University of Southampton

Faculty of Environmental and Life Sciences

School of Biological Sciences

Immunogen glycosylation in protein and RNA delivery systems

by

Himanshi Chawla

ORCID ID: 0000-0001-9828-6593

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

September 2023

University of Southampton

Abstract

Extensive glycosylation of viral glycoproteins is a key feature of the antigenic surface of viruses, yet glycan processing can be influenced by the assembly of the protein and production systems. These are important factors to consider in vaccine design, as non-native like assembly and the presence of artificial glycan holes can potentially influence the immune response. To address these concerns, various stabilizing approaches, including the incorporation of proline, disulphide-bonds, glycine-rich linker,and many more, have been employed in the design of recombinant protein immunogens to ensure correct assembly and enhanced expression. Additionally, post-expression purification is typically implemented in recombinant-based immunogens to capture properly folded protein and minimize non-native material.

A common strategy that has been employed in the stabilization of HIV-1 immunogens, the introduction of proline substitutions, has been adopted in SARS-CoV-2 vaccine development. This thesis reveals the influence of these mutations on SARS-CoV-2 Spike (S) protein conformations and glycan compositions. Furthermore, this study compares the glycan compositions presented by S protein and receptor binding domain (RBD). Additionally, a serology study of chemically treated S protein with kifunensine (ER-mannosidase I inhibitor), aimed at eliciting all oligomannose-type glycans, highlights that the immune response following SARS-CoV-2 infection, with respect to immunoglobin binding, is not dictated by the glycan processing state of the S protein. This is, however, not the case in HIV-1 vaccine design, in which glycans play a crucial role in structural integrity and act as an epitope for many broadly neutralizing antibodies.

In contrast to traditional protein-based immunogens, RNA-based vaccines have emerged as a highly effective delivery platform due to its rapid adaptability. However, one potential limitation is that the immunogen is entirely encoded by nucleotide sequence, and thus, there is no opportunity at the post-expression level to control immunogen assembly. In HIV-1 vaccine research, this is of particular importance because non-native epitopes can compromise the desired immune response, and native immunogen assembly is essential for the presentation of glycan-based epitopes targeted by broadly neutralizing antibodies. The work presented here reveals the archetypal protein architecture and glycosylation of the HIV-1 Env expressed via an RNA-based immunogen. Furthermore, this thesis explores the native-like signatures of the Env and the glycan compositions of the material derived from cell-types, muscle, and dendritic cells, which likely produce immunogens near the sites of intramuscular RNA injection. Finally, this study shows that nucleotide editing can enhance the glycan occupancy of RNA-derived immunogens. The data presented in this thesis emphasize the significance of considering glycan site occupancy and glycan heterogeneity when developing and assessing vaccine candidates.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: **Himanshi Chawla**

Title of thesis: **Immunogen glycosylation in protein and RNA delivery systems**

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

I confirm that:

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

   Chapter 1: Chawla H, Fadda E, Crispin M. Principles of SARS-CoV-2 glycosylation. *Curr Opin Struct Biol* 2022, 75:102402.

   Chapter 2 and Chapter 3: Chawla H, Jossi SE, Faustini SE, Samsudin F, Allen JD, Watanabe Y, Newby ML, Marcial-Juárez E, Lamerton RE, McLellan JS, et al. Glycosylation and Serological Reactivity of an Expression-enhanced SARS-CoV-2 Viral Spike Mimetic**.** *J Mol Biol* 2022, 434:167332

   Chapter 4: Allen JD, Chawla H, Samsudin F, Zuzic L, Shivgan AT, Watanabe Y, He W, Callaghan S, Song G, Yong P, et al. Site-Specific Steric Control of SARS-CoV-2 Spike Glycosylation. *Biochemistry* 2021, 60:2153–2169.

   Chapter 4: Konrath KM, Liaw K, Wu Y, Zhu X, Walker SN, Xu Z, Schultheis K, Chokkalingam N, Chawla H, Du J, et al. Nucleic acid delivery of immune-focused SARS-CoV-2 nanoparticles drives rapid and potent immunogenicity capable of single-dose protection. *Cell Rep* 2022, 38:110318.

   Signature: Date:20/11/23

Acknowledgements

I am deeply grateful to so many people who have helped and supported me throughout my DPhil, without them this thesis would not have been possible. It is with immense appreciation that I extend my thanks to Prof. Max Crispin for having faith in me and welcoming me into his lab. His invaluable support, guidance, and confidence in me to work on this interesting project has enabled me to become the independent researcher I am today. I really enjoyed being part of his group, it has been good four years.

I am grateful to everyone at the University of Southampton for making such a welcoming and supportive work environment to work. I must also thank Dr Paul Skipp for co-supervising me and Dr Mark Dixon, the building manager, for always being supportive so that all the labs run smoothly. A special thanks to Joel for all the guidance regarding mass spectrometry and glycan analysis when I started in the lab and his endless patience for all my doubts and willingness to help out. I would also like to specifically thank, Grace Hayes, for being really helpful regarding RNA production and protein expression.

I am really thankful to all the members in Crispin group for making it such a fun and friendly place to work, Abi, Lydia, John, Grace, Maddy, CharlieE, CharlieB, Al, Joel, Yasa, Hannah, Dylan, Dilys, Jacob and Wenwen. A particular mention to Lydia and Abi for making these four years amazing. I would also like to thank all my friends back home. Special mention to Kunal, for always cheering me up and for keeping me sane! Thank you for having infinite patience throughout all of this.

I would also like to thank members of the Irvine group, to help me with all the experimental work carried out during my three months secondment at Massachusetts Institute of Technology. I would like to extend my thanks to Dr. Namit Chaudhary for teaching me lipid nanoparticle production and handling mice models, which I was quite scared about in the start.

Furthermore, I am grateful for the people in the Shattock group based at Imperial College London, with a particular mention to Dr. Paul F McKay for helping me get started with setting up RNA work in the lab. I would also like to extend my thanks to all the collaborators who have made this thesis possible, notably the laboratories of Prof. Peter Bond, Dr. Alex G Ritcher, Dr. Adam F. Cunningham, Prof. McLellan, Prof. Irvine, Prof. Shattock and Prof. Sanders.

Finally, I would like to express my thanks to my parents, without them I couldn't have come this far. I appreciate all your support, your sacrifices, your enduring love, and everything you have done for me. Special thanks to Jiniya and Ayush for consistently being my rock! Jiniya, thank you for all the encouragements throughout my childhood; I always felt like I had someone by my side

Definitions and Abbreviations

2-AA.....2-aminobenzoic acid; a label for released glycans to enable quantification

ABO .....a system for classifying human blood on the basis of the presence or absence of two antigens on the red cell membrane: there are four such blood types (A, B, AB, and O)

ACE2.....Angiotensin converting enzyme 2

ACN .....Acetonitrile

AIDS.....Acquired immunodeficiency syndrome

AmpR .....Ampicillin resistance

ASA.....Accessible surface area

Asn .....asparagine residues

AUC .....Area under the curve

BG505 NFL.664.....single chain native flexibly linked Env trimer which is highly similar to BG505 SOSIP.664 with furin cleavage site replaced with glycine-serine rich linker.

BG505 SOSIP.664 .....Stabilising mutations introduced into the HIV-1 Envelope glycoprotein from the BG505 clade A/E isolate which maintains a native-like conformation as a soluble protein.

bNAb .....Broadly neutralizing antibody

C18 .....Octadecyl carbon chain

C2C12.....Mouse myoblast cell line

CAVD.....Consortium of AIDS Vaccine Development

CCR5.....C-C chemokine receptor type 5

CD .....Connector domain

CD4 .....Cluster of differentiation 4

CD4bs.....CD4 binding site

CDR.....complementarity-determining region

CH .....Central helix

ChAdOx nCoV-19 ..... chimpanzee (Ch) adenovirus-vectored (Ad) vaccine developed by University of Oxford (Ox) against novel coronavirus first identified in 2019.

CHO ..... Chinese Hamster Ovary Cells

CID ..... Collision induced dissociation

COVID-19 ..... Coronavirus Disease-2019

CRF ..... circulating recombinant forms

Cryo-em ..... Cryogenic electron microscopy

CT ..... Cytoplasmic tail

CWG ..... cloaking with glycans

CXCR4 ..... C-X-C motif chemokine receptor 4

DC2.4 ..... Mouse dendritic cell line

DMEM ..... Dulbecco's Modified Eagle Media

DNA ..... Deoxyribonucleic acid

DTT ..... Dithiothreitol

Easy-nLC 1200 ..... Nanoflow liquid chromatography instrument used for the separation of glycopeptides prior to analysis by mass spectrometry.

EAVI ..... European AIDS Vaccine Initiative

EDTA ..... Ethylenediaminetetraacetic acid

ELISA ..... Enzyme-linked immunoabsorbent assay

EndoH ..... Endoglycosidase H

Env ..... Envelope glycoprotein of HIV-1

ER ..... Endoplasmic reticulum

ERGIC ..... ER/Golgi intermediate compartments

ESI ..... Electrospray ionization

FBS ..... Fetal bovine serum

FDR ..... False discovery rate

Glc ..... Glucose

GNL ..... Galanthus nivalis lectin beads

GnT.....N-acetylglucosaminyltransferase I

GOI.....Gene of interest

GP.....Glycoprotein

gp120.....Receptor binding portion of HIV-1 Env

gp140.....HIV-1 Env glycoprotein with a truncated C-terminus to solubilize the protein.

gp160.....Precursor polypeptide of HIV-1 Env that exists prior to furin cleavage.

gp41.....Region of HIV-1 Env which is responsible for membrane fusion.

GPC.....glycoprotein complex

Group M.....main group

Group N.....new group

Group O.....outlier group

HA.....Hemagglutinin

HBS.....HEPES-Buffered Saline

HC.....Hospitalized convalescent

HCD.....Higher energy collision induced dissociation.

HEK293F.....Human embryonic kidney 293 Freestyle cells, used for recombinant protein production.

Hex.....Hexose

HILIC-UPLC.....hydrophilic interaction ultra-performance liquid chromatography

HIV-1.....Human Immunodeficiency virus

HR1.....Heptad repeat 1

HR2.....Heptad repeat 2

HRP.....Horseradish peroxidase

HxB2.....HIV-1 reference strain used for alignment of the diverse HIV-1 sequences

IAA.....Iodoacetamide

IAVI.....International AIDS Vaccine Initiative

IgG.....Immunoglobulin G

IgGAM.....Immunoglobulins G, A, M

IMP .....Intrinsic mannose patch

IN .....Integrase

ITU .....Intensive Treatment Unit

IVT .....Invitro transcription

K_D .....Dissociation constant

Kif .....Kifunensine

KS .....Kaposi's sarcoma

LC-MS .....Liquid chromatography-mass spectrometry

LC-ESI MS .....Liquid chromatography-electrospray ionization mass spectrometry

LNPs .....Lipid nanoparticles

M .....Membrane

Man .....Mannose

MBL .....Mannose-binding lectin

MD .....Molecular dynamics

MDFF .....Molecular dynamics flexible fitting

MPER .....Membrane-proximal external region

MRA .....multireference alignment

MSA .....multivariate statistical analysis

MSM .....Men who had sex with Men

N protein .....Nucleoprotein

Nab .....Neutralizing Antibody

NB-DGJ .....*N*-butyldeoxygalactonojirimicin

NB-DNJ .....*N*-butyldeoxynojirimycin

NeuAc .....Neuraminic acid

NFL .....Natively flexibly linked

NHC .....Non-hospitalized convalescent

NMR .....Nuclear magnetic resonance

NS-EM .....Negative stain-electron microscopy

nsp's .....Non-structural proteins

NTD.....N-terminal domain

NxS/T .....Amino acid sequence that defines site of attachment for an N-linked glycan

OD .....Optical density

ORFs.....Open reading frames

Ori.....Origin of replication

OSTase .....Oligosaccharlytransferase

p.p. .....percentage point

PBMCs .....Peripheral blood mononuclear cells

PBS.....Phosphate buffered saline

PBS-T .....Oxoid phosphate buffered saline with 0.1% Tween-20

PCP.....*Pneumocystis pneumonia*

PCR .....Polymerase chain reaction

PDB .....Protein database

PEImax .....Polyethylenimine "Max" reagent used for transient transfection of cells

PGT145 .....HIV-1 antibody which binds to quaternary epitope of the trimer

PL^{pro}.....papain-like protease

PNGaseF.....Peptide N-glycosidase F

PNGS .....Potential N-linked glycosylation site

PTM.....Post-translational modifications

PVDF.....Polyvinylidene fluoride

RBD.....SARS-CoV-2 receptor binding domain

RBM .....receptor binding motif

RdRp.....RNA dependent RNA polymerase

RIPA.....Radioimmunoprecipitation assay buffer

RM20E1 .....Antibody isolated from a macaque, which can only neutralize virus lacking the N611 glycan site

RNA .....Ribonucleic acid

rpm .....revolutions per minute

RSV.....Respiratory syncytial virus

RT .....HIV-1 reverse transcriptase

S protein ECD .....Spike protein ectodomain

S protein ..... SARS-CoV-2 Spike glycoprotein

saRNA.....self-amplifying ribonucleic acid also referred to as replicon.

SARS-CoV-2 2P S.....soluble spike protein with the incorporation of two proline residues at positions K986P and V987P

SARS-CoV-2 .....severe acute respiratory syndrome coronavirus-2

SDS-PAGE .....Sodium dodecyl sulphate–polyacrylamide gel electrophoresis

SEC .....Size Exclusion Chromatography

SGP .....Sub genomic promoter

siRNA .....small interfering RNA

SPR .....Surface plasmon resonance

ss .....single-stranded

TAMP .....Trimer associated mannose patch

TM.....Transmembrane region

TMPRSS2 .....Transmembrane protease, serine 2

UPLC .....Ultra performance liquid chromatography

UTR.....Untranslated region

V1/V2.....Variable loops 1 and 2 located at the apex of HIV-1 Env

VEEV.....Venezuelan equine encephalitis virus

VLP.....Virus-like particles

WHO .....World Health Organizations

WT .....Wildtype

XIC.....Extracted ion chromatogram

Chapter 1 Introduction

Contributions. Some parts of the section 1.2 of the introduction are published as a review in Current opinion of structural biology. The whole review is written by me and has been reviewed by other co-authors in the review, Dr. Max Crispin and Dr. Elisa Fadda. Figures 1.6 and 1.8 are provided by Dr. Elisa Fadda, University of Maynooth, Ireland. All other sections in the introduction are written by me primarily for this thesis.

1.1 Glycosylation in enveloped viruses

Enveloped viruses are encapsulated in lipid envelopes, which are decorated with densely glycosylated proteins [1]. These glycans help mask immunogenic viral components from the host immune system and facilitate attachment and viral entry into the host cells. Once inside, viruses hijack the host glycosylation and protein-folding machinery to assemble new viruses coated with host glycans. Numerous viral proteins have evolved to utilize the host glycosylation machinery to decorate their proteins with "self" glycan moieties. The analysis of viral proteins has revealed significant variation in the structure, density, and conservation of glycans. Some of the viral envelope proteins have been extensively investigated, including the envelope glycoprotein (Env) of human immunodeficiency virus-1 (HIV-1) [2,3], the hemagglutinin glycoprotein (HA) of influenza virus [4,5], the spike glycoprotein (S) of coronaviruses [6–8], the glycoprotein (GP) of Ebola virus [9,10], the glycoprotein complex (GPC) of Lassa virus [11,12], and the envelope glycoprotein (E) in viruses such as dengue, Zika, and other flaviviruses [13,14]. However, glycosylation extends beyond envelope proteins to encompass numerous secreted viral proteins. For example, the secreted GP of Ebola [15].

Understanding the glycosylation of these viral proteins and the immunological basis driving diversity is an important aspect to consider in the development of vaccine candidates. This will assist in how the glycosylation of vaccine candidates compares to the native viral glycosylation. Although studies of viral glycosylation are often constrained to analysis of recombinant proteins that are usually produced in cell lines with glycosylation capacity different from host cells [16]. Therefore,

the glycosylation pattern of recombinant protein might be different from the glycans presented on the viral protein [17]. However, certain glycan structures, influenced by the quaternary structure of the protein and the amino acid sequon, tends to remain conserved such as oligomannose-type glycans[18]. These conserved glycan compositions can be used to monitor the structural integrity of the protein, as in case of instability of the protein, the conserved residues will get more exposed and thus will be more accessible for glycan processing. It is important to understand these glycan structures as in some instances such as in HIV-1, glycan occludes half of the protein surface of the Env and are also targeted by numerous antibodies. However, in case of some viruses such as flaviviruses, the antigenic surface is covered by one, two, or no N-linked glycosylation sites [19]. In this thesis, the impact of vaccine design approaches are investigated for two key vaccine targets, HIV-1 and SARS-CoV-2.

1.1.1 Monosaccharides: Building blocks of glycans

The glycan structures are built using simplest form of sugar that is monosaccharides [20]. They are typically classified on basis of the number of carbon atoms they contain. Majority of the mammalian N-linked glycans consist of five atom rings, furanose, or six atom rings, pyranose (Figure 1.1 A). Monosaccharides usually exist in solution as an equilibrium mixture of cyclic and acyclic forms that is usually represented as chair and Fischer projections. The chair conformation of monosaccharides is shown in **Figure 1.1.A**. The precise nomenclature that is D or L configuration of each monosaccharide is determined by the specific arrangement of side groups, hydroxy group or non-OH group. The configuration is easily determined by using Fisher projection of the monosaccharide, if the OH group or non-OH group is on the right, then the overall configuration is D and if the OH group is on the left, the overall configuration is L [21]. Hexoses, encompassing six-carbon monosaccharides such as glucose, galactose, and mannose, are subject to this isomeric variation. Notably, glucose and mannose differ solely in the positioning of the hydroxyl group at carbon 2 (see Figure 1.1.A). Chemical modifications, such as the addition of an N-acetylamine at position 2, result in N-Acetyl hexosamine (HexNAc). Further complexities arise from additional modifications, such as the inclusion of sialic acid or neuraminic acid (NeuAc), which introduces extended side chains,

potentially conferring a negative charge to the monosaccharide and influencing the function of the attached glycan.

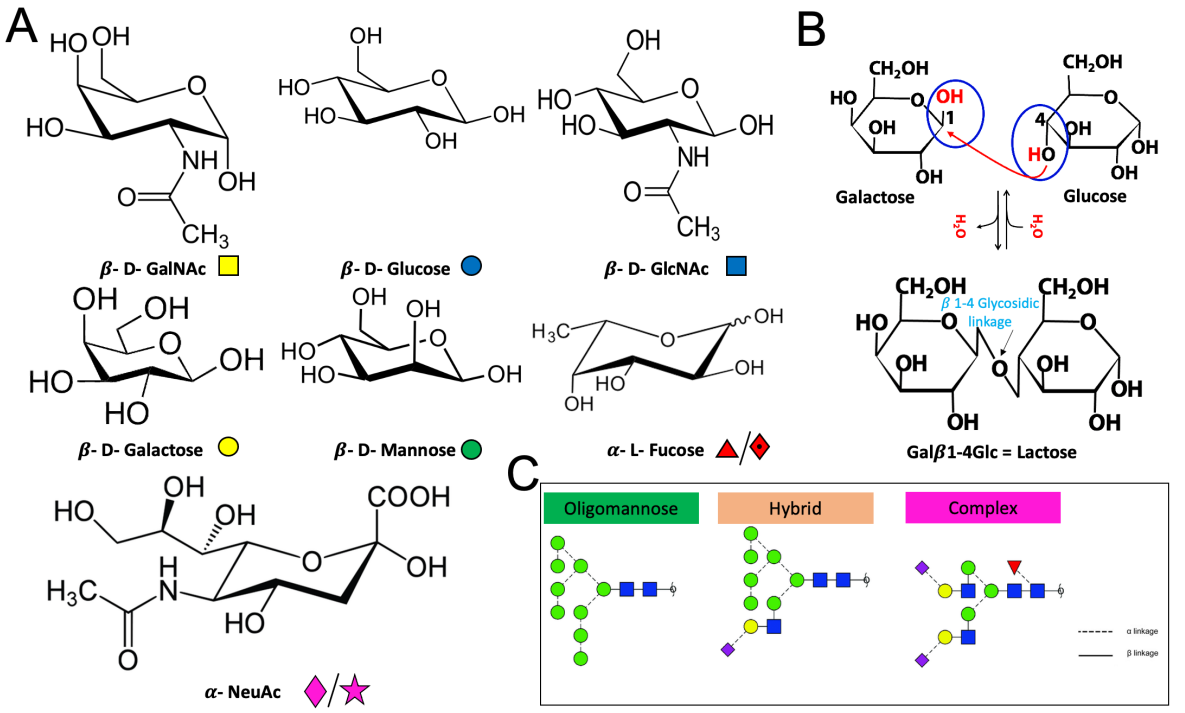


Figure 1 1. Monosaccharide structures, linkages, and schematic of glycan compositions. (A) Monosaccharides that commonly form N-linked glycans on human glycoproteins. Displayed symbols represent each monosaccharide, used for both the Symbol nomenclature for glycans (SNFG) and the Oxford classifications. In instances where two symbols are displayed, the left one corresponds to SNFG, while the right one represents Oxford. These structures are taken from the wikipedia. (B) The two monosaccharides are joined together by an oxide linkage formed by the loss of a water molecule. Here, an illustrative example of condensation reaction is showed between galactose and glucose leading to the formation of lactose [22]. The bond between two sugars is known as glycosidic bond. (C) Representation of most common glycan compositions found on N-linked glycans, oligomannose, hybrid and complex-type glycans.

Monosaccharides are linked to one another enzymatically via a glycosidic bond to create linear or branched oligosaccharides [20]. The donor sugar in an activated form is enzymatically attached via a glycosidic bond to a particular hydroxyl group on the acceptor sugar (**Figure 1.1.B**). Here, an illustrative example is represented between glucose and galactose linked via β- 1-4 glycosidic linkage leading to formation of lactose [22]. This process is repeated one by one, by highly specific glycosyltransferases until a precise glycan target is created. To avoid drawing the complete chemical structure, glycobiologists represent these glycans by a series of connected shapes. This thesis mainly discusses about three glycan compositions, oligomannose, hybrid and

complex-type glycans which are commonly found on N-linked glycosylation (**Figure 1.1.C**). These glycan compositions are usually synthesised via glycosyltransferases present in the ER and Golgi, which is described in detail in the next section.

1.1.2 Protein glycosylation

Protein glycosylation is highly specific and is orchestrated by a series of enzymes, usually located in the ER and Golgi-compartments, that recognize particular amino acid motifs on the target protein [23]. The resulting glycoprotein may have diverse glycan structures that can vary in composition and complexity. These glycan structures can influence the protein's stability, solubility, and interactions with other molecules [24].

The two most studied protein-directed glycosylation mechanisms in viral research, are those attached to asparagine residues, referred to as N-linked glycans, and those attached to serine and threonine, O-linked glycans. Prominent examples of viral glycoproteins coated with high number of N-linked glycans include HIV-1 [25], and high levels of O-linked glycans include Ebola [26]. In this thesis, I have primarily focused on N-linked glycosylation as it is prominently present within the viral glycoproteins, namely HIV-1 and SARS-CoV-2, while O-linked glycans constitute less than 1% in these viruses. The N-linked glycosylation pathway involves the covalent attachment of carbohydrate chains (glycans) to specific asparagine residues (Asn) within the consensus sequence Asn-X-Ser/Thr, where X can be any amino acid except proline [27]. While the potential position of the glycan is linked to the amino acid sequence of the protein, therefore encoded by the nucleic acid of the virus, but can change under selection pressures such as immune evasion, and their role in glycoprotein folding and assembly [28]. Viral glycoproteins utilize host cell's glycosylation machinery as they are able to hijack the cellular glycosylation pathway. The N-glycosylation pathway primarily occurs within the endoplasmic reticulum (ER) and Golgi apparatus of eukaryotic cells, and the overview of the pathway is shown in **Figure 1.2.**

The initial step of N-linked glycan biosynthesis involves the *en bloc* transfer of core oligosaccharide structure ($Glc_3Man_9GlcNAc_2$; Glc: glucose, Man: mannose, GlcNAc: N-acetylglucosamine) from dolichol phosphate to the amide groups of asparagine in potential N-linked

glycosylation sites (PNGS). This transfer is mediated by oligosaccharyltransferase (OST) enzyme complex present in the ER lumen [29]. Once the core oligosaccharide structure is transferred to asparagine residues, other glycan processing enzymes in the ER and Golgi starts trimming down and subsequently build up glycans, resulting in a variety of different classes of glycans, including oligomannose, hybrid, and complex-type N-glycan structures (**Figure 1.2**). Notably, in context of viruses, the glycan pathway may not be strictly linear, and viral particles may bud off early in the glycosylation pathway, which may result in the presence of unusual glycosylation on some viral glycoproteins [30]. In addition, glycosylation is dictated by the expression level of glycosyltransferases and glycosidases which is dependent on producer cell, along with other parameters such as metabolite availability [31].

The $Glc_3Man_9GlcNAc_2$ attached to asparagine residues are further trimmed by ER α-glucosidases I and II which trims to remove the two glucose residues, generating monoglucosylated ($Glc_1Man_9GlcNAc_2$) N-linked glycans, which interacts with the lectins, folding chaperones, calnexin and calreticulin [32]. The final glucose is cleaved by ER α-glucosidases II, leading to release of the glycoprotein from Calnexin/Calreticulin cycle which signals that the protein is correctly folded and can transit from ER to Golgi. Alternatively, misfolded proteins are reglucosylated by UDP-glucose glycosyltransferase [33]. Misfolded proteins undergo repeated cycles of ER chaperone, calnexin/calreticulin, mediating folding until they are either correctly folded or eventually degraded by the ER-associated degradation pathway. It has been reported that viruses spend a considerable amount of time within the ER, which allows time to form extensive network of disulfide bonds, which is essential for the formation of infectious virions [34,35] .

Following correct folding, the glycan processing pathway continues with the removal of mannose residues from $Man_9GlcNAc_2$ to $Man_8GlcNAc_2$, mediated by ER α-mannosidase I. This is accompanied by the transition from the ER to the Golgi where up to three mannose residues are trimmed by Golgi resident α-mannosidases to generate a $Man_5GlcNAc_2$ structure. Glycans spanning from $Man_9GlcNAc_2$ to $Man_5GlcNAc_2$ constitute oligomannose-type glycans. Compact folding of the protein usually result in the presence of these under-processed glycans [36].

The glycan sites which are more accessible further undergo processing by N-acetylglucosaminyltransferase I (GnT I) to hybrid-type glycans by addition of an N-acetylglucosamine residue onto the D1 arm of the $Man_5GlcNAc_3$. Fucosyltransferase can act upon the $Man_5GlcNAc_2$ which can further results into hybrid-type or/and complex-type glycans [37]. The absence of further mannosidase trimming onto the D3 and D2 arms will result in the formation of a hybrid-type glycan, which allows processing of the D3 arm while the D1 and D2 arms remain unaltered. Alternatively, further mannose-trimming by Golgi α-mannosidase II on the D1 and D2 arms will result in $Man_3GlcNAc_3$ glycan structures, which can be processed by various glycosyltransferases dependent on cell-type. This will lead to the generation of a variety of complex-type glycans. The repertoire of glycan structures produced by a given cell reflects its expression level of glycan processing enzymes.

The additional GlcNAc attached to $Man_3GlcNAc_3$ can be transferred, or the glycan becomes more branched. Following GlcNAc transfer, galactosyltransferase can act and add galactose to the GlcNAc, which can then act as an acceptor for sialic acid [38]. This range of different glycoforms is dependent on cell-type/tissue type and can lead to microheterogeneity on individual glycan sites across multiple proteins [39,40]. As mentioned above, the glycan pathway is not strictly linear and there is no defined exit point of the processing pathway, and mature glycoproteins can contain a mixture of glycans in many different processing states. For example, E1 and E2 protein of Hepatitis C virus which buds from the ER [41]. While there are many more N-glycan modifications possible during ER and Golgi transit, the modifications outlined above are the most common and relevant for the cell lines outlined in this thesis.

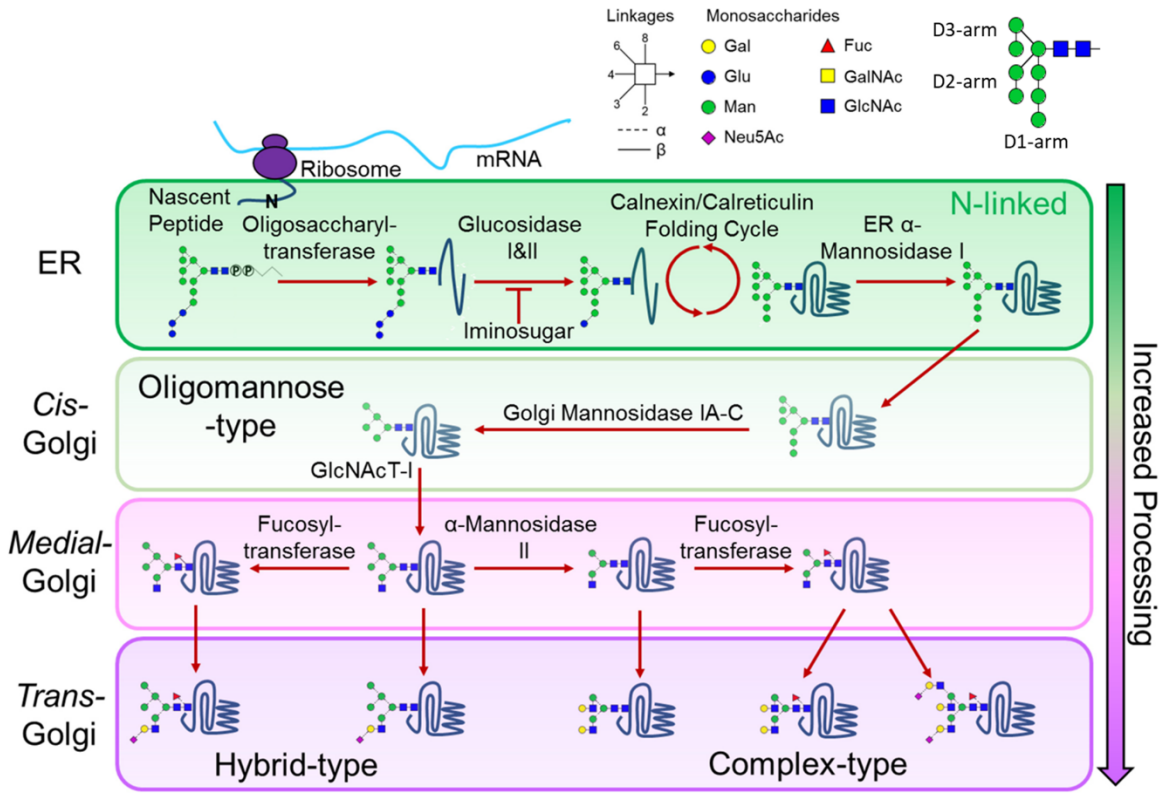


Figure 1.2. Mammalian N-linked glycosylation pathway. The pathway initiates when the N-linked glycan precursor ($Glc_3Man_9GlcNAc_2$) is covalently transferred from a dolichol phosphate to asparagine residue by oligosaccharyltransferase (OST). The terminal three glucose residues are removed by the ER resident glucosidase I and II. Once the protein is correctly folded in the calnexin and calreticulin cycle, the terminal mannose from the D2 arm is trimmed by ER mannosidase I. Subsequently, the protein transits from ER to Golgi where Golgi mannosidases are responsible for further mannose trimming, leading to $Man_5GlcNAc_2$. The transfer of N-acetylglucosamine residue via GlcNAc transferase -I (GnT-I) initiates the first branch of the N-glycan and processing to hybrid-type glycans. Once the Golgi resident mannosidases II trims the two remaining mannose residues, other cell-specific glycosyltransferases such as galactosyl-, fucosyl-, sialyl-transferase, generates a variety of complex-type glycans. Figure adapted from Watanabe *et al*., 2019 [1].

Unlike N-linked glycosylation, O-linked glycosylation is not identified by a specific sequon, which makes it much more difficult to predict this modification [42]. However, there is progress in improving methods to identify O-linked glycosites [43]; in mucin-like domains, they can be detected by the high levels of serine, threonine and proline [44]. Also, these glycans can be further elongated by competing glycosyltransferases to form up to eight core structure, which creates heterogeneity in O-linked glycosylation [1] and thus impedes its analysis [45,46]. A number of viral glycoproteins contain mucin-like domains with high population of serine and threonine residues that are heavily modified by O-linked glycosylation. For example, Ebola GP and sGP [47], herpes simplex virus gC protein [48].

While O-glycans are sparsely present on the viral glycoproteins HIV-1 and SARS-CoV-2, discussed in this thesis. There are a few subsets of HIV-1 isolates that have revealed the presence of O-linked glycans and their shielding effect from neutralizing antibodies [49]. Similarly, SARS-CoV-2 contains really a low number of O-linked glycans known till now [8,50]; however, one of the O-linked glycan (Thr678) is known to influence the cleavage of the viral glycoprotein and may influence viral infectivity [51,52].

The characterization of N- and O-linked glycosylation has been challenging due to the heterogeneity and complexity arising from glycan processing enzymes. However, the availability of potent mass spectrometric and chromatographic techniques, along with the commercial availability of proteases, have enabled the investigation of these post-translational modifications [53]. The experimental approaches, including matrix assisted laser desorption ionisation (MALDI), electrospray ionization (ESI) mass spectrometry, and liquid chromatography-mass spectrometry have allowed the characterization of viral glycome and site-specific glycan analysis [54].

1.1.3 Glycosylation heterogeneity in viral tropism and transmission

The heterogeneity of glycosylation usually reflects the repertoire of glycosyltransferases, the metabolic state and a myriad of other factors in the producing cell [55]. Similarly, any multicellular target host will exhibit diverse tissue-specific glycosylation, potentially impacting viral tropism [56]. These heterogeneities in glycosylation can influence the spread of viruses within a species. Such glycan differences are apparent within the human population as exemplified by the ABO blood group system [23]. Individuals with blood group A express fucosylglycoprotein α-N-acetylgalactosaminyltransferase that adds N-acetylgalactosamine, whereas those with Blood group B express fucosylglycoprotein galactosyltransferase that add galactose [57]. These variations in glycan structures can impact the presence of antibodies and, consequently, limit viral transmission from individuals of Blood group B to Blood group A. These effects have been observed in the infectivity of HIV-1 and other viruses [58,59]. The distribution of blood groups may exert selective pressures that influence the prevalence and spread of viral infections [60]. Viruses may evolve in

response to the glycan landscape of host populations, and this interplay between viral evolution and host glycan diversity can shape the dynamics of viral infection within a species.

Although the mammalian N-linked glycan pathway is mostly conserved, several significant differences can affect viral antigenicity [61]. For example, humans lack the galactose-α-1,3-galactose epitope, a common structure found at the terminals of mammalian glycans [62]. The galactose-α-1,3-galactose epitope is present on pig cells and the porcine endogeneous retrovirus which can potentially be transmitted to human cells during xenotransplanation (transplantation of pig organs into humans). The presence of this epitope on pig cells makes them susceptible to antibodies in human serum, leading to potential risks associated with transplantation [63]. Antibody-mediated immunity against these epitopes can impede viral infectivity across different species [64]. Additionally, O-linked glycans and glycolipids may display notable compositional variations between species [23].

Furthermore, the differences in glycan composition across different species are an important feature of zoonosis of envelope viruses, especially regarding the potential influence of host immunity on transmission [1,65,66]. For instance, the ability of avian influenza hemagglutinin (HA) to interact with both α2,3- and α2,6-linked sialic acid has been demonstrated to facilitate viral spillover from birds to humans [56,67]. Moreover, variations in the expression of these sialic acid structures between the upper and lower respiratory tracts in humans can influence the localization of influenza infection within an individual [68,69].

Similar role of sialic acids has been reported in transmission and tropism of the betacoronaviruses S protein. For example, the MERS-CoV S protein exhibits a preference for binding to α2,3-linked sialic acid receptors over α2,6-linked ones, which are abundant in key replication sites within the lower respiratory tract, notably the alveoli of the human lung [70]. Furthermore, investigations into SARS-CoV-2's interaction with sialic acids have revealed their involvement in facilitating infection, potentially influencing tissue tropism [52,71,72]. Notably, recent findings indicate that SARS-CoV-2 can invade the central nervous system, which may be potentially linked to the prevalence of gangliosides (containing sialic acid) in the human brain [72,73]. In addition, sialic acid present on one of the O-linked sites on SARS-CoV-2 S protein have been observed to play a role in furin cleavage and act as a limiting factor for spike protein processing.

Overall, it is important to understand the enormous diversity of glycans that exist in terms of glycan composition and structure, arising naturally and in experimental systems. Assessing this heterogeneity in the glycosylation of specific proteins is challenging in glycoproteomics and requires elaborate site-specific glycan analysis with isolated glycoproteins [74]. Understanding the mechanisms and regulation of these glycosylation processes is essential for advancing our knowledge of cell function and for the development of vaccine design. In this thesis, I have focused on understanding the assembly and glycosylation of two of the enveloped viruses, SARS-CoV-2 and HIV-1 to inform the immunogen design.

1.2 SARS-CoV-2 Spike glycosylation

1.2.1 The emergence of novel coronavirus

The World Health Organization (WHO) maintains vigilant surveillance over viral animal reservoirs, actively assessing potential viruses capable of zoonotic transmission to humans, thus bearing the propensity for pandemic propagation [75]. Among these, coronaviruses have coexisted within the human population for many years, and typically cause diseases such as the common cold [76]. In addition, SARS-CoV and MERS-CoV caused relatively high mortality and emerged in 2002 [77] and 2012 [78], respectively. While both of these outbreaks were relatively contained, the WHO categorized coronaviruses as a subset of animal-borne pathogens endowed with pandemic potential. In late 2019, the earliest cases of a novel respiratory illness were reported in Wuhan, Hubei Province, China. The disease was initially referred to as a pneumonia of an unknown cause. Chinese health authorities quickly identified a new coronavirus as the causative agent and shared its genetic sequence with the World Health Organization (WHO) and other global health agencies. A new virus was identified from airway epithelial cells of infected patients on January 12th, 2020 [79] and the WHO declared a global pandemic in March 2020.

1.2.2 SARS-CoV-2 viral genome and life cycle

SARS-CoV-2 belongs to the family Coronaviridae and possesses a positive-sense, single-stranded RNA genome. The genome is approximately 7000 nucleotides long and encodes for several structural

and non-structural proteins. The SARS-CoV-2 genome is organized into 12 functional open reading frames (ORFs) that encode viral proteins [80]. Notably, the S protein facilitates viral entry, the envelope (E) protein contributes to assembly, the membrane (M) protein is involved in virus-host interactions, and the nucleocapsid (N) protein packages the viral RNA [81].

Briefly, the SARS-CoV-2 life cycle initiates when the S protein on the surface of virus's surface encounters host cells adorned with angiotensin-converting enzyme 2 (ACE2) receptors on their surface [82] (**Figure 1.3**). The host serine protease, TMPRSS2, involved in priming of S protein for S1/S2 cleavage [83]. This is followed by the attachment and fusion between the viral and host cellular membranes, resulting in the release of viral genomic RNA into the cells. Once the viral RNA is inside the host cell, it hijacks the host cellular machinery and begins replicating its genome. Firstly, two viral replicase polyproteins, pp1a and pp1ab, will be translated. These are cleaved by the papain-like protease (PL^{pro}) and 3C-like protease (also known as the main protease) to form functional non-structural proteins such as Helicase and RNA dependent RNA polymerase (RdRp) [84]. The structural proteins, S, M, E and the N protein, are then translated by ribosomes that are bound by the endoplasmic reticulum (ER). These proteins then enter the ER and undergo post-translational modification as would host-cell glycoproteins. In addition, some of the accessory proteins are also encoded. The virion then buds into the ER Golgi intermediate complex (ERGIC) and is secreted via exocytosis.

Unlike HIV-1, the mutation rate of SARS CoV-2 is significantly lower. This is attributed to the presence of proofreading mechanisms within the RNA polymerase. This polymerase possesses 3′ to 5′ exonuclease activity, which diligently excises mis-incorporated nucleotides [85]. Furthermore, it is noteworthy that the period of time during which SARS-CoV-2 can be transmitted is brief, so there is little time for viral evolution to occur in a host prior to transmission [86]. In stark contrast, HIV-1 exhibits the capacity for enduring years of evolutionary progression while retaining the capability for continued transmission [87].

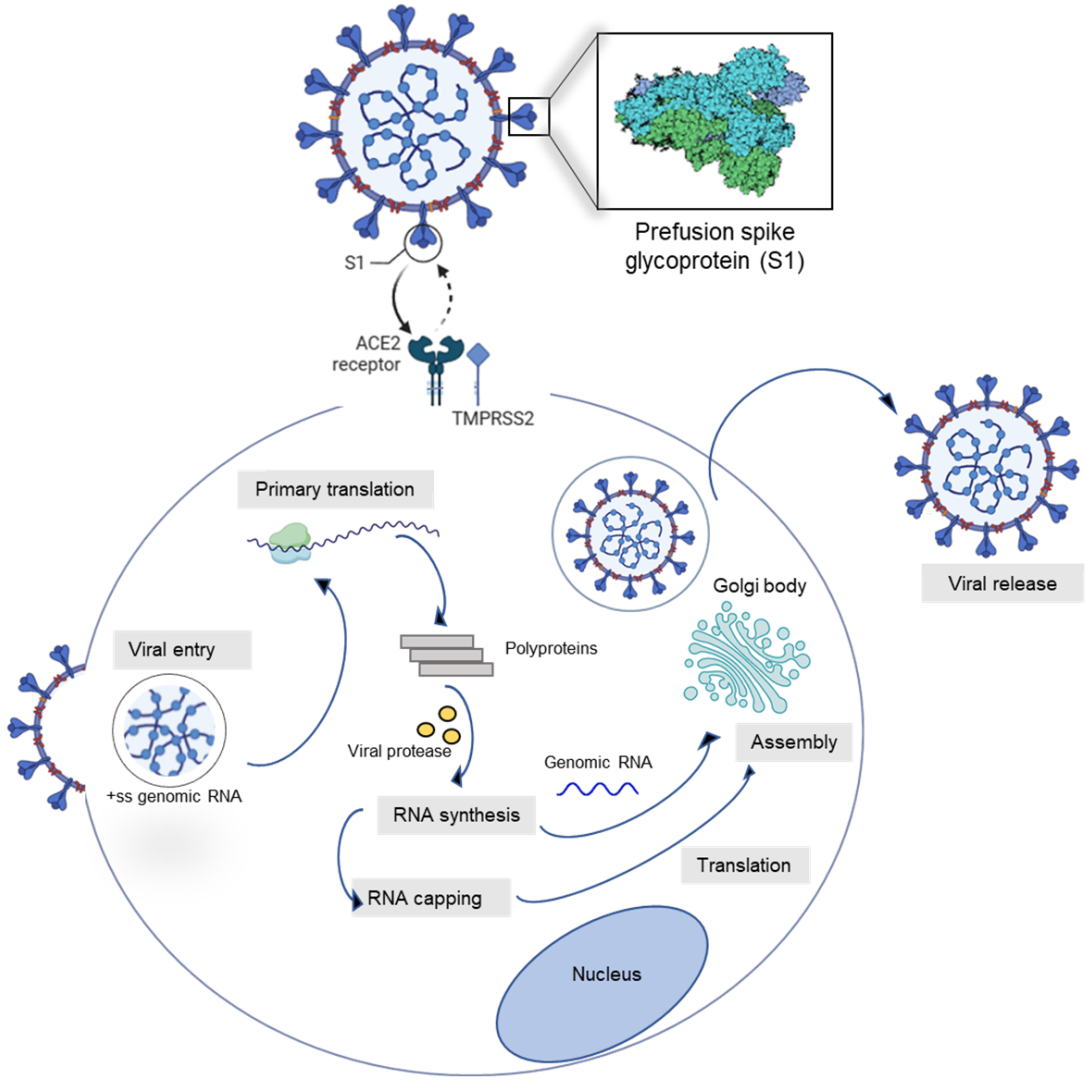


Figure 1.3. The life cycle of SARS-CoV-2 virion. Key stages of viral life cycle including viral entry, replication and transcription, assembly, and release. TMPRSS-2, transmembrane serine protease 2; +ss genomic RNA, positive single stranded genomic RNA.

1.2.3 Spike glycoprotein structure and function

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) S glycoprotein is important for the assembly and stability of the virus [88]. The entry of SARS-CoV-2 is primarily mediated by the interaction between S and the ACE2 receptor, expressed on the surface of the host cell [89,90]. The S glycoprotein is a trimeric class I fusion protein and comprises of two functional subunits, S1 and S2. The total length if S protein is 1273 aa and the size of 180-200 kDa [91]. The S protein is translated as a single polypeptide and is cleaved at the S1/S2 interface at furin cleavage site, RRAR by furin (**Figure 1.4**). SARS-CoV-2 S has a specific furin cleavage site which is not present in other SARS-like coronaviruses [92]. The S1 subunit is separated into the N-terminal domain (NTD) and

the receptor binding domain (RBD). The RBD in S1 subunit interacts with the ACE2 receptor [93]. The S2 subunit contains fusion peptide (FP), heptad repeats (HR1), central helix (CH), connector domain (CD), transmembrane domain (TM). The S2 subunit is responsible for viral fusion and entry. FP is a short segment of 15-20 conserved amino acids, mainly glycine or alanine, plays an important role in mediating membrane fusion by disrupting and connecting lipid bilayers of the host cell membrane [94]. The S protein is extensively glycosylated, typically encoding 66 N-linked glycosylation sites per trimer and a lower number of O-linked glycans (<1%) present at low occupancy [8,50,95]. The S1 subunit spans N17-N801 and the S2 subunit spans N1074-N1194 glycosylation sites. The S2 subunit displays low oligomannose-type glycans and are highly conserved among variants [96,97]. The high density of viral glycosylation has prompted investigations into the functional role of glycans in a wide range of settings, from its impact on protein folding to its influence on immunogenicity.

