R.W., van Anken, E., Nabatov, A. a, Liscaljet, I.M., Bontjer, I., Eggink, D., Melchers, M., Busser, E., Dankers, M.M., Groot, F., et al. (2008). The carbohydrate at asparagine 386 on HIV-1 gp120 is not essential for protein folding and function but is involved in immune evasion. Retrovirology *5*, 10.

Sanders, R.W., Derking, R., Cupo, A., Julien, J.-P., Yasmeen, A., de Val, N., Kim, H.J., Blattner, C., de la Peña, A.T., Korzun, J., et al. (2013). A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. PLoS Pathog. *9*, e1003618.

Sato, S., Ouellet, M., St-Pierre, C., and Tremblay, M.J. (2012). Glycans, galectins, and HIV-1 infection. Ann. N. Y. Acad. Sci. *1253*, 133–148.

Scanlan, C.N., Pantophlet, R., Wormald, M.R., Saphire, E.O., Stanfield, R., Wilson, I.A., Katinger, H., Dwek, R.A., Rudd, P.M., and Burton, D.R. (2002). The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of α1→2 mannose residues on the outer face of gp120. J. Virol. *76*, 7306–7321.

Scanlan, C.N., Offer, J., Zitzmann, N., and Dwek, R. a (2007). Exploiting the defensive sugars of HIV-1 for drug and vaccine design. Nature *446*, 1038–1045.

Scharf, L., Scheid, J.F., Lee, J.H., West, A.P., Chen, C., Gao, H., Gnanapragasam, P.N.P., Mares, R., Seaman, M.S., Ward, A.B., et al. (2014). Antibody 8ANC195 reveals a site of broad vulnerability on the HIV-1 envelope spike. Cell Rep. *7*, 785–795.

Scheid, J.F., Mouquet, H., Ueberheide, B., Diskin, R., Klein, F., Oliveira, T.Y.K., Pietzsch, J., Fenyo, D., Abadir, A., Velinzon, K., et al. (2011). Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. Science *333*, 1633–1637.

Schønning, K., Jansson, B., Olofsson, S., Nielsen, J.O., and Hansen, J.S. (1996). Resistance to V3-directed neutralization caused by an N-linked oligosaccharide depends on the quaternary structure of the HIV-1 envelope oligomer. Virology *218*, 134–140.

Schülke, N., Vesanen, M.S., Sanders, R.W., Zhu, P., Lu, M., Anselma, D.J., Villa, A.R., Parren, P.W.H.I., Binley, J.M., Roux, K.H., et al. (2002). Oligomeric and conformational properties of a proteolytically mature, immunodeficiency virus type 1 gp140 envelope glycoprotein. J. Virol. *76*, 7760–7776.

Shan, M., Klasse, P.J., Banerjee, K., Dey, A.K., Iyer, S.P.N., Dionisio, R., Charles, D., Campbell-Gardener, L., Olson, W.C., Sanders, R.W., et al. (2007). HIV-1 gp120 mannoses induce immunosuppressive responses from dendritic cells. PLoS Pathog. *3*, e169.

Shattock, R.J., and Rosenberg, Z. (2012). Microbicides: topical prevention against HIV. Cold Spring Harb. Perspect. Med. *2*, a007385.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V, and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. *1*, 2856–2860.

Shingai, M., Nishimura, Y., Klein, F., Mouquet, H., Donau, O.K., Plishka, R., Buckler-White, A., Seaman, M., Piatak, M., Lifson, J.D., et al. (2013). Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. Nature *503*, 277–280.

Silvestri, G., Sodora, D.L., Koup, R.A., Paiardini, M., O'Neil, S.P., McClure, H.M., Staprans, S.I., and Feinberg, M.B. (2003). Nonpathogenic SIV Infection of Sooty Mangabeys Is Characterized by Limited Bystander Immunopathology Despite Chronic High-Level Viremia. Immunity *18*, 441–452.

Sok, D., Laserson, U., Laserson, J., Liu, Y., Vigneault, F., Julien, J.-P., Briney, B., Ramos, A., Saye, K.F., Le, K., et al. (2013). The effects of somatic hypermutation on neutralization and binding in the PGT121 family of broadly neutralizing HIV antibodies. PLoS Pathog. *9*, e1003754.