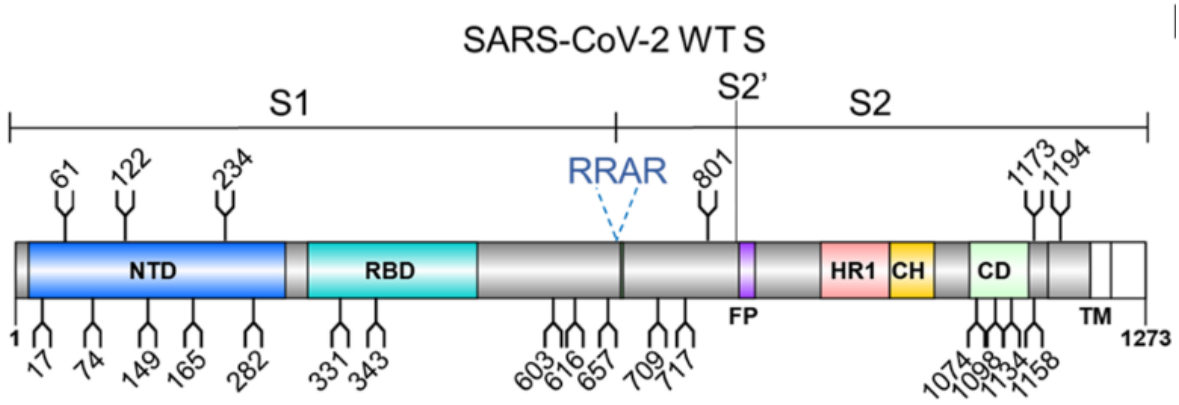


Figure 1.4. **Schematic of SARS-CoV-2 wildtype (WT) S protein.** NTD, N-terminal domain; RBD, receptor binding domain; FP, fusion peptide; HR1, heptad repeat 1; CH, Central helix; CD, connector domain; TM, transmembrane domain.

The S protein on the virus is a key factor involved in infection and is present in all human coronaviruses. The RBD of S1 subunit is involved in the receptor binding and cleavage of S1/S2 subunit mediates the fusion of virus with host cell membrane. RBD binds to ACE2, initiating conformational changes in the S2 subunit. This process involves the insertion of the FP into the target cell membrane, unveiling the prehairpin coiled-coil structure of the HR1 domain. This interaction culminates in the formation of a six-helix bundle (6-HB). This orchestrated molecular event brings

the viral envelope and cell membrane into close proximity, facilitating viral fusion and entry. The trimeric S protein is a key target for neutralizing antibody responses [98,99] and a key component of all currently approved vaccines [100–105]. S protein has also been used as reagent in serological testing [106,107], principally because of its primary role in viral infection and its exposure on the viral surface.

1.2.4 Spike protein in vaccine design

There is a wide range of S protein constructs across the currently approved vaccines, which rarely utilize native sequences, with exception of the Sinovac vaccine construct [102] which corresponds to a deactivated virus with a native S protein sequence. Other formats are based on a range of non-native sequences and even the effective adenoviral-based ChAdOx1 nCoV-19 (AZD1222) vaccine [108] encoding native S protein sequence is codon optimized to improve expression level in human cell lines. Other design features used within the trimeric S protein formats include C-terminal truncations to remove the hydrophobic transmembrane regions required for solubilization and the use of stabilizing mutations, such as in the 2P, HexaPro, and S-closed formats [109–111]. For example, the RNA-based Pfizer/BioNTech vaccine encodes a codon-optimized S protein with the 2P mutations [100]. The proline mutations are incorporated in these formats to maintain the S protein in pre-fusion conformation as there is enhanced presentation of neutralizing antibody (NAb) epitopes compared to the post-fusion conformation [112]. Notably, not all vaccine formats rely on trimeric S protein and the RBD-based constructs have shown promising immunogenicity [113] and safety [114].

The wide range of constructs at the basis of currently approved SARS-CoV-2 vaccines suggests that immunogenicity is not sensitive to precise glycan processing. In one extreme, robust immune responses have been generated from material derived from insects [115–117], plants [118,119], and mammalian expression systems [100,105]. Even in mammalian derived immunogens, wide range of cell-specific glycosylation is utilized from common recombinant expression systems such as HEK 293 to *in vivo* production by adenoviral and RNA-delivery systems [100,103]. In the latter case, it is anticipated that the bulk of production is located within phagocytic immune cells,

which would be expected to impact cellular-specific glycosylation [120]. Even the bulk change in glycosylation of S protein using chemical inhibitor, kifunensine which inhibits ER-mannosidaseI and arrest the glycan processing at oligomannose-type glycans , does not influence serological properties. Despite these observations, at site-specific level, glycans on S are critical for maintaining the quaternary structure [121] and thus indirectly influences immunogenicity and antigenicity. Furthermore, in HIV-1, immunogen glycosylation is known to enhance the humoral immunity by influencing antigen trafficking to immune cells [122]. Overall, the glycosylation of SARS-CoV-2 S protein has a weaker impact on the immunological properties compared to more densely glycosylated protein such as HIV-1 [123].

Beyond the impact of glycans on the immune response to immunogens, there is growing evidence that viral glycosylation contributes to the inflammatory response to viral material [1]. Thus, understanding the fine structure of SARS-CoV-2 S protein glycan shield on the virus and of different immunogens can contribute to the understanding of the role of glycosylation in natural infection and vaccine development. This section reviews the parameters that are structurally affecting the glycosylation of the S protein and assess the extent to which recombinant S protein mimics that of the native virus **(Figure 1.5)**.

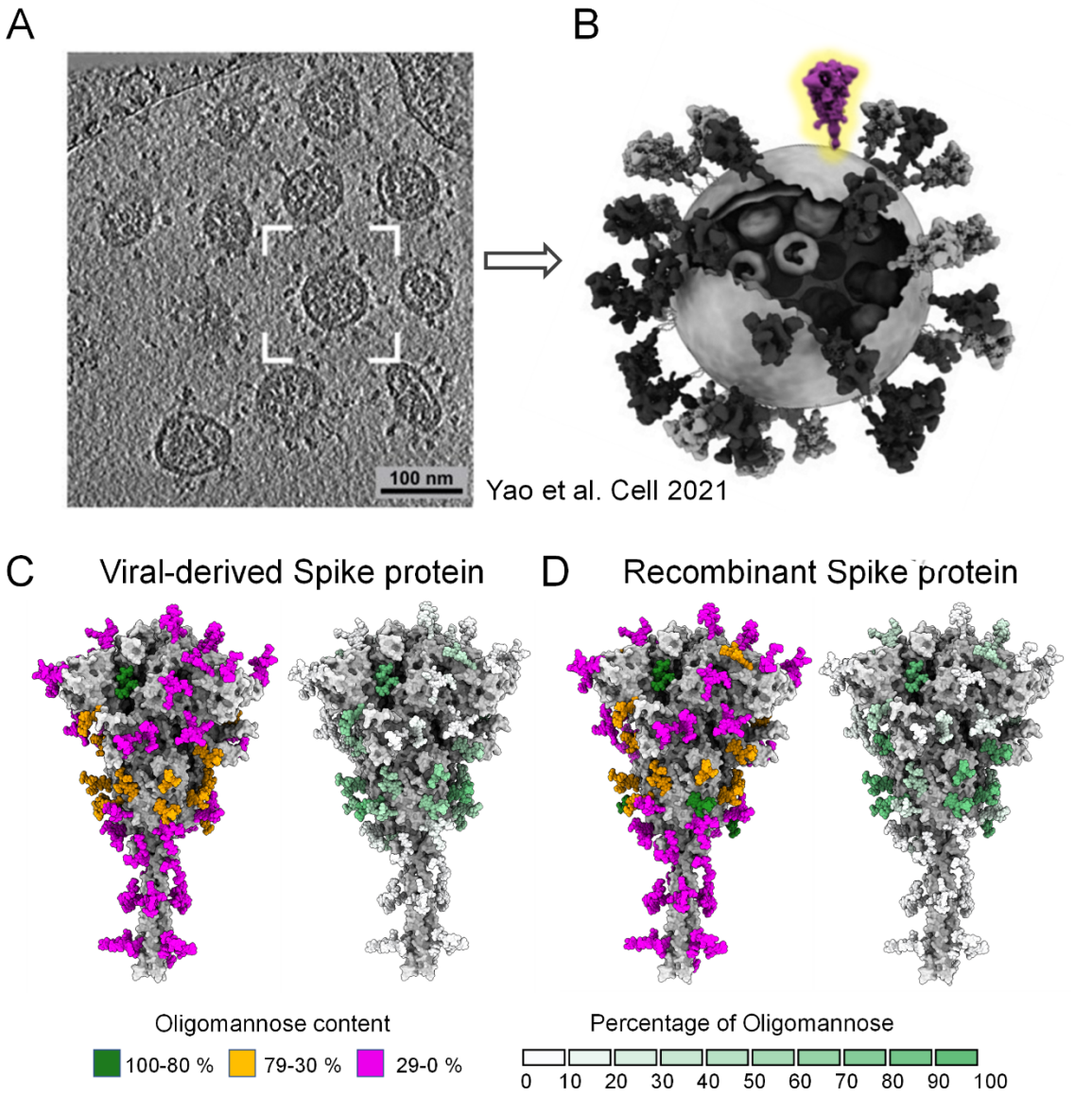


Figure 1.5. Glycosylation of viral-derived and recombinant spike protein.(A) Representation of cryo-electron tomograph (Cryo-ET) slice of SARS-CoV-2 virions reproduced from Yao et al. [124]. (B) Model of representative SARS-CoV-2 virus highlighting one S protein on the envelope, has been adapted to highlight a single spike from Yao et al. [124]. (C) Fine glycan processing of viral S representing content of oligomannose-type glycans and heat plot of percentage of oligomannose-type glycans of viral-derived S protein, using reported abundances [125,126]. (D) Glycan processing of recombinant S, values of which obtained from data produced earlier [125,126]. The glycan composition is categorized in three groups based on the abundance of oligomannose-type glycan content: green (100—80%), orange (79—30%) and magenta (29—0%). The heat plot represents the percentage of oligomannose-type glycans at each site from a scale of 0% (white) to 100% (green).

1.2.5 Role of N-linked glycans in protein folding and stability

The N-linked glycosylation of S protein is essential for folding and assembly in both recombinant formats and in viral settings. Glycosylation mediates interactions with the calnexin/calreticulin folding pathway and secretion checkpoint. Inhibitors targeting entry into the pathway, for example,

inhibition of α-glucosidases by iminosugar-containing/based compounds, such as *N*-butyldeoxynojirimycin (NB-DNJ) and *N*-butyldeoxygalactonojirimicin (NB-DGJ), have led to a marked reduction in viral infectivity [127]. Beyond the capacity for glycans to mediate chaperone interactions, they can also impart physicochemical stability. Furthermore, the high degree of N-linked glycosylation of the SARS-CoV-2 S protein suggests a selective pressure on the virus that has driven the density of the glycan shield. Among the many different key functions driving the evolution of the shield, glycans can be aiding S folding and structural integrity [121,128], facilitate the interaction with cellular factors [129,130], or be shielding underlying epitopes [121,131]. Because of this key functional role, the glycan shield evolved significantly along the phylogeny, and it is continually evolving [132], with the very recent loss of N370 glycosylation due to T372A mutation in SARS-CoV-2. Through molecular dynamics (MD) simulation, it has been observed that presence of both N234 and N370 results in tying the closed RBDs together, and likely hinders the RBD opening [132–134]. This possibly explains the absence of N370 glycan site in SARS-CoV-2. Similarly, we can speculate that SARS-CoV-2 glycosylation carries the signature of selective pressure from its host prior to zoonosis.

There is a direct relationship between the spatial accessibility of the glycosylation sites and the processing state [135,136]. Within the context of the S glycoprotein, MD simulations have shown that the N-glycan at position N234 is one of the least accessible, in both closed and open S conformations see **Figure 1.6**, which explains the high abundance of large oligomannose-type glycans at this site [8,121]. The N234 glycan is also one the most ordered glycans observed in cryo-EM structures [109] and this glycan has also been shown by MD simulations to stabilize the RBD open conformation and its dynamics [121,132]. MD simulations experiments [121] have shown that the N234 glycan is able to access and effectively fill the cleft left vacant by the opening of the RBD and thus supporting its stability. This function is also supported by the N165 glycan, located in proximity **(Figure 1.6)**. Indeed, deletion of these glycans through single point N234A, N165A and N343A mutations corresponds to a significantly reduced binding to the ACE2 receptor [121,128,137]. However, none of these single point mutations fully abrogate the binding suggesting that multiple residues are involved in RBD opening [121,128]. Further insight from MD and cryo-

electron micrograph (cryo-EM) studies have shown that glycosylation at N343 is essential for gating the RBD opening and closing [128,132]. The N234 and N165 shields the receptor binding motif (RBM) consistently in both 'close' and 'open' state whereas shielding by N343 decreases with RBD opening [128,132] **(Figure 1.6)**. Overall, the interaction of the clustered N165, N234, and N343 glycosylation sites with the surrounding protein helps stabilize the orientation and dynamics of the open RBD conformation allowing the different possible orientations and thereby supporting receptor recognition [121,128,132] **(Figure 1.6.B).**

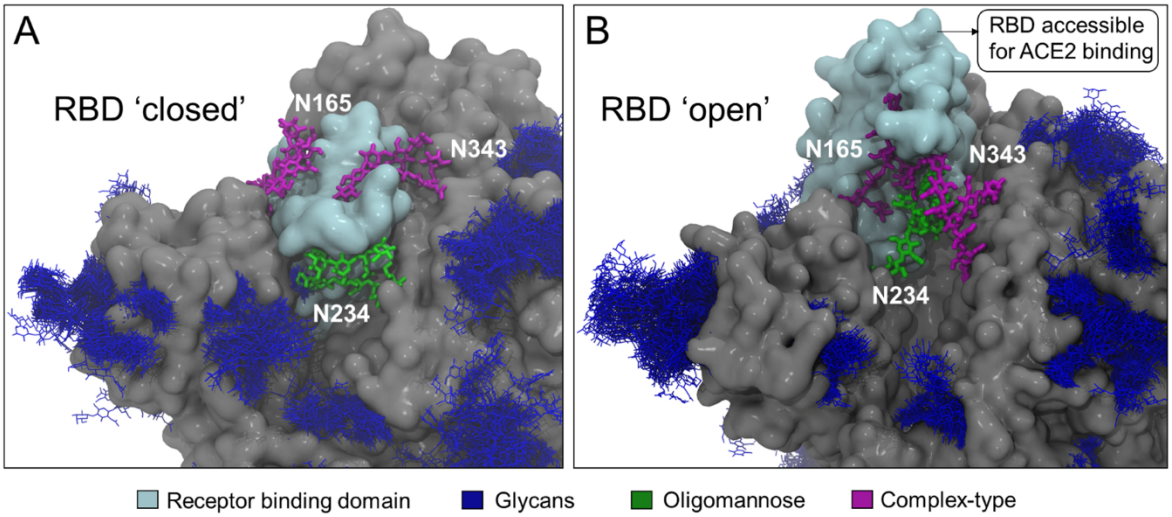


Figure 1.6. Role of glycans in stabilizing dynamics of RBD. Model representation of A) RBD 'closed' and B) RBD 'open' showing the interaction between neighbouring glycans, N165, N234 and N343. The RBD is accessible for ACE2 binding in 'open' state. The N343 glycan act as a "glycan gate" as it pushes the RBD from "down" to the "up" conformation. The N234 is modelled with $Man_9GlcNAc_2$ glycan represented in green, N165 and N343 are modelled with biantennary complex-type glycans [125,126]. The RBD is shown in cyan, and the remaining S protein highlighted in grey. These models are reproduced from previous studies [128,132]. This figure has been produced in collaboration with Elisa Fadda, Maynooth University, Ireland.

1.2.6 Role of spike protein structure in glycan maturation

The glycosylation processing of viral glycoproteins is highly dependent on the accessibility of the glycan sites to the glycan processing enzymes in the ER and Golgi apparatus [39,138]. The glycan maturation is strongly influenced by the glycosylation site accessibility and can sterically be impaired by the local architecture of the protein [39]. In contrast to N-linked glycans which are added co-translationally in the ER, O-linked glycans are added in the Golgi apparatus and their presence is influenced by both the intrinsic susceptibility of the sequence to modification but also the local steric accessibility. Importantly, the proximity of potential O-glycosylation sites to the furin cleavage site,

that generates the separate S1 and S2 subunits, means that their modification can impact viral maturation and infectivity [51].

One extreme example of steric constraints limiting glycan maturation within the S protein is the N234 site where limited enzymatic access results in almost exclusively oligomannose-type glycans being presented at that site [8,121,125] **(Figure 1.7.A)**. More broadly, comparison of S protein from wide variety of laboratories and expressed in different mammalian cell lines demonstrated consistent subpopulation of oligomannose glycans, indicating that this is an intrinsic property of the glycoprotein [125]. These mannose signatures were also observed in spike protein derived from infectious virus [125,126]. However, when only the S1 subunit is expressed, N234 is fully processed to complex-type glycosylation [126] consistent with the relaxation of steric restrictions to processing in the monomeric form [95,117,126] **(Figure 1.7.B)**.

Sialylation was another significant difference observed between S1 and trimeric S, the former expressing high glycan content of sialylated glycans with high branching compared to the latter which showed low sialylated glycan content including mostly mono-sialylated glycans [117]. Likewise, an increase in glycan content of O-glycosylation was observed in S1 recombinant protein compared to trimeric S. For example, T323 is more occupied in S1 recombinant protein, and T678 O-glycosylation was observed only on the isolated S1 subunit suggesting more accessibility to Golgi resident transferases [51,126] **(Figure 1.7.B)**. However, within the native trimeric sequence furin cleavage unmasks the T678 site which is modified around 25% with O-glycan structures [126]. Furthermore, when isolated RBD was expressed as a recombinant protein reveals high occupancy of O-glycosylation compared to trimeric S protein [95]. For example, T323 is fully occupied in monomeric RBD whereas this site shows low occupancy on trimeric S-protein [95]. In addition, the T470 O-glycosylation site was identified only on monomeric RBD, albeit at very low abundance [95] **(Figure 1.7.C)**. Similarly, the N-glycosylation sites, N331 and N343 of monomeric RBD revealed elevated branching and terminal processing when compared to trimeric S-protein [125]. These glycan characterizations of different viral protein formats highlight the impact of protein architecture on glycan structure. Although a wide variety of protein formats have shown to be

effective vaccines, the changes in glycosylation may impact the immunogen efficacy, as has been observed in case of HIV-1 [122].

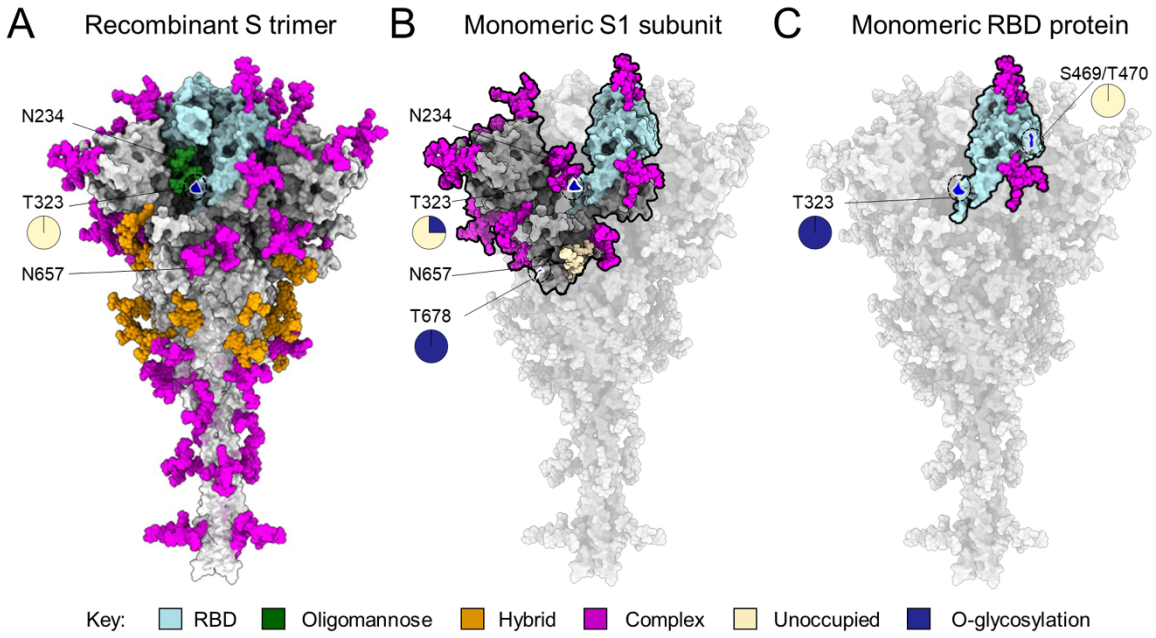


Figure 1.7. The influence of protein structure on glycan maturation.(A) Illustration of glycan composition of recombinant trimeric S protein of which values reproduced from Eldrid et al. [95].The N-linked glycosylation takes place at specific sequon, Asn-X-Ser/Thr (X is any amino acid except proline) whereas O-linked glycosylation is not dictated by specific sequon and occurs on serine and threonine in exposed regions. The N-linked glycosylation is presented in three categories on basis of oligomannose content as described in Figure 1, oligomannose (green), hybrid (orange) and complex-type (magenta) glycans. The O-linked glycosylation at T323 site (see magnification) on trimeric S is present at low levels (0.2%) of which values obtained from Eldrid et al. [95]. (B) The glycan composition of recombinant monomeric S1 subunit of which values reproduced from Wang et al. [117] and Brun et al. [126]. Most of the N-glycan sites on S1 subunit is highly processed represented in magenta except N657 which is unoccupied, represented in wheat color. The O-glycosylation is present on S1 subunit at sites, T323 and T678 (see magnification). (C) The glycan composition of monomeric RBD protein (cyan) which binds to main host receptor (ACE2). The N-glycan sites of RBD is highly processed represented in magenta of which values reproduced from Allen et al. [125]. The O-glycosylation was observed on monomeric RBD protein at sites T323 and S469/T470 [95].

1.2.7 Role of SARS-CoV-2 S glycans in viral entry

The initial step for SARS-CoV-2 infection in the host cell is to bind with the host cell receptor and enter the target cells. While additional factors can enhance entry (see below), entry is principally mediated by S-ACE2 interaction [50,83,93]. ACE2 is ubiquitously expressed on epithelial, endothelial, and blood cells [139]. The viral S protein and ACE2 receptor both are extensively glycosylated, with ACE2 exhibiting 14 N-linked glycans across the dimer [8,50,109]. O-glycosylation was found at low levels on both ACE2 and S protein [50,95,126]. Previous studies have determined the effect of N and O-glycosylation processing of both S and ACE2 on S-ACE2

binding [140,141]. The glycan engineering of ACE2 did not significantly impact the ACE2-S binding [140,141] whereas that of S glycosylation had modest influence on binding [141]. In contrast to these recombinant settings, glycan processing of S protein can influence viral entry [141].

Beyond the impact of glycan maturation on receptor recognition, modification of the furin cleavage by O-linked glycosylation can impact viral entry. For example, P681H and P681R mutation found in highly transmissible alpha and delta variants respectively, decreases O-glycosylation which potentially increases furin cleavage and may influence viral infectivity [51]. Together these results have highlighted the importance of S protein glycosylation on viral entry and abrogation of these glycans using inhibitors provides insight into intervention strategies to target the SARS-CoV-2 infection [142,143].

While ACE2 stands out as the most extensively characterized receptor for SARS-CoV-2, studies have revealed the virus's capability to infect cell lines lacking detectable ACE2 expression [144–146]. Recent research has identified TMEM106B, a lysosomal transmembrane protein, as an alternative receptor facilitating SARS-CoV-2 entry into ACE2-deficient cells [147]. TMEM106B directly interacts with the RBD of the S protein, and the substitution of E484D in the spike enhances its binding to the TMEM106B receptor. Despite the essential role of ACE2 in infection, carbohydrate-based attachment factors come into play, particularly when ACE2 expression is limited [148]. This phenomenon expands viral tropism beyond ACE2 expression alone. Notably, cell surface heparan sulfate (HS) serves as a recognition site for the SARS-CoV-2 S protein, distinct from the ACE2 binding site, enhancing infection [149–151]. HS binding to RBD promotes an open conformation, exposing RBD for ACE2 interaction and viral infection [149–151]. Additionally, sialylated glycolipids interact with RBD, facilitating viral entry [152], a phenomenon also exploited for diagnostic purposes [153].

Additionally, viral glycans can also be recognised by soluble and cellular lectins which play an important role in recognition. Nuclear magnetic resonance (NMR) studies and MD simulations have revealed the binding of human lectins such as galectins with N331 and N343 glycan sites of RBD [154]. Mannose-binding lectins (MBL) can recognize glycans on S protein, and have shown role *in vitro* in inhibiting SARS-CoV-2 infection by activation of lectin mediated complement pathway

[155]. Furthermore, there are cellular proteases such as transmembrane serine protease 2 (TMPRSS2) which enhance viral infectivity by cleaving both S and ACE2 protein [83]. These attachment factors mediate the viral S protein interaction with ACE2 and their inhibition reduces viral infectivity [83,150,151,156].

1.2.8 Antibody responses in SARS-CoV-2 infection

The antibody response is an important form of adaptive immunity against viral infections [157] . Majority of SARS-CoV-2 infection revealed seroconversion within 2 weeks post-symptom onset (PSO), producing IgG and IgM against the viral spike and nucleocapsid proteins. The rapidity of antibody responses suggests that the antibodies have undergone few somatic hypermutation [158]. This neutralizing antibody usually develops in extrafollicular phase (EF), when B cells quickly differentiate into plasma cells outside of follicles. Although the EF response give rise to neutralizing antibodies, these antibodies are probably not sufficient to control the SARS-CoV-2 infection [157,159]. As the prevention of reinfection depends on long-lived plasma and memory B cells, which requires robust germinal centre responses [160]. In the GC phase, lasting for weeks, B cells undergo somatic hypermutation and affinity-based selection, giving rise to predominantly isotype-switched, high-affinity plasma cells. These cells form a long-lived compartment in the bone marrow.

After SARS-CoV-2 infection, antibodies produced by early memory B cells may initially have few mutations [161]. However, over the subsequent months, the spike- or RBD-specific memory compartment continues to evolve, resulting in antibodies with increased somatic hypermutation, neutralizing potency, and breadth [98,99,162,163]. Spike-specific plasma cells persist in the bone marrow a year after infection, indicating the establishment of a long-lived plasma-cell compartment. Robust T cell responses, including spike-specific CD4+ T cells, are induced by SARS-CoV-2 infection, promoting antibody responses [164]. Studies in a macaque model and SARS-CoV-2-seropositive organ donors suggest that infection leads to the formation of functionally robust germinal centers (GCs) that may persist for months, possibly due to antigen persistence [165]. Cross-reactive memory B cells from previous exposure to seasonal coronaviruses may also contribute to

the GC response to SARS-CoV-2, supported by the higher somatic hypermutation levels in antibodies against S2, a highly conserved region among human coronaviruses [166].

1.2.9 Exploiting breaches within the glycan shield

Despite the extensive glycosylation and dynamics of the S protein, antibodies can still recognize breaches within the glycan shield of S protein and can neutralize the virus [121,167]. Previous studies have demonstrated the glycan composition and accessible surface area (ASA) of these glycans using liquid chromatography mass spectrometry (LC-MS) and MD simulations respectively [8,121,125,167]. Despite the extensive coverage of glycans on the head domain, the glycans are sufficiently sparse for there to be still extensive vulnerabilities to antibody binding [121,167] **(Figure 1.8.A)**. In contrast, the glycans in the stalk region have shown effective glycan shielding in MD simulation studies suggesting less accessibility for antibodies to bind this region [121,167]. Recent studies have shown the isolation of antibodies against the head region including RBD, Receptor binding motif (RBM), N-terminal domain (NTD) and S2 subunit from COVID-19 patients [163,168–170]. The antibodies against RBD such as C002 and SM211 have shown recognition of quaternary epitope on the RBD. The C002 neutralizing antibody (nAb) binds to a region which spans two RBD's, one in 'down' conformation adjacent to RBD in 'up' conformation **(Figure 1.8.B)**. The S2M11 antibody binds to two neighbouring RBDs and stabilize the close conformation of trimeric S protein **(Figure 1.8.B)**. Furthermore, Abs have also been elicited against NTD present on the head of the S protein [163,168,171]. One such antibody is 4A8 highlighted in Figure 4B, which recognizes the NTD in an identical region despite the dynamics of RBD and have also shown to stabilize the NTD region. The N149 glycan site which is present in NTD might also be involved in NTD-4A8 interaction [168] **(Figure 1.8.B)**.

In addition to the view of accessibility derived from static models, considering the dynamics of the system further enhances our understanding of the antigenic surface. Dynamic motions of the S protein can influence the glycan-glycan interaction and can bring the glycans of different domains in close contact. Using the network analysis approach, the glycans responsible for effective shielding were predicted [167,172,173]. The glycans, N234 and N165, involved in RBD dynamics

demonstrated stronger glycan-glycan interaction in the open state compared to the closed state as shown in Figure 2B. The glycans at sites N603 and N616 have been shown to play a central role in connecting the upper head with the lower head as these glycans are responsible for the proper shielding of the S protein [167]. Neutralizing antibodies exploit spaces between these shielded region where the glycan density is low or bind to epitopes containing glycans as has been observed in case of HIV-1 **(Figure 1.8.B)** [25,172].

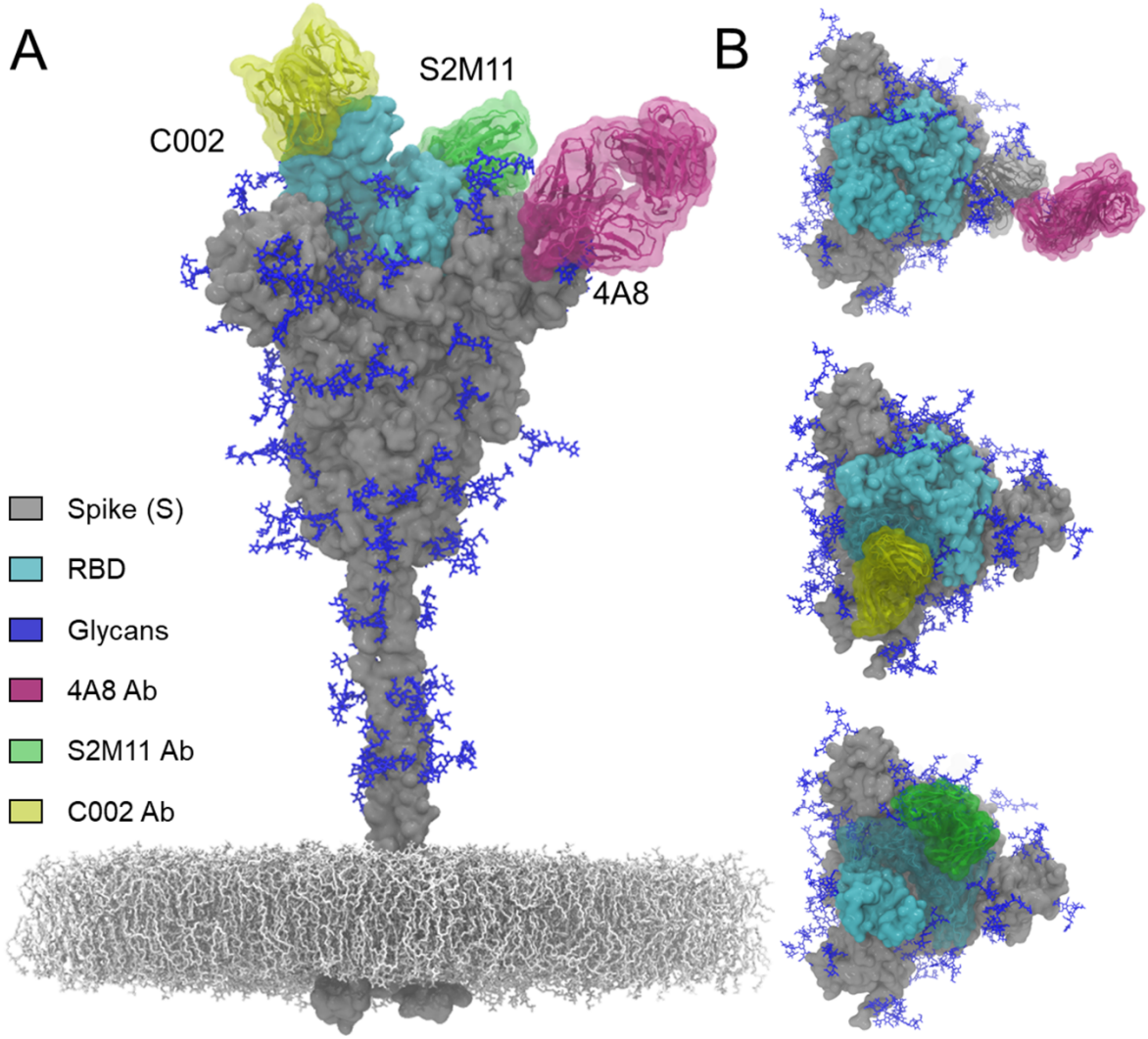


Figure 1.8. Antibody recognition within the breaches of S protein glycan shield. (A)Representation of trimeric S protein with glycans (blue) present on it [121]. The RBD present on the head region of the S is highlighted in cyan. The stalk region is effectively shielded with glycans. The mesh network represents the virion membrane on which S protein is embedded. (B) Illustration of antibodies binding within the glycan shield of S protein obtained by structural alignment (www.pymol.org). The 4A8 nAb (pink, PDB 7C2L) targets the NTD region of the S protein [168]. The C002 neutralizing antibody (nAb) (yellow, PBD 7K8T) is binding to RBD "up" conformation [170]. The S2M11 (green, PDB 7K43) nAb recognizes a quaternary epitope consisting of two neighbouring RBDs and stabilizes the trimeric S in closed state [169]. This figure has been provided by Elisa Fadda, Maynooth University, Ireland.

1.2.10 Further glycobiology perspectives in vaccine development and beyond

It has been established that underlying protein architecture has significant impact on glycosylation [39]. In one extreme trimeric viral spike exhibit oligomannose-type glycans whereas RBD domains are almost entirely devoid of such structures [8,121,125]. Given the known influence of glycan processing on immunogen targeting, it will be fruitful to understand the impact of protein formats on vaccine efficacy. Similarly, there is considerable opportunity to enhance the understanding of RNA-based [100] and adenoviral-based delivery system [103] by considering the range of vaccine materials derived from different cellular sources upon immunization [126,174]. It is already established that protein modification can influence antigen stability, and it is likely that such modifications will also impact antigen distribution and cellular presentation [122,175,176]. Overall understanding the interplay between protein architecture, production routes, and glycan processing will illuminate how viral immunogens work and aid in their optimization.

1.3 HIV-1 Env glycosylation

1.3.1 Emergence of HIV-1/AIDS

Acquired Immune Deficiency Syndrome (AIDS) was first recognized as a new clinical challenge in 1981. Initially, the unusual increase of diseases like Kaposi sarcoma (KS) [177] and *Pneumocystis pneumonia* (PCP) [178] were reported. The occurrence of these disease often acts as indicators of severe immunodeficiency. Men who had sex with men (MSM) were the initially affected population and were highly stigmatized. It was in 1982 that the human immunodeficiency virus (HIV) was identified as the true etiologic agent of AIDS. Although it was initially linked with the homosexual community, it was soon detected in intravenous drug users, individuals from Haiti that had recently entered the USA, and individuals who had received blood transfusions. These observations, together with the observation of possible mother-to child transmission, and the occurrence of the disease among a group of homosexual men in California, led to the proposal that the immunodeficiency syndrome was caused by an infectious agent. The initial description of HIV-1 was given by Barré-

Sinoussi and Robert Gallo in 1983 [179,180]. HIV is classified in the family Retroviridae, subfamily Lentivirinae, and genus Lentivirus.

The lentivirus genus is usually characterized by a long incubation period of infection and the ability to infect non-dividing cells [181]. As a retrovirus, HIV possesses an RNA-based viral genome. Once inside a host cell, RNA undergoes reverse transcription into DNA, which is subsequently incorporated into the host cell genome through the action of an integrase enzyme. Subsequently, the infected host cell engages in the translation of viral proteins, culminating in the production of additional viral particles. The immune deficiency caused by HIV stems from its ability to invade and destroy CD4+ T cells, progressively reducing their count over several years [182]. In the absence of intervention, these levels eventually drop beneath a critical threshold, compromising cell-mediated immunity. In individuals with compromised immunity, opportunistic infections like *Pneumocystis pneumonia* prove fatal, whereas in healthy individuals, the immune system effectively clears such infections, averting the development of severe symptoms.

Despite AIDS being a significant global health concern, the development of a vaccine capable of providing comprehensive protection against HIV infection remains elusive. This can be attributed, in part, to the genetic diversity that arises from the retrovirus's use of an error-prone reverse transcriptase during the conversion of viral RNA into DNA. This results in presence of many quasispecies of HIV in an infected host [183]. Consequently, the conservation of protein epitopes is compromised, rendering any strategy solely reliant on these epitopes susceptible to the rapid mutations of the HIV population, which quickly nullify selective pressures.

Phylogenetically, HIV-1 can be classified into group M (major), N (non-major, non-outlier) and O (outlier). Group N infections have been detected in only a few individuals in Cameroon, whereas group O is present in Cameroon and a few surrounding countries [184]. Group M, on the other hand, accounts for over 99% of infections and is the most variable, with extraordinary diversity in the envelope sequence among isolates. M-group (main) can be further classified into clades A, B, C, D, F, G, H, and J, alongside circulating recombinant forms (CRFs) [185]. Consequently, an effective vaccine must elicit protection against the diverse clades circulating worldwide.

1.3.2 Viral genome and life cycle of HIV-1

The RNA component of HIV-1 is 9.8 kb long and bears a 5′ cap, and a 3′ poly (A). The HIV-1 genome encodes several structural and non-structural proteins that perform a variety of roles to aid in viral replication. The virus contains two copies of RNA genome, which encodes for 9 genes (*gag, pol, env, tat, rev, nef, vif, vpr, and vpu*) [186]. These genes are flanked by long terminal repeats (LTRs) that control viral transcription and integration into the host genome. *gag, pol, and env* genes are primarily involved in the production of viral proteins. The *gag* encodes for structural proteins, matrix, capsid, and nucleocapsid that forms the core of the virus [187]. The *pol* encodes enzymes necessary for replication and integration, including reverse transcriptase (RT), integrase (IN), and protease (PR). The *env* encodes the viral glycoprotein (Env), which allows the virus to bind to and enter the target cells. All of these genes contribute to the integrity of the virus. The Env is translated as a single gp160 polypeptide, which is cleaved by the host proteases into external glycoprotein, gp120 and a transmembrane spanning protein, gp41 [188]. This cleavage occurs during the transit from the ER to the Golgi. These gp120-gp41 heterodimers assemble to form a trimer on the mature virion. It is extensively N-glycosylated and plays an important role in viral tropism and mediates infection.

During viral entry, the gp120 of Env attaches to the CD4+ (cluster of differentiation) receptor on the surface of target cells (T helper cells, macrophages, and dendritic cells). This interaction triggers a conformational change within gp120, allowing the virus to bind to a co-receptor, usually C-C chemokine receptor 5 (CCR5) or C-X-C motif chemokine receptor 4 (CXCR4) [189]. CCR5-tropic viruses are associated with most infections, and individuals with a homozygous defect in this gene are generally considered resistant to HIV-1 infection [190]. Upon the attachment with co-receptor, the gp41 undergoes conformational changes which results in extend and exposure of its fusion peptide, which then inserts into plasma membrane of target cells leading to the release of the capsid into the target cell cytoplasm [191]. The schematic of HIV-1 viral life cycle is described in **Figure 1.9**.

Once inside the host cell, the viral RNA genome is reverse transcribed into double stranded complementary DNA by reverse transcriptase (RT). This step is extremely error-prone and lead to

the introduction of mutations in the viral Env due to a lack of a proof-reading mechanism [192]. Following the transcription, the complementary DNA migrates to the nucleus and remains associated with multiple viral proteins, including integrase (IN), matrix (MA) and the accessory viral protein (vpr). Integrase assists in the integration of viral DNA into host genomes [193]. This process is usually associated with the latent infection of HIV-1 [194]. The viral DNA then uses the host translation and transcription machinery to produce new infectious virions. *gag, pol,* and *env* are the major genes which span the majority of the viral genome. The Gag and Pol get translated into the cytoplasm, whereas the Env enters the secretory machinery of the ER and Golgi and utilize host glycosylation machinery and thus gets heavily glycosylated. The new viral particles are assembled at the cell membrane, encapsulating RNA and proteins. The viral proteases cleave the newly synthesized viral proteins into their functional forms, and mature virions are then released from the host cells and spread to other cells.

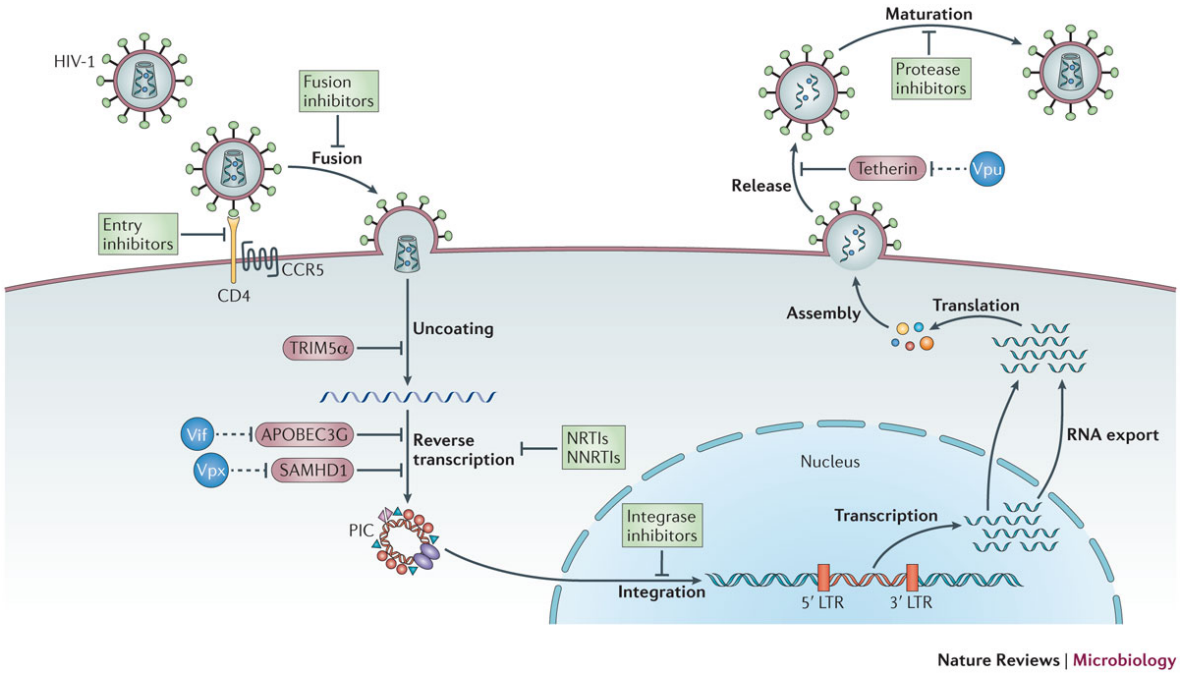


Figure 1.9. Schematic review of HIV-1 replication cycle. Image source: Nature Reviews Microbiology [195]. LTR, Long terminal repeats; PIC, Pre-integration complex; Vpu, Viral protein U; Vif, Viral infectivity factor; Vpx, virion-packaged protein.

1.3.3 HIV-1 Env Structure and function

The HIV-1 Env is a type I viral membrane fusion protein and is involved in the interaction of virus with host cells by mediating entry into target cells. Env is a metastable assembly of three

heterodimeric complexes composed of gp120 and gp41 subunits, which forms a trimer [196]. The gp120 comprises a highly variable surface, characterized by five variable loops interspersed with five constant regions (**Figure 1.10.A**). The gp41 subunit is more conserved and comprises an ectodomain, a membrane proximal external region (MPER), a transmembrane domain (TM), and a cytoplasmic tail (CT).

Due to the conformational dynamics and high density of N-glycans on HIV-1 Env, structural studies aiming to solve the structure have been relatively difficult. Despite the challenges, the majority of the structural data is produced using the variations of recombinant Env constructs. The first published construct featured the monomeric gp120 core in complex with CD4 [197]. This construct, along with numerous subsequent studies, unveiled the fine structural details of the Env, and receptor bound conformation of gp120 [197–200]. In order to recombinantly mimic the trimeric Env presenting epitopes of broadly neutralizing antibodies as found on native virion, solubilization and stabilization strategies were employed, using post-translational modifications and specific amino acid substitutions. A breakthrough discovery of recombinant Env has been the emergence of the 'SOSIP' format containing several subtle modifications to generate a stable and soluble trimer mimic [201]. The modifications include the introduction of an artificial disulfide bond between gp120 and the ectodomain of gp41 (SOS), and an isoleucine to proline mutation (I559P) in gp41 which stabilize the prefusion conformation of Env [202,203]. These immunogens are referred to as SOSIPs and are truncated at position 664 to generate a soluble construct of Env, SOSIP.664 (**Figure 1.10.B**). This approach was first successfully applied to the Env from the Clade A transmitter founder virus, BG505 [204]. Sanders et. al, additionally introduced the N332 site lacking from the BG505 isolate, as this site forms a target for numerous bnAbs [201]. BG505 SOSIP.664 is currently the most studied HIV-1 recombinant immunogens in vaccine development and has been elucidated at high structural resolution [173,205–207]. This stabilization approaches employed to BG505 strain has been applied to stabilize the Envs from multiple clades [25] .

The quaternary structure of the natively folded viral spikes requires the furin cleavage of the gp160 pro-protein, which occur in the trans-Golgi network [208]. Furin cleaves between the gp120 and gp41 subunits of the gp160 polypeptides, recognizing the amino acid motif R-X-R/L-R [209].

However, the low levels of endogenous furin result in large populations of uncleaved trimers, which have been evidenced from NS-EM data of viral spike mimetics [210,211]. These uncleaved trimers were found to adopt non-native "open" conformations [211–213]. Moreover, the lack of furin cleavage perturbs Env's structural integrity, and have implications for the post-translational modifications and antigenicity of candidate immunogens [18,214]. For example, the destabilization of the quarternary structure influence the binding of quarternary-specific bnAbs such as PGT145 and PGT151 [215,216]. Binley et. al, circumvented this problem by co-transfecting plasmid containing the furin gene concomitantly with Env [201]. The furin cleavage is also replaced by a more susceptible furin cleavage site (R6) (**Figure 1.10.B**).

Furthermore, additional strategies were utilized to bypass the furin cleavage site as it removes the co-transfection with furin. To address this, native flexibly linked (NFL) HIV Env trimers were designed which contains substitution at the furin cleavage site REKR with a long glycine-serine rich linker 2xG_4S peptide that covalently links gp120 and gp41 [217–219] (**Figure 1.10.B**). The NFL design contains all other mutations present in SOSIP, including the I559P substitution, disulfide bond between gp120 Ala_{501} to gp41 Thr_{605} and a truncation at D664 to create a soluble trimer. The crystal structure of this cleavage-independent HIV Env exhibits a native-like prefusion structure closely resembling BG505 SOSIP.664, with NFL displaying glycosylation and antigenic properties that mirror those of the SOSIP trimer [217].

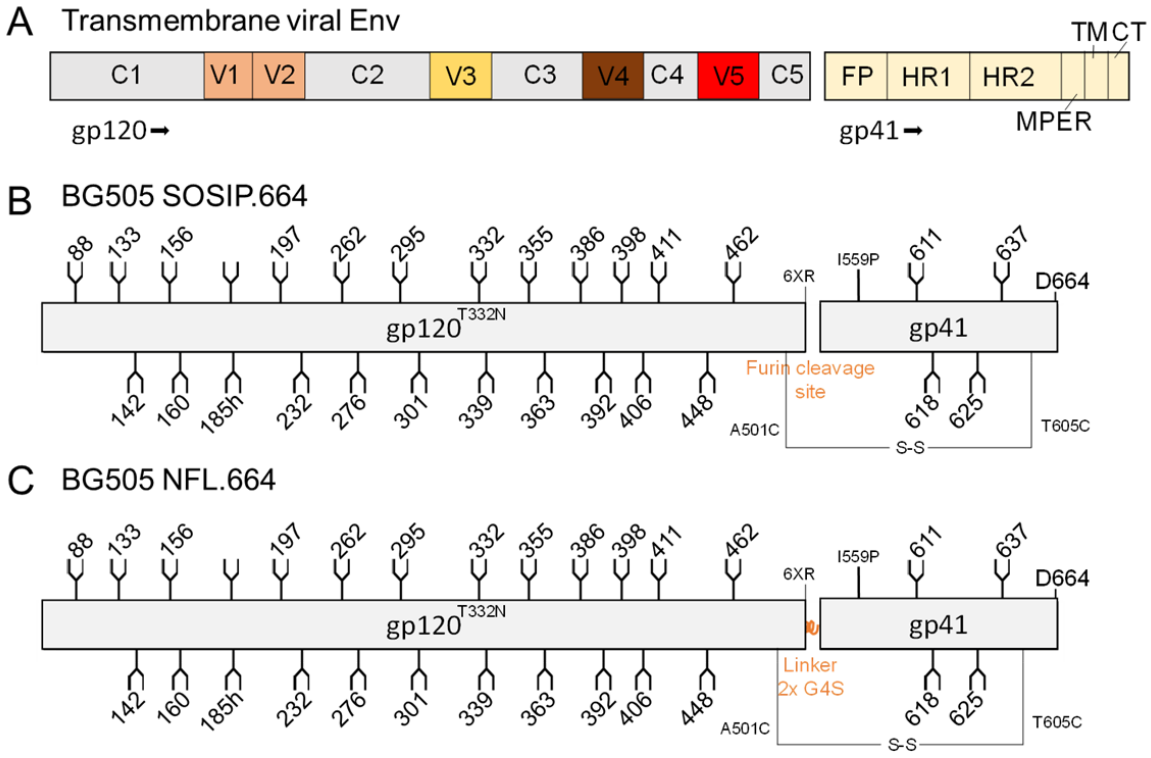


Figure 1 10. Schematic representation of Env structure. A) Primary structure of gp120 and gp41. Variable regions (V1-V5) of gp120 are colored, constant regions (C1-C5) are represented in grey. FP, fusion peptide; HR1, HR2, heptad repeat helices 1 and 2; MPER, membrane proximal external region; TM, transmembrane domain; CT, cytoplasmic tail. B) Schematic representation of the BG505 SOSIP.664 construct including PNGSs, and all modifications incorporated. C) Schematic representation of single chain natively flexibly linked Env, BG505 NFL.664, with all PNGSs and modifications.

1.3.4 Role of glycans on Env in maintaining structural integrity

HIV-1 Env is characterized by an unusually high presence of N-linked glycosylation sites, resulting in steric constraints within the glycan shield and, hence, a large abundance of under processed oligomannose-type glycans, which displays lower structural variation compared to host cell. The gp120 is heavily glycosylated, with the number of N-linked glycans ranges from 18-33, shielding the underlying protein from immune surveillance [28,220]. The glycan shield continually evolves to impose steric hinderance on newly produced antibodies, resulting in non-random shifts in the position of their N-linked glycans [221]. Although these N-linked glycans derive from the host glycosylation's machinery, studies on gp120 have unveiled an unusually high and conserved population of oligomannose-type glycans [222–225]. The high abundance of glycans creates steric hinderance for the actions of ER and Golgi α-mannosidases responsible for trimming the

$Man_9GlcNAc_2$, leading to the presence of clusters of oligomannose-type glycans, the 'intrinsic mannose patch' (IMP) [222,226] (**Figure 1.11**).

The IMP is a characteristic feature of gp120 that is conserved across all HIV-1 clades [222,226] and longitudinally during infection [227]. These glycans are known to remain highly conserved, as they are linked to the quaternary structure of the protein. One such example is N262, which is largely embedded in the protein cleft and presents a high abundance of less processed $Man_9GlcNAc_2$ structures [207,228,229]. In addition to N262, there are several N-glycan sites, including N293, N332, N339, N363, and N386, which form oligomannose clusters and have been observed to be highly conserved on recombinant gp120 monomers and native-like trimers [214,222].

In contrast to gp120, the oligomannose-type glycan content observed on native-like recombinant soluble [18], membrane-associated [230], and virion-derived Env [3], displays even higher oligomannose-type glycan content. This has led to the emergence of an additional 'trimer associated mannose patch' (TAMP) (**Figure 1.11**). These oligomannose-dominated glycan sites usually arise due to the glycan-glycan and glycan protein interactions in a natively folded trimer [231]. However, TAMP sites do not always constitute a complete processing state presenting oligomannose-type glycan content, but represents an elevation in under processed glycan compositions. The TAMP sites are a characteristic feature of cleavage-induced, native-like envelope folding and are not a feature of uncleaved or incorrectly folded trimers [214,232]. Therefore, the conservation of both IMP and TAMP sites can signify that the trimer is properly cleaved and assembled in a trimeric form, which is an important feature for maintaining the structural integrity of HIV-1 Env [18]. This highlights the importance of determining the site-specific glycan composition of immunogens, as it can be used to monitor the trimers with native-like conformation.

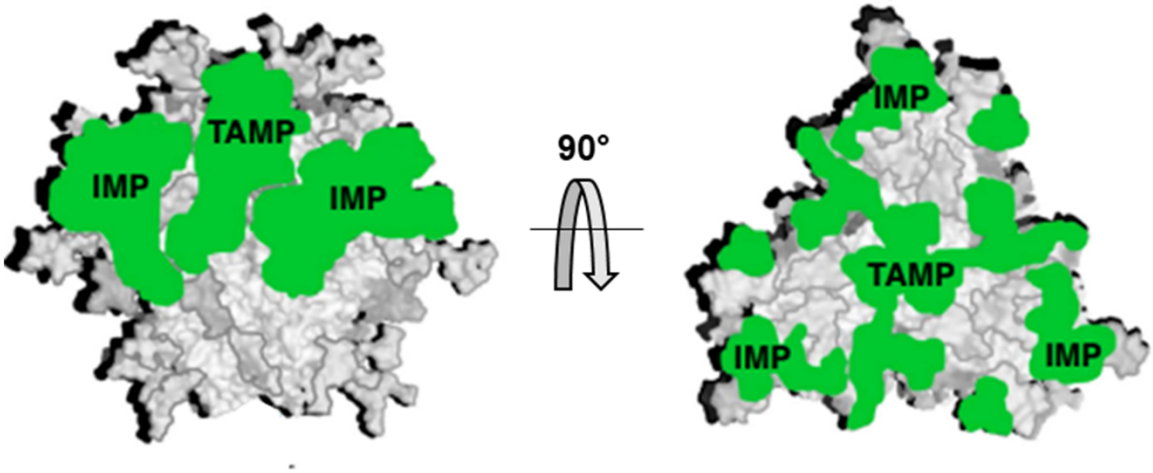


Figure 1.11. Localization of intrinsic mannose patch (IMP) and trimer associated mannose patch (TAMP) on Env. IMP and TAMP represents two regions of underprocessed oligomannose-type glycans that are key features of native-like Env trimers. Figure adapted from Behrens et al. 2017 [232].

1.3.5 Antibody response in HIV-1 infection

HIV-1 virion surface is covered with Env which makes it the sole target for elicitation of antibodies. The initial antibody response to Env manifests with in the first two weeks of infection but is non-neutralizing [233,234]. This is due to the metastable structure of the Env, the virus produces non-functional spike such as uncleaved gp160, shedding of gp120 and gp41 stumps [235]. These non-functional counterparts present improper assembly and glycan processing of the spike which is capable to divert the antibody response towards non-neutralizing antibodies. The abundance of non-neutralizing antibody response thus dilutes the neutralizing antibody response elicited by the functional spikes.

After the first 2-3 months, the infection starts progressing and the autologous neutralizing antibodies starts to appear. Autologous neutralizing antibodies are usually potent; however, they are very strain specific and drive the selection of virus and results in escape mutants. In some chronically infected patients, cross neutralizing antibodies have been observed which has a significant breadth and neutralization potential [234,236]. Cross-neutralizing antibodies are the ones which can neutralize diverse HIV-1 isolates and should be elicited by an effective vaccine. Moore et al., observed the development of broadly cross neutralizing antibody, PGT128, targeting N332 glycan

epitope which was absent on the initial infecting virus [237]. This represents one of the examples of the evolution of antibodies with escape mutants to control the viral infection .

To assess the overall breadth and potency of the neutralizing antibody reactivity , a panel of pseudoviruses from all major circulating subtypes were distinguished based on their sensitivity to antibody-mediated neutralization [238]. The viruses were categorized in four subgroups on basis of patterns of sensitivity as shown in **Table 1.1**. This tiered classification is to define the relative susceptibility of viruses to NAbs. The Tier1 A strains are mostly lab-adapted strains and are highly sensitive whereas Tier1B viruses represent mostly circulating, though neutralization sensitive [239]. Tier1 viruses seems to present more 'open' structure which makes it more susceptible. In contrast, Tier2 viruses presents more compact structure of Env based on the neutralization resistance to patient sera. The major goal in development of vaccine design is to elicit NAbs that can elicit Tier 2 viruses.

Table 1.1 Categories of virus subgroups on the basis of sensitivity to antibody neutralization.

Order of Viruses	Virus resistance	Antibody-mediated neutalization
Tier 1A	Vey sensitive	very high
Tier 1B	Sensitive	above average
Tier 2	Monderate resistance	Moderate
Tier 3	Highly resistant	Low

There are various factor which hinder the development of NAbs including poor accessibility of the conserved region such as CD4bs, high accessibility of variable regions such as V1/V2 and V3 and massively glycosylated surface of Env which occludes half of the protein surface [240]. Along with these factors, the variability in HIV-1 Env due to the presence of error-prone transcriptase and the evolution of viruses to escape from antibodies creates extreme antigenic diversity, which poses a significant challenge for the development of effective vaccines.

Broadly neutralizing antibodies (BnAbs) are a class of antibodies that have the ability to neutralize the wide range of HIV-1 strains. The maturation of BnAbs takes several years as it requires multiple rounds of viral escape and antibody development [241] . BnAbs presents unusual features such as unusually long or short CDRH3 antigen binding loops, and high levels of somatic hypermutation [240]. The long CDRH3 region is usually found on antibodies targeting V1/V2 region of the Env such as PGT145 which contains HCDR3 region of 33 residues, required to penetrate the

N160 glycan at the apex of the trimer to contact both protein and glycan residues [242]. Owing to these distinctive attributes, traditional vaccine approaches face challenges in eliciting broadly neutralizing antibodies (bnAbs). To address this, a germline-targeting vaccine priming strategy has been developed, aiming to activate specific precursor B cells targeting conserved residues on the Env such as CD4bs [243,244]. This approach involves selecting productive somatic mutations resembling bnAbs and generating memory B cells, which can be subsequently boosted to prompt further productive mutations.

1.3.6 Sites of vulnerability on Env for antibody binding

The generation of effective B cell responses is hindered during infection due to various structural attributes of the virion surface. One such factor is the heavily glycosylated viral spike, which is weakly immunogenic as it is derived from host cell's glycosylation machinery, and the high density of glycans limits the accessibility of antibodies to neutralizing epitopes [231]. Nevertheless, there are structural features of the quaternary trimeric Env that are less accessible, such as CD4 binding regions and inter-protomer regions (MPER) [245]. Furthermore, the large spacing between envelope spikes is unfavourable for B cell activation, which requires antigen crosslinking of B cell receptors [246].

Although the glycan shield constantly evolves to evade the immune system, the number of PNGSs on HIV-1 Env remains relatively constant. HIV-1 Env usually presents 18-30 PNGSs site, which cover half of the mass of the trimeric Env surface. Despite shielding the viral spike, these glycans also serve as an epitope for many BnAbs [246]. There are few regions known to be conserved on Env that presents sites of vulnerability for antibody binding, and the majority of these regions constitute bnAb epitopes involving partial or complete interaction with glycans presented on Env. These regions include the mannose patch, the CD4 binding site, the MPER of gp41, the V3 region, the V1/V2 region, and the interprotomer regions between gp120 and gp41 (**Figure 1.12**) [247]. One such example of glycan recognizing antibody is PGT128 which binds to the high-mannose glycan at N332 site, where antibody interacts with all the Man_9 glycans. A complex structure with an engineered outer domain of gp120 illustrated that PGT128 binds with GDIR motif in V3 loop and interacts not only with N332, but also with core sugars of N301site [248,249] (**Figure 1.13.A**).

PGT128 also indirectly contacts the N262 glycans as it binds to N301 glycan which in turn make contact with N262 [206].

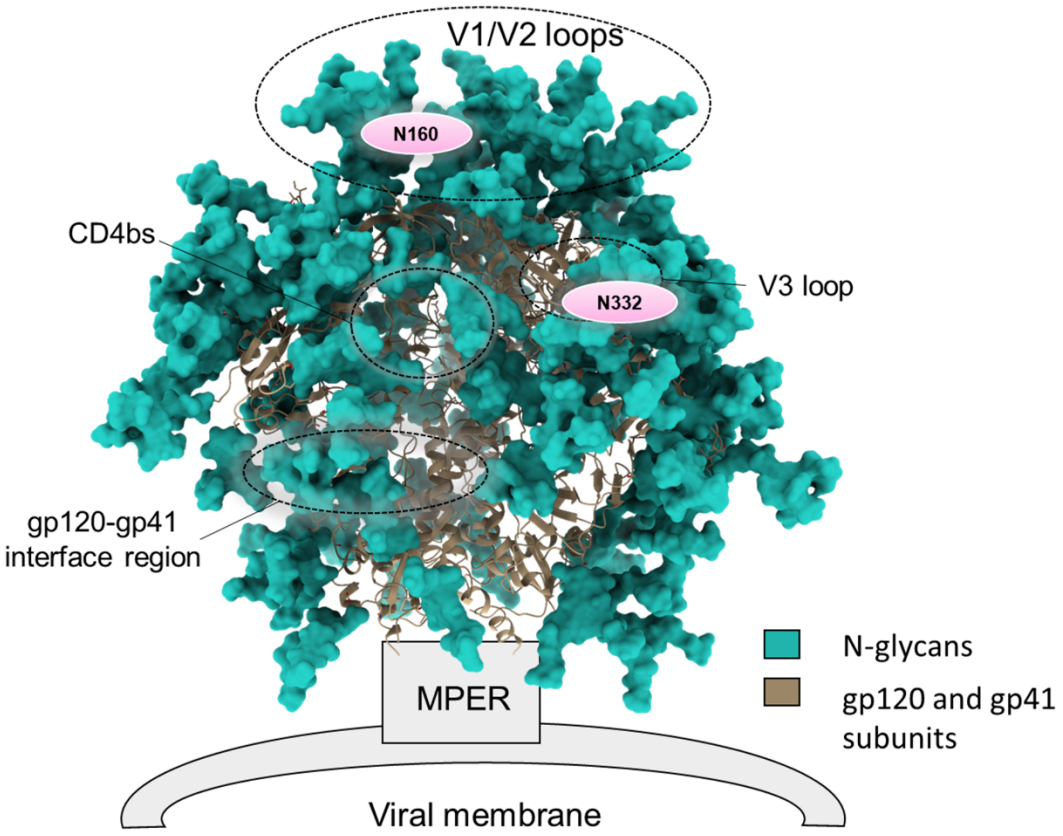


Figure 1.12. Regions on HIV-1 Env spike targeted by bNAbs. The model highlights five regions which are usually the target of bnAbs on Env spike: The CD4bs (e.g., VRCO1, CH103); the V1/V2 loop and glycan sites, N156 and N160 (e.g. PG9, PG16); the V3 loop and glycan N332 (e.g. PGT128, PGT125); the MPER (e.g. 4E10) and the gp120-gp41 interface region (PGT145, PGT151). The model is based on fully glycosylated trimer (PDB ID 5C7K). N-glycans highlighted in cyan.

Other important sites of vulnerability for bnAbs are trimer apex which is characterized by the conserved N156 and N160 glycans that project towards upwards of the trimeric surface. The bnabs which targets this region are PG9/PG16, PGT145/PGDM1400, CAP256.VRC26 and others. Both PG9 and PG16 binds with the C-strand of V1/V2 loop and with the oligomannose-type glycans at N156, N160 and sialylated hybrid N156, respectively [250,251] (**Figure 1.13.B**). PGDM1400 bnAb binding is highly quarternary specific and it binds to cationic residues in the gp120 V1/V2 loop and the oligomannose-type glycans of N160 present in V2 loop of the Env [252]. CAP256.VRC26 also

binds to V1/V2 apex region similarly to PG9/PG16, oligomannose-type glycans at N160 region, hybrid at N156, or bi, tri, tetra- antennary structures containing terminal sialic acids [253].

Furthermore, these bnAbs serves as a great tool in vitro for identifying the antigenicity of immunogen candidates and have also been utilized in the affinity purification of recombinant Env [36,254]. For example, PGT145, which targets a trimer apex region, has been utilized to affinity purify the properly folded trimers. PGT145 affinity columns are routinely used to improve the efficiency of HIV-1 Env protein expressed recombinantly [254]. The cryo-EM structure of BG505 SOSIP.664 with PGT145, revealed that the long, anionic, beta hairpin of heavy-chain complementarity-determining region 3 (CDR3) penetrates through the N160 glycan triad, interacting primarily with one N160 [242] (**Figure 1.13.C**). These examples illustrate the importance of fine processing in HIV-1 Env for bnAb binding.

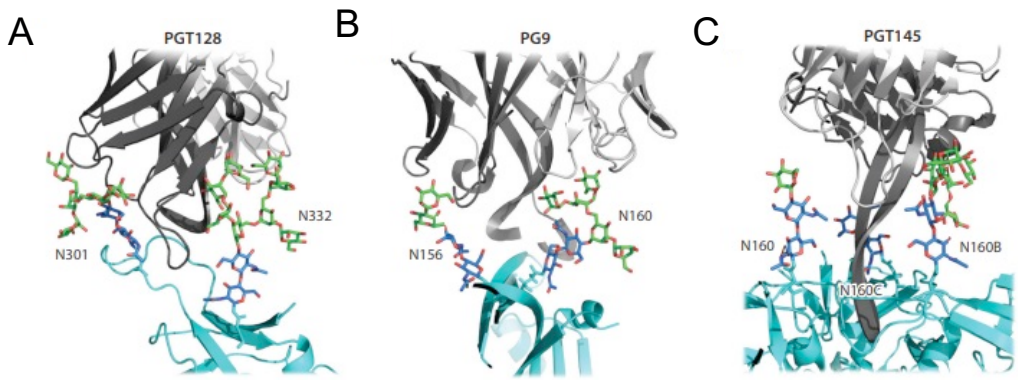


Figure 1.13. Antibody penetration of Env glycan shield. A) PGT128 binding with N301 and N332 glycans. B) In PG9, an apex bnAb, CDRH3 is penetrated between N156 and N160. C) PGT145, another apex bnAb, inserts its long CDRH3 between three symmetrically related N160 glycans at the threefold axis. The figure has been adapted from Crispin et. al., *Annual Review of Biophysics, 2018* [249].

1.3.7 Glycan occupancy

The HIV-1 Env is heavily glycosylated in terms of both sequence coverage and size of oligosaccharide residues, and the presence of these glycans are important for structural integrity and the immune response. However, the absence of these conserved PNGS or underoccupancy at PNGS could lead to the elicitation of non-neutralizing antibody response [255,256]. Therefore, it is

important to have PNGS glycan occupancy conserved on immunogen candidates mimicking the viral spike to facilitate the immuno-focusing to more relevant epitopes.

Interestingly, virally derived and membrane bound trimers have revealed higher glycan occupancy compared to recombinant SOSIP.664 counterparts [3,257,258]. Site-specific glycan analysis has revealed overall high glycan occupancy, but V1/V2 and gp41 sites were most susceptible to under-occupancy, thus creating artificial glycan holes [257–259]. A recent study has revealed the underoccupancy at position N611, due to the presence of antibody that only targets virus lacking N611 which was further confirmed with cryo-EM and site-specific glycan data [259–261].

Enhancing the immune response toward target epitopes can involve strategically filling glycan holes [262,263], whereas certain vaccine design strategies remove PNGS to concentrate on specific regions to attract antibodies [263–265]. Despite the distinctiveness of both of these approaches, intact glycan shields have shown greater potential for eliciting neutralizing antibodies compared to those with multiple holes [266]. Although the mechanism of underoccupancy at PNGSs is not well understood, it has been hypothesized that codon-optimization or the introduction of an aromatic amino acid or nucleotide modifications can potentially enhance the occupancy [3,267,268].

A recent study demonstrated the enhanced glycan occupancy on HIV-1 Env, by substituting NXS to NXT, because oligosaccharyltransferase (OST) has a higher affinity toward NXT [3,269–271]. The threonine showed high affinity to OST as the side chains of threonine can form H-bonds to Trp-463, Trp 464, and Asp-465 of OST and additional Vander wall interaction between methyl group of the threonine and Ile-572 of OST. In contrast, side groups of serine showed relatively less interaction with the OST. OST is an enzyme complex involved in the transfer of core oligosaccharide from dolichiol phosphate to either NXS or NXT. Studies on model glycoproteins have previously shown that NXT sequons have a higher probability of getting glycosylated than NXS [271]. This may be a useful strategy to enhance glycan occupancy on nucleotide-based immunogens as it involves the editing of nucleotides of target protein. The glycan transfer via oligosaccharyltransferase is shown in **Figure 1.14**.

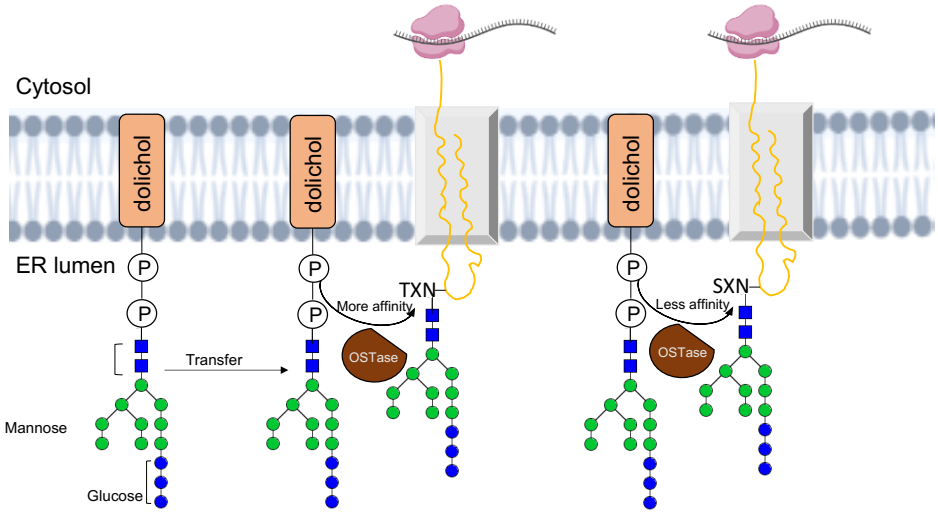


Figure 1.14. Glycan transfer from dolichol phosphate to asparagine residue via Oligosaccharyltransferase (OSTase). The OSTase is involved in transfer of core saccharide from dolichol-P to asparagine residue of NXS or NXT. The OSTase is known to have more affinity towards NXT than NXS [271].

1.3.8 Cell-specific glycan compositions of Env

The prevalence of oligomannose-type glycans on Env stems from protein-induced steric constraints, while host cell glycosylation predominantly influences the glycoprotein's more exposed segments. This processing, unique to cell types, relies on the host cell's expressed glycosidases and glycosyltransferases [120]

A comparative analysis of HIV-1 Env derived from HEK and CHO cells have revealed a highly similar level of oligomannose-glycan content, while the complex-type glycan presents considerable cell-specific differences [18]. As the CHO derived material showed higher sialic acid content than HEK-293 F expressed Env. The abundance of sialic acid is known to influence the antigenicity and immunogenicity of Env [272].

Similarly, the glycosylation analysis between the expression of virally-derived gp120 produced in PBMCs and in HEK cells revealed cell-specific differences [210]. Specifically, sialylated glycans obtained from PBMC-produced material predominantly exhibited α2-6-linkages, while material from HEK sources showed exclusive α2-3-linked structures. Notably, α2-6-linked sialylated glycans have previously been associated with anti-inflammatory effects, suggesting potential immunological implications tied to these structures [273].

Furthermore, viral glycoproteins expressed in two primary target cells for the infection of HIV-1 Env, PBMCs and monocyte-derived macrophages have also revealed similar biochemical changes [120]. The gp120 produced in MDM indicated the presence of a broad and diffuse band on SDS-PAGE compared to gp120 produced in PBMC's. Glycoside analysis further revealed that the MDM gp120 shows the presence of N-linked carbohydrates containing lactosaminoglycans, which was not observed in PBMC's derived gp120. Notably, there is negligible presence of lactosaminoglycans structures in recombinant SOSIP.664 [274]. This divergence in glycan processing highlights that recombinant trimers may reliably recapitulate only the mannose patch of native trimers and the cell-dependent variation arises in complex-type structures may limit the utility of targeting epitopes containing complex-type glycans [120].

These fine changes in glycan processing have emerged as an important factor in both viral infectivity and in the formation of bnAb epitopes. As some bnAbs, such as PGT121, rely on the recognition of both oligomannose-type and complex-type glycans, as depicted on glycan arrays [275]. Others, like PGT151, rely on tri- and tetra-antennary structures at the gp120-gp41 junction [276]. Meanwhile, the involvement of α2-6-linked sialylated hybrid and complex-type glycans has been linked to apex-targeting bnAbs like PG16 [251,277]. Notably, interactions between Env's sialic acid residues and Siglecs are pivotal in macrophage infection, which expresses low levels of cell-surface CD4 [278]. Therefore, it is important to understand the factors influencing the complex-type glycosylation, as it may help in refining the immunogen design, including optimal cell line selection for the expression of vaccine candidates.

1.3.9 Considerations for immunogen manufacture

All the factors mentioned above, such as Env structure, glycan occupancy, and cell-type expression, exert varying degrees of influence on Env glycosylation processing. Currently, the extent to which Env immunogens should mimic the native viral spike remains uncertain. While studies on the immunogenicity of native-like trimers have shown the elicitation of autologous Tier-2 nAbs and heterologous Tier-1 nAbs [260,279–282]. Weakly neutralized heterologous Tier-2 antibodies have been reported in rabbits [283]. Except cows, the potent bnAbs required for a protective vaccine have

not yet been generated via immunization with native-like trimers [284,285]. The elicitation of bnAbs against the envelope spike will inevitably be challenging, as their epitopes are broad but poorly immunogenic due to their conserved nature. Otherwise, these epitopes would have been naturally selected against by immune pressure.

One of the leading approaches aimed at eliciting bnAbs entail targeting B cell lineages with highly engineered immunogens but this strategy also includes the polishing with a final immunogen displaying native-like features [243,264,286,287]. One such example of vaccine candidate is the engineered outer domain germline targeted version 8 (eOD-GT8) 60-mer nanoparticle, which has reported to elicit HIV-specific B cell, and antibody responses in humans. In a randomized phase I clinical trial (IAVI-G001), vaccine priming with eOD-GT8 60-mer adjuvanted with $AS01_B$ showed a favourable safety profile, with 97% of vaccine recipients induced VRC01-class bnAb precursors [288,289]. Moreover, the vaccination with eOD -GT8 60-mer also developed polyfunctional T helper cells, potentially enhancing vaccine efficacy, including after subsequent booster vaccines [290]. These clinical studies utilizing germline-targeting vaccine priming, encourage the development of boosting regimens to induce bnAbs. There are currently two follow-up clinical trials ongoing, utilizing an mRNA-based platform, IAVI G002 seeks to assess the safety and immunogenicity of eoD-GT8 mer mRNA vaccine (mRNA-1644) and IAVI G003 involves the priming with mRNA-1644 followed by a booster with the Coreg28v2 60-mer mRNA vaccine (mRNA-1644v2-Core) in HIV-1 uninfected adults [291]. However, the assembly and glycosylation of the Env derived from the mRNA immunogens is not known yet.

Although the extent to which glycosylation influence the design of vaccine is not yet known. In particular, the influence of expression cell-type and glycan underoccupancy has emerged as an important consideration in vaccine design, as under-occupancy may induce unfavourable, immunodominant autologous nAb epitopes [259,266]. In addition, complex-type glycan processing, determined by the glycosyltransferases present in the producer cell, can also play a role in shaping the elicited antibody response.

1.4 Strategies for vaccine design

1.4.1 Recombinant protein and nucleic acid-based immunogens.

Advancements in protein-based vaccines, along with subsequent refinements in expression, purification, and formulation methodologies, have facilitated the production of highly purified proteins on a scalable and cost-effective basis. Recombinant proteins are commonly produced in secreted form by eliminating the transmembrane and internal regions that tether the virus to the cell membrane. Secreted, soluble proteins make the purification process easier and can potentially generate high yields [292]. However, the truncation of the protein, in some cases, destabilizes the protein and the development of efficient in vitro production system to express sufficient quantities can slow iterative testing of new immunogen designs. Thus, the development of an effective nucleic acid-based platform for vaccine candidates would be of significant interest for preclinical and clinical development.

Nucleic acid-based vaccines have rapid adaptability, as they do not require complex GMP production pipelines as required in the case of recombinant protein design [293]. Moreover, this will also assist in the development of multi-immunogen vaccine strategies, as it would be expensive in the case of recombinant-based mass protein production. In addition, native-like transmembrane protein can be delivered via the nucleic-acid based platform.

Common platforms for nucleic acid-based vaccines include the delivery of plasmid DNA or the direct injection of mRNA to express antigens into the host cells. DNA vaccines require delivery to the nucleus, whereas the RNA-based approach can be directly introduced into the cytosol for desired protein expression. However, the delivery of DNA to the nucleus is a rate-limiting step, and there is a potential risk of integration into the host chromosomal DNA [294]. DNA-based vaccines have also performed poorly in human clinical trials [295]. In contrast to this, RNA-based vaccines have emerged as a highly effective delivery platform and have the potential to express proteins in in vivo studies and in humans [296–300].

However, there are still some challenges associated with the mRNA-based platform. One such factor is the short half-life of mRNA and the requirement for storage at -80°C, which poses the complexity in the manufacturing and delivery of vaccine candidates [301]. Another factor is the limited track record of mRNA vaccines, as their long-term efficacy and safety are still not well understood. Furthermore, there is limited control at the post expression level in the case of the mRNA-based platform compared to recombinant protein-based vaccines, in which desired protein can be selected utilizing post-purification methods.

1.4.2 Conventional mRNA and Self-amplifying mRNA

The approval of mRNA-based vaccines for countering the SARS-CoV-2 pandemic underscores the versatility of both conventional mRNA and self-amplifying mRNA (saRNA) as adaptable immunotherapy platforms for addressing infectious diseases. In preclinical studies, both conventional and self-amplifying RNA have shown protective immunization against multiple infectious diseases including influenzae, RSV, Rabies, Ebola, SARS-CoV-1, SARS-CoV-2, and HIV-1 [297,302–306].

Both conventional and self-amplifying mRNA share the essential elements of eukaryotic mRNA: a 7-methyl guanosine cap (m7Gp3N; N, any nucleotide) at the 5′ region, a 5'UTR, an open reading frame (ORF), a 3′ UTR, and a polyA tail containing 40—120 adenosine residues [296,307]. These components are amenable to modification, including the introduction of chemically modified nucleotides, sequence optimization, 5′ cap structure, and polyA tail elongation, in order to improve the efficiency of mRNA translation and reduce immunogenic properties [308–310]. Both types of mRNA can be produced in a cell-free system, using an enzymatic transcription reaction. Manufacturing of mRNA vaccines requires the replacement of the encoding sequence with target antigens against different diseases, while preserving the physiochemical characteristics of the RNA molecule.

A conventional mRNA vaccine carries only the coding sequence, flanked by the regulatory elements, while self-amplifying RNA encodes for RNA-dependent RNA polymerase (RDRP) along with the desired gene of interest. Self-amplifying RNAs are engineered based on the RNA genome

of alphaviruses, which possess inherent self-replication capabilities due to the presence of RDRP. These self-amplifying mRNAs are often referred to as 'replicons,' and they facilitate sustained antigen expression at lower doses compared to the conventional mRNA, primarily because of their ability to self-amplify within cells [305,311]. Moreover, replicons can encode multiple antigens within the same replicon, enabling the expression of both the target antigen and an immunomodulatory biological molecule, such as an adjuvant [312,313].

In conclusion, both conventional and self-amplifying systems serve as a excellent platforms for vaccine design, each offering its unique advantages. Conventional mRNA-based strategies are straightforward and accommodate the delivery of smaller RNA molecules compared to replicons. To address this limitation, a trans-replicon system has recently been developed, separating the backbone containing NSPs from the gene of interest on distinct RNA molecules [314]. The mechanism of antigen expression via conventional mRNA and self-amplifying RNA based platform is shown in **Figure 1.15**.

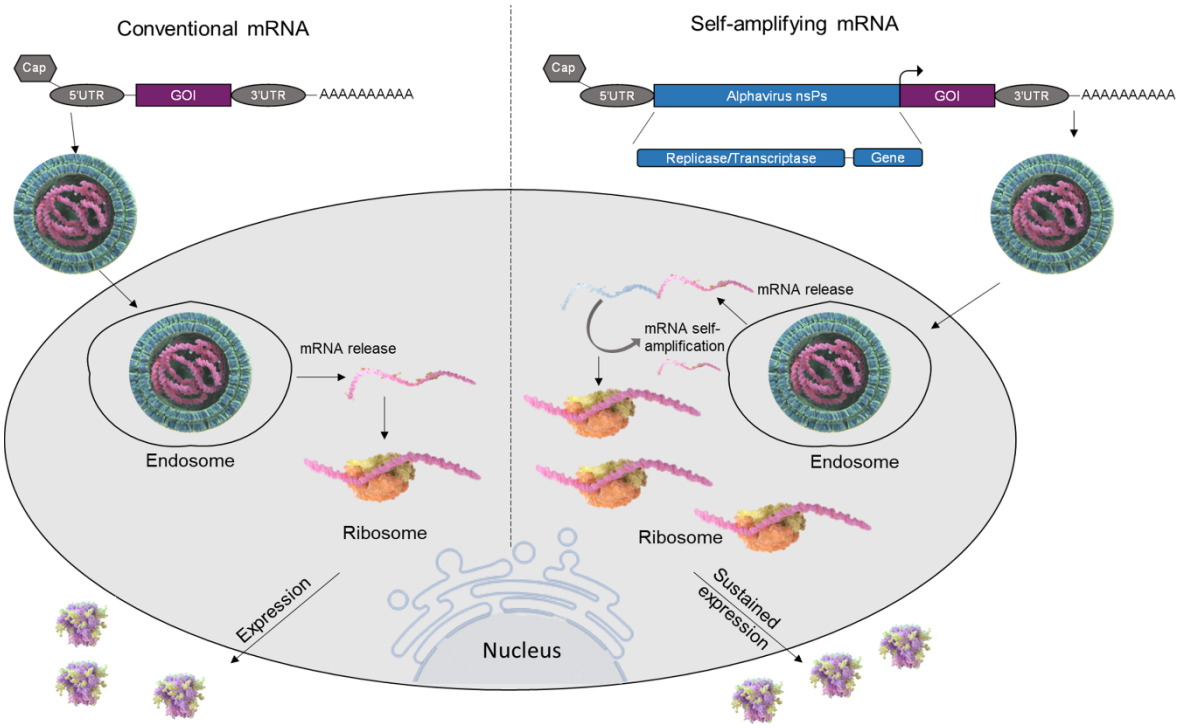


Figure 1.15. Schematic representation of mechanism of antigen expression via conventional mRNA and self-amplifying mRNA (Replicons). GOI, gene of interest; nsps, non-structural proteins; mRNA, messenger ribonucleic acid; UTR, Untranslated region.

1.4.3 Progress in improving mRNA-based vaccines

An RNA-based platform offers a significant advantage in terms of rapid deployment compared to traditional protein-based vaccines. However, one potential limitation is that the vaccine candidate is entirely encoded by the nucleotide sequence, and there is no opportunity for control of immunogen assembly at post expression level, which is commonly employed in traditional protein-based vaccines. These limitations can influence antigen expression, leading to issues such as improper folding of the protein, post-translational modifications, and antigen presentation. Nevertheless, these limitations may be addressed through nucleotide modifications, the addition of chemically modified nucleotides, and improvements in delivery methods [310,315–317]. In mRNA-based vaccines, three key attributes play a vital role: mRNA stability, antigen expression, and mRNA delivery to the target cells [296].

Several strategies have emerged over the course of extensive research and investment in the development of protein-based vaccines to stabilize antigenic proteins. These strategies encompass the codon optimization, the introduction of proline residues, the incorporation of glycine-serine-rich linkers at the place of furin cleavage sites, the introduction of disulfide bonds, the utilization of two-component nanoparticle systems, and the integration of foldon domains. The foldon domain is a small β- propeller-like trimeric structure found at the C terminus of bacteriophage T4 fibritin, which has been utilized to promote trimerization of viral glycoproteins [318–320]. These approaches have been widely studied in recombinant-protein based vaccines and have been exploited in the design of nucleic acid-based vaccine design [109,110,112,217,321]. One such example is of clinically approved mRNA vaccines against COVID-19, mRNA-1273 and BNT162b2, encoding SARS-CoV-2 S protein containing two proline substitutions to enhance the protein stability [100,322]. Furthermore, both of these mRNA molecules undergo chemical modification through the substitution of uridine with pseudouridine, a modification aimed at enhancing mRNA translation efficiency [308].

Due to the rapid adaptability, mRNA-based vaccines offer the potential to explore multiple antigens at the same time. This enables the design of a diverse panel of mRNA antigens for comparative analysis of immunogenicity and to improve the vaccine candidates. For example, a

recent study designed a panel of Respiratory syncytial virus class I fusion (RSV F) mRNA antigen to assess immunogen assembly, immunogenicity, protection, and safety of protein expressed in animal models. This comprehensive antigen panel comprised different forms of the RSV F protein, including secreted, membrane-associated, prefusion-stabilized, and non-stabilized structures [309].

Another important attribute in the mRNA-based platform is achieving effective delivery to the specific target cells and produce sufficient amounts of proteins of interest. A range of materials can be employed to encapsulate RNA, including lipids, lipid-like substances, polymers, and protein derivatives. Lipid nanoparticles, in particular, have undergone extensive investigation and have successfully advanced to clinical use for delivering small molecules, siRNA drugs, and mRNA [317]. However, once the RNA encapsulated in LNPs reaches the target cells, LNPs can be internalized by multiple mechanisms, including macropinocytosis and clathrin-mediated and caveolae-mediated endocytosis [323]. The endocytic pathway requires the endosomal escape for effective delivery, which relies on the properties of the nanoparticles and the cell type. Nevertheless, improvements in nanoparticles can be achieved through strategies such as optimizing pKa of ionizable lipids, tailoring lipidic tails and modulating the type and ratio of lipids [324,325].

The two authorized COVID-19 vaccines, mRNA-1273 and BNT162b2, utilize lipid nanoparticles for delivering antigen mRNA, while numerous other lipid nanoparticle-mRNA formulations are actively advancing through clinical evaluation, promising breakthroughs in combating viral infections, cancer, and genetic diseases. [326–330]

1.4.4 Post-translational modification in RNA based immunogens

Post-translational modifications are covalent modifications that occur on proteins following their synthesis by ribosomes during the translation process. These post translational modifications (PTMs) are crucial for protein's structure, function, localization, and interactions with other molecules. Some important PTMs in viral immunogens include glycosylation, phosphorylation, acetylation, ubiquitination, and proteolytic cleavage. Understanding these modifications in vaccine candidates is important as they can impact the presentation of viral antigens to the immune system and thus influence the development of protective immune response.

One of the advantages of the mRNA-based vaccine approach is that the target protein is synthesized and modified by the host's cellular processes. In contrast, traditional recombinant proteins require production in various production systems, yeast, plants or bacteria that can result in altered posttranslational modification leading to immune recognition. However, recombinant protein allows control at the post-expression level, such as purification of properly assembled protein, fully glycosylated protein, and elimination of non-native like constituents. These non-native-like constituents can arise from incomplete proteolytic cleavage or by shedding of immunogens during the processing of proteins [321,331]. For example, S1 subunit shedding in case of SARS-CoV-2 [126]. These constituents can potentially elicit non-neutralizing antibody response [36]. To overcome this, post-expression purification can be employed to purify the properly assembled protein from whole expressed material.

Another consideration in viral vaccine design is glycosylation of viral glycoproteins. As glycans play an important role in stabilizing the viral protein and serve as an epitope for the antibody response [1]. Therefore, it is important to have these glycans conserved as they are present on the native virus in vaccine candidates for an effective immune response. However, in certain cases, a robust immune response can be achieved even with heterogeneity in the population of expressed immunogens. The precision in fine glycan processing is particularly relevant in the case of viruses like HIV-1, where the immune response is heavily influenced by glycans. This indicates the significance of glycosylation patterns in achieving desired outcomes in vaccine design and development.

1.5 Aims and hypothesis

As novel immunogen platforms are being developed, it is important to characterize the glycan shield. Here, an established LC-MS based approach is used to study next-generation immunogen approaches for two critical vaccine targets: HIV-1 and SARS-CoV-2. Both HIV-1 and SARS-CoV-2 viral spike share common class I prefusion conformation. There are various approaches which has been employed to stabilize the viral spike in prefusion conformation, as this conformation is crucial for presentation of epitopes for neutralizing antibodies [201,332] . One such approach is an introduction

of prolines which impedes the switch to post-fusion conformation [333]. This approach has been well-studied in HIV-1 and have been employed successfully in the design of stabilizing HIV-1 Env trimer [203].

During the COVID-19 pandemic, S protein of SARS-CoV-2 was stabilized using the same approach, involving the introduction of prolines to stabilize the S protein resulting in S2P (two proline substitutions), which has been widely utilized in vaccine design [109]. However, low expression of recombinant S2P protein prompted the development of expression-enhanced S protein containing four more additional proline mutations [110]. This expression enhanced variant is known as HexaPro was developed by Prof. Jason Mclellan's laboratory (University of Austin, Texas). **Chapter 3** of this study focuses on understanding the influence of these proline mutations on protein conformation and glycan composition through MD-simulation studies and LC-MS analysis.

The glycan composition of SARS-CoV2-2 S presents similar glycan composition as on HIV-1 Env with the presentation of oligomannose-type glycans. However, the abundance of oligomannose-type glycans on S protein was relatively low, signifying the denser glycosylation of HIV-1 Env compared to S protein and highlights one of the reasons why it is possible for SARS-CoV-2 infections to be cleared by the immune system and not HIV-1. In addition, **Chapter 3** includes a serological analysis of both HexaPro and S2P protein, as the S proteins have been extensively employed in serological studies to confirm the COVID-19 infections.

Furthermore, it is important to compare the glycosylation of recombinant and viral SARS-CoV-2 S protein. In **Chapter 4**, a comparative analysis of SARS-CoV-2 S protein is performed on variants of recombinant S protein developed towards the stabilization of S protein, including S2P, HexaPro and Close S, with data obtained from viral-derived S protein. In addition, this study compares the glycan composition of monomeric RBD proteins which have been utilized in preclinical studies and clinical studies [113,114], with recombinant and viral S protein [125]. The glycan composition data of RBD reveals the presence of more processed glycan structures as compared to trimeric viral spike. This is similar to the glycan composition pattern observed on HIV-1 Env gp120 subunits and trimeric Env [18]. This signifies the role of protein architecture in dictating

the glycan compositions, as the sites which are shielded in trimeric viral spike becomes accessible in monomeric subunits, thus gets more processed by glycan processing enzymes.

As glycosylation is highly dependent on three-dimensional architecture, it can be used to determine the protein assembly and manufacturing processes. One such feature is the preservation of oligomannose-type glycans which is linked to the quaternary structure of the protein and can be used as a sensitive reporter of native-like protein architecture. **Chapter 5** utilizes these native-like signatures to determine the immunogen assembly of HIV-1 Env derived from mRNA-based platform and compared with recombinant and viral HIV-1 Env. Prior to the execution of this work, no study had been performed to determine the protein assembly and glycan composition of Env derived from mRNA-based platform. It is important to know, as RNA-based platform eliminates the opportunity to control the immunogen assembly at post-expression level which is usually employed in case of traditional protein-based vaccines. This work demonstrates conservation of native-like oligomannose composition on Env derived from RNA replicon. However, there were few sites displaying higher glycan processing compared to as observed on native-like viral Env. This is potentially due to the presence of non-native like Env which is consistent with the NS-EM data obtained in this study. In this study, *Galanthus nivalis* lectin beads were employed to purify the whole Env material, whereas typical recombinant-based protein production utilizes post-production purification methods such as size exclusion chromatography or affinity chromatography to purify the trimeric material.