Sok, D., Doores, K.J., Briney, B., Le, K.M., Saye-Francisco, K.L., Ramos, A., Kulp, D.W., Julien, J.-P., Menis, S., Wickramasinghe, L., et al. (2014). Promiscuous glycan site recognition by antibodies to the high-mannose patch of gp120 broadens neutralization of HIV. Sci. Transl. Med. *6*, 236ra63.

Stambach, N.S., and Taylor, M.E. (2003). Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells. Glycobiology *13*, 401–410.

Stephens, E., Maslen, S.L., Green, L.G., and Williams, D.H. (2004). Fragmentation characteristics of neutral N-linked glycans using a MALDI-TOF/TOF tandem mass spectrometer. Anal. Chem. *76*, 2343–2354.

Sterjovski, J., Churchill, M.J., Ellett, A., Gray, L.R., Roche, M.J., Dunfee, R.L., Purcell, D.F.J., Saksena, N., Wang, B., Sonza, S., et al. (2007). Asn 362 in gp120 contributes to enhanced fusogenicity by CCR5-restricted HIV-1 envelope glycoprotein variants from patients with AIDS. Retrovirology *4*, 89.

St-Pierre, C., Manya, H., Ouellet, M., Clark, G.F., Endo, T., Tremblay, M.J., and Sato, S. (2011). Host-soluble galectin-1 promotes HIV-1 replication through a direct interaction with glycans of viral gp120 and host CD4. J. Virol. *85*, 11742–11751.

Stremlau, M., Owens, C.M., Perron, M.J., Kiessling, M., Autissier, P., and Sodroski, J. (2004). The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. Nature *427*, 848–853.

Stremlau, M., Perron, M., Lee, M., Li, Y., Song, B., Javanbakht, H., Diaz-Griffero, F., Anderson, D.J., Sundquist, W.I., and Sodroski, J. (2006). Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. Proc. Natl. Acad. Sci. U. S. A. *103*, 5514–5519.

Swanson, M.D., Winter, H.C., Goldstein, I.J., and Markovitz, D.M. (2010). A lectin isolated from bananas is a potent inhibitor of HIV replication. J. Biol. Chem. *285*, 8646–8655.

Tan, J., and Sattentau, Q.J. (2013). The HIV-1-containing macrophage compartment: a perfect cellular niche? Trends Microbiol. *21*, 405–412.

Tomaras, G.D., Ferrari, G., Shen, X., Alam, S.M., Liao, H.-X., Pollara, J., Bonsignori, M., Moody, M.A., Fong, Y., Chen, X., et al. (2013). Vaccine-induced plasma IgA specific for the C1 region of

the HIV-1 envelope blocks binding and effector function of IgG. Proc. Natl. Acad. Sci. U. S. A. *110*, 9019–9024.

Travers, S. a (2012). Conservation, Compensation, and Evolution of N-Linked Glycans in the HIV-1 Group M Subtypes and Circulating Recombinant Forms. ISRN Aids *2012*, 823605.

Trimble, R.B., and Tarentino, A.L. (1991). Identification of Distinct Endoglycosidase (Endo) Activities in Flavobacterium meningosepticum : Endo F1, Endo F2, and Endo F3. Endo F1 and Endo H hydrolyze only high mannose and hybrid glycans. J. Biol. Chem. *266*, 1646–1651.

Trkola, A., Purtscher, M., Muster, T., Ballaun, C., Buchacher, A., Sullivan, N., Srinivasan, K., Sodroski, J., Moore, J.P., and Katinger, H. (1996). Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. J. Virol. *70*, 1100–1108.

Trkola, A., Kuster, H., Rusert, P., Joos, B., Fischer, M., Leemann, C., Manrique, A., Huber, M., Rehr, M., Oxenius, A., et al. (2005). Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. Nat. Med. *11*, 615–622.

Tsai, C.-C., Emau, P., Jiang, Y., Tian, B., Morton, W.R., Gustafson, K.R., and Boyd, M.R. (2003). Cyanovirin-N gel as a topical microbicide prevents rectal transmission of SHIV89.6P in macaques. AIDS Res. Hum. Retroviruses *19*, 535–541.