Furthermore, to determine the influence of different cell-types, **Chapter 5** presents the comparison of the glycan composition of Env expressed in different production systems, including recombinant protein expressed via transient transfection of plasmid DNA, Env expressed via RNA replicon in HEK293F, mouse muscle and dendritic cells. HEK 293F cells are usually employed in in vitro for expression of recombinant proteins. To explore the influence of potential cell-type expressing immunogen in case of intramuscular injection, this study explored the expression of Env in muscle and dendritic cells. Moreover, the site-specific analysis of Env revealed glycan underoccupancy at few sites. To improve the glycan occupancy this study utilized the nucleotide editing approach which has been previously employed on HIV-1 recombinant Env [268].

Overall, this thesis discusses about the protein assembly and glycan compositions of Env/spike produced via recombinant protein and mRNA-based platform. Site-specific glycan composition analysis has been used to characterize the oligomannose-type glycans which suggests the structural integrity of the protein. Furthermore, MD simulation, serological testing and NS-EM have been employed in collaboration to characterize the Env/spike and get more insight about the protein binding, conformation, and architecture.

Chapter 2 Materials and Methods

2.1 Plasmids constructs for DNA and RNA replicon

The plasmids encoding SARS-CoV-2 S HexaPro and SARS-CoV-2 S2P were provided by Prof. Jason Mclellan (The University of Texas at Austin). The plasmid for RNA replicon, BG505 NFL.664 was provided by Prof. Darrell J Irvine (Massachusetts Institute of Technology, USA). The plasmids for SARS-CoV-2 were grown in E.coli. and glycerol stocks were made. The plasmids were then isolated using the Plasmid Maxi or miniprep kits (Qiagen) according to the manufacturer's protocol. The plasmids of RNA replicon were prepped in the same way, and then RNA was transcribed using invitro transcription kit described later in this chapter.

2.2 Recombinant protein expression and purification

The viral glycoproteins and the cell line used for expression in this thesis is summarized in **Table 2.1**. Some of the recombinant proteins were provided by the collaborators, that is, RBDg5.1, RBDg5.1 24mer, RBDg8.2. The SARS-CoV-2 expressed in Vero cells is expressed by Sai Li lab, I just used the raw files for the analysis provided by them and is also available on the Massive server. All other viral glycoproteins used in this thesis are expressed by me and plasmid DNA provided from respective labs are mentioned in the **Table 2.1**.

Table 2.1. Summary of analysed viral glycoproteins.

	Construct	Cell	Chapter	Source and material provided
SARS-CoV-2	SARS-CoV-2 S HexaPro	HEK 293F	3	Jason Mclellan- plasmid DNA
	SARS-CoV-2 S2P	HEK293F	3	Jason Mclellan- plasmid DNA
	SARS-CoV-2 S	Vero cells	4	Sai Li - Data raw files
	RBD monomer	HEK 293F	4	Raiess Andrabi - plasmid DNA
	RBDg5.1	ExpiF293	4	Daniel Kulp - protein
	RBDg 5.1 24mer	ExpiF293	4	Daniel Kulp - protein
	RBDg 8.2	ExpiF293	4	Daniel Kulp - protein
HIV-1	BG505 NFL.664	HEK 293F	5	Richard Wyatt - plasmid DNA
	saRNA BG505 NFL.664	HEK 293F, C2C12, DC2.4	5	Darrell J Irvine - RNA
	saRNA BG505 NFL.664 mut	HEK293F	5	Made in the lab by Site-directed mutagenesis

2.2.1 Recombinant S protein expression in HEK 293F cells

For expression of prefusion S ectodomain of SARS-CoV-2 HexaPro construct, the base construct of SARS-CoV-2 S 2P (GenBank: MN908947) having proline substitutions at residues 986 and 987, a "GSAS" substitution at furin cleavage site (residues 682–685) and C-terminal foldon trimerization motif, HRV3C protease recognition site, Twin-Strep-tag and octa-histidine tag cloned in mammalian vector pαH was used. The HexaPro construct was created by addition of four proline substitution (residues 817, 892, 899, 942) in 2P base construct as described previously [110]. Plasmid encoding S protein was used to transiently transfect FreeStyle 293-F cells (Thermo Fisher) using polyethylenimine (PEI). Cells were maintained at a density of 0.2–3.0 × 106 cells/mL at 37 °C, 8% CO2 and 125 rpm shaking in FreeStyle 293F media (Fisher Scientific). Transfection mix was prepared in Opti-MEM (Fisher Scientific) using two solutions, DNA (310 μg/l) and PEI max reagent (1 mg/mL, pH 7) in a ratio of 1:3 in 25 mL of Opti-MEM respectively, followed by addition of DNA solution to the PEI mix and incubated for 30 minutes at room temperature. Cells were transfected at a density of 1 × 106 cells/mL and incubated at 37 °C, 8% CO2 and 125 rpm shaking. Culture was harvested after 7 days post transfection and the media was separated from the cells by centrifugation at 4,000 rpm for 30 minutes. The supernatant was filtered using 0.22 μm pore size filter (Merck) followed by S protein purification using 5 mL His Trap FF column connected to Akta Pure system (GE Healthcare).

2.2.2 Kifunensine treatment of SARS-CoV-2 S protein

HEK 293F cells were transfected at a density of 1×10^6 cells/mL and incubated at 37°C, 8% CO_2 and 125 rpm shaking. To elicit oligomannose-type glycans on S glycoprotein, 20 μM kifunensine was added at the time of transfection. Culture was harvested after 7 days post transfection and the media was separated from the cells by centrifugation at 4,000 rpm for 30 mins.

2.2.3 Ni^{2+} affinity purification and Size exclusion chromatography

His-tagged (6× Histidine residues) glycoproteins were purified using a 5 mL HisTrap column (GE Healthcare). Prior to loading the sample, the column was washed with washing buffer (50 mM

Na2PO4, 300 mM NaCl) at pH 7. The sample was loaded onto the column at a speed of 2 mL/min. The column was washed with washing buffer (10 column volumes) containing 50 mM imidazole and eluted in 3 column volumes of elution buffer (300 mM imidazole in washing buffer). The elution was concentrated by a Vivaspin column (100 kDa cut-off) to a volume of 1 mL and buffer exchanged to phosphate buffered saline (PBS). Further, purification of target S protein fraction was carried out using size-exclusion chromatography using a Superdex 200 16 600 column (GE healthcare). The target fraction was concentrated in 100 kDa vivaspin (GE healthcare) to a volume of 1 mL.

2.2.4 PGT145 antibody expression

Plasmid encoding heavy and light chains of PGT145 antibody were transfected in HEK 293F cells using polyethyleneimine. Transfection mix was prepared in OptiMeM using three solution, heavy chain DNA (155 μg/L), light chain (155 μg/L), and PEI max reagent (1 mg/mL, pH 7) in a ratio of 1:3 (DNA: PEI max) in 25 mL of Opti-MEM respectively, followed by addition of DNA solution to the PEI mix and incubated for 20 minutes at room temperature. The cells were cultured and harvested in same way as described above. The supernatant was then purified using protein A affinity chromatography.

2.2.5 Protein A affinity purification

The PGT145 used for the generation of affinity columns was purified using a 5 mL HiTrap Protein A affinity column (GE Healthcare). The column was equilibrated using a 20 mM $NaPO_4$, pH 7.5 buffer. The column was loaded at a flow rate rate of 1 mL/min. Once all the sample has pass through, the antibody was eluted using a 0.1 M citric acid buffer, pH 3. The acidic buffer was used to favour the antibody dissociation from the column. The pH was immediately neutralised by the addition of 1 M Tris-HCl buffer, pH 9.0, and immediately buffer-exchanged with PBS using 100kDa cut-off vivaspin column.

2.3 RNA replicon synthesis and expression in mammalian cells

2.3.1 Replicon production and purification of HIV-1 env

RNA replicon was derived from alphavirus VEE's genome as described previously [334]. The VEEV genome is modified, and the structural genes are replaced with env gene, BG505 NFL.664 (PDB:6P62). Briefly, plasmid DNA encoding for replicon was grown in *Escherichia coli* and purified using the Plasmid QIAprep spin maxi kit (QIAGEN). The plasmid backbone sequence contains the Ampicillin marker, pBR322 ori, nsp1—4 sequons of VEE's genome. The replicon sequence is flanked by the T7 promoter on the 5′end and MluI restriction sites at 5′ end of the T7 promoter and 3′ end encoded poly (A) sequence. To generate the template for in vitro transcription (IVT) reaction, first the plasmid DNA was linearized via the digestion of these plasmids using MluI (Thermo Fisher Scientific) and the linearized DNA was purified by the DNA purification kit (QIAGEN). Linearized DNAs were then transcribed using mMessage IVT kit and reaction was set up according to manufacturer's protocol. The RNA was then cleaned using megaclear kit (Thermo Fisher Scientific) to get rid of any buffers or impurifies carried away during the IVT reaction. Before transfection, RNA was run on the gel using RNA gel electrophoresis to check the purity of RNA. The RNA concentration and purity was verified by spectrophotometry using NanoDrop One and was then stored at -80°C until further use. The RNA with absorbance ratio, $A_{260/280}$ ratio of ~1.8 were used for transfections.

2.3.2 Culturing of mammalian cell lines for replicon expression

HEK 293F cells were cultured in freestyle media and are maintained at a density of 0.2×10^6 cells per ml at 37°C, 8% CO_2, 125 rpm. The cells were transfected at a density of 1×10^6 cells per mL.

The mouse muscle cell line (C2C12) was cultured in Dulbecco's Modified Eagles Medium supplemented with 10% fetal bovine serum (FBS) as recommended in manufacturer's protocol (American type culture collection, Catalogue no. CRL-1772). The cells were maintained in T75 adherent flasks at 2×10^6 cells per mL at 37°C, 5% CO_2.

The dendritic cell lines (DC2.4) was maintained in RPMI-1640 media supplemented with 10% FBS, 1× L-glutamine, 1× non-essential amino acids, 1x HEPES buffer, 0.0054× β-mercaptoethanol according to manufacturer's protocol (Sigma-Aldrich, Catalogue no. SCC142). The cells were maintained in T75 adherent flasks at 2×10^6 cells per ml at 37°C, 5% CO_2.

2.3.3 In Vitro Transfection of RNA replicon

To assess the assembly and glycosylation of Env, HEK cells were transfected with RNA. Further to investigate the cell derived changes in replicon expression of Env, C2C12 and DC2.4 were transfected with RNA. The RNA replicon transfection was carried out using two methods, neon transfection system (Thermo Fisher Scientific) and TransIT mRNA kit (Mirus Bio). Using the neon transfection system, the cells were electroporated with RNA using 100 µL neon tip, according to manufacturer's protocol. Briefly, the required number of the cells, based on the plate size used for transfection, were pelleted, and washed with PBS. Subsequently, cells were prepared for electroporation, by mixing 1 µg RNA per 1×10^6 cells in serum free media, OptiMem. Followed by electroporation at 1200 V, 30 ms and then the cells were ejected from the electroporation tip in 6 well plates containing 2 mL media for culturing of cells. Each mL contains electroporated solution of 1×10^6 cells mixed with 1 µg RNA. The supernatant was harvested after 48 hrs and supplemented with new media and harvested again after 24 hrs to attain high yield of protein expression.

Using the TransIT-mRNA kit transfection was carried out as directed in manufacturer's protocol. Briefly, suspension HEK 293F cells were grown to 1—1.5 × 10 cells/mL in 25 mL Freestyle media. The adherent cells were seeded at 2×10^6 cells in 12mL of media in T75 flask and were transfected at 80% confluency. The transfection solution was composed of 100 µL of Reduced sera free media, OptiMem, 1 µg of RNA replicon, 2 µL of the mRNA boost reagent and 2 µL of the Trans-IT mRNA reagent for 1 mL of seeded cells. The transfection mixture was then incubated at room temperature for 2—5 mins. Followed by addition of transfection complexes dropwise to the cells. The cells were incubated at 37°C for 2 hrs for suspension cells and 48 hrs for adherent cells. The supernatant was harvested after incubation and the cells were pelleted down, followed by filtering of the supernatant via 0.22 µm pore size sterile filter.

2.3.4 RNA replicon expressed HIV-1 Env purification

The supernatant containing soluble Env was collected and purified by affinity chromatography *Galanthus nivalis* lectin (GNL) beads (Vector Labs), which preferentially binds to structures containing α-1,3 mannose residues such as many viral glycoproteins, in this study, HIV-1 Env. Briefly, 10 column volumes of wash buffer, 1× PBS was used to wash the beads and then supernatant containing Env was applied, and the solution was allowed to drain through gravity. After application of sample, the beads were washed again with 3 columns of wash buffer to remove unbound materials. Lastly, the protein was eluted using 3—5 columns of elution buffer, 0.5 M α-methyl mannoside, pH 4. The eluted protein was then concentrated using 100 kDa Amicon filter.

For purification of trimer specific Env, the supernatant containing Env was purified by PGT145 antibody [242]. The antibody was added at the concentration of 10 μg/mL to the supernatant, the mixture was incubated overnight at 4°C. Followed by purification of PGT145 bound trimer specific Env using protein A affinity chromatography.

2.3.5 Site-directed mutagenesis of replicon construct, BG505 NFL.664

To improve the occupancy, threonine substitutions were introduced into the env gene leading to NXS to NXT transition at N-linked glycan sites, N185e, N197, and N611. These sites were sequentially mutated by three separate polymerase chain reactions (PCRs) using three sets of forward and reverse primers carrying mutations for different sites. The primers were designed using Agilent quick change primer design and manufacturer's PCR conditions (Thermo Fisher Phusion polymerase) with 10 ng/μL env gene in a final volume of 20 μL. The reaction was then set on thermal cycler at 98°C initial denaturation and then 18 cycles with set conditions as 98°C for 20 seconds, 58—68 °C for 30 seconds, 72°C for 6 mins, followed by final extension at 72°C for 5 mins. The PCR product was digested with dpnI, followed by transformation in XL-10 gold cells. The plasmid DNA was sequenced and subsequently used in the PCR reactions for substitution of all three sites.

2.4 Glycan analysis of viral glycoproteins

2.4.1 Glycopeptide mapping using LC-MS Orbitrap Eclipse

In this thesis, I have mainly focused on the site-specific N-linked glycan analysis of viral glycoproteins. The primary method used in this thesis were first developed by Behrens et al [335] which were further developed to investigate glycan occupancy [3,268] . All the analyses in this thesis were obtained using an Orbitrap Fusion Tribrid mass spectrometer connected to an EasySpray nano LC [336].

To prepare the samples for glycopeptide analysis, first the native viral proteins were unfolded using denaturant, urea and disulfide bonds were reduced using the reducing agent, dithiothreitol (DTT). The reduced thiol residues can reform disulfide bonds, and to prevent this, they were capped using Iodoacetoamide. The excess IAA was quenched by incubation of the reaction with DTT. This is followed by buffer exchange of the reaction with Tris buffer using vivaspin column to remove all the excess reagents. After the buffer exchange, the protein was digested using different protease enzymes to achieve single N-linked sequon on the glycopeptides in the site-specific analysis and it allows the detection of glycan composition at each PNGS site. Utilizing proteases with specific specificity prevents total protein degradation, yielding a pool of peptides detectable through mass spectrometry at positions . In this thesis, I have primarily used three proteases, trypsin, chymotrypsin and α lytic protease (ALP). The cleavage site of these proteases is described below in **Table 2.2**.

Table 2.2 Specificity of proteases for cleavage of proteins.

Proteases	Amino acid cleavage site
Trypsin	Arginine (R) and lysine (K)
Chymotrypsin	Phenylalanine (F), Tryptophan (W) and Tryrosine (Y)
α lytic protease	Threonine (T), Alanine (A), Serine (S) and Valine (V)

Once the digested glycopeptides are prepared, prior to analysis of mass spectrometry, liquid chromatography is utilized to separate the glycopeptides to reduce the complexity of the data and

detection of all the glycopeptides. The LC method herein used is reverse phase LC, where the hydrophobic stationary phase interacts with the hydrophobic regions of the glycopeptides, so glycopeptides with higher proportion of hydrophobic residues will elute later . The hydrophobic stationary phase used in this case is an octadecyl carbon chain (C18) fused to silica. To enhance the separation, 280 min gradient of two solvents (formic acid and acetonitrile) is used that is beginning with a high concentration of polar solvents, 95% formic acid (FA) and 5% Acetonitrile (ACN) and then increasing the concentration of a non-polar solvent that is ACN during the separation as described in **Fig. 2.1**. The LC is used in-line with ionization of glycopeptides which is followed by mass spectrometer analysis. To allow the ionization of glycopeptides, they are resuspended in 0.1% formic acid which promotes the formation of ions and are ionized using Electrospray ionization (ESI) approach. ESI is a commonly used method as it allows the conversion of molecules in a solution into ions at an atmospheric pressure. In ESI, a fine capillary needle, subjected to high voltage, generates an aerosol known as the Taylor cone, from which a jet of liquid drops emerges [336,337]. This jet gradually evaporates, enhancing the charge on the droplets. Once the droplet charge surpasses a critical threshold called the Rayleigh limit, an explosive dissociation occurs, resulting in a stream of charged ions for the spectrometer.

Once the glycopeptides are ionized, the ions will be filtered in the quadrupole mass filter on the basis of *m/z* scan range which can be set up in the method. For given dc and ac voltages between the four rods, only ions of a certain m/z ratio in this case, scan range of 375 – 1500 *m/z*, pass through the quadrupole filter and all the other ions will be deflected from their original path [338]. The filtered ions will then pass through the Orbitrap mass analyzer, which consists of an outer barrel-like electrode and a coaxial inner spindle-like electrode which traps ions in an orbital motion around the spindle [336]. The ions orbit around the central electrode and induce a current in the outer electrode, the angular frequency with which they orbit is related to the *m/z* of the ion. Fourier transform function converts the angular frequency into a *m/z* readout. Orbitrap analyzer is quite sensitive and allows the accurate assignment of m/z for glycopeptide ions. Following the MS1 detection of all the glycopeptide ions, the ions pass through the Ion routing multipole (IRM) which serves a cell for

higher-energy collision dissociation (HCD) and also transfers the ions between orbitrap mass analyzer and linear ion trap mass analyzers [339] (**Figure 2.1**).

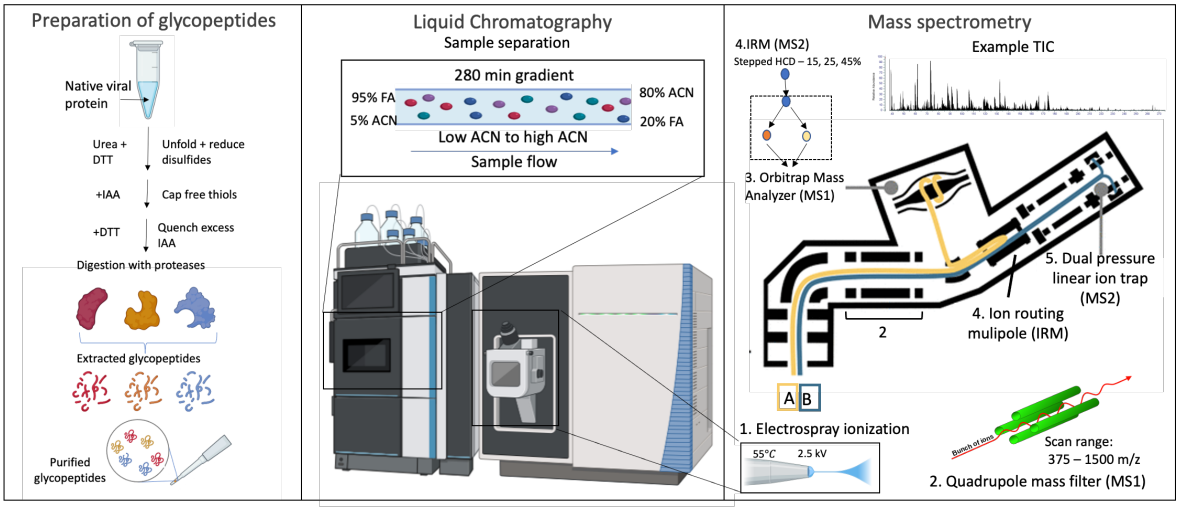


Figure 2.1. Glycopeptide analysis using the LC-MS. DTT, dithiothreitol; IAA, iodoacetoamide; FA, formic acid; ACN, Acetontirile. The orbitrap mass analyzer schematic is adapted from Heil et al. [339].

In the HCD, both the peptide and the attached glycan gets fragmented and results in the generation of a, b, and y ions (**Figure 2.2.A**). In HCD fragmentation, the N-linked glycan itself gets fragmented from the peptide backbone and the peptide fragments can be used to sequence the glycopeptide of interest [340] (**Figure 2.2.B**). MS2 data obtained from this can then be searched against a protein sequence of interest to determine the position of PNGS in the glycopeptide. HCD has higher collision efficiency and therefore is suitable for fragmentation of N-linked glycans where the glycosylation occurs at specific sequon i.e., NXT and NXS. However, in case of O-linked glycosylation, where the glycan can attach to multiple serine and/or threonine residues that serve as a potential glycosite, it is difficult to determine the position of amino acid in glycopeptide. To analyze the O-linked glycans, Electron transfer dissociation (ETD) is used in which only peptide backbone gets fragments and the glycan remain attached to the peptide backbone. ETD generates c and z fragments that retain the intact glycan (**Figure 2.2 A**). ETD-based fragmentation also helps in identification of glycosylation where the specific sequon of amino acid is not known for glycosylation such as tyrosine sulfation. In this thesis, I have used stepped HCD, in which total precursor ion accumulation time per scan is divided into multiple (usually three) equal parts, and ions

accumulated in each separate event are fragmented at different HCD collision energies [340], in this case, 15, 25, and 40 are being used. Precursor ions generated from this are then analyzed together in one MS/MS Scan.

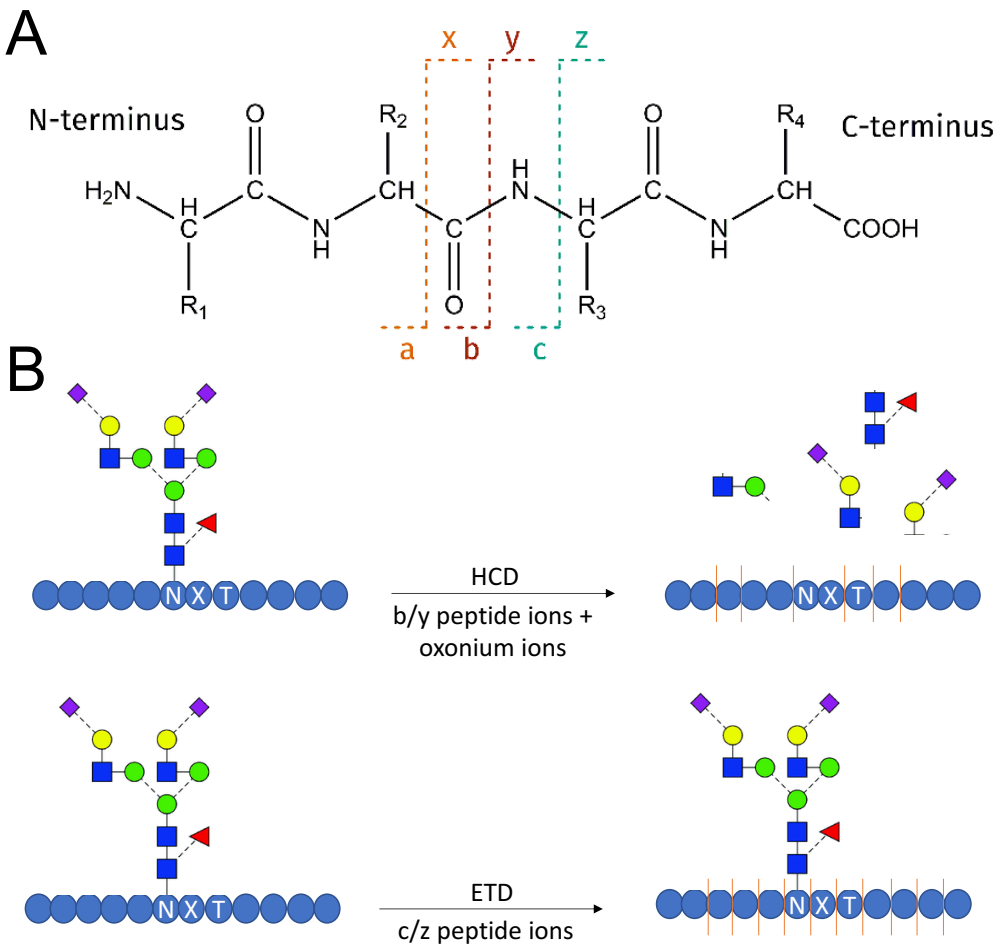


Figure 2.2 Peptide fragmentation patterns resulting from HCD and ETD fragmentation. A) Schematic of potential peptide fragmentation of a peptide consisting of four amino acids with side chains R1 to R4. The schematic is taken from Hale et al. [341]. B) Representation of fragmentation of peptide and N-linked glycans in HCD-based and ETD-based fragmentation.

2.4.2 Liquid chromatography-mass spectrometry of glycopeptides

HIV-1 env/SARS-CoV-2 S proteins were denatured for 1 hr in 50 mM Tris/HCl, pH 8.0 containing 6 M of urea. Next, the sample was reduced and alkylated by adding 5 mM dithiothreitol (DTT) and 20 mM iodoacetamide (IAA) and incubated for 1 hr in the dark, followed by a 1 hr incubation with 20 mM DTT to eliminate residual IAA. The alkylated Env glycoproteins were buffer exchanged into 50 mM Tris/HCl, pH 8.0 using Vivaspin columns (3 kDa) and three aliquots were digested separately overnight using trypsin (Mass Spectrometry grade, Promega), chymotrypsin (Mass Spectrometry

Grade, Promega) or alpha lytic protease (Sigma Aldrich) at a ratio of 1:30 (w/w) at 37°C. The next day, the peptides were dried and extracted using C18 Zip-tip (MerckMilipore) or Oasis HLB cartridges (Waters) according to manufacturer's protocol. Notably, glycopeptides desalted using HLB cartidges showed enhance yield in glycopeptides compared to C18 Zip-tip. Following the extraction, the peptides were dried again, re-suspended in 0.1% formic acid, and analyzed by nanoLC-ESI MS with an Ultimate 3000 HPLC (Thermo Fisher Scientific) system coupled to an Orbitrap Eclipse mass spectrometer (Thermo Fisher Scientific) using stepped higher energy collision-induced dissociation (HCD) fragmentation. Peptides were separated using an EasySpray PepMap RSLC C18 column (75 µm × 75 cm). A trapping column (PepMap 100 C18 3 µm particle size, 75 µm × 2 cm) was used in line with the LC prior to separation with the analytical column. The LC conditions were as follows: 280-mins linear gradient consisting of 4—32% acetonitrile in 0.1% formic acid over 260 mins followed by 20 mins of alternating 76% acetonitrile (ACN) in 0.1% formic acid and 4% ACN in 0.1% formic acid, used to ensure all the sample had eluted from the column. The flow rate was set to 300 nL/min. The spray voltage was set to 2.5 kV and the temperature of the heated capillary was set to 40 °C. The ion transfer tube temperature was set to 275 °C. The scan range was 375—1500 m/z. Stepped HCD collision energy was set to 15, 25 and 45% and the MS2 for each energy was combined. Precursor and fragment detection were performed using an Orbitrap at a resolution MS1 = 120,000. MS2 = 30,000. The AGC target for MS1 was set to standard and injection time set to auto which involves the system setting the two parameters to maximize sensitivity while maintaining cycle time. Full LC and MS methodology can be extracted from the appropriate Raw file using XCalibur Freestyle software or upon request.

Glycopeptide fragmentation data were extracted from the raw file using Byos (Version 4.0; Protein Metrics Inc.). The glycopeptide fragmentation data were evaluated manually for each glycopeptide; the peptide was scored as true-positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the glycan identified. The MS data was searched using the Protein Metrics 305 N-glycan library with sulfated glycans added manually. The relative amounts of each glycan at each site as well as the unoccupied proportion were determined by comparing the extracted chromatographic areas for different glycotypes with an identical peptide sequence. All

charge states for a single glycopeptide were summed. The precursor mass tolerance was set at 4 ppm and 10 ppm for fragments. A 1% false discovery rate (FDR) was applied. The relative amounts of each glycan at each site as well as the unoccupied proportion were determined by comparing the extracted ion chromatographic areas for different glycopeptides with an identical peptide sequence. Glycans were categorized according to the composition detected. HexNAc(2)Hex(9–3) was classified as M_9 to M_3. Any of these compositions that were detected with a fucose are classified as FM. HexNAc(3)Hex(5–6)Neu5Ac)(0–4) was classified as Hybrid with HexNAc(3)Hex(5–6)Fuc(1)NeuAc(0–1) classified as Fhybrid. Complex-type glycans were classified according to the number of processed antenna and fucosylation. Complex glycans are categorized as HexNAc(3)(X), HexNAc(3)(F)(X), HexNAc(4)(X), HexNAc(4)(F)(X), HexNAc(5)(X), HexNAc(5)(F)(X), HexNAc(6+)(X) and HexNAc(6+)(F)(X). Core glycans are any glycan smaller than HexNAc(2) Hex(3).

2.4.3 In-gel release of N-linked glycans of HexaPro S

Gel bands corresponding to the HexaPro S and glycan engineered HexaPro S protein (kifunensine-treated) were excised and washed three times with alternating 1 ml acetonitrile and water, incubating and shaking for 5 minutes following addition of each wash solution. All liquid was removed following the final wash stages and N-linked glycans were released in-gel using PNGaseF, (2 µg enzyme in 100 µL H_2O) (New England Biolabs) at 37 °C overnight. Following digestion, the liquid was removed from the gel bands and placed into a separate Eppendorf. The gel bands were then washed twice with 100 µl MilliQ H2O and this was pooled with the original solution. The extracted glycans were then dried completely in a speed vac at 30 °C.

2.4.4 Fluorescent labelling of N-linked glycans

The released glycans were subsequently fluorescently labelled with procainamide using 110 mg/ml procainamide and 60 mg/ml sodium cyanoborohydride in a buffer consisting of 70% DMSO, 30% acetic acid. For each sample, 100 µl of labelling mixture was added. Labelling was performed at 60 °C for 2 hours. Excess label and PNGaseF were removed using Spe-ed Amide-2 cartridges

(Applied Separations). First, the cartridges were equilibrated sequentially with 1 ml acetonitrile, water and acetonitrile again. Then 1 ml of 95% acetonitrile was added to the procainamide-released glycan mixture and applied to the cartridge, allowing the cartridge to drain by gravity flow. After the mixture has emptied the cartridge, two washes using 97% acetonitrile were performed. To elute the labelled glycans 1 ml HPLC grade water was added to the cartridges and the elution collected. The elution was then dried completely using a speed vac, before resuspending in 24 µl of 50 mM ammonium formate.

2.4.5 HILIC-UPLC analysis

A 6 µl aliquot of the resuspended glycans were mixed with 24 µl of acetonitrile and analysed on a Waters Acquity H-Class UPLC instrument with a Glycan BEH Amide column (2.1 mm × 150 mm, 1.7 µM, Waters), with an injection volume of 10 µl. A gradient of two buffers; 50 mM ammonium formate (buffer A) and acetonitrile (buffer B) was used for optimal separation. Gradient conditions were as follows: initial conditions, 0.5 ml/min 22% buffer A, increasing buffer A concentration to 44.1% over 57.75 minutes. Following this the concentration of buffer A was increase to 100% at 59.25 minutes and held there until 66.75 minutes and the flow rate was dropped to 0.25 ml/min, to fully elute from the column. Finally, the %A was reduced to 20% to prepare for subsequent runs. Wavelengths used for detection of the procainamide label were: excitation 310 nm, emission 370 nm. Data were processed using Empower 3 software (Waters, Manchester, UK). The relative abundance of oligomannose-type glycans was measured by digestion with Endoglycosidase H (per sample in 20 µl volume) (Endo H; New England Biolabs). A 6 µl aliquot of labelled glycans was combined with 1 µg endoH to a final volume of 20 µl. Digestion was performed overnight at 37 °C.

Digested glycans were cleaned using a 96-well PVDF protein-binding membrane (Millipore) attached to a vacuum manifold. Prior to application to the membrane, 100 µl HPLC-grade H2O was added to each sample. Following equilibration with 150 µl ethanol, and 2 × 150 µl HPLC-grade H2O, the sample was added to the 96-well plate and the flow-through was collected in a 96-well collection plate. Each well was then washed twice with HPLC-grade H2O to a final elution volume of 300 µl. The elution was then dried completely at 30 °C. Prior to analysis the sample was resuspended in 6

µl ammonium formate and 24 µl acetonitrile and analysed. The EndoH is specific for glycans that are mannosylated on the D1 and D2 arms found only in oligomannose-type and hybrid-type glycans. This allows the identification of percentage of oligomannose-type glycans in the total glycan pool.

2.5 Binding assays

2.5.1 Western blot

Proteins were resolved by SDS-PAGE gel and transferred onto a PVDF or nitrocellulose membrane using the Trans-Blot® Turbo™ Transfer system (BioRad). The membranes were blocked using 5% skim milk powder (Sigma Aldrich) in PBS + 0.5% Tween, overnight at 4°C or for 1 hr at room temperature. The membranes were washed thrice in in 0.5% Tween-PBS and the primary antibody (1:1,000 – 3,000) was incubated in 0.5% Tween-PBS for 2 hr at room temperature. The membranes were washed again to remove any non-specific binding and incubated with the secondary antibody (goat anti-human IgG-HRP, 1:2,000; Abcam) in 0.5% Tween-PBS for 1 hour. The membranes were washed again thrice in 0.5% Tween-PBS and then Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific) were used to stain the proteins of interest. The images were captured with the Fuji Intelligent Dark Box ii and Image Reader 1000 Pro software.

2.5.2 Determination of affinity using surface plasmon resonance (SPR)

Analysis of SARS-CoV-2 HexaPro binding with ACE2 protein was analyzed using a Biacore T200 (Cytiva/GE Healthcare). The proteins were buffer exchanged in the running buffer used for the SPR, HEPES Buffered Saline (HBS-P+) (0.01 M HEPES pH 7.4, 0.15 M NaCl, 0.005% v/v Surfactant P20). Prior to injection of $NiCl_2$, metallic contaminants were removed via a pulse of 350 mM ethylenediaminetetraacetic acid (EDTA) at a flow rate of 30 µL/min for 1 min. Followed by loading of Ni^{2+} at a flow rate of 10 µL/min for 1min. SARS-CoV-2 S protein (50 µg/mL), ligand was injected at 10 µL/min for 240 s. Flow cell 2-1 was used in which one of the cells was used as a control for determination of non-specific binding to the chip. Control cycles were performed by flowing the analyte (ACE2 protein) over the control cell having absence of ligand (S protein); negligible binding was indicated. The analyte was serially diluted ranging from 200 nM to 3.125 nM in triplicated along

with HBS P+ buffer only as a control and were injected at a flow rate of 50 μL/min. Association and dissociation time was set as 300 s and 600 s respectively. After each cycle, the chip was regenerated with a pulse EDTA (350 mM) for 1 min at a flow rate of 30 μL/min. The 1:1 binding model was used for fitting the resulting data using Biacore Evaluation Software (GE Healthcare) and subsequently fitted curves were used to calculate K_D. Paired t tests were used and graph plotting was performed using GraphPad prism v8.1.

2.5.3 Serum ELISA methodology (Chapter 4)

All sera were obtained by centrifugation of whole blood at 3500 rpm for 5 mins, then stored at -20°C until use. Antibodies to S glycoprotein were detected using an in-house developed, high-sensitivity ELISA, as previously described [107]. In short, Nunc 96-well plates (ThermoFisher) or high binding plates (Corning) were coated with 2 μg/mL 2P or HexaPro or kifunensine HexaPro and blocked with 2% bovine serum albumin (BSA) (Sigma Aldrich) (w/v) in PBS-T (Oxoid phosphate buffered saline with 0.1% Tween-20, Sigma Aldrich). Serum was initially diluted 1:40 in PBS-T and then serially diluted. Secondary antibodies (combined horse radish peroxidase-conjugated mouse anti-human IgG, A and M monoclonal antibodies) were diluted in PBS-T (anti—IgG R-10 1:8,000, anti–IgA MG4.156 1:4,000, and anti-IgM AF6 1:2,000; Abingdon Health). Signal was developed using TMB-Core (Bio-Rad) for between 6 and 12 mins then stopped with 0.2 M H_2SO_4 (Sigma-Aldrich). Optical density (OD) at 450 nm was detected using the Dynex DSX automated liquid handler (Dynex Technologies, USA). Signal: noise ratio (S:N ratio) was calculated by dividing the individual optical density (OD) values from PCR+ serum samples (signal) by the average OD from the pre-2019 negative controls (noise). Statistical significance was assessed using a RM 2-way ANOVA with Geisser-Greenhouse correction, followed by Sidak's multiple comparisons test, using GraphPad Prism version 8.

2.6 Patient Sample collection and ethical approval (Chapter 4)

Serological responses to 2P and HexaPro forms of SARS-CoV-2 S glycoprotein were analyzed in samples from acutely unwell intensive treatment unit (ITU) patients with SARS CoV-2, convalescent

individuals who have had mild disease and normal control sera from pre-2019. We have previously shown that severity of disease affects the quantitation of antibody [107] and so a spectrum of samples were used under ethics gained to aid assay development (NRES Committee West Midlands - South Birmingham 2002/201 Amendment Number 4, 24 April 2013) and from a Convalescent health care worker cohort (London - Camden & Kings Cross Research Ethics Committee 20/HRA/1817). Hospitalized subjects also provided nasopharyngeal swabs which were guanidine isothiocyanate inactivated, then analyzed by revers-transcriptase PCR directed against the SARS-CoV-2 ORF1ab and N genes (Viasure, CerTest Biotec). Pre-2019 negative control sera were obtained as part of a University of Birmingham study, reference ERN_16-178. All study participants gave written, informed consent and samples were fully anonymized.

2.7 Integrative modelling and molecular dynamics simulation

Three S protein models were built using Modeller version 9.21 [342]: i) 2P with one RBD in the "up" conformation, ii) HexaPro with one RBD in the "up" conformation, and iii) HexaPro with two RBDs in the "up" conformation. For 2P, the ECD was modelled using the cryo-EM structure of SARS-CoV-2 2P S ECD in the open state (PDB: 6VSB) [109]. The ECD of the one-RBD-up HexaPro was modelled using the cryo-EM structure of SARS-CoV-2 HexaPro S with one RBD up (PDB: 6XKL) [110]. For the two-RBD-up ECD, we performed molecular dynamics flexible fitting (MDFF) [343], whereby the atomic coordinates of the HexaPro one-RBD-up structure was fitted to the electron density map of HexaPro two-RBD-up (EMDB: EMD-22222) [110]. The initial structure was prepared in VMD [344] and MDFF was performed in vacuum using NAMD version 2.11 [345] and the CHARMM36 force field [346]. The MDFF simulation was performed until convergence using a range of scaling factors from 0.3 to 40 with secondary structure and domain restraints applied to the protein. The stalk and the transmembrane (TM) domain of all three models were built using the NMR structure of SARS-CoV HR2 domain (PDB: 2FXP) [347] and the NMR structure of HIV-1 gp-41 TM domain (PDB: 5JYN) [348], respectively, while missing loops in the NTD and the C-terminus of the ECD were modelled based on a higher resolution cryo-EM structure of S protein in the closed state (PDB: 6XR8) [349]. The same modelling protocol previously described to build a

full-length model of the wild-type SARS-CoV-2 S protein was used [350]. CHARMM-GUI Glycan Reader and Modeller [351] was used to add oligomannose (Man-9) glycans to all 22 glycosylation sites. The glycosylated S protein models were then embedded into a pre-equilibrated model of the endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membrane [352] built using CHARMM-GUI Membrane Builder [353]. The system was solvated with TIP3P water molecules and neutralized with 0.15 M NaCl salt. Stepwise energy minimization and equilibration simulations with decreasing amount of positional and dihedral restraints were conducted following the standard CHARMM-GUI protocols [354]. For each S protein model, a 200 ns production simulation was performed. The Nosé-Hoover thermostat was used to maintain the temperature at 310 K [355,356], while a semi-isotropic coupling to the Parrinello-Rahman barostat was used to maintain the pressure at 1 atm [357]. The electrostatic interactions were calculated using the smooth particle mesh Ewald method with a real-space cut-off of 1.2 nm [358], and the van der Waals interactions were truncated at 1.2 nm with a force switch smoothing applied between 1.0 and 1.2 nm. The simulations employed a 2 fs integration time step with the LINCS algorithm applied on all covalent bonds involving hydrogen atoms [359]. All simulations were run using GROMACS 2018 [360] and the CHARMM36 force field [346]. ASA calculation was performed using the GROMACS tool *gmx sasa*.

2.8 Negative-stain electron microscopy and 2D class averaging

3 μL of protein at 0.02 mg/mL was applied to a 400 mesh Cu grid, blotted with filter paper, and stained with 2 % uranyl formate. Micrographs were collected on a Thermo Fisher Tecnai Spirit microscope operating at 120 kV with a FEI Eagle CCD (4k) camera (2.06 Å/pixel; 52,000x magnification) using Leginon automated image collection software [361]. Particles were picked using DogPicker [362] and 2D classification was processed with iterative multivariate statistical analysis (MSA)/multireference alignment (MRA) [363].

2.9 Model construction

The structural models used for presentation of N-linked glycans of BG505 NFL.664 is modelled on BG505 SOSIP.664 Env structure (PDB ID: 5C7K) as both of the structures are highly similar and

superimpose on each other [206,217]. The structural models used for presentation of N-linked glycans of SARS-CoV-2 S is modelled on full-length model of S protein (PDB ID: 6XKL).

Chapter 3 Glycosylation and serological reactivity of an expression-enhanced SARS-CoV-2 viral spike mimetic

Contributions.

Chapter 3 involves the glycan and serological analysis of SARS-CoV-2 spike, HexaPro. External collaborators from the laboratory of Dr. Peter J. Bond (National University of Singapore, Singapore) contributed to this study by providing MD simulations of HexaPro S and S2P protein outlined in this chapter and Appendix A.1. The MD-simulation section, 3.4 is written by Firdaus Samsudin and Peter J bond. Plasmid for HexaPro S and S2P was kindly provided by Prof. Jason Mclellan (University of Austin, USA). In addition, external collaborators from laboratory of Dr. Alex G Ritcher and Dr. Adam F Cunningham (University of Birmingham) did the serological analysis. I drafted the result section and figures for serological analysis. The work presented in this chapter is published in the *Journal of Molecular Biology.* The paper is primarily written by me and has been reviewed by all the co-authors in the paper.

3.1 Introduction

Recombinant viral glycoproteins are an important resource for vaccine development, diagnostics and as research reagents. Viral glycoprotein glycosylation can influence an extensive range of properties including immunogen trafficking [122], immunogenicity [364,365], antigenicity [366,367] and serum clearance rates [368]. Importantly, recombinant viral spike glycosylation can be influenced both by features of the glycoprotein sequence, such as glycan density and protein architecture [230,232,369] and an array of expression conditions, such as producer cell type and expression conditions [370–373]. It is therefore important to define the glycosylation of recombinant viral glycoproteins and monitor changes that may occur during target optimization and the development of manufacturing procedures [374,375]. As carbohydrates on viral proteins can influence the immune response, it is important to look at the binding of sera antibodies to antigens possessing distinct glycoforms [376]. This study investigates the antigenic properties of glycoprotein reagents developed

in response to the coronavirus infectious disease 2019 (COVID-19) pandemic, focused on the viral S glycoprotein [109,110].

The causative agent of COVID-19, Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), is a positive-sense single-stranded RNA virus that has caused significant morbidity and mortality throughout the world [377,378]. Like other coronaviruses, SARS-CoV-2 utilizes the S glycoprotein for recognition of the host cell entry receptor and subsequent membrane fusion, which is mediated by the S1 and S2 subunits, respectively [1]. The S protein is a trimeric class I fusion protein and is post-translationally cleaved into S1 and S2 subunits using the host cellular protease, furin [379]. The S1 subunit possesses an N-terminal domain (NTD) and a receptor-binding domain (RBD), also known as domain A and B, respectively [380].

The exposed position of the RBD enables binding to the angiotensin converting enzyme 2 (ACE2) receptor [381] and, as a result, it is the main target of anti-SARS antibodies during infection [98,99,382–385]. Due to this phenomenon, combined with its high recombinant protein yields, several antibody tests have been developed using RBD as a tool to test for previous SARS-CoV-2 infection [386–389]. One disadvantage of using RBD as an antigenic bait for testing is that it may not capture the entire antibody response to the S protein as it lacks the full trimeric structure [107]. In addition to RBD, other antibody tests use the nucleoprotein (N protein) as antigenic bait to detect prior SARS-CoV-2 infection, such as the Abbott test [390,391]. Similarly to the RBD, N protein will not capture the complete antigenic surface and therefore may not reveal the full immune response to SARS-CoV-2 infection. As the S protein is the prime target of neutralizing antibodies, the native-like trimeric S glycoprotein may facilitate the presentation of a more complete range of epitopes for antibody testing [392,393]. A vaccine against COVID-19 must reproduce or enhance, the immune response to infection. In response to COVID-19 infection, the immune system is capable of inducing a strong, long lived, neutralizing antibody response [394,395]. Previous studies have shown the isolation of antibodies from the COVID-19 infected individuals and mapped the antibody response onto the S protein surface. Notably, the RBD of SARS-CoV-2 is an immunodominant epitope [99,382–385,396]. Furthermore, the antibody response against the RBD was highly diverse, with the

humoral immune system capable of utilizing a wide range of different heavy chain genes to target the RBD [383].

Prefusion stabilization strategies have been employed to stabilize the S protein as neutralizing antibody epitopes are predominantly presented on the prefusion conformation [201,332,397,398]. A common strategy is the introduction of proline substitutions which impedes the switch to the post fusion conformation [333]. As the proline residues are supposed to block the formation of long helices that are present in the intermediate and post-fusion conformations [112]. For SARS-CoV-2, the expression of a stable, soluble form of the S-protein was originally achieved by truncation at the transmembrane domain and the incorporation of two proline residues (K986P and V987P) [109] (**Figure 3.1**; SARS-CoV-2 2P S, henceforth termed "2P"). Despite the utility of the 2P construct for structural analysis [109] and serological testing [107,399,400], the low expression levels prompted the development of an expression enhanced version containing four additional prolines (**Figure 3.1**; SARS-CoV-2 HexaPro, henceforth termed "HexaPro") [110]. Hsieh et al observed that multiple proline substitution boosted the expression and thermostability of the protein and the combination of beneficial four proline substitutions incorporated in HexaPro showed 10-fold increase in expression compared to S2P variant [110]. HexaPro exhibits native-like protein architecture, antigenic properties, and contains the twenty-two N-linked glycosylation sites of the native viral spike [8,110,125]. Additionally, HexaPro has shown promising results as a vaccine candidate in mice immunization, resulting in high-titre neutralizing antibodies [162,401,402]

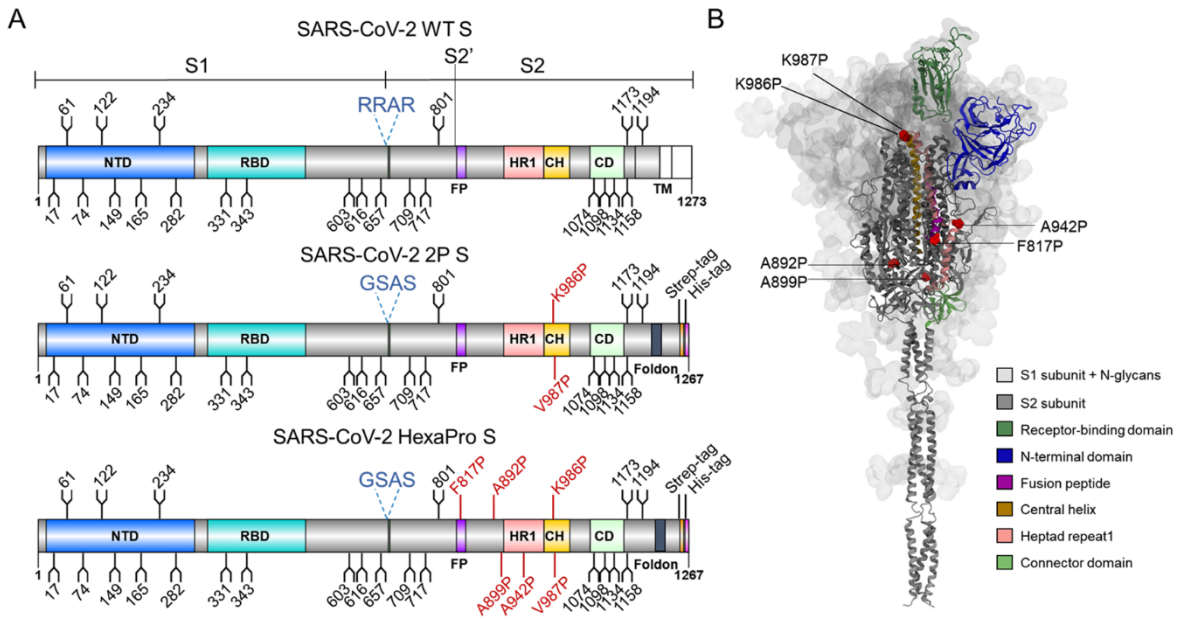


Figure 3.1. Representation and characterization of recombinant SARS-CoV-2 spike protein. A) The protein domains are represented as N-terminal domain (NTD), receptor-binding domain (RBD), fusion peptide (FP), heptad repeat 1 (HR1), central helix (CH), connector domain (CD), and transmembrane domain (TM). The fusion cleavage site is illustrated as dashed lines (blue). N-linked glycosylation sequons (N-X-S/T, where X≠P) are shown as branches. SARS-CoV-2 WT presents S1 and S2 domain with furin cleavage site (RRAR) and transmembrane domain at C-terminal end. SARS-CoV-2 2P prefusion stabilized protein with proline substitutions at residues 986 and 987 and, a "GSAS" mutation at furin cleavage site (residues 682-685). HexaPro prefusion stabilized protein of SARS-CoV-2 with a "GSAS" mutation at furin cleavage site and six proline substitutions, highlighted in red. B) Structural representation of HexaPro S protein illustrating six proline substitutions (red spheres) in SARS-CoV-2 ectodomain (PDB ID: 6XKL). The S1 subunit along with N-glycans are shown as transparent molecular surface. The S2 subunit is shown in dark grey. Different domains present in S1/S2 subunits are highlighted in respective colors in ribbon diagram of only one protomer.

This study compares the protein expression, binding analysis, and glycan compositions of expression enhanced variant, HexaPro with S2P variant of recombinant S protein. To investigate the differences in glycan composition among HexaPro and S2P, liquid chromatography-mass spectrometry (LC-MS) was used. Furthermore, MD-simulation was conducted to compare the protein conformations of both variants. HexaPro S protein, known for its stability and enhanced recombinant expression, can serve as a valuable serological reagent. This study further validates HexaPro's utility through serological binding analysis with sera from infected individuals. Overall, this information underscores HexaPro's stability and its high resemblance to 2P, demonstrating HexaPro's potential as an effective immunogen and a serological testing reagent.

3.2 Expression and binding analysis of SARS-CoV-2 HexaPro S protein

For characterization of prefusion stabilized SARS-CoV-2 S protein, we transiently transfected plasmid encoding SARS-CoV-2 S protein containing the HexaPro stabilizing mutations in Human embryonic kidney (HEK) 293F cells. To ensure the analysis of only native-like trimeric protein, the supernatant was first purified using nickel-affinity chromatography followed by size exclusion chromatography (SEC) (**Figure 3.2.A**). To functionally characterize the binding of the expressed protein with ACE-2, the binding affinity (expressed here using the dissociation constant, K_D) of HexaPro with a soluble recombinant ACE2 was determined using surface plasmon resonance (SPR) which was repeated three times analytically to determine the average K_D (**Figure 3.2.B**). A residual plot of the data revealed minimal deviation between observed values and calculated values using a 1:1 binding model between SARS-CoV-2 HexaPro S protein and ACE2 (**Figure 3.2.C**). The K_D values are comparable to that previously reported for 2P [140] (**Figure 3.2.D**).

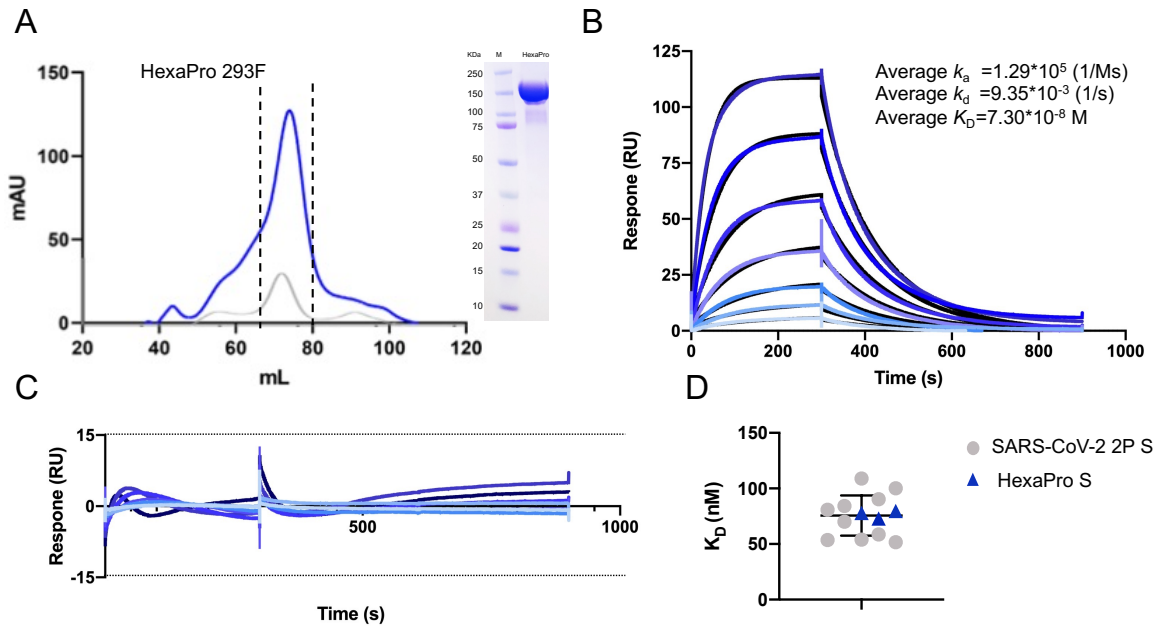


Figure 3.2. Production of purified S protein and determination of S protein binding with soluble ACE2. A) SEC of affinity-purified recombinantly expressed S protein, HexaPro peaks shown in blue and S2P peaks shown in grey.. Dashed lines indicate fractions collected for subsequent use. B) SPR of HexaPro (ligand) with soluble ACE2 receptor (analyte). Dark blue to light blue lines represent the serial dilutions of ACE2 protein from 200 nM to 3.125 nM, respectively. Black lines are fitted values of the respective concentration to illustrate the best fit to a 1:1 binding model. Three repeats were performed and averaged to determine the k_a, k_d and K_D values. C) Residual plot illustrating the deviation of the fitted data to the raw values of the experimental data at different concentrations. D) Representation of K_D values determined using various repeats, grey dots illustrate the binding of 2P with ACE2 (values reproduced from [140]) and the binding of HexaPro with ACE2 are shown as blue triangles. The mean of K_D values of 2P is plotted as a black line and the error bars represent +/- standard deviation calculated using GraphPad Prism.

3.3 Determination of impact of proline mutations on Spike glycosylation

To determine the effect of stabilizing mutations on glycosylation, the site-specific glycan compositions of SARS-CoV-2 HexaPro was determined. In preparation for glycopeptide analysis, purified three replicates of the soluble ectodomain of the SARS-CoV-2 protein truncated at the transmembrane domain were independently expressed. This construct was used to determine the high resolution cryo-electron microscopy (cryo-EM) structure [110]. To generate glycopeptide samples derived from these batches of protein, three different protease enzymes, trypsin, chymotrypsin, and alpha-lytic protease were employed. These proteases cleave the protein chain at specific amino acids and were selected to generate glycopeptides that contain a single glycosylation sequon (N-X-S/T-X, where X-any amino acid except proline). Liquid chromatography-electrospray ionization mass

spectrometry (LC-ESI MS) was used to quantify the oligomannose-type glycans, complex-type glycans and the proportion of unoccupied potential N-glycosylation sites (PNGS) across all 22 N-linked glycan sites on the HexaPro protomers (**Figures 3.3.A and 3.3.B**).

To determine the potential impact of more extensive modifications on the prefusion stabilized S protein, glycan profile of SARS-CoV-2 S protein of 2P was compared with HexaPro. Interestingly, the overall processing states of the recombinant S protein were conserved across both the versions, with few variations at the site-specific level (**Figures 3.3.B and 3.4.A**). Site-specific analysis of these recombinant proteins suggests that the levels of oligomannose-type glycans are consistent with native S protein on infectious virus and also with other coronaviruses [7,8,124–126].

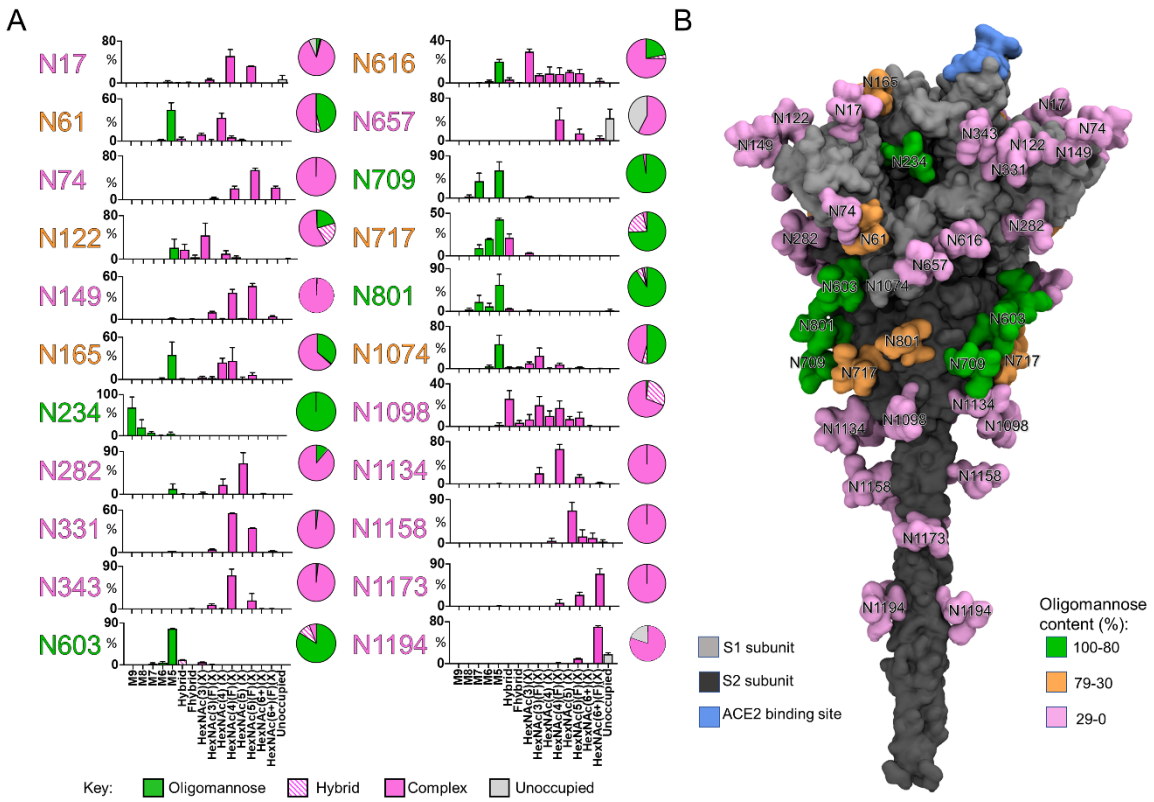


Figure 3.3. Site-specific glycosylation of expression-enhanced recombinant trimer of SARS-CoV-2 S protein (HexaPro). A) Relative quantification of the N-linked glycosylation sites of trimeric S protein, produced in HEK293F cells. The bar graph represents the mean of three independently expressed replicates with error bars representing the standard error of the mean. The color codes in the schematic illustrates the processing state of glycans from least processed to most processed, oligomannose (green), hybrid (dashed pink), and complex glycans (pink). The proportion of unoccupied N-linked glycan sites are represented in grey. The pie charts summarize the quantification of these categories. The N-linked glycan site labels are colored based on the oligomannose-type glycan content, green (80-100%), orange (30-79%) and magenta (0-29%). B) The model was constructed using the prefusion structure of trimeric SARS-CoV-2 S glycoprotein as described in Materials and Methods. The S1 and S2 subunits are shown as light and dark grey, respectively. The glycans sites are categorized as high-mannose type glycans (green), hybrid glycans (orange), and complex-type glycans (pink). The ACE2 receptor binding site is shown in blue.

In addition to the information obtained from studying the populations of oligomannose-type glycans at individual glycosylation sites, understanding the processing of complex-type glycans is also informative when considering immunogen design, reagents for serological studies, and for understanding the extent to which recombinant glycoprotein can be used as mimics of the functional viral spike. For example, the epitope of neutralizing antibody S309, which targets the S protein of SARS-CoV-1 and SARS-CoV-2, contains a fucosylated glycans at N343 [403]. HexaPro is 99% fucosylated at the N343 site, with almost all the glycans bearing fucose residues (**Table 3.1.A**). Furthermore, sulfated N-linked glycans have been detected on viral glycoproteins and they could potentially play role in immune regulation, as in influenza [404,405]. The presence of these sulfated groups in haemagglutinin (HA) protein on the surface of influenzae virus dramatically enhance the binding with the receptor and fully glycosylated HA maximizes the interaction which could markedly impact the viral replication, receptor binding, antigenicity, and elicitation of innate immunity [404,405]. We detected sulfation at several N-glycosylation sites on HexaPro (N74, N149, N1194) and observed a similar abundance of sulfation across both S proteins (2P and HexaPro) (**Figure 3.4**), which is in accordance with analysis of other SARS-CoV-2 S proteins [50,406] (**Table 3.1.A and B**).

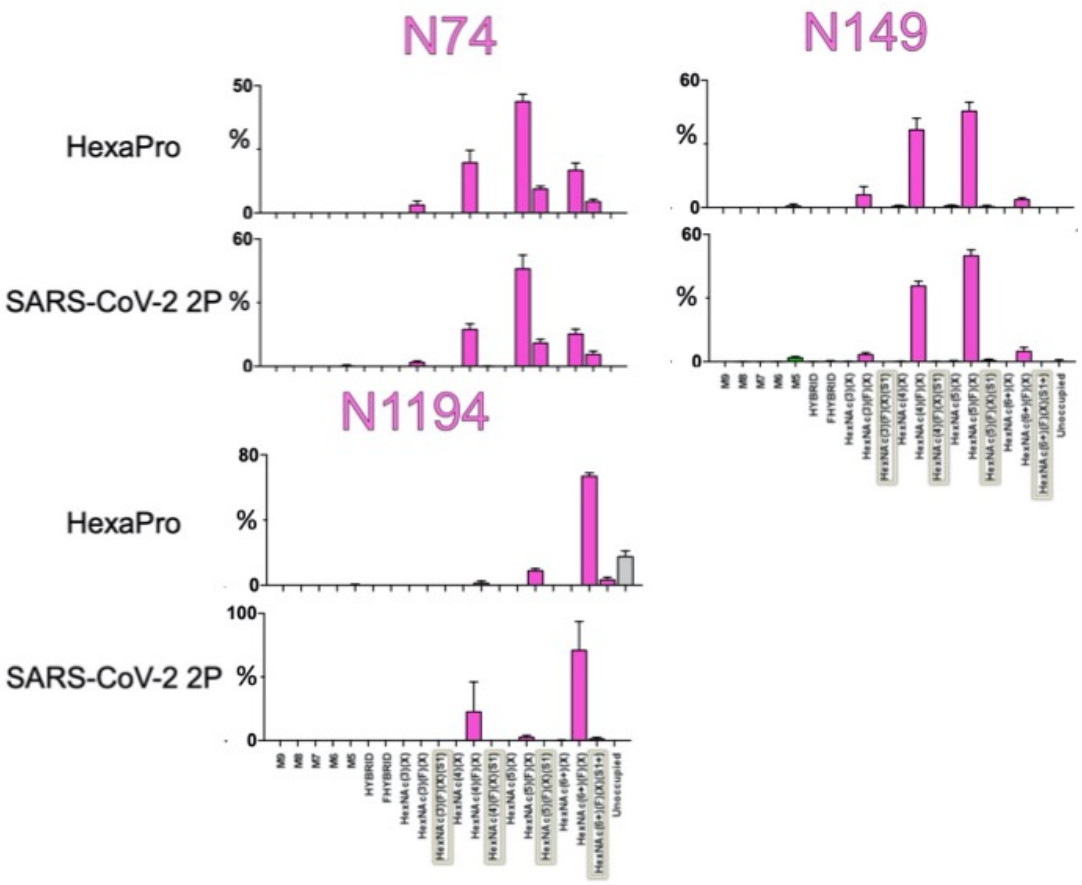


Figure 3.4. Representation of extensive site-specific analysis of N-linked glycosylation sites of SARS-CoV-2 2P and HexaPro. The recombinant proteins were digested using multiple proteases, analysed by LC-MS, and the resulting data was searched using library containing sulfo groups. The bar graph illustrates the values from three biological repeats. The sites represented here are the only ones which has shown the sulfation in the analysis. The abbreviation used here are M, mannose; X, Hex & Neu5Ac; F, Fucose, S, sulfation.

Table 3.1. Abundance of glycoform observed across SARS-CoV-2 HexaPro S and 2P. The left-hand table represents the average of glycan composition at each N-linked glycan site obtained from the reported values of three biological replicates. The right-hand table is representing the global averages of glycan compositions at all the N-linked glycan sites. The SARS-CoV-2 S 2P data (B) is reanalysed to include sulfated glycans from the previously published analysis [8,407].

A

	N17	N61	N74	N122	N149	N165	N234	N282	N331	N343	N603	N616	N657	N709	N717	N801	N1074	N1098	N1134	N1158	N1173	N1194
M9	0	0	0	0	0	0	68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M8	0	0	0	0	0	0	19	0	0	0	0	0	0	3	1	4	0	0	0	0	0	0
M7	0	0	0	0	0	0	7	0	0	0	2	0	0	35	9	20	0	0	0	0	0	0
M6	0	2	0	0	0	1	1	0	0	0	4	2	0	0	20	10	4	0	0	0	0	0
M5	3	45	0	20	1	35	5	11	2	0	78	20	0	59	43	56	46	2	0	0	0	0
Hybrid	0	3	0	17	0	1	0	0	0	0	10	3	0	0	21	5	2	26	0	0	0	0
Fhybrid	1	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	2	3	0	0	0	0
HexNAc(3)(X)	0	9	4	44	0	3	0	2	0	0	5	29	0	2	4	2	9	6	0	0	0	0
HexNAc(3)(F)(X)	6	1	2	0	9	2	0	0	5	9	1	7	0	0	0	0	25	20	20	0	0	0
HexNAc(4)(X)	0	33	0	10	1	24	0	20	0	0	0	9	0	0	0	0	1	10	0	5	0	0
HexNAc(4)(F)(X)	51	5	32	4	37	26	0	0	56	72	0	8	40	0	0	0	8	18	65	0	7	2
HexNAc(5)(X)	0	2	4	0	1	2	0	65	0	0	0	10	0	0	0	0	0	7	0	67	0	0
HexNAc(5)(F)(X)	32	0	35	0	46	7	0	0	35	18	0	9	14	0	0	0	0	3	8	13	14	24	9
HexNAc(6+)(X)	0	0	7	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	10	0	0
HexNAc(6+)(F)(X)	0	0	15	0	4	0	0	0	2	0	0	2	4	0	0	0	0	0	1	3	69	71
Unoccupied	7	0	0	1	0	0	0	0	0	0	0	0	42	0	0	2	0	0	0	0	0	18

Total	
M9	3
M8	1
M7	3
M6	2
M5	19
Hybrid	4
Fhybrid	1
HexNAc(3)(X)	6
HexNAc(3)(F)(X)	5
HexNAc(4)(X)	5
HexNAc(4)(F)(X)	20
HexNAc(5)(X)	7
HexNAc(5)(F)(X)	12
HexNAc(6+)(X)	1
HexNAc(6+)(F)(X)	8
Unoccupied	3

%	N17	N61	N74	N122	N149	N165	N234	N282	N331	N343	N603	N616	N657	N709	N717	N801	N1074	N1098	N1134	N1158	N1173	N1194
Mannose	3	46	0	21	1	36	100	11	2	0	84	22	0	98	74	91	50	2	0	0	0	0
Hybrid	1	3	2	21	0	1	0	0	0	1	10	3	0	0	22	5	4	29	0	0	0	0
Complex	89	50	98	58	98	63	0	89	98	99	6	75	58	2	4	2	46	69	100	100	100	81
Unoccupied	7	0	0	1	0	0	0	0	0	0	0	0	42	0	0	2	0	0	0	0	0	18
Fucosylation	90	6	85	7	96	35	0	0	98	99	1	27	58	0	1	0	37	47	100	17	100	78
Sialylation	21	5	30	5	34	7	0	15	11	74	0	0	8	0	1	0	2	50	2	63	3	35

Total	
Mannose	29
Hybrid	5
Complex	63
Unoccupied	3
Fucosylation	45
Sialylation	17

B

	N17	N61	N74	N122	N149	N165	N234	N282	N331	N343	N603	N616	N657	N709	N717	N801	N1074	N1098	N1134	N1158	N1173	N1194
M9	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M8	0	0	0	0	0	0	44	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0
M7	0	0	0	0	0	0	14	0	0	0	0	0	0	7	17	10	0	1	0	0	0	0
M6	0	1	0	0	0	0	4	0	0	0	2	1	0	26	25	17	2	2	0	0	0	0
M5	3	68	1	56	2	0	4	1	1	2	62	5	0	59	32	49	55	6	0	0	1	0
Hybrid	0	8	0	17	0	3	1	0	0	0	3	3	0	0	21	7	8	19	0	0	0	0
Fhybrid	2	1	0	2	0	1	0	0	0	2	0	1	0	0	1	0	1	5	0	0	0	0
HexNAc(3)(X)	0	5	0	3	0	0	0	4	0	0	4	10	0	0	2	3	5	14	0	1	0	0
HexNAc(3)(F)(X)	6	1	2	5	4	0	1	0	4	12	4	2	3	0	1	1	5	7	15	0	0	0
HexNAc(4)(X)	0	6	0	3	0	14	0	11	0	0	0	18	0	0	0	1	2	19	0	48	0	0
HexNAc(4)(F)(X)	32	4	18	10	36	29	1	6	53	46	23	37	39	5	1	8	12	9	44	0	4	23
HexNAc(5)(X)	0	6	0	3	0	20	0	30	0	0	0	6	0	0	0	0	0	5	0	44	0	0
HexNAc(5)(F)(X)	53	1	58	1	51	32	0	46	38	37	1	13	37	1	0	2	8	7	23	0	25	3
HexNAc(6+)(X)	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	3	0	6	0	0
HexNAc(6+)(F)(X)	3	0	22	0	5	0	0	0	3	0	0	2	4	0	0	0	1	4	18	0	70	73
Unoccupied	1	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0

Total	
M9	1
M8	2
M7	2
M6	4
M5	19
Hybrid	4
Fhybrid	1
HexNAc(3)(X)	2
HexNAc(3)(F)(X)	3
HexNAc(4)(X)	6
HexNAc(4)(F)(X)	19
HexNAc(5)(X)	6
HexNAc(5)(F)(X)	19
HexNAc(6+)(X)	1
HexNAc(6+)(F)(X)	10
Unoccupied	1

%	N17	N61	N74	N122	N149	N165	N234	N282	N331	N343	N603	N616	N657	N709	N717	N801	N1074	N1098	N1134	N1158	N1173	N1194
Mannose	4	69	1	56	2	0	97	1	2	2	65	6	0	94	74	77	57	9	0	0	1	0
Hybrid	2	9	0	19	0	5	1	0	0	3	3	4	0	0	22	7	9	24	0	0	0	0
Complex	94	22	99	25	97	95	2	99	98	96	32	90	83	6	4	16	33	67	100	100	99	100
Unoccupied	1	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0
Fucosylation	95	6	100	18	96	63	2	52	98	98	29	56	83	5	3	11	27	31	100	0	99	100
Sialylation	20	4	41	5	31	18	0	14	22	4	0	1	2	0	2	5	6	30	7	39	4	69

Total	
Mannose	28
Hybrid	5
Complex	66
Unoccupied	1
Fucosylation	53
Sialylation	15

In all formats of SARS-CoV-2 S expressed in mammalian cells, a higher proportion of complex-type glycans compared to oligomannose-type glycans were observed [8,50,125,126,261,404]. Moreover, the complex-type glycans somewhat obscure immunogenic surfaces and constitute a shield to evade the immune system [131,226,249], albeit not at a level observed in many other viral envelopes [1,7]. Out of 22 N-linked glycan sites on each protomer,

HexaPro S presents more than 50% highly processed complex-type glycans on 15 N-linked sites which is comparable to the 2P expressed in different laboratories [8,125]. The underoccupancy at the glycosylation sequon at N657 is present on both HexaPro (**Figures 3.3.A** and **Table 3.1.A**) and 2P (**Table 3.1.B**). The glycan site at the C-terminus, N1194, is fully occupied in the case of 2P (however, there is an elevation of unoccupied PNGS (18%) at N1194 on in the HexaPro S protein (**Figure 3.4**).

The oligomannose-type glycan content of the glycans of the HexaPro protein (29%) (**Table 3.1.A**) is lower when compared to other viral glycoproteins including HIV-1 Env (60%) and LASV GPC (49%) [11,335]. This is consistent with earlier observations using 2P protein which indicated that SARS-CoV-2 S is less shielded with glycans, which may be beneficial for the elicitation of neutralizing antibodies [8]. Overall, there is a similar level of oligomannose-type glycans across both 2P and HexaPro (**Table 3.2**).

Table 3.2. **Comparison of composition of glycoforms across variants of SARS-CoV-2 S protein.** The table illustrates the difference between HexaPro and 2P at each N-linked glycan site from the average of three biological repeats of each variant. The top left-panel represents the average of the glycan composition difference between HexaPro and 2P at individual N-linked glycan sites. The top right-panel represents the global average difference between HexaPro and 2P at all N-linked glycan sites. The lower left-panel represents the sum of M9-M5 as mannose, sum of hybrid and Fhybrid as Hybrid, sum of HexNAc(3)(X) to HexNAc(6+) (F)(X) as complex, and unoccupied across all N-glycan sites. The lower-right panel represents the sum of global average composition difference between HexaPro and 2P. The increase in abundance of glycan composition in HexaPro is represented in range from yellow to green whereas the increase in abundance of glycan composition in 2P is shown in range from yellow to red".

Comparison	N17	N61	N74	N122	N149	N165	N234	N282	N331	N343	N603	N616	N657	N709	N717	N801	N1074	N1098	N1134	N1158	N1173	N1194
M9	0	0	0	0	0	0	38	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M8	0	0	0	0	0	0	-25	0	0	0	-1	0	0	2	1	4	0	0	0	0	0	0
M7	0	0	0	0	0	0	-7	0	0	0	2	0	0	28	-7	10	0	-1	0	0	0	0
M6	0	1	0	0	0	1	-3	0	0	0	2	1	0	-26	-5	-7	2	-2	0	0	0	0
M5	0	-23	0	-36	-1	35	1	10	0	-2	16	14	0	0	11	6	-9	-5	0	0	0	0
Hybrid	0	-4	0	0	0	-3	-1	0	0	0	7	0	0	0	1	-2	-5	7	0	0	0	0
Fhybrid	-1	-1	2	2	0	-1	0	0	0	-2	0	-1	0	0	-1	0	1	-2	0	0	0	0
HexNAc(3)(X)	0	4	4	41	0	3	0	-1	0	0	1	19	0	2	2	-1	4	-7	0	-1	0	0
HexNAc(3)(F)(X)	1	0	-1	-5	6	2	-1	0	1	-4	-4	5	-3	0	-1	-1	19	13	5	0	0	0
HexNAc(4)(X)	0	28	0	7	0	10	0	9	0	0	0	-9	0	0	0	-1	-1	-9	0	-43	0	0
HexNAc(4)(F)(X)	19	1	14	-7	1	-3	-1	-6	3	25	-23	-29	1	-5	0	-8	-5	9	21	0	3	-22
HexNAc(5)(X)	0	-4	4	-3	1	-18	0	35	0	0	0	4	0	0	0	0	0	1	0	23	0	0
HexNAc(5)(F)(X)	-21	-1	-23	-1	-5	-26	0	-46	-3	-19	-1	-4	-23	-1	0	-2	-5	1	-11	14	-1	6
HexNAc(6+)(X)	0	0	7	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	-3	0	4	0	0
HexNAc(6+)(F)(X)	-3	0	-7	0	-1	0	0	0	-1	0	0	0	0	0	0	0	0	-4	-16	3	-1	-3
Unoccupied	6	0	0	1	0	0	0	0	0	0	0	0	26	0	0	2	0	0	0	0	0	18

Total	
M9	2
M8	-1
M7	1
M6	-2
M5	0
Hybrid	0
Fhybrid	0
A1	3
FA1	1
A2/A1B	-1
FA2/FA1B	1
A3/A2B	1
FA3/FA2B	-7
A4/A3B	0
FA4/FA3B	-2
Unoccupied	2

%	N17	N61	N74	N122	N149	N165	N234	N282	N331	N343	N603	N616	N657	N709	N717	N801	N1074	N1098	N1134	N1158	N1173	N1194
Mannose	-1	-23	0	-36	-1	36	3	10	0	-2	19	16	0	4	-1	14	-7	-7	0	0	0	0
Hybrid	-1	-5	2	2	0	-4	-1	0	0	-2	7	-1	0	0	-1	-2	-5	5	0	0	0	0
Complex	-5	28	-2	33	2	-32	-2	-10	0	4	-26	-15	-26	-4	1	-14	12	2	0	0	0	-19
Unoccupied	6	0	0	1	0	0	0	0	0	0	0	0	26	0	0	2	0	0	0	0	0	18

Total	
Mannose	1
Hybrid	0
Complex	-3
Unoccupied	2

Thus this study reveals that the site-specific glycosylation of the expression enhanced version of SARS-CoV-2, HexaPro, is highly similar to the glycosylation of 2P and native S protein as presented on the virus [8,124,125]. Also, this study confirm earlier observations [110] that both forms of the recombinant protein have indistinguishable binding properties (K_D) to the receptor, ACE2, indicating the functional form of protein is intact. However, there are some differences in glycan composition which could suggest differences in the conformational properties between the HexaPro and 2P (**Table 3.2**). This motivated us to extend the analysis of protein conformational flexibility by performing MD simulations, and to characterize its functional behaviour by performing serological testing. This builds upon previous observations of cryo-EM and a small-scale serological evaluation [110].

3.4 Differences in oligomannose content between 2P and HexaPro

The glycan at a structural site, N234, which has been shown to stabilize the RBD up conformation in the trimeric structure of the S protein, is principally oligomannose-type, and is conserved across both the constructs of recombinant S protein (**Figures 3.3.A** and **3.3.B**). Also, the oligomannose-type glycans of N234 likely arises from steric clashes with the protein component, which in turn limits the ability of glycan processing enzymes to act, as it is sandwiched between N-terminal and receptor-binding domains (**Figure 3.5.B**). Overall, the oligomannose content is highly similar across both 2P and HexaPro (**Table 3.2**). However, the $Man_9GlcNAc_2$ content is higher in the case of HexaPro compared to 2P, indicating the reduced accessibility to glycan processing enzymes [8] (**Table 3.2**). This may be attributed to the presence of the two "up" conformation observed in HexaPro's RBD, potentially influencing the accessibility to glycan processing enzymes [110]. Also, at the site-specific level there are several other sites which showed changes in oligomannose content across 2P and HexaPro (**Figures 3.5.A, B and Table 3.2**). Differences in glycan processing states were observed at glycan sites N61, N122, N165, N603, N616 and N801. The major differences were observed at N165 and N122, which are in close proximity to the RBD (a roughly 50 percentage point difference in both cases). The N165 site possesses a higher abundance of oligomannose-type glycans on HexaPro shown in dark blue whereas the N122 site possesses a lower abundance of oligomannose-type glycans on HexaPro shown in red (**Figure 3.5.B**).

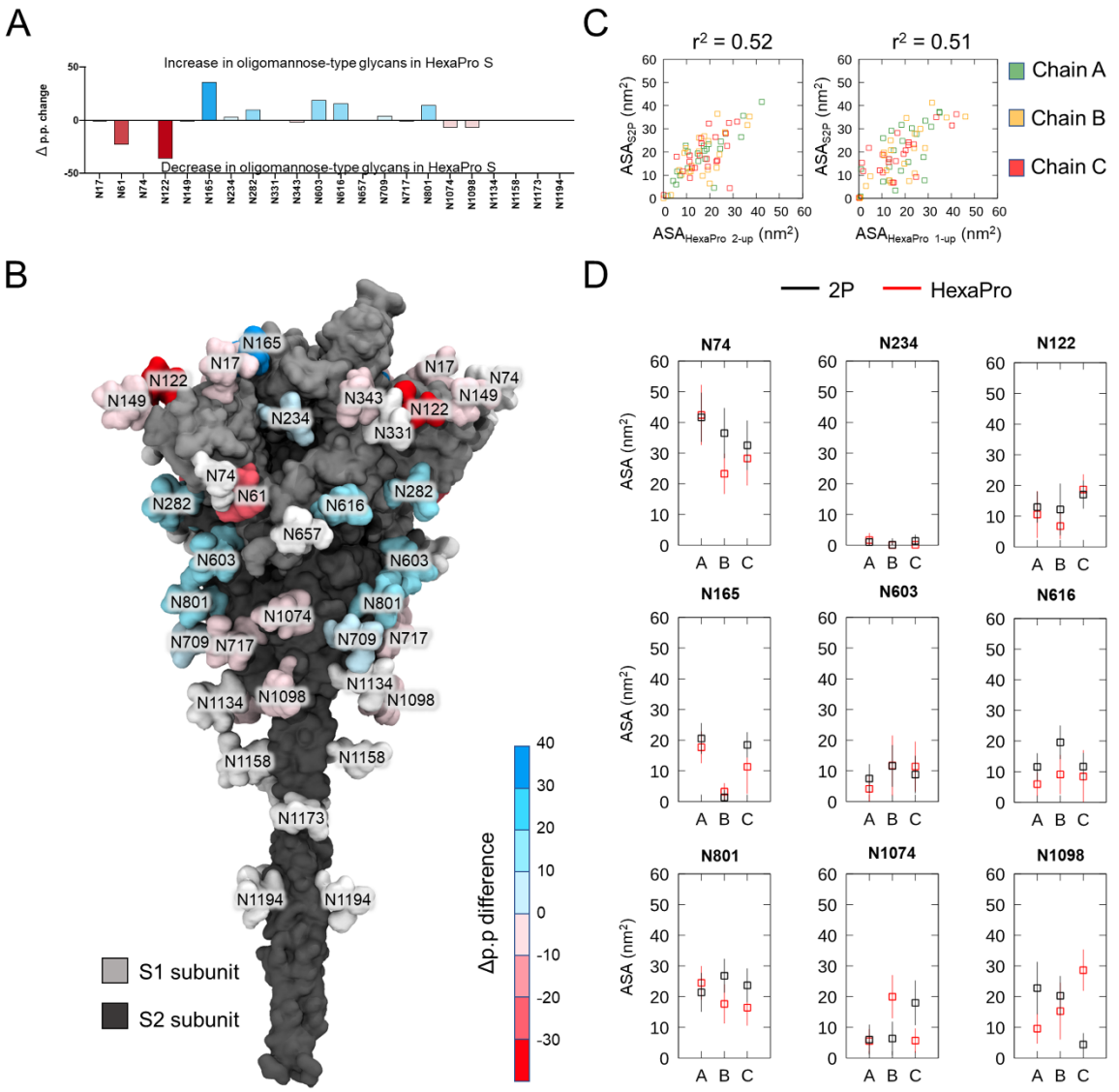


Figure 3.5. **Comparison of glycan composition across prefusion-stabilized SARS-CoV-2 S protein** A) The percentage point change in oligomannose-type glycan content between SARS-CoV-2 S protein, HexaPro and 2P. The percentage point (p.p.) difference on the y-axis is the arithmetic difference between the percentiles of oligomannose-type glycans between the two populations (here defined as: p.p. = % HexaPro – % 2P). Positive values (blue) indicate a higher abundance of oligomannose-type glycans in HexaPro relative to 2P. Negative values (red) indicate a lower abundance of oligomannose-type glycans in HexaPro relative to 2P. B) A full length model of SARS-CoV-2 S protein with N-glycans colored based on the percentage point change values. The scale represents the differences in oligomannose-type glycans observed in HexaPro when compared to 2P protein. Colors correspond to p.p. values in Panel A. The model was constructed using prefusion structure of trimeric SARS-CoV-2 S glycoprotein, as detailed in Materials and Methods. C) Correlation of ASA values between HexaPro and 2P S protein. The average ASA values of all glycans from three replica simulations of two-RBD-up HexaPro (left side) and one-RBD-up (right side) structures plotted against the average ASA values from simulations of the respective 2P structures. D) The average ASA values for 2P (black) and HexaPro two-RBD-up (red) with error bars showing standard deviations along the trajectories and across three repeat simulations. The displayed sites are those with changes in the oligomannose content across both versions. N74 (high ASA values) and N234 (lowest ASA values) were used as a reference. The chain (A, B, or C) of the trimeric S protein is indicated along the x-axis.

To investigate the molecular basis of the observed changes in oligomannose-type glycan content across both the S proteins, a series of triplicate 200 ns MD simulations of the HexaPro two-RBD-up, HexaPro one-RBD-up and 2P one-RBD-up constructs were performed. The HexaPro two RBD-up model was generated by fitting the structure of HexaPro one-RBD-up to the electron density map of HexaPro two-RBD-up using molecular dynamics flexible fitting (MDFF) (details in Methods section). A potential N-linked glycosylation sites were glycosylated with $Man_9GlcNAc_2$ glycans, as these represent the primary substrate for glycan processing enzymes and hence may be used to predict glycan processing as previously described [125].

First, the correlation between accessible surface area (ASA) for all glycans between HexaPro and 2P was investigated. There is a moderate correlation ($r^2 = 0.5$) between both one-RBD-up and two-RBD-up HexaPro when compared to the 2P, suggesting there is some degree of difference in glycan accessibility between HexaPro (two-RBD-up & one-RBD-up) and 2P (one-RBD-up) (**Figure 3.5.C**). To understand whether this difference is due to the protein architecture, or due to the stochastic nature of the simulations, the ASA for each of the glycosylation site between HexaPro and 2P was compared. Furthermore, the arithmetic difference between the ASA values of HexaPro (two RBD-up) and 2P (one RBD-up) at each glycan site in all three chains was calculated (**Appendix A.1**). The error bars represent standard deviations throughout the trajectories which capture the variation caused by sampling. A positive value represents lower accessibility in HexaPro, which could correlate to reduced glycan processing and a higher abundance of oligomannose-type glycans. The difference at most of the sites was small and with substantial standard deviations, suggesting that most of the differences arise from the stochastic sampling of the glycans.

Then, glycan sites which displayed alterations in oligomannose content were additionally investigated in the site-specific glycan analysis. For example, on HexaPro, N165 displays an increase of almost 50 percentage point in oligomannose-type glycan compositions when compared with 2P (**Figure 3.5.A**). As the N165 site is in close proximity to the RBD region, its glycan processing state may be influenced by the orientation of the RBD. When comparing the model generated with two RBD domains in the up configuration, to one-RBD-up, the steric environment of the N165 and N122 sites were expected to change. However, the simulations comparing HexaPro two-RBD-up versus

one-RBD-up showed little changes in solvent accessibility (**Figure 3.5.D**). This is likely because in the HexaPro (two-RBD-up) simulation, the additional RBD in the initial up configuration tends to revert to a down state, and the final snapshot in all simulation replicas is similar to the 2P (one-RBD-up) simulation (**Appendix A.2**). It is noteworthy that in our initial model, derived from the experimental electron density map of the HexaPro two RBD-up conformation, the RBD in chain C resembles an intermediate state between the up and down states, rather than a fully open conformation. The changes in ASA values we observed are only significant at a few sites and minor with variations across the three chains, indicating that the differences are due to sampling.

Finally, protein-structural analyses were carried out to compare the dynamics of 2P and HexaPro. Similar profiles of root-mean square fluctuations (RMSF) were noted between the HexaPro (two-RBD-up and one-RBD-up) and 2P (one-RBD-up) simulations, except for the RBD in chain C, due to the up to down conformational changes described above (**Appendix A.3.A**). The principal motion measured during the simulations also reveals similar dynamics between these two S proteins (**Appendix A.3.B**). Hence, the simulations suggest that despite local glycan perturbation, the two S protein versions have very similar dynamics, and our computational analysis does not provide evidence to support a steric explanation for the differences in glycosylation.

3.5 Conservation of serological reactivity across recombinant SARS-CoV-2 S protein

To compare the serological reactivity of the recombinant 2P and HexaPro S protein, we tested the binding of different immunoglobulin isotypes in sera from subjects with or without a prior SARS-CoV-2 infection to these viral antigens. This extends the observations presented by Hsieh et. al [110] by using a larger and geographically distinct donor group and by examining a range of antibody isotypes. Sera from three groups of subjects from the United Kingdom were analyzed: hospitalized subjects (HS) which included individuals that were admitted to hospital and had RT-PCR confirmed SARS-CoV-2 infection; non-hospitalized convalescent (NHC) subjects, who were tested positive by clinically validated antibody test [107] and were not hospitalized and a negative control group, from whom sera was taken before 2019 (Pre-19). As expected, strong IgG, IgA and IgM responses were

detected to both S glycoproteins in all hospitalized subjects with severe disease (**Figure 3.6.A and 3.7 A**).

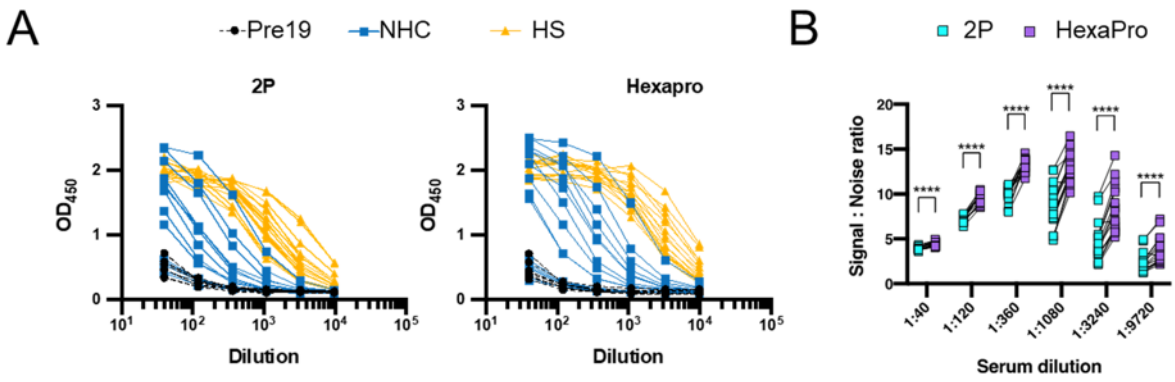


Figure 3.6. Antibody responses detected in an ELISA using 2P or HexaPro Spike as the target antigen. A) Absorbance values of sera serially diluted 3-fold from Pre19 (black circles, dashed lines), NHC (blue squares) or HS (yellow triangles) donors. B) Signal: Noise ratio calculated using individual PCR+ HS sera OD values as signal and the mean Pre19 sera as noise at multiple serum dilutions. Each point represents an individual signal: noise ratio from one serum, with a line connecting the same serum sample tested against 2P (cyan) or HexaPro (violet). This work has been carried out in collaboration with Dr. Alex G. Ritcher, and Dr. Adam F. Cunningham (University of Birmingham).

In contrast to the strong responses observed in severe cases, IgG, IgA and IgM responses were observed in the NHC subjects, and in some instances these responses were not above those of control sera (**Figure 3.7.A**). There was minimal binding of IgG to S glycoprotein by Pre19 sera. Both 2P and HexaPro showed comparable serological reactivity, with a slightly increased level of binding of patient, but not control, sera to HexaPro in the NHC sera (**Figure 3.7.A**). Overall, the signal: noise ratio was superior for HexaPro compared to 2P, particularly as sera were diluted (**Figure 3.6.B**), but overall, a key conclusion is that the HexaPro was not inferior to the 2P glycoprotein.