Tsai, C.-C., Emau, P., Jiang, Y., Agy, M.B., Shattock, R.J., Schmidt, A., Morton, W.R., Gustafson, K.R., and Boyd, M.R. (2004). Cyanovirin-N inhibits AIDS virus infections in vaginal transmission models. AIDS Res. Hum. Retroviruses *20*, 11–18.

Tsuchiya, K., Ode, H., Hayashida, T., Kakizawa, J., Sato, H., Oka, S., and Gatanaga, H. (2013). Arginine insertion and loss of N-linked glycosylation site in HIV-1 envelope V3 region confer CXCR4-tropism. Sci. Rep. *3*, 2389.

Tyms, A.S., Berrie, E.M., Ryder, T.A., Nash, R.J., Hegarty, M.P., Taylor, D.L., Mobberley, M.A., Davis, J.M., Bell, E.A., Jeffries, D.J., et al. (1987). Castanospermine and other plant alkaloid inhibitors of glucosidase activity block the growth of HIV. Lancet 1025–1026.

Utachee, P., Nakamura, S., Isarangkura-Na-Ayuthaya, P., Tokunaga, K., Sawanpanyalert, P., Ikuta, K., Auwanit, W., and Kameoka, M. (2010). Two N-linked glycosylation sites in the V2 and C2 regions of human immunodeficiency virus type 1 CRF01_AE envelope glycoprotein gp120 regulate viral neutralization susceptibility to the human monoclonal antibody specific for the CD4 binding domain. J. Virol. *84*, 4311–4320.

Valladeau, J., Ravel, O., Dezutter-Dambuyant, C., Moore, K., Kleijmeer, M., Liu, Y., Duvert-Frances, V., Vincent, C., Schmitt, D., Davoust, J., et al. (2000). Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules. Immunity *12*, 71–81.

Vallari, A., Holzmayer, V., Harris, B., Yamaguchi, J., Ngansop, C., Makamche, F., Mbanya, D., Kaptué, L., Ndembi, N., Gürtler, L., et al. (2011). Confirmation of putative HIV-1 group P in Cameroon. J. Virol. *85*, 1403–1407.

Veazey, R.S., Marx, P. a, and Lackner, a a (2001). The mucosal immune system: primary target for HIV infection and AIDS. Trends Immunol. *22*, 626–633.

Wagner, N., Löhler, J., Kunkel, E.J., Ley, K., Leung, E., Krissansen, G., Rajewsky, K., and Müller, W. (1996). Critical role for beta7 integrins in formation of the gut-associated lymphoid tissue. Nature *382*, 366–370.

Walker, B. (2007). Elite control of HIV infection implications for vaccines and treatments. Top. HIV Med. *15*, 134–136.

Walker, L.M., Phogat, S.K., Chan-Hui, P.-Y., Wagner, D., Phung, P., Goss, J.L., Wrin, T., Simek, M.D., Fling, S., Mitcham, J.L., et al. (2009a). Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science *326*, 285–289.

Walker, L.M., Phogat, S.K., Chan-Hui, P.-Y., Wagner, D., Phung, P., Goss, J.L., Wrin, T., Simek, M.D., Fling, S., Mitcham, J.L., et al. (2009b). Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science *326*, 285–289.

Walker, L.M., Huber, M., Doores, K.J., Falkowska, E., Pejchal, R., Julien, J.-P., Wang, S.-K., Ramos, A., Chan-Hui, P.-Y., Moyle, M., et al. (2011a). Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature *477*, 466–470.

Walker, L.M., Huber, M., Doores, K.J., Falkowska, E., Pejchal, R., Julien, J.-P., Wang, S.-K., Ramos, A., Chan-Hui, P.-Y., Moyle, M., et al. (2011b). Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature *477*, 466–470.

Walker, L.M., Sok, D., Nishimura, Y., Donau, O., Sadjadpour, R., Gautam, R., Shingai, M., Pejchal, R., Ramos, A., Simek, M.D., et al. (2011c). Rapid development of glycan-specific, broad, and potent anti-HIV-1 gp120 neutralizing antibodies in an R5 SIV/HIV chimeric virus infected macaque. Proc. Natl. Acad. Sci. U. S. A. *108*, 20125–20129.

Wang, W., Nie, J., Prochnow, C., Truong, C., Jia, Z., Wang, S., Chen, X.S., and Wang, Y. (2013). A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. Retrovirology *10*, 14.