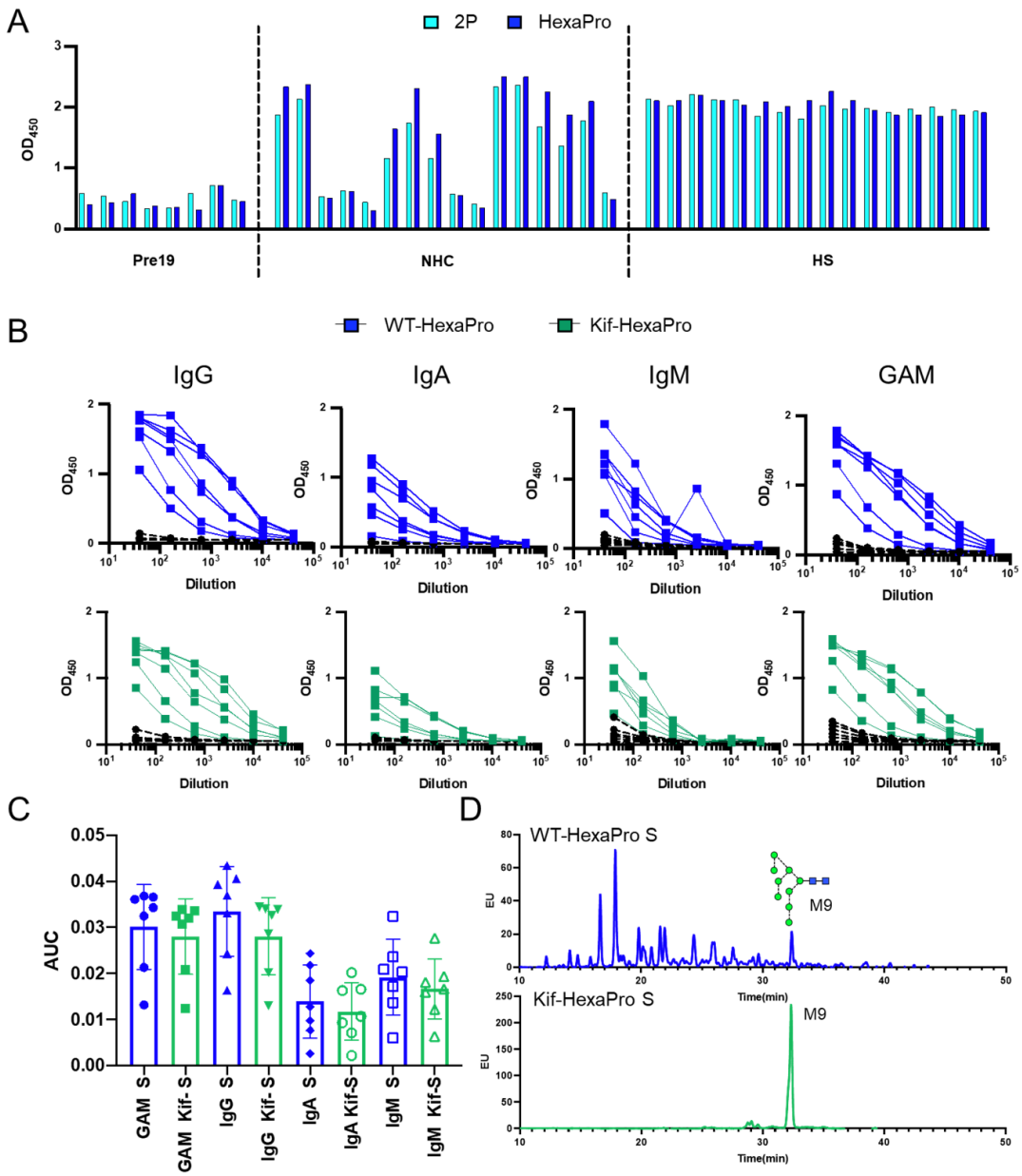


Figure 3.7. Antibody binding to spike glycoproteins. Individual serological responses from pre-2019 donors (Pre19, *n* = 8), non-hospitalized convalescent donors (NHC, *n* = 16) or PCR+ hospitalized subjects (HS, *n* = 16) as determined by ELISA using HRP-labelled combined anti-IgG, IgA and IgM. A) Absorbance values of sera serially diluted from a starting dilution of 1:40 against 0.1 µg 2P (cyan bars) or HexaPro (blue bars). B) Absorbance values of sera serially diluted from a starting dilution of 1:40 from Pre-19 (black circle, dashed lines), NHC of HexaPro protein (blue squares) and Kifunensine-treated HexaPro (green squares) as determined by ELISA using HRP-labelled combined anti-IgG, IgA, IgM, and GAM. C) Area Under the Curve (AUC) of responses shown in figure B. The blue bars representing the AUC of HexaPro with IgG, IgA, IgM and GAM. The green bars representing the AUC of kif-treated HexaPro with different immunoglobulins. The mean ± standard deviation from the mean (SD) is plotted. D) HILIC-UPLC profile of N-linked glycans from WT (wildtype) HexaPro and Kifunensine-treated HexaPro produced in HEK 293F cells and purified by Ni^{2+} column followed by SEC. The blue peaks representing glycan spectra of WT-HexaPro. The green peaks representing glycan spectra of kifunensine-treated HexaPro showing only $Man_9GlcNAc_2$ (M9) glycans. This work has been carried out in collaboration with Dr. Alex G. Ritcher, and Dr. Adam F. Cunningham, (University of Birmingham).

Although the small difference in glycosylation between 2P and HexaPro did not result in diminished antibody binding to HexaPro, it is conceivable that wider batch-to-batch variations could impact consistency of serological reactivity. Sera were subsequently tested for antibody binding to the glycoprotein featuring significantly engineered glycosylation. HexaPro was expressed in the presence of kifunensine (Kif) which results in oligomannose-type glycans, and the altered processing state was confirmed using hydrophilic interaction ultra-performance liquid chromatography (HILIC-UPLC) analysis of fluorescently labelled pool of N-linked glycans (**Figure 3.7.D**). The WT HexaPro chromatogram shows the presence of diverse glycans in contrast to Kif-treated HexaPro, where $Man_9GlcNAc_2$ (M9) glycans were predominant (**Figure 3.7.D**). Furthermore, comparison was made between IgG, IgA, IgM, and combined GAM antibody binding to HexaPro with no glycan engineering (referred to as 'wildtype') and Kif-treated HexaPro. The serological response detected using SARS-CoV-2 positive sera were similar whether kif-treated HexaPro or WT HexaPro was used in the assay (**Figure 3.7.B**). Negligible binding was observed with pre-19 sera (shown in black) with both kif-treated and WT HexaPro. Area Under the Curve (AUC) calculations confirmed that both WT S and Kif-treated S protein were bound similarly by IgG, IgA, IgM and IgGAM (**Figure 3.7.C**). This suggests that the immune response elicited following SARS-CoV-2 infection, with respect to immunoglobulin binding, is not dictated by the glycan processing state of the S protein, as converting the glycans at every site from their native-like compositions, does not impact the detection sensitivity. In contrast, in some glycosylated viruses such as HIV-1, kif-treatment moderately impairs Env processing and function [408].

3.6 Discussion

This study investigated a range of biophysical, glycan composition and serological binding properties of HexaPro, an expression-enhanced version of SARS-CoV-2 S protein. HexaPro contains six proline mutations which lead to its high expression of protein and stabilization compared to S2P variant, and promising immunogenic properties [110,162,401,402,409]. The comparable affinity of HexaPro to the ACE2 receptor with the earlier version, 2P which has been used in several vaccine studies [100,101,410,411], and extensively as a serological reagent [106,107,399]. Furthermore, the

impact of these additional proline mutations on glycan composition was determined using LC-ESI MS and compared it with the 2P. It was interesting to note that overall, the glycosylation was highly similar between both versions, except at a few sites. To further explore phenomena that could be directing these changes we performed MD simulations to decipher the conformational properties of both versions.

The structural site, N234 showed fully oligomannose-type glycans in both the versions suggesting that the central integrity of the protein architecture is unperturbed [121,124,125]. Overall, the oligomannose content was similar in both the S proteins, however, the $Man_9GlcNAc_2$ content was higher in HexaPro, suggesting less accessibility to glycan processing enzymes. This motivated us to look at the accessible surface area (ASA) of oligomannose-type glycans in 2P and HexaPro using MD simulations. Significant differences were not observed in ASA values at N-glycan sites that exhibited alterations in oligomannose content as determined through LC-MS. Additionally, remarkably similar protein dynamics were noted in both versions. This suggests that both of the proteins have very similar conformational properties. However, the simulations sample a short timescale while glycans are processed over a much longer timescale and these simulations do not provide evidence for the steric changes predicted from the differential glycosylation observed in the LC-MS data. It is conceivable that the changes in oligomannose-type glycan content could instead be due to high expression of HexaPro resulting in changes in productive enzyme:substrate recognition events. Overall, however, these results suggest that there are limited structural and dynamic differences between 2P and HexaPro.

S proteins are being deployed in serological testing and have proven to be effective in confirming prior infection of SARS-CoV-2 in infected patients [107,412,413]. Due to its high expression, HexaPro could widen the availability of S protein for serological testing. This study was aimed to investigate the binding of HexaPro with sera IgG, IgA, and IgM to better understand its interaction with antibodies induced in COVID patients, in order to maximize the potential of HexaPro in these applications. Highly similar antibody binding is also observed with both 2P and HexaPro, indicating that the inclusion of mutations in HexaPro does not appear to impact immune recognition. The results also reveal highly similar reactivity with glycoengineered HexaPro possessing all

oligomannose-type glycans at all PNGS, indicating that sera binding is not readily impacted by the fine processing of the glycans of the S protein. Moreover, since the level of antibody binding was not significantly reduced after glycan engineering of HexaPro, the data may be interpreted as indicating that after natural infection most antibodies do not target epitopes that can be influenced by variations in glycan processing. If so, since antibodies from infected individuals can neutralise infection in vitro, it suggests that protection from infection is not associated with the targeting the terminal region of glycans. It is noted, however, that the primary glycan at N234 remains unchanged by kifunensine. Overall, these observations suggest that variations in the S protein glycosylation of SARS-CoV-2 will not impact the serological assessments currently being performed across the globe.

Chapter 4 Site-specific glycosylation of Virion-Derived SARS-CoV-2 Spike is mimicked by a soluble trimeric immunogen.

Contributions.

Chapter 4 compares the glycosylation of the virion SARS-CoV-2 S protein, recombinant S protein and monomeric RBD proteins. LC-MS data for virion SARS-CoV-2 S protein was prepared by Prof. Sai Li and Dr. Yutong Song of Tsinghua University, which I have reanalysed as part of this study to allow the direct comparison with recombinant protein. The monomeric RBD plasmids were designed and expressed by the laboratory of Prof. Dennis Burton at The Scripps Research Institute. The Close-S protein was designed and expressed by the laboratory of Dr. Hans Langedijk, Janssen Vaccines and Prevention. RBD nanoparticles are designed and expressed by the laboratory of Dr. Daniel Kulp, the Wistar Institute. The glycan analysis of RBD nanoparticles is published and full list of author contribution can be found in Kylie et. al., 2022 [414]. Some part of this work is published and is a first co-author share between me, Dr. Joel D Allen and Dr. Firdauss Samsudin [125]. The manuscript is primarily written by Joel and me and is reviewed by all the collaborators, of which introduction and viral glycan analysis is included in this chapter.

4.1 Introduction

The Coronavirus Disease 2019 (COVID-19) pandemic has prompted the development of an unprecedented array of vaccine candidates against the causative pathogen, SARS-CoV-2. All approaches aim to deliver molecular features of the virus in order to induce immunity. The viral spike glycoprotein, also referred to as S protein, has emerged as the principal focus of vaccine design efforts as antibodies against this target can offer robust immunity [98,171,299,300,415,416]. Encouragingly, neutralization can readily occur despite the extensive array of N-linked glycans distributed across the viral spike consistent with numerous vulnerabilities in this so-called glycan shield [8]. Despite these observations, glycosylation has emerged as an important parameter in vaccine development for SARS-CoV-1 and SARS-CoV-2 [417,418]. The glycosylation processing

state can influence immunogen trafficking in the lymphatic system [122], influence the presentation of both native and unwanted cryptic epitopes [419], and reveal to what extent immunogens recapitulate native viral architecture. The presence of oligomannose-type glycans on surface of viral protein can interact with the mannose-binding lectin which will either traffick the viral protein to the follicular dendritic cells or activate the complement pathway which can influence the immunogen trafficking and potentially enhance the accumulation of viral protein in the lymphatic system for germinal center activation [122].

Evidence is also emerging that glycosylation can somewhat influence the interaction between SARS-CoV-2 and its target receptor, ACE2 [50,140,141,420]. Each protomer of the trimeric SARS-CoV-2 S protein contains at 22 N-linked glycosylation sequons that direct the attachment of host glycans to specific Asn residues. This extensive glycosylation is important in lectin-mediated protein folding and direct stabilization of the protein fold [421]. In addition, certain glycans are incompletely matured during biogenesis and can lead to the presentation of immature glycans terminating with mannose residues that can act as ligands for innate immune recognition [11,210,232].

Despite the focus on the S protein in vaccine development efforts, there has been considerable divergence in the mechanisms of delivery. In one approach, nucleic acid encoding the spike is delivered through mRNA or with a viral-vector [105,299,300,304,322,416,422]. The resulting S protein is assembled and glycosylated by host tissue. In a contrasting approach, the S protein can be recombinantly manufactured either as recombinant protein using mammalian or insect cell lines or using inactivated virus-based approaches which allows detailed characterization of the immunogen prior to delivery [102,109,110,392,423]. Immunogen glycosylation can be influenced by factors specific to the manufacturing conditions such as cell type or cell culture conditions [370,372], however, construct design and protein architecture can also have substantial impact. For example, under-processed oligomannose-sites can occur at sites sterically hidden from the host mannosidase by the tertiary or quaternary architecture including obfuscation by neighboring protein and glycan structure [121,232]. Immunogens displaying native-like architecture recapitulate these sites of oligomannose glycosylation. Conversely, immunogen design can adversely impact the presentation of native-like glycosylation. Importantly, despite the differences in biosynthesis of S protein in

virions and from mammalian expression systems they seem to generate broadly similar glycosylation [124]. However, the success of a broad range of different vaccine platforms exhibiting different S protein glycosylation, indicate that native-like glycosylation is not a prerequisite for a successful vaccine. Despite this observation, understanding S protein glycosylation will help benchmark material employed in different serological and vaccine studies and help define the impact of this extensive feature of the protein surface.

The flexible and heterogeneous nature of N-linked glycosylation necessitates auxiliary methodologies in addition to cryo-electron microscopy or X-ray crystallography to characterize this key part of the S protein structure. Site-specific glycan analysis employing liquid chromatography-mass spectrometry is a widely used approach to obtain this information [1,16,335,374,424]. As research into the structure and function of the SARS-CoV-2 S protein has progressed, more details of the glycan shield of S protein have become apparent. Analyses on recombinant trimeric S protein revealed divergent N-linked glycosylation from host glycoproteins with the presence of under-processed oligomannose-type glycans at several sites [8,50,117]. Comparative analyses with monomeric and trimeric S proteins have revealed site-specific differences in glycosylation with regards to both oligomannose-type glycans and the presentation of sialic acid [117]. Analysis of S protein from insect cells demonstrated that oligomannose-type glycans were conserved on trimeric S protein, notably at N234 [425]. In addition, molecular dynamics (MD) studies have proposed that the N234 site plays a role in stabilizing the receptor binding domain (RBD) in an exposed "up" conformation [121]. The presence of larger under-processed oligomannose-type glycans, such as $Man_9GlcNAc_2$, on both mammalian and insect-derived S protein provides an indication that the structure of the S protein is driving the presentation of these glycans. Subsequent studies have investigated the presentation of N-linked glycans on S protein produced for vaccination, notably the Novavax full length S protein and S protein isolated following administration of the ChAdOx-nCoV-19 vector [425,426]. The observed glycan signatures were broadly in agreement with previous analyses. However, these studies involved the truncation of glycan structures using glycosidase treatment – which is useful for categorizing glycans into high mannose or complex-type glycans and determining PNGS occupancy on low amounts of material – but does not allow for the identification

of changes in terminal glycan processing such as sialylation. The glycan processing of the two N-linked glycan sites located on the RBD has also been investigated for monomeric RBD. These sites present high levels of complex-type glycans [427,428] and as the majority of antibodies raised against SARS-CoV-2 S protein target the RBD it is important to fully characterize the structure of the RBD, including the presentation of glycans.

N-linked glycans are highly dynamic and can substantially vary in chemical composition within a single batch of protein. Whilst glycosylation is heterogeneous, it is important for therapeutic and vaccine design to understand whether similar glycoforms arise across different protein expression platforms from different sources to ensure the antigenic surface of the S protein remains consistent when used as an immunogen or in serological assays [399,429–431].This is particularly important as glycan processing can be impacted by adverse protein conformations [432].

This chapter describes the extent to which recombinant proteins reproduce the glycan composition of viral derived SARS-CoV-2 S protein produced from cultured Vero cells which was obtained from a previous study [124]. An identical analytical approach was used to determine the glycan composition at each site, illustrating conservation across recombinant protein and virion derived material. Furthermore, the glycosylation of monomeric RBD recombinant protein was investigated, as it has previously been explored as a subunit vaccine and a candidate for serological testing [386,389]. The site-specific glycan analysis of the two sites located in the RBD of SARS-CoV-2, N331 and N343, when comparing monomer and trimer reveals a broad consensus in glycan processing, with some modest change in the processing of complex-type glycans. A similar comparative analysis of MERS-CoV RBD was performed to investigate whether the RBD glycans of MERS-CoV were under the same constraints as for SARS-CoV-2. The observed differences contrasted the glycan processing of SARS-CoV-2 RBD monomer when compared with S protein, with trimeric MERS-CoV RBD glycan sites presenting restricted glycan processing, likely due to conformational masking of these sites, either by proximal glycans or nearby protein clashes, on trimeric MERS-CoV S protein. This contrasts with glycan processing of SARS-CoV-2 RBD monomer when compared with S protein, with trimeric MERS-CoV RBD glycan sites presenting restricted glycan processing, likely due to conformational masking of these sites, either by proximal

glycans or nearby protein clashes, on trimeric MERS-CoV S protein [7,397]. Taken together, the results reveal the conserved structural N-glycan sites in S protein when compared with native viral spike. Understanding S protein glycosylation will aid in the analysis of the vaccine and serological work of the global COVID-19 response.

4.2 Glycan analysis of virion SARS-CoV-2 S protein

This study reveals the glycan compositions of N-glycan sites of the viral SARS-CoV-2 S protein, which have been reanalyzed from the raw files published previously [124] (**Figure 4.1**). The reanalysis was performed using similar analytical settings used for the analysis of recombinant HexaPro S protein described above. There are slight differences at site-specific level in glycan compositions analyzed across both the analytical settings, but there is overall similar glycosylation content and occupancy. This is potentially due to the parameters used for calculating the total glycan compositions. This study utilized the average of extracted ion chromatogram (XIC) to calculate the percentage of glycan compositions present at each N-linked glycan site. Yao et al., used the intensities of each glycan type in identical site were combined and analyzed for proportion [124]. The data has been reanalyzed to corroborate the similar analysis parameters to determine glycan compositions across viral and recombinant S protein. This will also help in investigating native-like signatures on recombinant S protein.

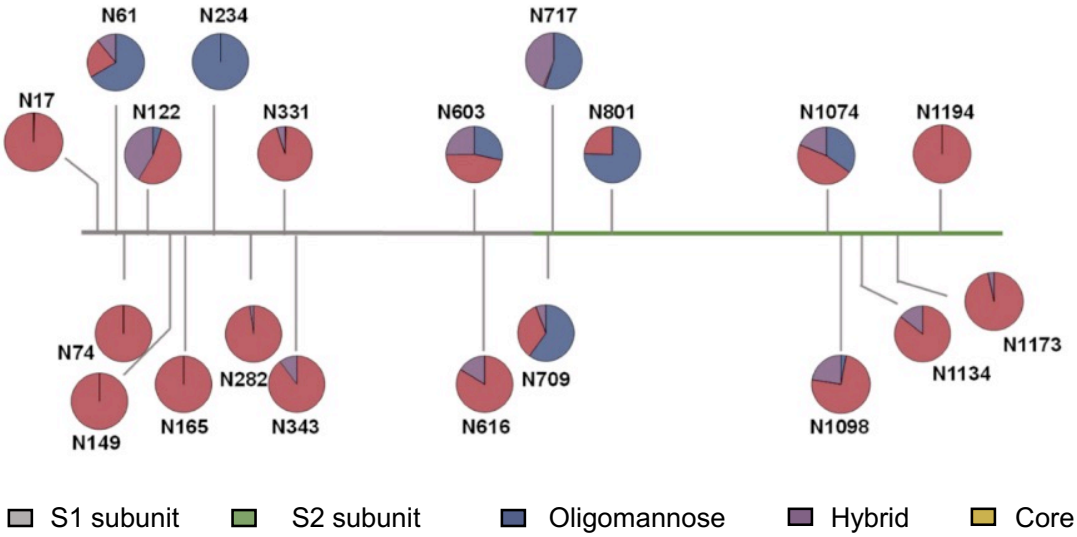


Figure 4.1. Site-specific glycan compositions of viral S protein. The percentage of glycan composition analyzed using average of extracted ion chromatogram (XIC) obtained at identical glycan site. The data is obtained in this study and is reanalyzed from the raw files obtained from previous study [124].

To investigate the abundance of glycans on S1 and S2 subunit of viral S protein, the glycan compositions were categorized into three categories, 1) oligomannose-type, 2) complex-type and 3) hybrid-type glycans. The glycan analysis of viral S protein showed high levels of glycan occupancy at all sites analyzed (**Figure 4.1**). This is an important feature to monitor in vaccine design, as glycan underoccupancy can potentially elicit non-neutralizing antibodies. Another parameter to observe is the conservation of presence of under processed oligomannose-type glycans, as the presence of these glycans are usually linked to quaternary structure of the protein, changes in the abundance of these glycans reflect the changes in the fine structure of the glycoprotein. These glycans are usually common on viral glycoproteins due to either glycan- or protein-mediated steric clashes with ER and Golgi resident enzymes, terminating the glycan processing pathway. For example, N234 site presenting high abundance of less trimmed glycan structures, $Man_9GlcNAc_2$, due to steric hinderance and location of this site on trimeric spike protein (**Figure 4.2, Table 4.1**). Additionally, N234 glycan site is involved in structural stability of the trimeric spike along with N165 site [8,121], therefore it is important to have these sites depicting conserved glycan compositions for native-like immunogen design. The N165 presents high proportion of complex-type glycans due to higher accessibility area (ASA) available for glycan processing enzymes, supported by MD-simulation data [125]. The N165

and N234 glycans, has also been known to modulate RBD's conformational dynamics by maintaining the up configuration necessary for ACE2 recognition [121,137,433].

Overall, the viral spike protein presents high-abundance of complex-type glycans and hybrid-type glycans (**Figure 4.2**). The N-glycan sites presents on RBD region, N331 and N343, and S2 subunit sites, N1098, N1134, N1173 and N1194 are primarily populated with highly processed glycan structures.

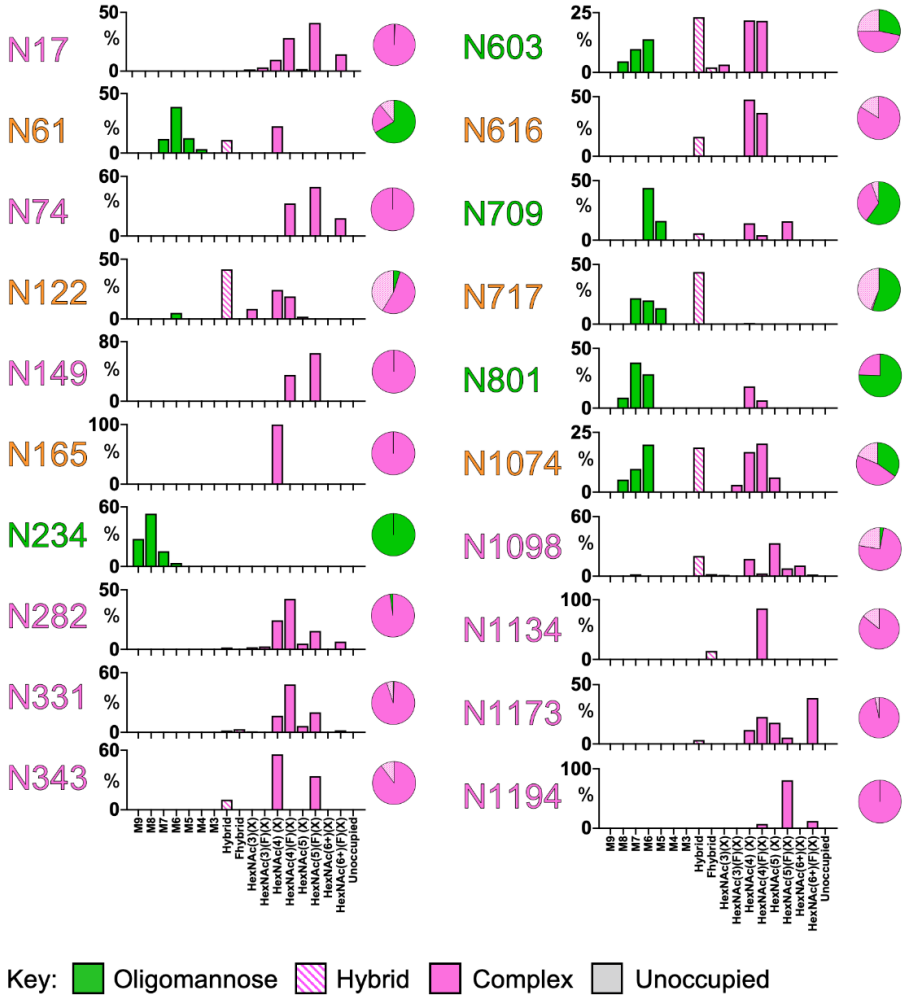


Figure 4.2. Site-specific N-linked glycan compositions of the viral S protein. The graphs summarize the quantitative mass-spectrometric analysis of the glycan population present at individual N-linked glycosylation sites simplified into categories of glycans. The oligomannose-type glycan series (M9 to M5; $Man_9GlcNAc_2$ to $Man_5GlcNAc_2$) is colored green, afucosylated and fucosylated hybrid-type glycans (hybrid and F hybrid) are dashed pink, and complex glycans are grouped according to the number of antennae and presence of core fucosylation (A1 to FA4) and are colored pink. Unoccupancy of an N-linked glycan site is represented in gray. The pie charts summarize the quantification of these glycans. Glycan sites are colored according to oligomannose-type glycan content, with the glycan sites labeled in green (80 to 100%), orange (30 to 79%), and pink (0 to 29%).The raw files have been obtained from the Sai Li's lab, single data files – no repeat.

4.3 Glycan composition across recombinant and Virion S protein

The spike (S) protein has emerged as the principal focus of vaccine design efforts [103,299,300] as it is involved in initiating viral infection by mediating host-cell entry [81].The abundance of N-linked glycans across the S protein is a potential source of heterogeneity and a display of native-like antigens is a central tenet in the elicitation of protective immunity. This study compares the glycan composition observed on S protein expressed recombinantly containing stabilized two proline mutations[8,407] with viral S protein expressed in Vero cells [124,125]. To compare the variability in glycosylation at site-specific level, glycan composition is broadly categorized into three categories: Oligomannose-type, complex-type and unoccupied. Sites with an abundance of underprocessed glycans are highly similar across both recombinant and viral S protein. These sites include, N61, N122, N234, N603, N709, N717, N801, and N1074 (**Figure 4.3**). This analysis revealed a broad consensus of glycan processing regarding high mannose glycans, with localized variations occurring. One example of remarkable homogeneity in oligomannose-type glycans is the N234 glycan site, which presents high abundance of under processed glycans across both recombinant and viral S proteins (**Figure 4.3.A and B**). This is important to note as conservation of glycan under processing reveals an information pertaining to the quaternary structure of the glycoprotein.

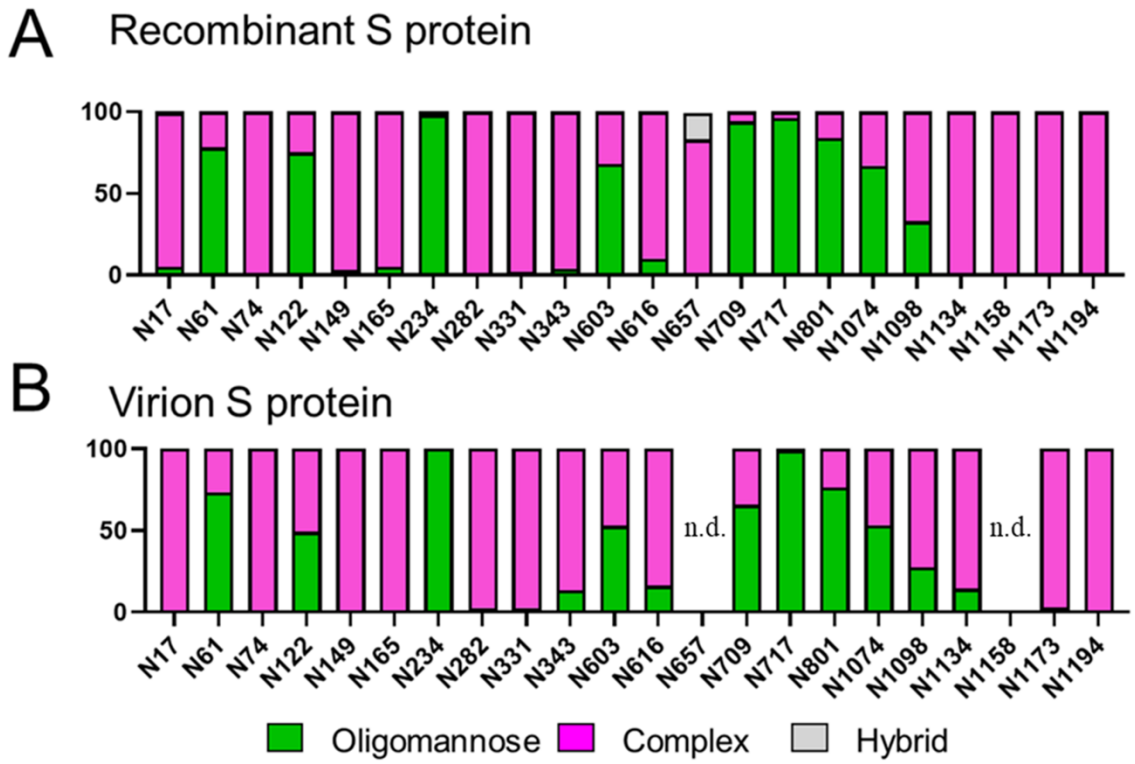


Figure 4.3. Site specific glycan compositions of recombinant and virally derived S protein. A) Glycan analysis of recombinant S protein with two proline substitutions for stabilization. B) Virally derived site-specific analysis was performed using data acquired by Yao et al. [124]. Data for sites N657 and N1158 could not be obtained and are not presented. The bar charts represent the relative proportions of glycoforms present at each site. The proportions of oligomannose- and hybrid-type glycans are colored green. Processed complex-type glycans are colored pink, and the proportions of unoccupied sites are colored gray. n.d., not determined.

Overall, the glycan analysis reveals highly similar oligomannose-type glycans and high glycan occupancy across both recombinant and viral S protein. Except N657 with slight underoccupancy on recombinant S protein. However, there was variability observed in complex- and hybrid-type glycan compositions. For example, N165 and N1074 presented diverse glycan processing states (**Table 4.1**). The elaboration of complex-type glycans with different monosaccharides can influence the function of the glycoprotein to which they are attached; for example, sialylation can extend the half-life of a glycoprotein in the body by masking the galactose that can otherwise be recognized by the hepatic asialoglycoprotein receptor (ASGPR) [434,435]. The starkest difference observed between the recombinant and viral derived S protein is the lack of sialic acid residues across not only N165 and N1074 but also all PNGSs of virally derived material (**Table 4.1**). The fucosylation of virally derived S protein is also lower across majority of sites compared to recombinant S protein.

Table 4.1. Site-specific glycan compositions of virally derived and recombinant S protein. The site-specific compositions across all PNGS of S protein, virally derived (A) and recombinant S2P protein (B). The glycan compositions are categorized as outlined in the materials and methods for the S protein.

A

Viral derived	N17	N61	N74	N122	N149	N165	N234	N282	N331	N343	N603	N616	N657	N709	N717	N801	N1074	N1098	N1134	N1158	N1173	N1194
M9Glc	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M9	0	0	0	0	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M8	0	0	0	0	0	0	53	0	0	0	4	0	0	0	0	0	5	1	0	0	0	0
M7	0	11	0	0	0	0	15	0	1	0	9	0	0	0	22	12	10	1	0	0	0	0
M6	0	37	0	4	0	0	4	0	0	0	12	0	0	44	20	31	20	1	0	0	0	0
M5	0	12	0	0	0	0	0	0	0	0	6	0	0	16	13	0	0	1	0	0	0	0
M4	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hybrid	0	10	0	37	0	0	0	2	2	0	20	16	0	6	44	33	19	22	0	0	3	0
Fhybrid	0	0	0	8	0	0	0	0	0	13	2	0	0	0	0	0	0	2	14	0	0	0
HexNAc(3)(x)	0	0	0	7	0	0	0	2	1	0	3	0	0	0	0	0	0	3	0	0	0	0
HexNAc(3)(F)(x)	0	0	1	0	0	0	0	2	2	0	6	0	0	0	0	0	3	1	0	0	0	0
HexNAc(4)(x)	0	21	0	25	0	100	0	24	16	0	19	47	0	14	1	24	17	16	0	0	12	0
HexNAc(4)(F)(x)	34	0	35	17	57	0	0	42	47	5	19	36	0	4	0	0	20	2	86	0	22	0
HexNAc(5)(x)	8	0	0	2	0	0	0	5	7	11	0	0	0	0	0	0	6	38	0	0	18	0
HexNAc(5)(F)(x)	57	5	48	0	43	0	0	16	20	62	0	0	0	16	0	0	1	6	0	0	8	100
HexNAc(6+)(x)	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	7	0	0	0	0
HexNAc(6+)(F)(x)	0	0	16	0	0	0	0	7	2	0	0	0	0	0	0	0	0	0	0	0	38	0
Unoccupied	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Core	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Oligomannose	0	63	0	4	0	0	100	0	1	0	31	0	0	60	55	44	35	4	0	0	0	0
Hybrid	0	10	0	45	0	0	0	3	2	13	22	16	0	6	44	33	19	24	14	0	3	0
Complex	100	27	100	51	100	100	0	97	96	87	47	84	0	34	1	24	47	73	86	0	97	100
Unoccupied	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Fucose	92	5	100	25	100	0	0	67	71	80	27	36	0	20	0	0	24	10	100	0	68	100
NeuAc	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	1	0	0	0	0
Sulfation	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

B

Recombinant S2P	N17	N61	N74	N122	N149	N165	N234	N282	N331	N343	N603	N616	N657	N709	N717	N801	N1074	N1098	N1134	N1158	N1173	N1194
M9Glc	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M9	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M8	0	0	0	0	0	0	44	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0
M7	0	0	0	0	0	0	14	0	0	0	0	0	0	7	17	10	0	1	0	0	0	0
M6	0	1	0	0	0	0	4	0	0	0	2	1	0	26	25	17	2	2	0	0	0	0
M5	3	68	1	56	2	0	4	1	1	2	62	5	0	59	32	49	55	6	0	0	1	0
M4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hybrid	0	8	0	17	0	3	1	0	0	0	3	3	0	0	21	7	8	19	0	0	0	0
Fhybrid	2	1	0	2	0	1	0	0	0	2	0	1	0	0	1	0	1	5	0	0	0	0
HexNAc(3)(x)	0	5	0	3	0	0	0	4	0	0	4	10	0	0	2	3	5	14	0	1	0	0
HexNAc(3)(F)(x)	6	1	2	5	4	0	1	0	4	12	4	2	3	0	1	1	5	7	15	0	0	0
HexNAc(4)(x)	0	6	0	3	0	14	0	11	0	0	0	18	0	0	0	1	2	19	0	48	0	0
HexNAc(4)(F)(x)	32	4	18	10	36	29	1	6	53	46	23	37	39	5	1	8	12	9	44	0	4	23
HexNAc(5)(x)	0	6	0	3	0	20	0	30	0	0	0	6	0	0	0	0	0	5	0	44	0	0
HexNAc(5)(F)(x)	53	1	58	1	51	32	0	46	38	37	1	13	37	1	0	2	8	7	23	0	25	3
HexNAc(6+)(x)	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	3	0	6	0	0
HexNAc(6+)(F)(x)	3	0	22	0	5	0	0	0	3	0	0	2	4	0	0	0	1	4	18	0	70	73
Unoccupied	1	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0
core	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mannose	4	69	1	56	2	0	97	1	2	2	65	6	0	94	74	77	57	9	0	0	1	0
Hybrid	2	9	0	19	0	5	1	0	0	3	3	4	0	0	22	7	9	24	0	0	0	0
Complex	94	22	99	25	97	95	2	99	98	96	32	90	83	6	4	16	33	67	100	100	99	100
Unoccupied	1	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0
Fucose	95	6		18		63	2	52	98	98	29	56	83	5	3	11	27	31	100	0	99	100
NeuAc	20	4		5		18	0	14	22	4	0	1	2	0	2	5	6	30	7	39	4	69
Sulfation	0	0	17	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

Several factors could be influencing these changes. The first is that the membrane tether afforded to the viral Env brings the glycans into the proximity of the glycan processing enzymes, and the second is that the producer cell had greater expression levels of the glycosyltransferase enzymes involved in glycan processing. The SARS-CoV-2 S protein in a viral context is likely under similar

constraints; however, the glycan processing is distinct. One factor that could be important is the early budding of the SARS-CoV-2 virion into the ER/Golgi intermediate complex (ERGIC) that may distance the spike protein from glycosyltransferases present in the trans-Golgi compared to Env, which remains attached to the membrane, proximal to glycosyltransferases [436]. The choice of the producer cell and culture condition of the virus, in this case Vero cells, which are derived from the kidney cells of Cercopithecus aethiops (African green monkey), may also influence the glycan processing as the expression levels of glycosyltransferase enzymes may account for the weakened attachment of sialic acid and fucose observed on viral S protein compared to recombinant S protein. Previous studies have highlighted the low abundance of sialic acid moieties on Vero cells; [437] however, these changes could also occur due to the culture conditions used for viral production. As these changes in glycosylation likely are not impacting the immunogenicity of the viral spike mimetics, as evidenced by the high efficacy of several vaccines, these observations remain important when considering how the virus may be interacting with the immune system via lectin interactions and may be informative when considering antigenic tests and purifications using glycan binding reagents. Despite the observed differences, we note that the site-specific glycosylation of the virally derived material outlined here is consistent with previously reported glycan analysis from virally derived S protein produced in Calu-3 lung epithelial cells [126].

4.4 Site-specific glycan analysis of Closed-S protein

There have been many advancements in the design of S protein in efforts towards stabilizing the S protein in its prefusion conformation. As it has been known that this conformation contains the epitopes for neutralizing antibodies and thus holds more potential as a vaccine immunogen [98,109,110]. Comparing the structure of SARS-CoV-2 S-2P [109,392] with that of the native virus [438,439] reveals that the former adopts a more open conformation, with one or more of the RBDs in the ‘up’ state. While antibodies have been reported to bind to both the ‘up’ and ‘down’ states of the RBD, those that target the conserved epitopes on the 'down' state are known to exhibit the highest neutralizing potency [98,396,440]. This section delves into the glycosylation of the stabilized S protein in the ‘Closed’ state, where all the RBDs are in the ‘down’ state. To achieve stabilization of

the prefusion S-protein, four substitutions—D614N, A892P, A942P, and V987P were introduced to create S-Closed. [111]. Previous studies on 2P revealed that V987P substitution is present in the hinge loop at the C-terminus of HR1 improve the stability of the trimer in prefusion conformation. There are two additional proline mutations in Close S to reduce the RBD exposure and stabilize the S-2P prefusion trimers. The A942P and A892P substitutions stabilize the S2 HR1 region that undergoes an extensive conformation change during fusion. Additionally, the substitution D614N in the head region, improved the thermal stability and less RBD exposure. This construct does not contain a heterologous trimerization domain that is typically required in soluble S designs [109,392,397]. Closed-S protein exhibits high expression, thermal and freeze-thaw stability. To confirm the influence of mutations on structural assembly and glycosylation, site-specific analysis of Closed S trimer was determined using LC-MS.

To understand the glycosylation across all PNGS, the glycan data has been categorized as; oligomannose-type, hybrid-type, complex-type and unoccupied. The glycan analysis of Closed-S protein revealed high glycan occupancy at all PNGS except N657, N1134 and N1194. The N234 site located on the S protein of SARS-CoV-2 has been empirically demonstrated to primarily exhibit oligomannose-type glycans [8]. This conservation of oligomannose-type glycan presentation is also observed in relation to the Closed-S protein (**Figure 4.4**). Correspondingly, the N709 site, which is recognized for its characteristic presentation of predominantly oligomannose-type glycans, manifests a significant prevalence of more extensively processed complex-type glycan compositions. However, this could be potentially due to majority of the substitutions at the S2 domain of Closed-S protein. This S protein presents high abundance of complex-type glycans at majority of the N-glycan sites (**Figure 4.4**). This is consistent with previous versions of S protein [8,125,426] and viral S protein [124,125] as described in above sections of this chapter. Despite few differences at site-specific level, overall, the glycan composition across Closed-S protein is highly similar to viral S protein.

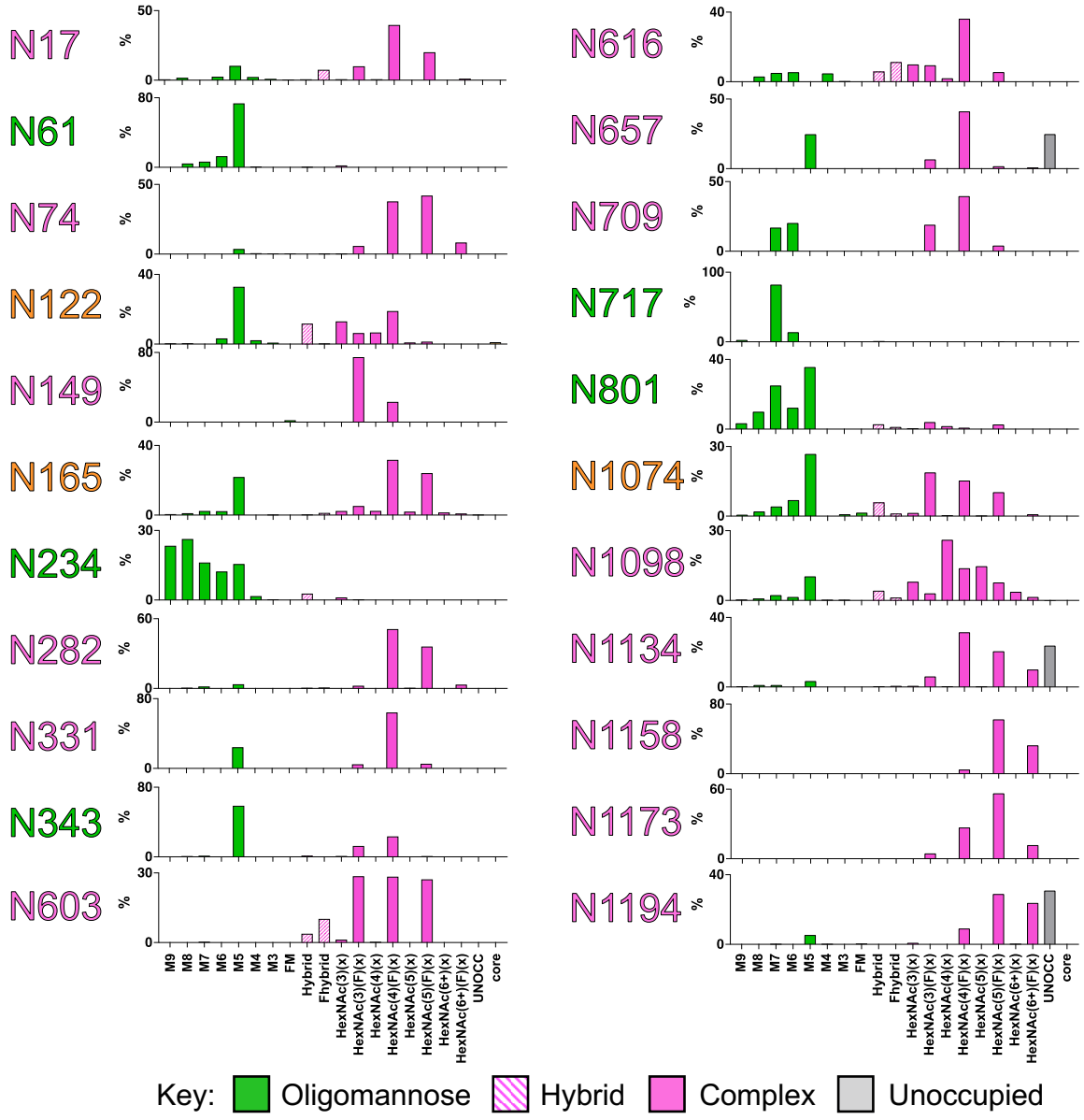


Figure 4.4. **Site-specific glycan analysis of SARS-CoV-2 Closed-S.** The graphs summarize the quantitative mass-spectrometric analysis of the glycan population present at individual N-linked glycosylation sites simplified into categories of glycans. The oligomannose-type glycan series (M9 to M5; $Man_9GlcNAc_2$ to $Man_5GlcNAc_2$) is colored green, afucosylated and fucosylated hybrid-type glycans (hybrid and F hybrid) are dashed pink, and complex glycans are grouped according to the number of antennae and presence of core fucosylation (A1 to FA4) and are colored pink. Unoccupancy of an N-linked glycan site is represented in gray. Close S protein was provided by Dr. Hans Langedijk, Janssen Vaccines and Prevention.