Ware, F.E., Vassilakos, a, Peterson, P. a, Jackson, M.R., Lehrman, M. a, and Williams, D.B. (1995). The molecular chaperone calnexin binds Glc1Man9GlcNAc2 oligosaccharide as an initial step in recognizing unfolded glycoproteins. J. Biol. Chem. *270*, 4697–4704.

Wei, X., Decker, J.M., Wang, S., Hui, H., Kappes, J.C., Wu, X., Salazar-Gonzalez, J.F., Salazar, M.G., Kilby, J.M., Saag, M.S., et al. (2003). Antibody neutralization and escape by HIV-1. Nature *422*, 307–312.

Welsch, S., Groot, F., Kräusslich, H.-G., Keppler, O.T., and Sattentau, Q.J. (2011). Architecture and regulation of the HIV-1 assembly and holding compartment in macrophages. J. Virol. *85*, 7922–7927.

West, A.P., Galimidi, R.P., Foglesong, C.P., Gnanapragasam, P.N.P., Huey-Tubman, K.E., Klein, J.S., Suzuki, M.D., Tiangco, N.E., Vielmetter, J., and Bjorkman, P.J. (2009). Design and expression of a dimeric form of human immunodeficiency virus type 1 antibody 2G12 with increased neutralization potency. J. Virol. *83*, 98–104.

West, A.P., Scharf, L., Horwitz, J., Klein, F., Nussenzweig, M.C., and Bjorkman, P.J. (2013). Computational analysis of anti-HIV-1 antibody neutralization panel data to identify potential functional epitope residues. Proc. Natl. Acad. Sci. U. S. A. *110*, 10598–10603.

White, T. a, Bartesaghi, A., Borgnia, M.J., Meyerson, J.R., de la Cruz, M.J. V, Bess, J.W., Nandwani, R., Hoxie, J. a, Lifson, J.D., Milne, J.L.S., et al. (2010). Molecular architectures of trimeric SIV and HIV-1 envelope glycoproteins on intact viruses: strain-dependent variation in quaternary structure. PLoS Pathog. *6*, e1001249.

Van Wilgenburg, B., Moore, M.D., James, W.S., and Cowley, S. a (2014). The productive entry pathway of HIV-1 in macrophages is dependent on endocytosis through lipid rafts containing CD4. PLoS One *9*, e86071.

Willey, R.L., Smith, D.H., Lasky, L. a, Theodore, T.S., Earl, P.L., Moss, B., Capon, D.J., and Martin, M. a (1988). In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity. J. Virol. *62*, 139–147.

De Witte, L., Nabatov, A., Pion, M., Fluitsma, D., de Jong, M. a W.P., de Gruijl, T., Piguet, V., van Kooyk, Y., and Geijtenbeek, T.B.H. (2007). Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. Nat. Med. *13*, 367–371.

Witvrouw, M., Fikkert, V., Hantson, A., Pannecouque, C., Keefe, B.R.O., McMahon, J., Stamatatos, L., Clercq, E. De, and Bolmstedt, A. (2005). Resistance of human immunodeficiency virus type 1 to the high-mannose binding agents cyanovirin N and concanavalin A. J. Virol. *79*, 7777–7784.

Wolinsky, S.M., Wike, C.M., Korber, B.T.M., Hutto, C., Wade, P., Rosenblum, L.L., Kunstman, K.J., Furtado, M.R., and Muñoz, J.L. (1992). Selective transmission of human immunodeficiency virus type-I variants from mothers to infants. Science *255*, 1134–1137.

Wood, N.T., Fadda, E., Davis, R., Grant, O.C., Martin, J.C., Woods, R.J., and Travers, S. a (2013). The influence of N-linked glycans on the molecular dynamics of the HIV-1 gp120 V3 loop. PLoS One *8*, e80301.

World Health Organisation, T. Global health sector strategy on HIV/AIDS 2011-2015.

Wu, X., Yang, Z.-Y., Li, Y., Hogerkorp, C.-M., Schief, W.R., Seaman, M.S., Zhou, T., Schmidt, S.D., Wu, L., Xu, L., et al. (2010). Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science *329*, 856–861.

Wu, Y., West, A.P., Kim, H.J., Thornton, M.E., Ward, A.B., and Bjorkman, P.J. (2013). Structural Basis for Enhanced HIV-1 Neutralization by a Dimeric Immunoglobulin G Form of the Glycan-Recognizing Antibody 2G12. Cell Rep. *5*, 1443–1455.

Wuhrer, M., Catalina, M.I., Deelder, A.M., and Hokke, C.H. (2007). Glycoproteomics based on tandem mass spectrometry of glycopeptides. J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. *849*, 115–128.

Wyatt, R., Kwong, P.D., Desjardins, E., Sweet, R.W., Robinson, J., Hendrickson, W.A., and Sodroski, J.G. (1998). The antigenic structure of the HIV gp120 envelope glycoprotein. Nature *393*, 705–711.

Yang, W., Shah, P., Eshghi, S.T., Yang, S., Ao, M., Rubin, A., Jackson, J.B., and Zhang, H. (2014). Glycoform analysis of recombinant and human immunodeficiency virus envelope protein gp120 via higher energy collisional dissociation and spectral-aligning strategy. Anal. Chem. *86*, 6959–6967.

Yasmeen, A., Ringe, R., Derking, R., Cupo, A., Julien, J.-P., Burton, D.R., Ward, A.B., Wilson, I. a, Sanders, R.W., Moore, J.P., et al. (2014). Differential binding of neutralizing and non-neutralizing antibodies to native-like soluble HIV-1 Env trimers, uncleaved Env proteins, and monomeric subunits. Retrovirology *11*, 41.

Yen, P.-J., Herschhorn, A., Haim, H., Salas, I., Gu, C., Sodroski, J., and Gabuzda, D. (2014). Loss of a conserved N-linked glycosylation site in the simian immunodeficiency virus envelope glycoprotein V2 region enhances macrophage tropism by increasing CD4-independent cell-to-cell transmission. J. Virol. *88*, 5014–5028.

Yoder, A., Yu, D., Dong, L., Iyer, S.R., Xu, X., Kelly, J., Liu, J., Wang, W., Vorster, P.J., Agulto, L., et al. (2008). HIV envelope-CXCR4 signaling activates cofilin to overcome cortical actin restriction in resting CD4 T cells. Cell *134*, 782–792.

Zaia, J. (2010). Mass spectrometry and glycomics. Omi. a J. Integr. Biol. *14*, 401–418.

Zanetti, G., Briggs, J. a G., Grünewald, K., Sattentau, Q.J., and Fuller, S.D. (2006). Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ. PLoS Pathog. *2*, e83.

Zhang, M., Gaschen, B., Blay, W., Foley, B., Haigwood, N., Kuiken, C., and Korber, B. (2004). Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. Glycobiology *14*, 1229–1246.

Zhou, T., Georgiev, I., Wu, X., Yang, Z.-Y., Dai, K., Finzi, A., Kwon, Y. Do, Scheid, J.F., Shi, W., Xu, L., et al. (2010). Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science (80-.). *329*, 811–817.

Zhu, P., Liu, J., Bess, J., Chertova, E., Lifson, J.D., Grisé, H., Ofek, G. a, Taylor, K. a, and Roux, K.H. (2006). Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature *441*, 847–852.

Zhu, T., Mo, H., Wang, N., Nam, D., Cao, Y., Koup, R., and Ho, D. (1993). Genotypic and phenotypic characterization of HIV-1 patients with primary infection. Science (80-.). *261*, 1179–1181.

Zhu, T., Korber, B.T., Nahmias, a J., Hooper, E., Sharp, P.M., and Ho, D.D. (1998). An African HIV-1 sequence from 1959 and implications for the origin of the epidemic. Nature *391*, 594–597.

Zhu, X., Borchers, C., Bienstock, R.J., and Tomer, K.B. (2000). Mass spectrometric characterization of the glycosylation pattern of HIV-gp120 expressed in CHO cells. Biochemistry *39*, 11194–11204.

Zolla-Pazner, S. (2004). Identifying epitopes of HIV-1 that induce protective antibodies. Nat. Rev. Immunol. *4*, 199–210.

Zou, Z., Chastain, A., Moir, S., Ford, J., Trandem, K., Martinelli, E., Cicala, C., Crocker, P., Arthos, J., and Sun, P.D. (2011). Siglecs facilitate HIV-1 infection of macrophages through adhesion with viral sialic acids. PLoS One *6*, e24559.