4.5 Influence of mutations on glycan compositions of Closed S

This section discusses about three constructs, 2P, HexaPro and Closed-S protein which have shown high stability and enhanced expression. S-2P which contains two proline substitutions (K986P and V987P) has shown highly similar glycan shield as viral S protein and has already been extensively

used in the clinically approved vaccines [100,299]. HexaPro S protein contains 6 proline substitutions and has shown 32-fold higher protein expression than S2P [110]. However, there is some difference in protein conformation across both proteins in cryo-EM studies, as S2P exposes only one RBD-up conformation whereas some proportion of HexaPro S protein has shown RBD in two-up conformation. Despite these differences in protein conformations, the glycan composition across both of these proteins is highly similar [407]. In preclinical studies, HexaPro has shown promising results as a SARS-CoV-2 vaccine candidate [401,402]. In contrast to SARS-CoV-2 S2P and HexaPro, Close-S exhibits all three RBD's in down state arresting the first step in the conformational change which may resemble more closely to native viral spike **(Figure 4.5)** [438,439]. Closed-S protein contains four substitutions (D614N, A892P, A942P, and V987P) with low surface exposure to create closed trimer [111]. The D614 mutation helped in enhancing stability which may be explained by a decrease in premature shedding of S1 [441].

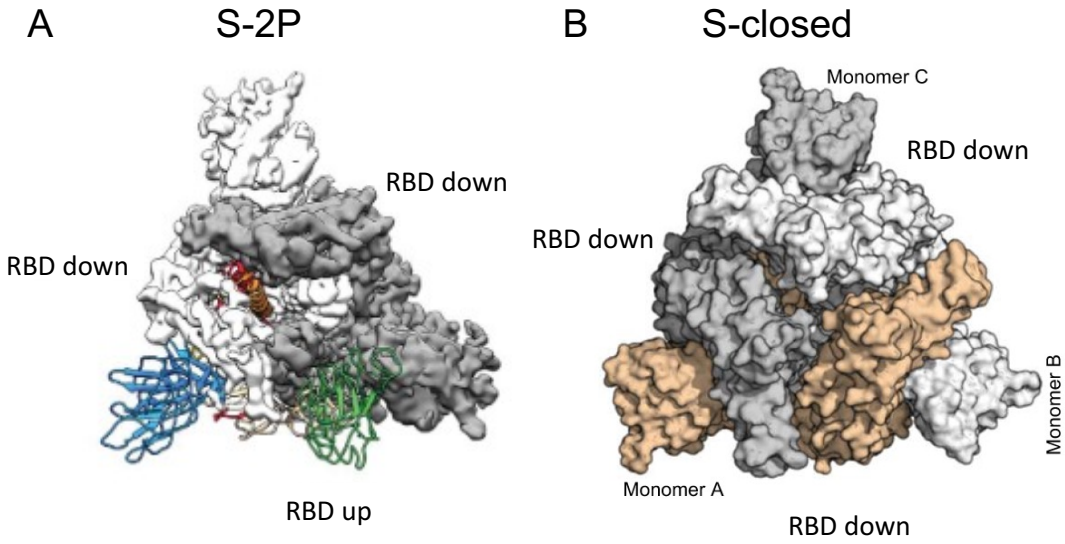


Figure 4.5: Schematic representation of cryo-EM structures of S-2P and S-closed. The PDB accession code for S-2P: 6CRZ [109] and S-closed: 7A4N [111].

Based on the cryo-EM studies, all these variants have shown high stability and trimeric conformation. This study compares the percentage point (p.p.) change in oligomannose content across these three constructs (**Figure 4.6**). The oligomannose-content was used for comparison, as it is associated with the quaternary structure of the protein. Firstly, the p.p. difference in oligomannose-type glycan content across Closed-S and 2P was compared. The glycan sites, N343 and N709

revealed the major changes ≥ 50 p.p. in oligomannose-type glycan content between Closed S and 2P (**Figure 4.6 A**). N343 site showed increase in oligomannose content in Closed S whereas N709 showed decrease in oligomannose content in Closed-S. Notably, both N331 and N343 glycan sites which are present in RBD regions, displayed an increase in oligomannose-content in Closed-S. This, however, potentially reveals that RBD region is less exposed in Closed-S due to steric hinderance and closed trimeric structure compared to the 2P in which one RBD remains exposed. Subsequently, a comparison was made regarding the glycan composition across the Closed-S and HexaPro proteins, both of which have exhibited enhanced expression in comparison to 2P. The Closed-S showed increase in p.p. difference in oligomannose-content at several sites, N61, N122, N331, N343, N657, and N717, compared to HexaPro (**Figure 4.6 B**). This reveals that Closed-S protein is quite intact structurally and has less exposed glycan sites compared to HexaPro. However, N603 and N709 showed ≥ 50 p.p. decrease in oligomannose-content in Closed-S compared to HexaPro (**Figure 4.6 B**). This change in oligomannose-content is potentially due to the mutation at D614N in Closed-S, which may have enhanced the accessibility of N603 and N709 sites for glycan processing.

Furthermore, the difference in glycan composition across HexaPro and 2P was compared. The N61, N122 and N164 glycan sites revealed the ≥ 30 p.p. changes in oligomannose-type glycan content between 2P and HexaPro (**Figure 4.6 C**). The N165 displayed an increase in p.p. difference in oligomannose-content glycan compositions when compared with 2P. However, this may be due to the close proximity of this site to the RBD region, its glycan processing may be influenced by the orientation of the RBD. This is consistent with the previously analyzed MD simulation of two RBD-up and one RBD-up conformation, which confirms that the steric environment of N165 and N122 were expected to change [407].

However, it is noteworthy from this analysis that glycan composition of N234 is not affected by the mutations in these variants of S protein. The conservation of oligomannose-glycan content at the N234 site signifies the structural property inherent to the S protein.

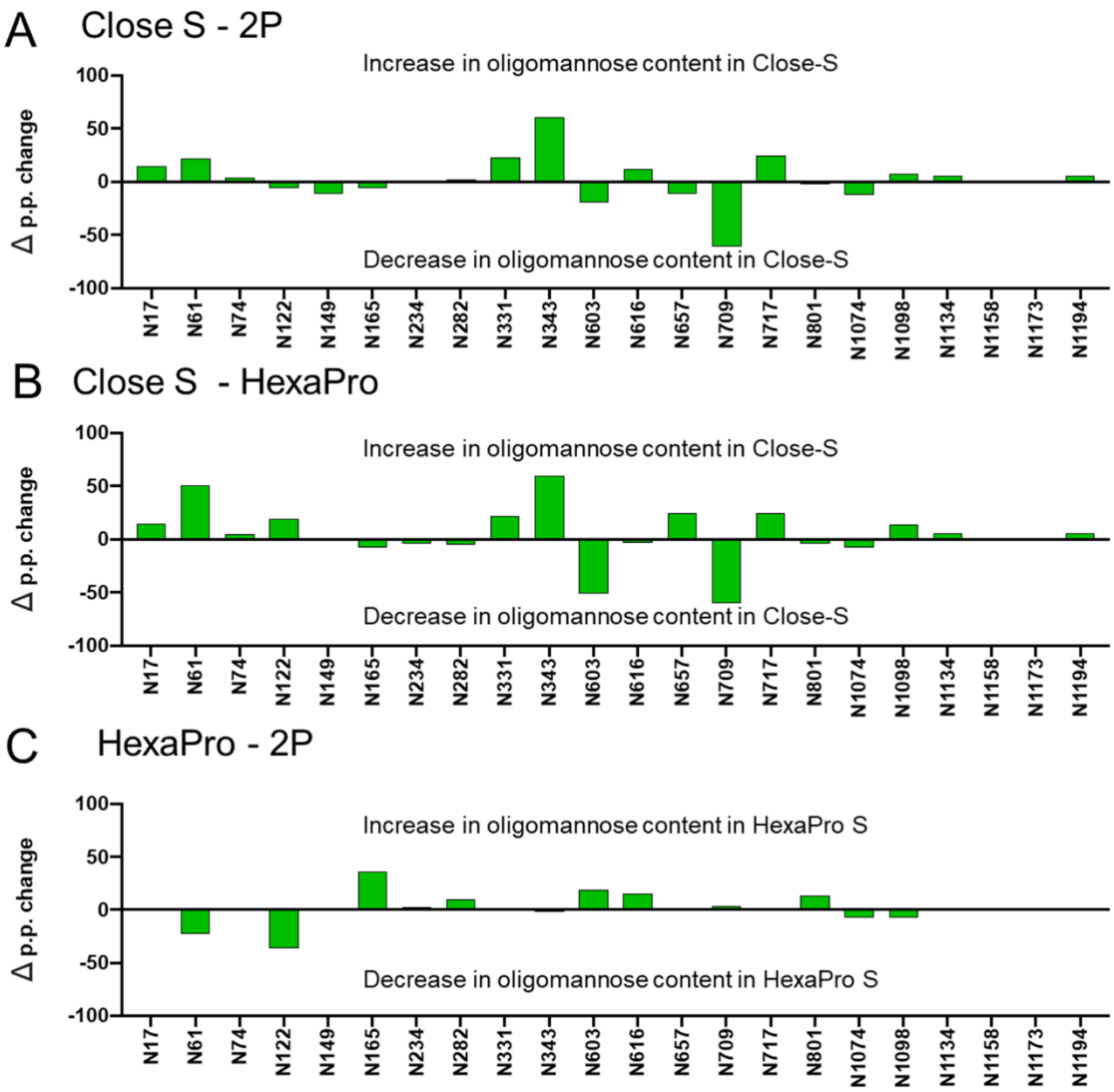


Figure 4.6. Comparative glycan analysis of Closed-S protein with S2P and HexaPro. The bar chart represents the percentage point difference in oligomannose-type glycan between Closed-S and 2P (A), Closed-S and HexaPro (B), HexaPro and 2P (C). Closed S protein was provided by Dr. Hans Langedijk, Janssen Vaccines and Prevention.

4.6 The expression of monomeric RBD constructs impacts glycan processing

Glycan analysis analysed in previous sections demonstrates that S protein glycosylation is influenced by quaternary protein architecture, and other factors. This section compares the glycosylation of soluble recombinant monomeric RBD with that of recombinant and viral derived trimeric S protein. The recombinant RBD was expressed and purified for site-specific glycosylation of the two glycan sites located in the RBD, N331 and N343, to compare with those observed on recombinant S protein

and viral derived S protein reported previously [8,124]. Overall, the N331 and N343 glycans across all expression formats were highly processed, with little to no oligomannose-type glycans detected (**Figure 4.7.A** and **B;** and **Table 4.2**). For the RBD sites presented on viral derived S protein, the complex-type glycans observed were similar to those observed at N282 with most of the PNGS occupied by bi- and tri-antennary glycans (HexNAc(4) and HexNAc(5)) (**Figure 4.7.A**). The majority of these glycans were fucosylated, however, a large proportion of the complex type-glycans on viral derived S lacked fucosylation (24% N331 and 20% N343). As with other sites on the viral derived S protein, minimal levels of sialylation were observed and the majority of glycans possessed at least one galactose residue.

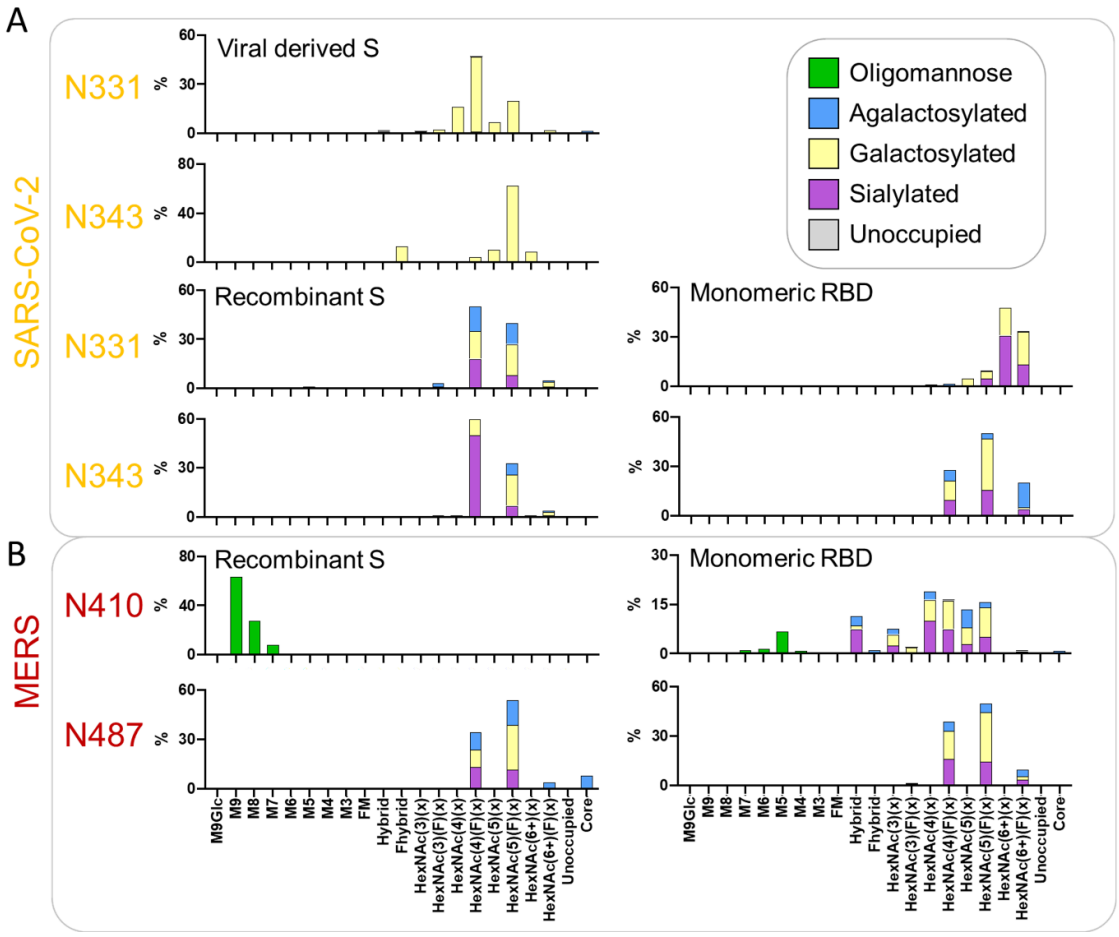


Figure 4.7. Comparative analysis of the glycosylation of the two PNGS on the RBD for viral derived S protein, recombinant S protein and monomeric RBD. (A) Detailed site-specific glycan compositions of the two sites located in the RBD of SARS-CoV-2. Recombinant S protein data is reproduced from Chawla et al. [407] and data for the viral S protein RBD sites was obtained from Yao et al. [124]. Site-specific glycan data is presented as outlined in Figure 2. **(B)** Site-specific compositions for N-glycan sites located in the RBD of MERS-CoV when expressed as part of a soluble recombinant S protein compared to RBD-only. Data for the MERS-CoV S protein were obtained from a previous study [7]. Site-specific glycan data is presented as outlined in Figure 2.

In comparison, the two RBD sites of recombinant soluble S protein are highly fucosylated, with close to 100% of the glycans at N331 and N343 containing at least one fucose (**Figure 4.7.B**). As with the viral-derived S protein the majority of the glycans are bi- and tri-antennary complex-type glycans. The recombinant S protein RBD sites are also more sialylated than N331 and N343 on viral-derived S protein, with 28% of N331 glycans containing at least one sialic acid and 60% of those on N343 (**Table 4.2**). Interestingly, when RBD is expressed as a monomer there are additional subtle changes compared to viral derived and recombinant S protein. The most prominent change is in the glycan branching; whereas recombinant and viral derived RBD glycan sites possess low quantities of tetra-antennary glycans, approximately one third of the glycans at N331 and one fifth of the glycans at N343 consist of these larger branched structures (**Figure 4.7**). When compared to recombinant S protein the monomeric RBD sites also possess low levels of biantennary glycans. Despite these changes both recombinant trimeric S protein and monomeric RBD sites have high levels of fucosylation. These results suggest that complex-type glycans are under differential control hierarchies where certain forms of glycan processing could be influenced by the structural presentation of glycan sites. The attachment of fucose, galactose and sialic acid for these RBD sites appears to be controlled by more global phenomena, such as the producer cell, as the attachment of these monosaccharides is similar when comparing monomeric RBD and trimeric recombinant S protein which were produced in identical cell lines. The branching of the RBD sites on monomeric RBD is greater than that of trimeric S protein, both viral and recombinant, and suggests that the quaternary structure of the glycoprotein may have a small role on the elaboration of complex-type glycans of the RBD. These results are similar to previous analyses that have compared trimeric S protein with monomeric S1 [117].

Table 4.2. Site specific glycan composition of RBD sites, N331 and N343

The site-specific composition of the two N-linked glycan sites, N331 and N343, are displayed, with the categories identical to Table S1. The values for N331 and N343 for recombinant and virion are reproduced from Table S1, the monomeric RBD values are those represented as bars in Figure 3.

N331

	Viral derived S	Recombinant S	Monomeric RBD
M9Glc	0%	0%	0%
M9	0%	0%	0%
M8	0%	0%	0%
M7	1%	0%	0%
M6	0%	0%	0%
M5	0%	1%	0%
M4	0%	0%	0%
M3	0%	0%	0%
FM	0%	0%	0%
Hybrid	2%	0%	0%
Fhybrid	0%	0%	0%
HexNAc(3)(x)	1%	0%	0%
HexNAc(3)(F)(x)	2%	3%	0%
HexNAc(4)(x)	16%	0%	1%
HexNAc(4)(F)(x)	47%	49%	1%
HexNAc(5)(x)	7%	0%	5%
HexNAc(5)(F)(x)	20%	40%	10%
HexNAc(6+)(x)	0%	0%	48%
HexNAc(6+)(F)(x)	2%	6%	34%
Unoccupied	0%	0%	0%
Core	1%	0%	0%

N343

	Viral derived S	Recombinant S	Monomeric RBD
M9Glc	0%	0%	0%
M9	0%	0%	0%
M8	0%	0%	0%
M7	0%	0%	0%
M6	0%	0%	0%
M5	0%	0%	0%
M4	0%	0%	0%
M3	0%	0%	0%
FM	0%	0%	0%
Hybrid	0%	0%	0%
Fhybrid	13%	0%	0%
HexNAc(3)(x)	0%	0%	0%
HexNAc(3)(F)(x)	0%	1%	1%
HexNAc(4)(x)	0%	1%	0%
HexNAc(4)(F)(x)	5%	61%	28%
HexNAc(5)(x)	11%	0%	0%
HexNAc(5)(F)(x)	62%	32%	50%
HexNAc(6+)(x)	9%	1%	0%
HexNAc(6+)(F)(x)	0%	3%	20%
Unoccupied	0%	0%	0%
Core	0%	0%	0%

N331

Viral derived S

Sialylated	Galactosylated	Agalactosylated	Mannose	Unoccupied
1.70%	94.46%	3.84%	0.67%	0.00%

Recombinant S

Sialylated	Galactosylated	Agalactosylated	Mannose	Unoccupied
27.59%	39.26%	31.66%	1.37%	0.00%

Monomeric RBD

Sialylated	Galactosylated	Agalactosylated	Mannose	Unoccupied
49.28%	47.45%	3.27%	0.47%	0.00%

N343

Viral derived S

Sialylated	Galactosylated	Agalactosylated	Mannose	Unoccupied
0.00%	100.00%	0.00%	0.00%	0.00%

Recombinant S

Sialylated	Galactosylated	Agalactosylated	Mannose	Unoccupied
59.09%	31.61%	9.31%	0.00%	0.00%

Monomeric RBD

Sialylated	Galactosylated	Agalactosylated	Mannose	Unoccupied
29.77%	44.02%	26.21%	0.30%	0.00%

To further explore differences in glycosylation of RBD PNGS, the similar comparative glycan analysis was performed on MERS-CoV. The site-specific glycan analysis of recombinant MERS-CoV S protein has been reported previously [7]. For the comparative analysis of trimeric recombinant protein and RBD, the MS files obtained in the previous study were searched using an identical version of the analysis software using the same glycan libraries. This analysis revealed differences at the glycan sites present on MERS RBD when expressed monomerically compared to recombinant trimeric soluble S protein. One of the sites, N487, is similar between the two platforms, presenting glycoforms typical of sites populated by complex-type glycans (**Figure 4.7.B**). In contrast N410 is occupied by exclusively oligomannose-type glycans when present on recombinant trimeric S protein. When monomeric MERS RBD is expressed, the processing of the N410 site is markedly

increased, with complex-type glycans predominating the glycan profile, although a subpopulation of oligomannose-type glycans remain (**Figure 4.7.B**). Modelling the N410 glycan onto published structures of MERS S protein reveals that the N410 glycan is protected from processing by mannosidase enzymes such as ER α-mannosidase I when buried within the trimer, but when presented on monomeric RBD it is readily accessible to glycan processing [125]. These observations further highlight how the quaternary structure of a glycoprotein is a key determinant of the glycan processing state and demonstrates how glycan sites can provide information about protein folding and quaternary structure.

4.7 Glycan composition analysis of SARS-CoV-2 RBD nanoparticles

Self-assembling nanoparticle immunogens are known to enhance the immunological response, as it enhances the uptake by dendritic cells and in trafficking to the draining lymph nodes [122,442]. This section discusses about glycan composition of the self-assembling SARS-CoV-2 RBD nanoparticles. The constructs which are analysed in this study are RBD g5.1, RBD g5.1 24mer and RBD g8.2. These constructs are developed by incorporation of additional glycans and feasibility of addition of N-linked glycans was modelled using advanced algorithm, cloaking with glycans (CWG) [414]. This led to the potential RBD constructs with three, five and eight additions of N-linked glycans on basis of glycan distance map. However, RBD g5.1 and RBD g8.2 showed high neutralizing antibody response and less binding to non-neutralizing antibodies [414] .

To determine the glycan composition analysis of RBD g5.1 and RBD g8.2, single site-specific glycan analysis was performed utilizing LC-MS. To convey the occupancy and processing states at each PNGS, the abundances of each glycan was categorized as: oligomannose-type (high mannose, $Man_3GlcNAc_2$ to $Man_9GlcNAc_2$, including fucosylated mannose), hybrid- and complex-type divided in three subgroups; agalactosylated (contains no galactose), galactosylated (containing atleast one galactose), and sialylated (containing atleast one sialic acid), and unoccupied (no glycan). Core glycans were also included in the analysis, which represents truncated glycan groups i.e., compositions smaller than $HexNAc_2Hex_3$.

Both RBD g5.1 and RBD g5.1 24mer nanoparticles are decorated with six N-glycan site. Three protease enzymes, trypsin, chymotrypsin and alphalytic protease were used to generate the glycopeptides. The glycan analysis reveals highly similar glycan composition across both RBD g5.1 and 24mer protein, except N331 and N428 sites showing low glycan occupancy in RBD g5.1 compared to RBD g5.1 24mer nanoparticle (**Figure 4.8**). The N460 glycan site has shown high underoccupancy in both the immunogens, RBD g5.1 and RBD g5.1 24mer. Comparison of RBD g5.1 and 24mer indicates high abundance of processed type glycans at glycan site N331, N354, N428 of RBD g5.1 24 mer. Overall, both the immunogens have shown high presence of processed hybrid- and complex-type glycans (**Figure 4.8**).

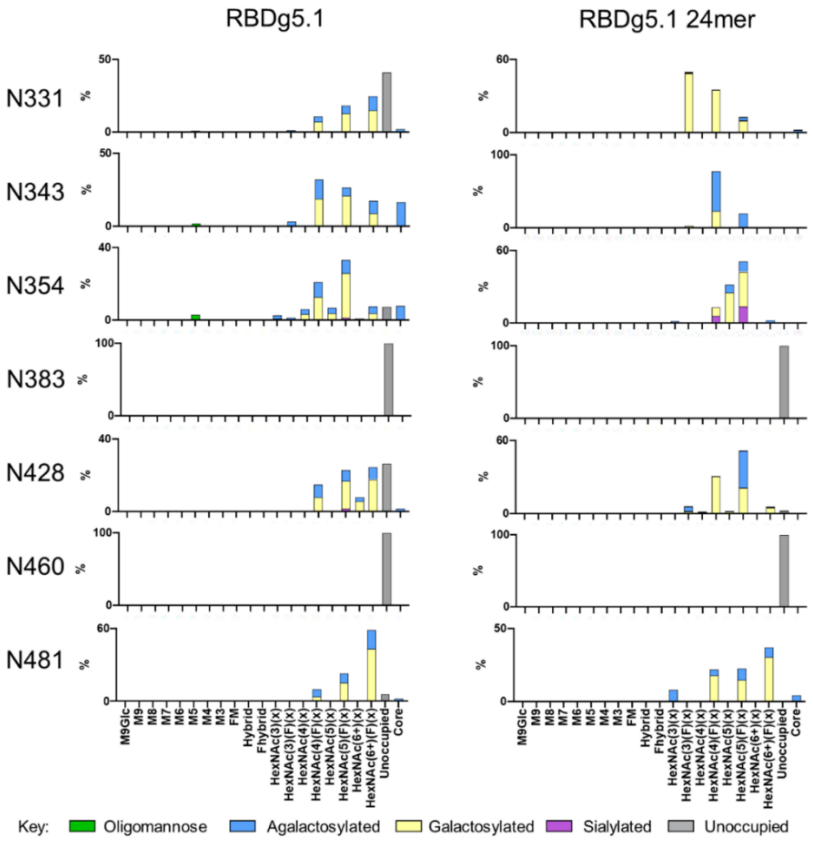


Figure 4.8. Glycan compositions of RBD monomer and nanoparticles. The bar graph represents site-specific glycan analysis at each PNGS present in RBD g5.1 and RBD g5.1 24-mer. Glycans were categorized and colored according to the detected compositions. Oligomannose-type glycans (M9–M4) are colored green. Hybrid-type glycans, those containing three HexNAcs and at least five hexoses, were colored as for complex-type glycans, because one arm can be processed in a similar manner. Complex-type glycans were categorized according to the number of HexNAc residues detected and the presence or absence of fucose. Core glycans represent any detected composition smaller than HexNAc2Hex3. For hybrid- and complex-type glycans, bars are colored to represent the terminal processing present. Blue represents agalactosylated, yellow galactosylated (containing at least one galactose), and purple sialylated (containing at least one sialic acid). The proportion of unoccupied PNGSs is colored gray. Recombinant protein for RBD g5.1 and 24mer was provided by Dr. Daniel Kulp (Wistar Institute, USA).

Furthermore, the glycan composition of RBD g8.2 was analyzed, which also seem to elicit strong humoral response [414]. The recombinant RBD g8.2 protein is decorated with nine N-linked glycan sites. The N-glycan site, N343, N354, N428 and N481 are showing high abundance of hybrid- and complex-type glycans (**Figure 4.9**). The N-glycan sites, N354 and N481, revealed higher glycan processing and ~50% of sialic acid. However, there are several N-glycan sites which have shown high underoccupancy, N331, N360, N389, N479 and N536. Perhaps, this could be due to the spatial conformation of the protein causing steric hinderance and resistance for the ER and Golgi processing enzymes to access the N-glycan sites. Interestingly, g5.1 has more occupied glycans than g8.2, suggesting g5.1 may be superior for focusing antibody responses.

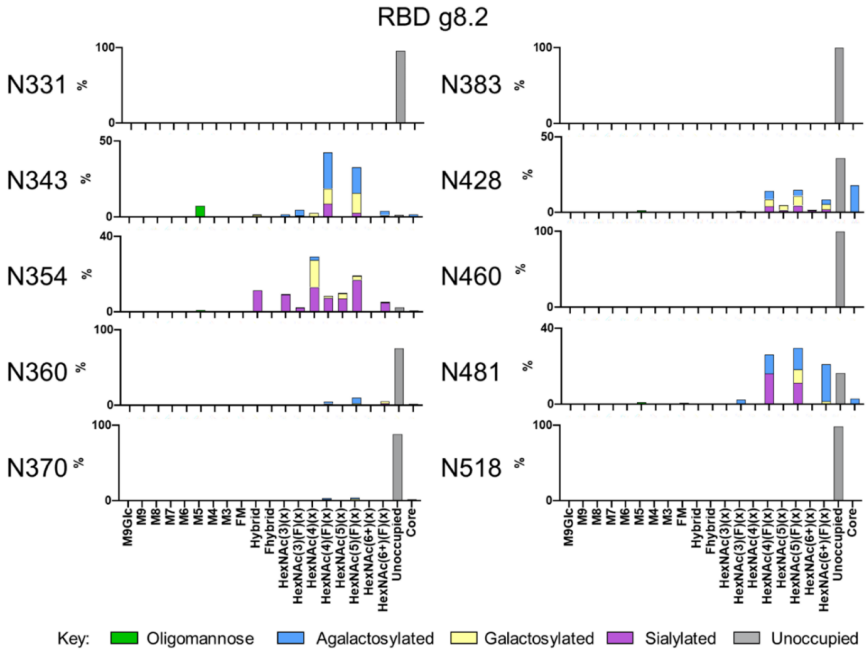


Figure 4.9. Site-specific glycan analysis of RBD g8.2. The bar graph represent glycan composition at each PNGS of RBD g8.2**.** Glycans were categorized and colored according to the detected compositions. Oligomannose-type glycans (M9–M4) are colored green. Hybrid-type glycans, those containing three HexNAcs and at least five hexoses, were colored as for complex-type glycans, because one arm can be processed in a similar manner. Complex-type glycans were categorized according to the number of HexNAc residues detected and the presence or absence of fucose. Core glycans represent any detected composition smaller than HexNAc2Hex3. For hybrid- and complex-type glycans, bars are colored to represent the terminal processing present. Blue represents agalacto-sylated, yellow galactosylated (containing at least one galactose), and purple sialylated (containing at least one sialic acid). The proportion of unoccupied PNGSs is colored gray. Recombinant protein for RBD g8.2 was provided by Dr. Daniel Kulp (Wistar Institute, USA). No replicates were done for this experiment.

4.8 Discussion

The global impact of COVID-19 has resulted in laboratories across the globe producing recombinant spike protein for vaccine design, antigenic testing and structural characterization. Whilst the site of glycan attachment is encoded by the viral genome, the processing of the attached glycans can also be influenced by a wide range of exogenous phenomena, including recombinant host systems and processes for production in cell culture. Using immediately available materials, this study compares the glycosylation of Vero-produced virus preparations for Spike proteins with those of HEK-293 derived recombinant Spike proteins. It is acknowledged that these preparations are likely to differ from Spike proteins as produced across a range of cells in infected patients.

However, this study shows that when a glycan site is located in regions of the SARS-CoV-2 S protein which are not readily accessible the glycan site will possess high levels of under-processed oligomannose-type glycans, a phenomenon that will likely be of a general nature. This conclusion is derived from the observation that the under-processing of glycans at N234 site was consistently oligomannose-type across different constructs of S protein, S2P, Closed S, HexaPro and are also conserved on Vero cell derived S protein derived from infectious SARS-CoV-2. Whilst the presence of oligomannose-type glycans are rare on the majority of host glycoproteins their abundance on SARS-CoV-2 is lower than that of other viral glycoproteins and means that the glycan shield density of this virus is low when compared to HIV-1 Env and Lassa GPC. This is likely to mean that the immunodominant protein epitopes remain exposed. The processing equivalence between recombinant and viral derived S protein indicates that recombinant S protein glycosylation for vaccinations will likely mimic those occurring in human infections and will remain antigenically comparable. Usually the complex-type glycosylation is driven more by other parameters including the producer cell and culturing conditions.

One observed exception to this was that the expression of monomeric RBD as opposed to trimeric S protein does increase glycan branching on the two glycan sites located on SARS-CoV-2 RBD. This suggests that the glycan processing of complex-type glycans, in addition to oligomannose-type glycans, may also be under structural constraints, albeit to a much lesser extent.

Self-assembling RBD nanoparticles also revealed higher glycan occupancy than their RBD counterparts which would be an important factor to consider for nucleic acid delivery of immunogens. Overall, these results demonstrate that whilst N-linked glycosylation is highly diverse at certain regions of the S protein, there is a broad consensus of glycan processing with regards to oligomannose-type glycans between virus and immunogen S protein. This is something which cannot be taken for granted, as when comparing recombinant and viral derived HIV-1 Env, the reduced glycan occupancy of immunogens can induce an immune response that is incapable of neutralizing the virus. The reproducibility of S protein glycosylation in different constructs of S protein, is of significant benefit for stabilized and enhanced expression of immunogen design, serology testing and drug discovery and will mean that the glycans of SARS-CoV-2 are unlikely to provide a barrier to combatting the COVID-19 pandemic.

Chapter 5 Signatures of native-like glycosylation in RNA replicon-derived HIV-1 immunogens

Contributions.

Chapter 5 outlines the assembly and glycan analysis of the HIV-1 Env derived from RNA replicon. The construct for RNA replicon, BG505 NFL.664, was kindly provided by the laboratory of Professor Darrell J. Irvine (Massachusetts Institute of Technology, USA). Negative stain-electrom micrographic (NS-EM) studies were used to confirm the immunogen integrity of Env derived from RNA replicon and was performed by the laboratory of Prof. Andrew Ward (The Scripps Institute, USA).

5.1 Introduction

The response to the recent pandemic has highlighted the value of adenoviral [443] and RNA-based immunogens [296]. A major advantage is the speed at which they can be deployed in comparison to traditional protein-based immunogens [311,444,445]. However, one potential limitation is that the viral immunogens must be entirely encoded by the nucleotide sequence and there is no opportunity for post-expression purification that is typically implemented in the production of recombinant-based immunogens [254,274,446]. Therefore, there is considerable interest in understanding how RNA-based and adenoviral-based immunogens are assembled *in situ* [126,426] and how nucleotide editing can help shape this process [268]. In HIV-1 vaccine research, understanding immunogen assembly is particularly important because it is thought that in order to elicit a broadly neutralizing antibody (bnAb) response [246] the immunogen must display native-like assembly [260,283,447–449], post-translational modifications [286,375,450], and minimal amounts of non-native material [279,280,283,451,452]. One of the leading HIV vaccine strategies aiming at eliciting bnAbs does however entail targeting particular B cell lineages with highly engineered immunogens but this strategy also includes immunization with a final ‘polishing immunogen displaying native-like features’ [265,287,453,454].

The HIV Envelope glycoprotein (Env) has emerged as the main focus for HIV-1 vaccine design and there have been considerable developments in the production of native-like Env glycoproteins [201–203,452,455–458]. Manufacturing of viral spike has proved challenging as the viral glycoprotein is metastable and requires complex post-translational maturation. In a native setting, the *env* gene is translated as a gp160 polypeptide which is heavily glycosylated and is proteolytically cleaved via furin into gp120 and gp41 heterodimers which trimerize to form a spike on the virus surface [205,459]. A breakthrough strategy in recombinant production involved solubilizing the protein by adding a stop codon at position 664 to prevent the translation of the transmembrane domain, disruption of the fusion mechanism of gp41 by the insertion of I559P mutation, and the use of stabilizing disulfide bonds, collectively referred to as SOSIP.664 [201,202,260]. This format is dependent on furin cleavage, and despite optimization of the furin cleavage site can still exhibit incomplete cleavage and resulting material usually requires post-expression purification [254].

The dependence on protease cleavage can be bypassed using single-chain constructs, such as the natively flexible linked format (NFL.664) [217–219]. The assembly and glycosylation of NFL.664 and SOSIP.664 are known to be highly similar and are of particular interest in nucleotide-based immunogen design [217]. However, as these recombinant formats generally require post-production purification, we are interested in the assembly of RNA-derived immunogens where this process is precluded and where there is consequently an elevated potential for the production of non-native-like material. Glycosylation is highly sensitive to three-dimensional macromolecular architecture [125,214] and we can utilize site-specific analysis of potential N-linked glycosylation sites (PNGS) to monitor immunogen assembly [374,460].

One such important signature in immunogen assembly is the preservation of underprocessed oligomannose-type glycan structures. Previous studies have demonstrated the presence of an extended cluster of oligomannose-type glycans, known as the intrinsic mannose patch (IMP), at the outer domain of gp120 [232]. The IMP sites are known to be highly conserved [461] and a target of numerous broadly neutralizing antibodies [246]. However, this region is known to be a characteristic of the monomeric gp120 and previous analyses of recombinant native-like trimers and virion-derived Env have demonstrated the presence of an additional so-called ‘trimer-associated mannose patch’

(TAMP) on the native-like trimer [214,231]. The presence of these glycans at protomer interfaces is linked to the quaternary structure of the protein and they usually arise due to steric occlusion of glycan processing enzymes [18]. These glycans often contribute to the epitopes of many broadly neutralizing antibodies and can also influence immunogen trafficking [122,215]. In contrast to oligomannose-type glycans which are predominantly dependent on the macromolecular architecture, complex-type glycans are known to be substantially influenced by the host cell [3,210]. Fine glycan processing in complex-type glycans, such as changes in sialylation or fucosylation can also influence antibody binding [120,275].

This study utilized glycopeptide analysis as a probe to understand the assembly and protein architecture-dependent glycan signatures of Env derived from RNA-based immunogens. The production of immunogens arising from the alphavirus-based replicon system was examined [462]. This system exhibits self-amplification within the target cell via the transcription of non-structural proteins which copies the replicon RNA in addition to desired proteins, in this case BG505 NFL.664 Env protein. Although RNA replicons are an effective platform for delivering viral glycoprotein [304,305,313,463], the requirement for high-fidelity folding in HIV-1 has prompted us to examine the assembly and processing of the Env protein. This study further investigate how assembly can be influenced by nucleotide editing.

5.2 Expression and characterization of single-chain replicon HIV-1 immunogen

In order to investigate the entire population of Env products produced by replicon-mediated expression we analyzed the supernatant captured by a broad-acting lectin, *Galanthus nivalis* lectin which binds to α-1,3 mannose linkage found on most of the viral glycoproteins. The resulting Env protein was analyzed by negative stain-electron micrography (NS-EM) and liquid chromatography-mass spectrophotometry (LC-MS).

The replicon system derived from the Venezuelan equine encephalitis virus (VEEV) was used for the expression of HIV-1 Env. The VEEV genome encodes for non-structural proteins (nsp1-4) including the replicase (RNA dependent RNA polymerase complex) and the structural genes,

separated by sub genomic promotor (SGP), which drives the transcription of structural proteins of the RNA, thus has a potential for sustained expression. The VEEV genome was used as a backbone as described previously [313] with structural genes replaced with single-chain natively flexible linked HIV-1 Env trimer of the clade A strain of HIV-1, BG505 (**Figure 5.1.A**). The design of the so-called BG505 NFL.664 trimer has been described previously [217,218] and is summarized in Figure 5.1A.

To express the Env protein using the VEEV replicon, plasmid DNA containing the VEEV's genome and BG505 NFL.664 genes was transcribed into RNA using an in vitro transcription (IVT) reaction, followed by the clean-up of RNA as described previously [304,313]. Transfection of purified RNA in human embryonic kidney (HEK) 293F cells was harvested and further purified using *Galanthus nivalis* lectin (GNL). This purification method was selected as GNL binds to the core of all N-linked glycans and thereby facilitates the analysis of the whole Env protein fraction, regardless of its native configuration. As the replicon-based immunogen delivery eliminates post-expression purification, GNL beads were used to purify the whole Env material. SDS-PAGE was run to confirm the presence of protein in purified material (**Figure 5.1.B**). Furthermore, NS-EM was utilized to determine the proportion of Env protein derived from replicon expression displaying native-like conformation. NS-EM analysis revealed that 83% of the total protein structure is trimeric and 17% constitutes non-native material (**Figure 5.1.C**). Non-native material has been known to elicit non-neutralizing antibody responses [279,451].

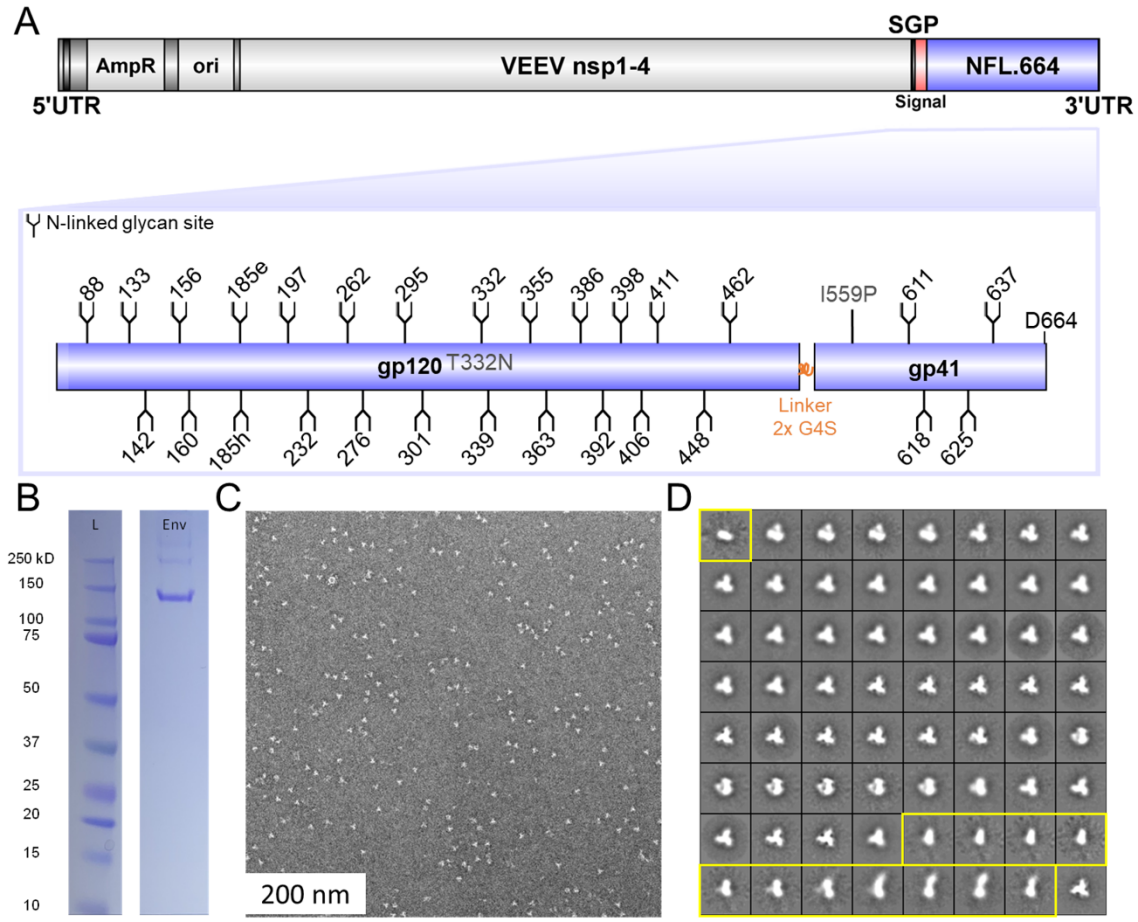


Figure 5.1. Expression of replicons encoding soluble single chain HIV-1 immunogen (BG505 NFL.664) A) Schematic representation of the BG505 NFL.664 replicon constructs. The domains of the construct illustrated are: Venezuelan equine encephalitis vector (VEEV), non-structural proteins (nsp's), sub genomic promoter (SGP), ampicillin resistance (AmpR), origin of replication (ori), Untranslated region (UTR). **B)** SDS-polyacrylamide gel electrophoresis analysis of the replicon expressed BG505 NF.664 Envelope in human embryonic kidney (HEK) 293F cells purified using *Galanthus nivalis* (GNL) lectin beads. **C)** Representative of NS-EM image of BG505 NFL.664 Env particles secreted by HEK 293F cells transfected with replicon RNA. **D)** 2D class averages of the particles. Amongst all of the particles imaged, 17% fall within class averages corresponding to non-native conformations shown in yellow. The NS-EM analysis was performed by Wen-Hsin Lee and Dr. Gabriel Ozorwoski (The Scripps Research Institute, USA).

5.3 Site-specific glycan characterization of replicon expressed HIV-1 Env protein.

To characterize the glycosylation of replicon-expressed Env, the site-specific glycan compositions were resolved using three or more biological replicates of Env expressed in mammalian cells, HEK 293F (**Figure 5.2.A and B**). To quantify the glycan compositions at each site, we digested glycopeptides using three protease enzymes, trypsin, chymotrypsin, and alpha lytic protease (ALP). These proteases were utilized to generate the glycopeptides containing a single N-linked glycan

sequon and the resulting glycopeptides were analyzed using LC-MS. The glycan compositions across the 28 N-glycan sites were ascertained; nonetheless, there were certain sites for which replicates could not be obtained, resulting in their presentation without error bars (**Figure 5.2.B**). To display the glycan compositions on HIV-1 Env, the glycan composition was modelled onto the trimeric Env based on the percentage of oligomannose-type glycan present at each site (oligomannose-type, 100-80%; mixed, 79—30%; and complex-type glycans, 29—0%; **Figure 5.2.A**).

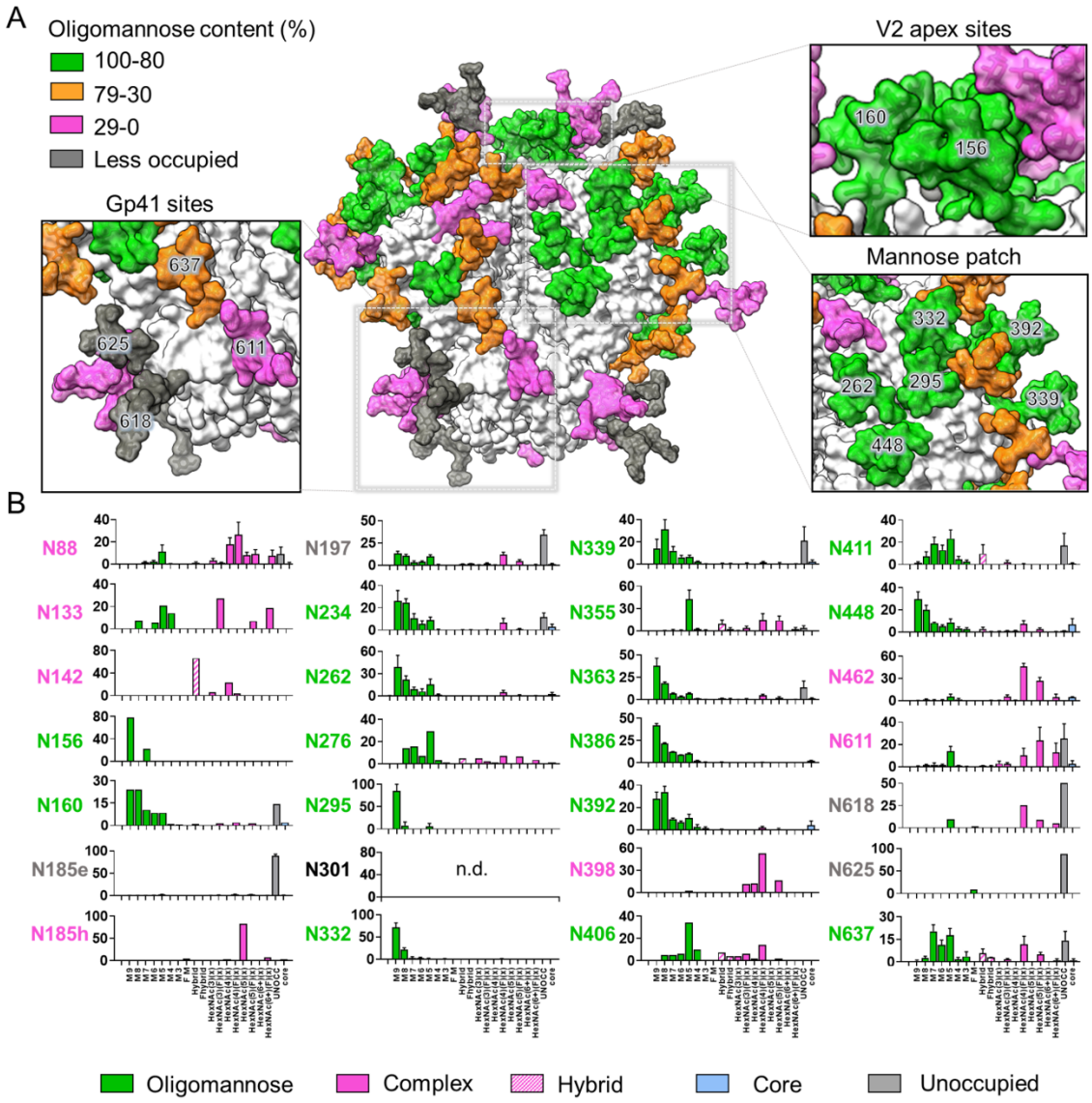


Figure 5.2. Glycosylation on the replicon expressed soluble NFL Env trimer. A) Representative glycans are modelled onto the HIV-1 Env trimer (PDB ID 5C7K) [206]. All the 28 N-glycans are mapped on the model based on abundance of oligomannose content as defined by the key. Less or underoccupied sites are highlighted in gray. B) Site-specific analysis of the replicon expressed BG505 NFL.664. Three or more biological replicates of BG505 NFL.664 were produced in HEK293F to generate the glycan compositions. Bars represent the mean ± the standard error of the mean of the biological replicates of expressed protein analyzed. Oligomannose-type glycans $Man_3GlcNAc_2$ to $Man_9GlcNAc_2$ are represented as M3 to M9 and fucosylated glycans are represented as F.

Prior studies demonstrated that changes in the abundance of underprocessed glycan structures can indicate changes in the fine structure of the protein. To determine the presence of native-like glycan signatures, the glycosylation patterns were compared across RNA replicon expressed Env in this study (**Figure 5.2A and 5.2B**; **Table 5.1**) with previously analyzed virion-derived Env [3,222] and recombinant Env protein [2,18,217,335]. The glycan analysis revealed a highly similar

oligomannose-type and hybrid-type glycan content, with localized variations occurring. A high proportion of oligomannose-type glycans within the IMP and TAMP was observed, aligning with a native-like architectural pattern [18,214]. For example, N332 within IMP displayed 100% oligomannose-type glycans with a predominance of $Man_{8-9}GlcNAc_2$ glycan structures (**Figure 5.2.A and 5.2.B**).

In contrast to oligomannose-type glycans, complex-type glycans are typically not significantly influenced by the protein structure of Env but rather by the host cell used for expression [210]. The complex-type glycans were observed at several sites including N88, N133, N142, N185h, N301, N398, N462 and the gp41 sites N611 and N618 (**Figures 5.2.A and 5.2.B**). This is consistent with the glycan composition observed in recombinant production via plasmid DNA described in later sections of this study, Figure 5.6, and with previous observations of native viral spike and recombinant Env protein [2,3,217,335].

Table 5.1. Glycan composition analysis of replicon expressed env in HEK 293F cells. The upper table shows the categorized glycan compositions at each N-linked glycan site with the reported value the mean of three or more biological replicates. The global averages are shown in the right-hand table. The lower table further categorizes the glycan compositions into oligomannose-, hybrid-, and complex-type as well as the percentage of glycan compositions containing at least one fucose or one sialic acid residue and core-type glycan structures.

	N88	N133	N142	N156	N160	N185e	N185h	N197	N234	N262	N276	N295	N301	N332	N339	N355	N363	N386	N392	N398	N406	N411	N448	N462	N611	N618	N625	N637
M9Glc	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M9	0	0	0	78	24	0	0	13	26	39	0	86	0	72	14	0	38	42	28	0	1	1	30	0	0	0	0	1
M8	0	7	0	0	24	0	0	10	25	22	14	8	0	22	31	0	18	22	34	0	5	7	20	2	1	0	0	3
M7	2	0	0	22	10	0	0	3	11	9	15	0	0	3	12	1	7	12	10	0	5	19	8	1	2	0	0	21
M6	2	6	0	0	9	0	0	4	5	6	7	0	0	2	6	1	3	9	7	0	6	13	5	1	1	0	0	11
M5	11	21	0	0	9	2	1	10	9	15	29	6	0	1	6	42	7	10	11	3	34	23	9	6	14	10	0	17
M4	0	14	0	0	1	0	0	1	0	2	3	0	0	0	2	2	1	2	2	1	10	4	3	2	1	0	0	2
M3	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	1	0	1	1	0	0	2	3	0	0	0	0	3
FM	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	9	0
Hybrid	1	0	66	0	1	0	0	1	0	0	5	0	0	0	1	10	1	1	0	0	7	10	3	0	1	0	1	6
Fhybrid	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	3	1	0	0	1	4	0	0	0	1	0	0	3
HexNAc(3)(x)	3	0	6	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	4	0	1	1	3	0	1	0
HexNAc(3)(F)(x)	1	27	0	0	2	1	0	2	0	0	2	0	0	0	1	4	1	0	0	11	6	2	1	5	2	0	0	2
HexNAc(4)(x)	18	0	23	0	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	12	2	0	1	0	0	0	0	0
HexNAc(4)(F)(x)	27	0	4	0	2	3	0	12	7	5	7	0	0	1	2	15	5	0	2	52	14	0	7	46	10	25	1	12
HexNAc(5)(x)	8	0	0	0	0	0	82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HexNAc(5)(F)(x)	9	7	1	0	2	2	1	5	1	1	6	0	0	0	1	14	2	0	1	17	2	0	3	26	24	9	0	5
HexNAc(6+)(x)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HexNAc(6+)(F)(x)	7	18	0	0	0	7	1	0	0	3	0	0	0	0	0	2	0	0	0	1	0	0	0	5	13	5	0	1
UNOCC	9	0	0	0	15	89	0	34	12	0	0	0	0	0	21	4	14	0	0	0	0	17	1	0	25	50	87	14
core	1	0	0	0	2	1	3	1	3	2	1	0	0	0	2	0	1	2	4	0	0	1	7	4	3	0	0	1

Total	
M9Glc	0
M9	18
M8	10
M7	6
M6	4
M5	11
M4	2
M3	1
FM	1
Hybrid	4
Fhybrid	1
HexNAc(3)(x)	1
HexNAc(3)(F)(x)	3
HexNAc(4)(x)	2
HexNAc(4)(F)(x)	9
HexNAc(5)(x)	3
HexNAc(5)(F)(x)	5
HexNAc(6+)(x)	0
HexNAc(6+)(F)(x	2
Unoccupied	14
core	1

%	N88	N133	N142	N156	N160	N185e	N185h	N197	N234	N262	N276	N295	N301	N332	N339	N355	N363	N386	N392	N398	N406	N411	N448	N462	N611	N618	N625	N637
Mannose	15	47	0	100	77	3	5	41	77	92	69	100	0	99	72	48	75	97	93	5	61	70	77	11	19	12	10	58
Hybrid	1	0	66	0	1	0	0	3	0	0	5	0	0	0	1	13	2	1	0	1	10	10	3	1	1	0	1	8
Complex	73	53	34	0	5	6	93	20	8	6	25	0	0	1	3	35	8	1	3	94	29	3	12	84	52	39	2	19
Unoccupied	9	0	0	0	15	89	0	34	12	0	0	0	0	0	21	4	14	0	0	0	0	17	1	0	25	50	87	14
Fucosylation	44	53	5	0	5	6	12	21	9	6	19	0	0	1	4	38	9	1	3	83	26	2	11	83	50	40	10	22
Sialylation	25	8	93	0	0	1	10	6	4	3	10	0	0	0	0	13	0	0	0	57	12	1	2	33	9	9	0	4
core	1	0	0	0	2	1	3	1	3	3	0	0	0	0	2	0	1	2	4	0	0	1	6	4	3	0	0	1

Total	
Mannose	51
Hybrid	5
Complex	29
Unoccupied	14
Fucosylation	21
Sialylation	11
core	1

Another parameter which influences the elicitation of antibodies is glycan site occupancy, as so-called 'glycan holes' have the potential to elicit non-neutralizing antibodies that can dilute the desired immune response [259]. The glycan-hole antibodies can be neutralizing but are usually strain-specific, autologous neutralizing antibodies [464]. In our analysis, several sites on replicon expressed Env are showing a population of unglycosylated peptides containing PNGS, N185e, N197 and within gp41 at sites, N611, N618, and N625 (**Figures 5.2.A** and **5.2.B**). Similar underoccupancy has been observed previously within recombinant native-like trimers and has resulted in non-neutralizing antibody responses [3,261,268,465]. For example, RM20E1 antibody isolated from a BG505 SOSIP.664-immunized macaque is directed to the glycan hole at N611 [259,268]. The application of approaches to enhance glycan occupancy is demonstrated in later sections of this study.

Overall, our glycosylation analysis reveals the glycan signature characteristic of a native-like trimer on Env derived from replicon RNA. As an increasing number of RNA-based vaccine

candidates are being developed, glycan analysis offers a route to probe immunogen integrity and to give insight into the scope for improving immunogen assembly.

5.4 Non-native material exhibits near native-like glycosylation

While the majority of the glycosylation sites within an intrinsic mannose patch are known to display under processed oligomannose-type glycans, any changes in glycan processing at these sites can provide an insight about structural changes in the trimeric structure of Env**.** Thus, glycans at these sites can be used as a sensitive reporter to signify the presence of non-native constituents in the expression. From NS-EM data, it was revealed that a certain proportion of the Env material expressed by replicon RNA contains non-native material. To confirm whether the presence of non-native material impacts immunogen glycosylation, the site-specific glycan analysis of replicon-expressed Env purified via GNL with the Env purified via trimer-specific antibody, PGT145 were compared [254]. To detect the differences in glycan processing, the glycan processing of oligomannose-type glycans at the IMP were considered, as these sites are under less influence from quaternary structure and their glycan processing states are known to be indicative of native-like conformations within the gp120 outer domain [18,232].

The glycosylation sequon at N262 is known to be highly conserved as the protein-proximal stem of the glycan is largely buried in a protein cleft [207,228,229]. Due to steric hindrance, this site is exclusively dominated by oligomannose-type glycans on HIV-1 Env and presents prominently $Man_9GlcNAc_2$ glycan structures [2,3,214,217]. Similar glycan compositions were observed in Env purified by GNL and PGT145; however, within our GNL-purified Env, a minor population of glycans exhibiting elevated trimming, $Man_{5\text{-}9}GlcNAc_2$, was identified (**Figures 5.3.A and 5.3.B**). This is consistent with the presence of a non-native fraction in the GNL-purified Env which is eliminated when purified with a trimer-specific antibody. The N332 site presents less processed oligomannose-type glycans in both GNL and PGT145 purified Env which is consistent with previous findings [217,335]. We observed higher glycan processing at other IMP sites, N339, N363, N386, and N448 on GNL purified Env (**Figure 5.3.A and 5.3.B**). The glycosylation composition of all the TAMP sites of Env produced by PGT145 purification could not be obtained. However, in GNL

purified material native-like glycosylation was observed at the TAMP sites, thereby demonstrating native-like glycan signatures independent of a quaternary architecture-specific purification stage (**Figure 5.2**).

To investigate the alterations in glycosylation across all N-glycan sites on the Env protein, the percentage point differences in $Man_9GlcNAc_2$ presentation by Env purified via PGT145 and GNL were plotted. (**Figure 5.3.C**). $Man_9GlcNAc_2$ abundance was selected as a reporting parameter to unveil distinctions in glycan processing. A prevalence of less processed $Man_9GlcNAc_2$ glycan content was noted at multiple sites, including N197, N234, and N392, within Env purified by PGT145 in comparison to GNL purification (**Figure 5.3.C**). This suggests that the total material in GNL-purified Env does contain some constituents which allow the material to be more processed and thus reflect changes in structural conformation consistent with the NS-EM observations. These non-native constituents do get eliminated when we purify the protein with PGT145 antibody which binds only to Env which presents a native-like quaternary trimeric structure. Overall, our findings suggest that the elimination of post-production purification in replicon expressed Env exhibits some proportion of Env displaying slightly enhanced glycan processing. This reflects the presence of some non-native constituents which is consistent with the NS-EM results.

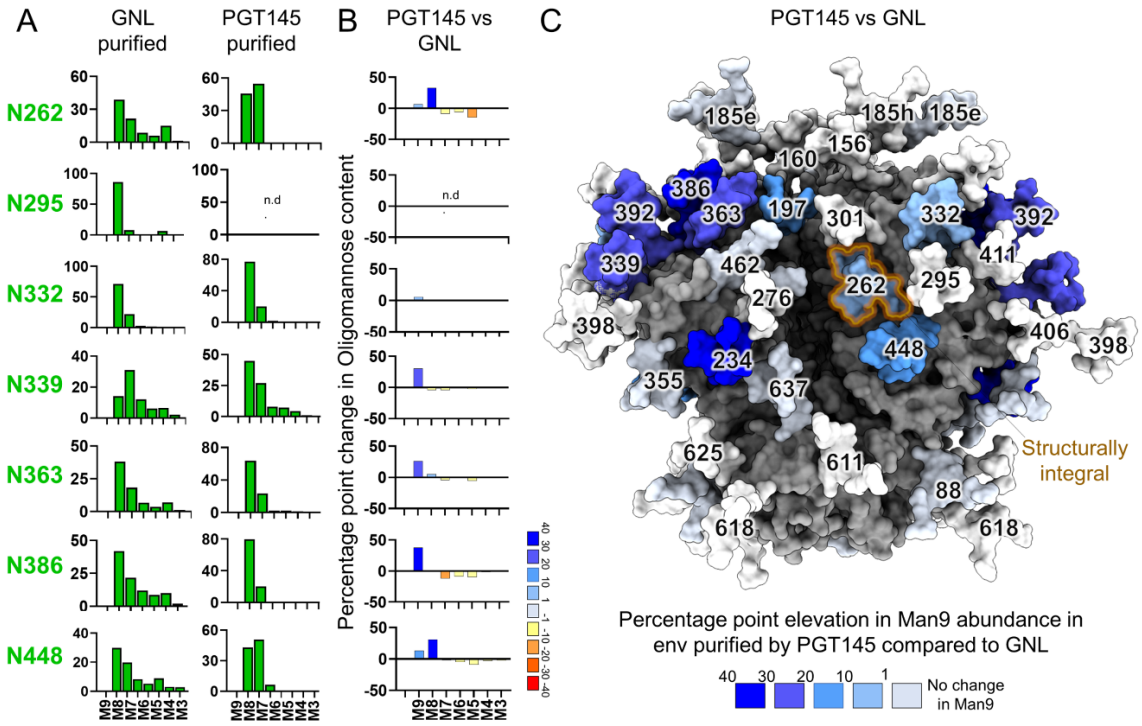


Figure 5.3. Comparison of glycan composition of replicon expressed BG505 NFL.664 Env purified using GNL and PGT145 affinity purification. A) Representation of glycan composition of the intrinsic mannose patch sites on Env purified via GNL and PGT145. Oligomannose-type glycans $Man_3GlcNAc_2$ to $Man_9GlcNAc_2$ are represented as M3 to M9. B) Percentage point change in oligomannose-type glycan content on Env purified via PGT145 versus GNL. The elevation in glycan composition on Env purified by PGT145 compared to GNL is represented in light to dark blue, respectively whereas the decline in glycan composition is shown in yellow to red, respectively. C) A map of the differences in $Man_9GlcNAc_2$ (Man9) abundance between Env purified by PGT145 and GNL affinity purification plotted on a model based on PDB 5C7K. The structural integral glycan site, N262 is shown with orange outline. The sites which were not determined are shown in white; n.d., not determined.

5.5 Cell-directed glycosylation changes in Env protein expressed via replicon

In this section, the glycan composition of the replicon expressed Env was subsequently analyzed in various cell lines, encompassing mouse muscle (C2C12) and dendritic (DC2.4) cell lines. These cell lines were selected due to their potential role as predominant cell types at the site of intramuscular RNA injection in host cells. Firstly, the Env was expressed using different production systems and the expression of protein was analyzed by western blot (**Figure 5.4**). No Env expression was observed in primary dendritic cells, which can be attributed to their inherent difficulty in transfection.

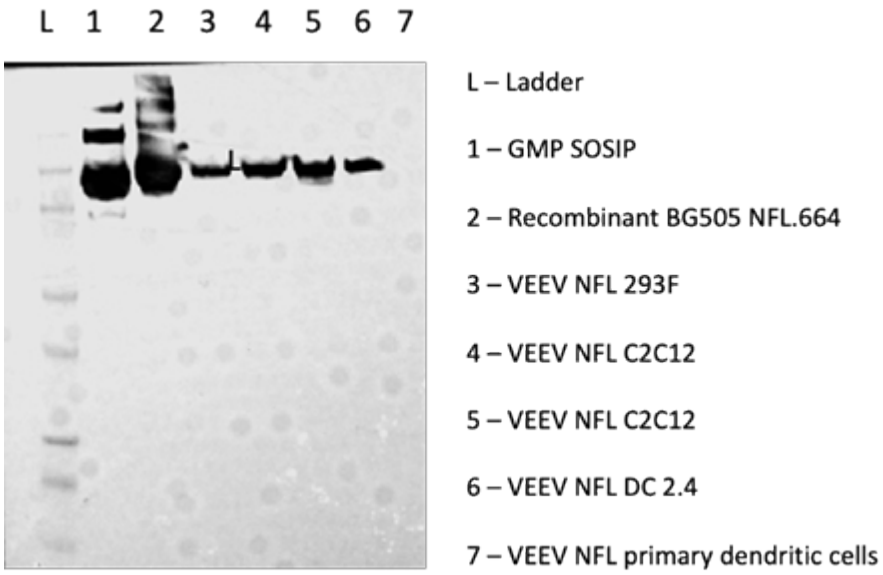


Figure 5.4. Western blot of Env expressed in different production systems. The blot was stained using primary antibody, human VRCO1 and secondary antibody, Goat pAb Hu IgG-HRP. Detection of western blot was carried out using Amersham ECL reagent.

To understand the glycan composition of replicon expression at cell-specific level, Env was expressed in different cell lines, including HEK 293F, DC2.4, and C2C12. Three biological replicates were performed across these different cell lines to generate site-specific glycan compositions (**Figure 5.5**). Highly similar oligomannose-type glycans were observed across these cell lines, characterized by the presence of conserved IMP and TAMP glycosylation sites. These observations are consistent with the previous finding that HIV-1 Env glycosylation is heavily influenced by the quaternary structure of the protein as glycan processing is conserved at many sites despite changes in the cell line used for expression or culture conditions [232]. However, divergent glycan processing of hybrid- and complex-type glycans on replicon expressed Env were detected in different cell lines. For example, Env produced in C2C12 and DC2.4 cells presented elevated levels of processing of complex-type glycans compared to those from Env expressed in HEK 293F cells (**Figure 5.5 and 5.6**). High glycan occupancy and processing at gp41 sites on Env expressed in C2C12 cells were observed (**Figure 5.5B**). Whilst there are differences in glycosylation at a site-specific level between Env derived from different cell lines, there were no substantial differences in Env glycan compositions when comparing material, from the same cell-type, whether expressed via plasmid

DNA or replicon RNA (HEK 293F; **Figure 5.6**). This highlights that Env glycan composition is largely independent of the immunogen delivery platform investigated in this study.

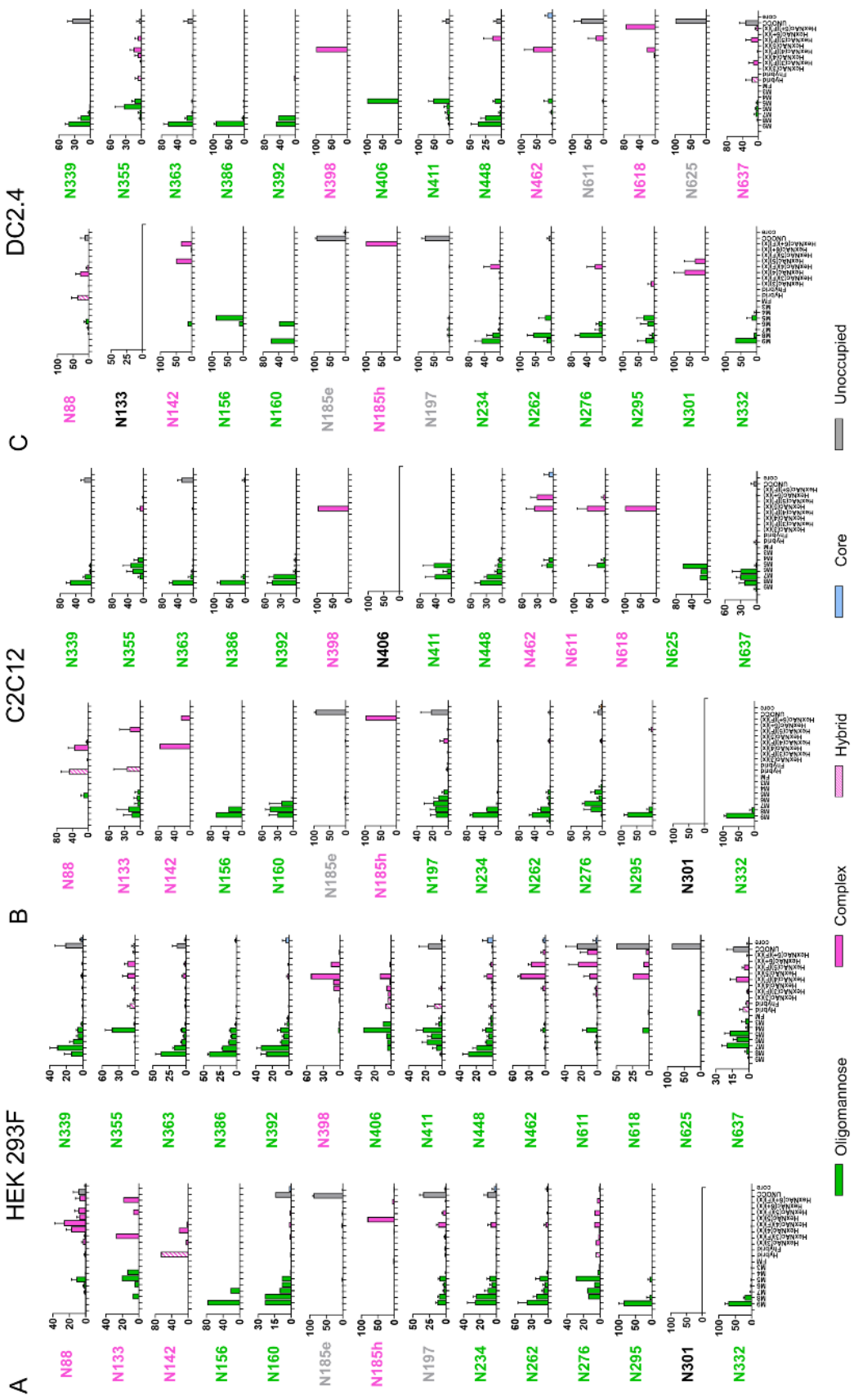


Figure 5.5. Glycan analysis of Env derived from replicon expressed in different cell-types A) HEK 293F, B) C2C12, and C) DC2.4 cells. The bar graph represents the three biological replicates (+/- SEM) of expanded glycan compositions observed across all the PNGS.

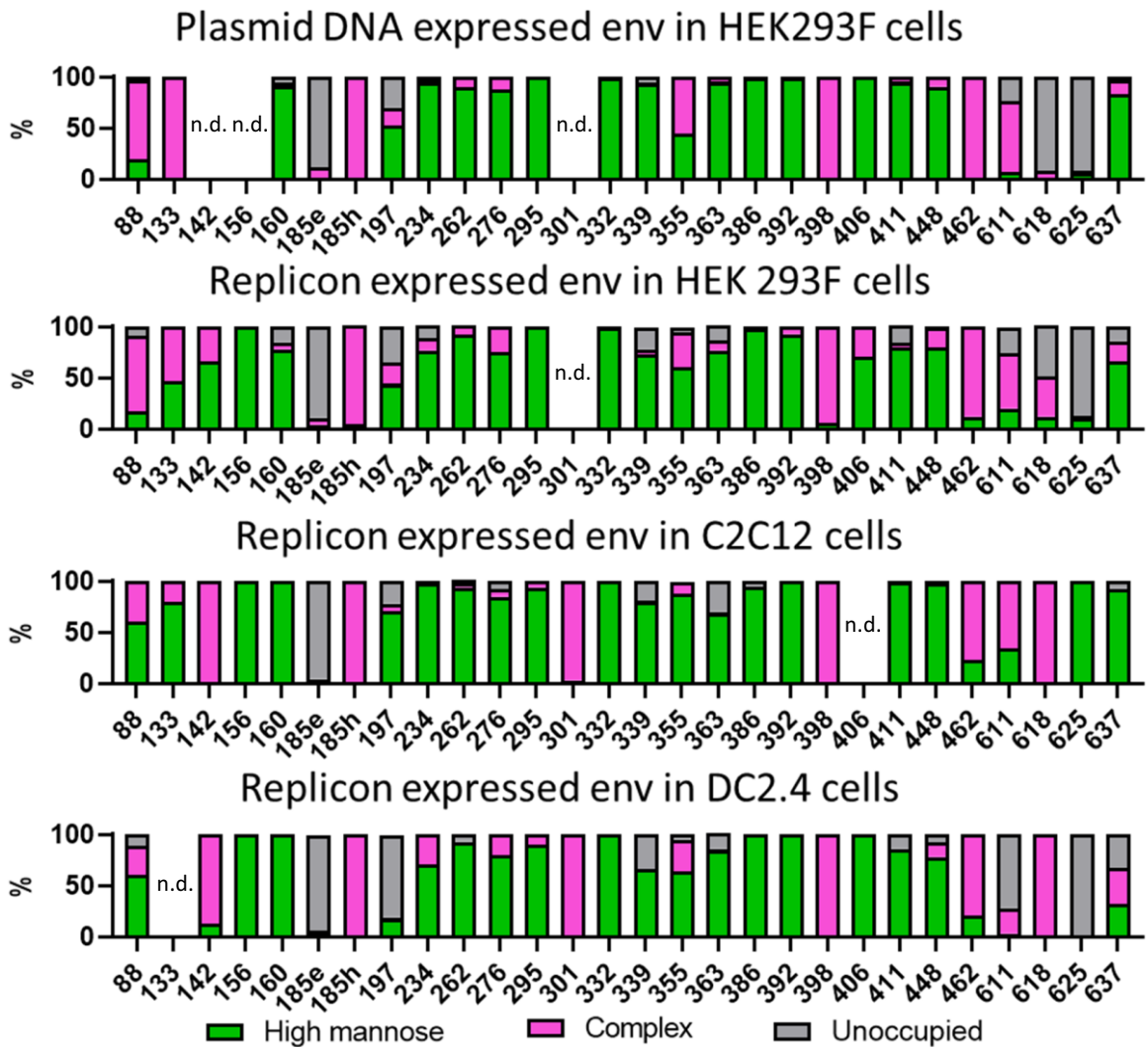


Figure 5.6. Site-specific glycan analysis of Env produced from replicon RNA across different production systems. The bar graphs represent the glycan composition observed at each N-glycan site with the average of three or more biological replications produced in different production systems. The glycan compositions are represented in three categories, underprocessed glycan structures which includes both oligomannose-type and hybrid-type glycans, complex-type which includes highly processed glycan compositions and core-truncated structures, and unoccupied which represents no glycan on N-glycan site. n.d., not determined.

Furthermore, in order to highlight the differences in site-specific glycan processing, a selection was made of one site characterized by a high prevalence of under-processed oligomannose-type glycans (N332), another exhibiting a mixture of complex- and oligomannose-type glycans (N355), and a third displaying predominantly complex-type glycans (N88). At the N332 glycosylation site, a predominance of less processed $Man_{8\text{-}9}GlcNAc_2$ glycan compositions in Env derived from all four expression systems was observed. However, an elevated levels of the more processed $Man_5GlcNAc_2$ glycan was noted in dendritic cell-derived Env, although this glycan constituted only ~10% of the population and $Man_{8\text{-}9}GlcNAc_2$ remained the most abundant glycan species.

The N355 and N88 glycosylation sites presented a diverse mixture of oligomannose-, hybrid-, and complex-type glycans across all expression systems. However, N355 showed a higher abundance of oligomannose-type glycans compared to N88, which presents 80–90% complex-type glycans (**Figure 5.7**). The population of oligomannose-type glycans observed at both these sites are more processed than those of N332, consistent with lower steric exclusion. Furthermore, the population of hybrid- and complex-type glycans observed at these sites are quite diverse. Compared to N355, the N88 site showed greater amounts of hybrid-type glycans on Env derived from muscle and dendritic cell lines (**Figure 5.7**). The presence of hybrid-glycans is noteworthy as they can form targets for bnAbs [277].

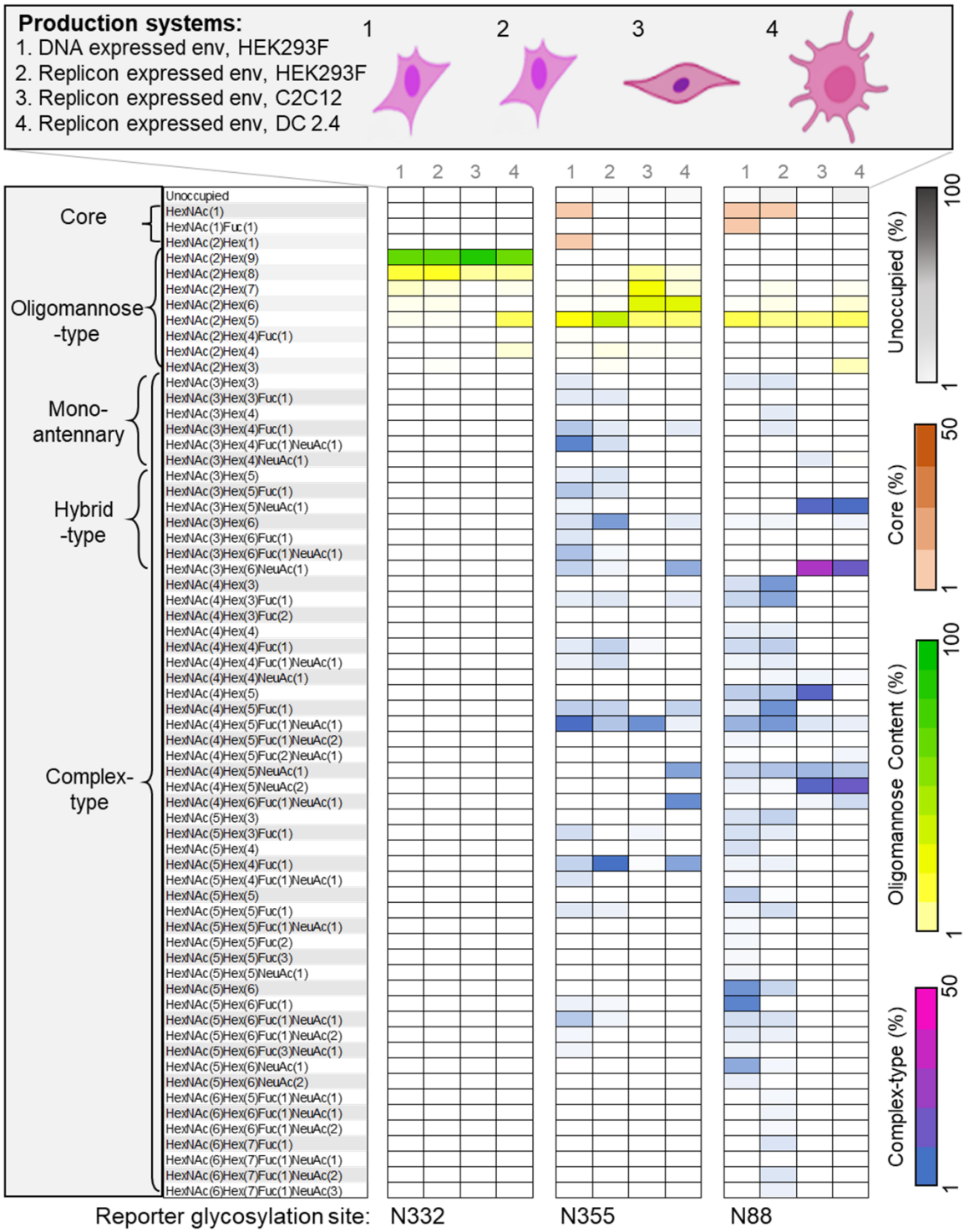


Figure 5.7. Differential glycan composition on Env expressed in different cell lines. Glycan composition at glycan sites N88, N355 and N332. Four different production systems were used as described in the key, including expression in three different cell lines, HEK293F, C2C12 and DC2.4 cells. All the glycan composition observed in the site-specific analysis are shown on the left, simplified in distinct categories represented as, core, oligomannose-type, and complex-type glycans. Some compositions annotated as complex-type glycans can exhibit isomers formally constituting hybrid-type glycans. Represented glycan compositions are colored according to the scale provided in the right for each category of glycan composition.

In contrast to oligomannose-type glycans, the attachment of different monosaccharides within complex-type glycans is more heavily influenced by the host cell [466], culture conditions or supplements provided for protein expression [467]. For example, in contrast to Env derived from C2C12 and DC2.4 cells, at N355 and N88 on material produced in HEK 293F cells, a high percentage of fucosylated species (~40%) and a lower percentage of sialylated species (10-20%) were observed (**Figure 5.8**). These observations are consistent with previous studies of glycoprotein derived from HEK 293F cells which have shown high levels of fucosylation but low sialylation [468,469].

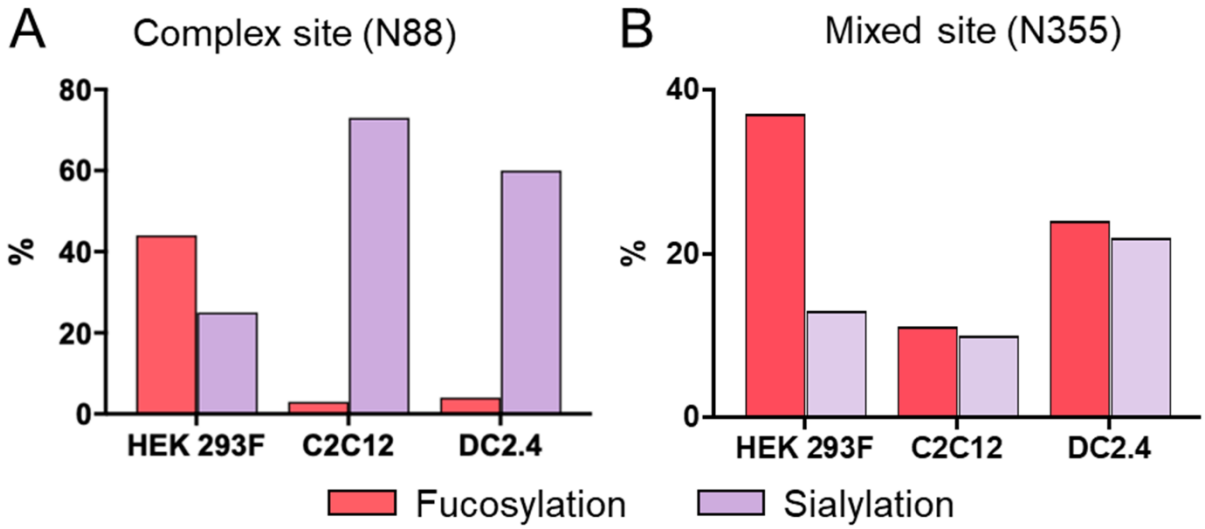


Figure 5.8.**Complex-type glycan abundances in replicon expression of env observed across different cell lines.** The levels of fucosylation and sialylation observed on env expressed in HEK 293F, C2C12 and DC2.4 cells of the two N-linked glycan sites A) Complex site, N88 and B) Mixed site, N355, are displayed. The bar graphs represent the percentage of fucosylation, and sialylation represented in pink and purple, respectively.

Overall, these findings are consistent with earlier observations that cell origin can impact glycosylation. For example, high sialylation is prevalent on Env derived from peripheral blood mononuclear cells (PBMCs) [210]. These features of glycan heterogeneity are important to note as they can influence the function of proteins, such as in immune modulation [466,470]. Whilst the glycosylation of complex and mixed sites on Env is heavily influenced by producer cells, the occupancy at all glycan sites tends to be more conserved, although we do detect some variations (**Figure 5.6**). However, these findings underscore the influence of the producer cells on Env glycosylation, it will be important to extend these observations into *in vivo* settings [215,275,277].

5.6 Targeted repair of glycan occupancy

RNA-based immunogens eliminate the possibility of post-expression purification of the desired product and thus improvements in immunogen assembly can only be achieved at the nucleotide level. In this study, various features of replicon expressed Env glycosylation were observed, which could elicit non-neutralizing antibody. One of the features is glycan underoccupancy which is an important parameter to consider in the case of HIV-1 immunogen design. In site-specific analysis data, there were a few N-glycan sites which presented low occupancy, N185e, N197, N611, N618, and N625 (**Figures 5.2.A and 5.2.B**). To improve the glycan occupancy, an approach previously applied by Derking et al. [268] on recombinant HIV-1 Env to augment the glycan occupancy was employed. The rationale behind this strategy is that the enzyme involved in glycan attachment from the dolichol phosphate to the asparagine site, oligosaccharyltransferase (OST) [471,472] has a higher affinity towards NXT than NXS glycan sequon [268,269]. The same strategy was sought to be applied to replicon-expressed Env, and three sites presenting low occupancy and featuring serine at the third position of the N-glycan sequon (N185e, N197, and N611) were selected for targeted enhancement of glycan occupancy (**Table 5.2**).

Table 5.2. The table representing PNGS across the Env amino acid sequence of BG505 NFL.664
The N-linked glycosylation usually occurs on NXS, or NXT is shown on the right side of the table present on each N-glyan site. Hxb2 sequence number is used as a reference.

Hxb2	PNGS	Aminoacid sequon
58	N88	NXT
103	N133	NXT
107	N142	NXT
118	N156	NXS
122	N160	NXT
152	N185e	NXS to NXT
155	N185h	NXS
167	N197	NXS to NXT
204	N234	NXT
232	N262	NXS
246	N276	NXT
265	N295	NXT
271	N301	NXT
301	N332	NXS
308	N339	NXT
324	N355	NXT
332	N363	NXS
355	N386	NXS
361	N392	NXT
367	N398	NXS
374	N406	NXT
379	N411	NXS
416	N448	NXT
430	N462	NXT
585	N611	NXS to NXT
592	N618	NXS
599	N625	NXT
611	N637	NXT

The N-linked glycosylation usually occurs on NXS, or NXT is shown on the right side of the table present on each N-glyan site. Hxb2 sequence number is used as a reference.

Sequential mutagenesis was carried out at each site, and all three site-specific mutations were introduced into a single plasmid for the transcription of RNA, resulting in the expression of modified

Env, herein referred to as the 'mutant'. The RNA construct with the original, unaltered gene sequence for the BG505 NFL.664 Env is designated here as the 'wild-type' (WT). In this section, site-specific glycan analysis was performed on the mutated Env and compared it with the WT Env to determine any improvement in glycan occupancy. The N185e site contains mostly unoccupied glycans (~90%) on WT Env. However, when serine adjacent to N-linked glycan site was substituted with threonine, the glycan occupancy was restored to 50% (**Figures 5.9A. and 5.9.B**). Similarly, the N197 and N611 sites were less occupied on WT Env but when altered, are almost fully occupied (**Figures 5.9.A and 5.9.B**). Subsequently, a comparison was made between the percentage point changes in glycan composition of the wild-type (WT) and the mutant variants, revealing a notable increase in glycan occupancy at all three glycan sites in comparison to the WT Env. (**Figures 5.9.C and 5.9.D**). Furthermore, to explore the effect of these mutations on all the PNGS including these edited sites, was investigated by mapping the percentage point changes in glycan occupancy observed on WT Env compared to mutant Env (**Figure 5.9.D**). It was observed that N625, N637, and N234, which are situated in proximity to the modified site N611, exhibited greater occupancy on the mutated Env compared to WT Env (**Figure 5.9.D**). However, the N618, which is adjacent to the N625 site, decreased in occupancy on mutated Env are more occupied compared to WT Env that display low glycan occupancy. Overall, there is a significant increase in glycan occupancy suggesting nucleotide editing can be used to improve RNA-based immunogens. It is anticipated that this strategy will prove useful in the development of lead immunogens by helping reduce off-target antibody responses.

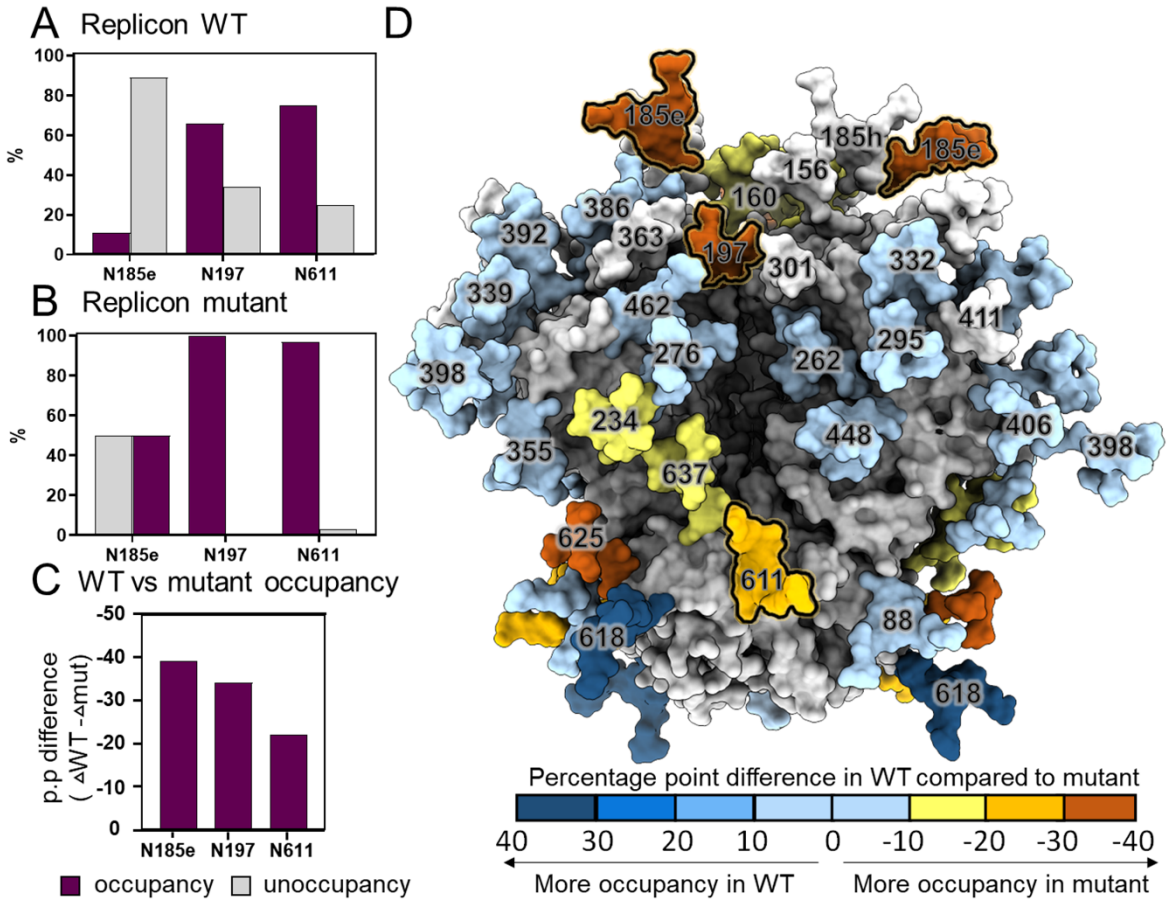


Figure 5.9. Changes in glycan occupancy by nucleotide editing. A) Bar graph representing glycan occupancy on wildtype Env, B) mutant Env with NXS to NXT mutation and C) wildtype versus mutant Env. The three sites, N185e, N197, N611 which were NXS on wildtype have been edited to NXT. The percentage glycan occupancy is shown in violet and unoccupancy at glycan sites in grey. D) A model of HIV-1 Env based upon PDB ID 5C7K, mapped with the percentage point differences in glycan occupancy in WT Env compared to mutant Env. Not determined glycan sites are shown in white.

5.7 Discussion

This study is aimed to understand the assembly and glycan composition of the single-chain soluble trimeric HIV-1 Env derived from RNA replicon. Since, this approach eliminates the possibility of control at post-expression level in in vivo settings compared to traditional protein-based vaccines, we analyzed the entire population of Env expressed by RNA replicon captured by broad-acting lectin, *Galanthus nivalis* lectin (GNL) beads. From NS-EM studies, this study revealed that the majority of the Env expressed by replicon were trimeric but there was a slight proportion of Env products constituting non-native material. Furthermore, we determine the site-specific glycan compositions across all the PNGS to investigate the native-like signatures typically observed on native-like viral Env.

One such feature is conservation of oligomannose-type glycans at IMP and TAMP regions. These regions were observed to be conserved, displaying a substantial prevalence of oligomannose-type glycans on Env derived from the replicon. These regions serve as an epitope for many broadly neutralizing antibodies and additional insights regarding the conservation of these epitopes on Env expressed via replicon could be gleaned through ELISA and FACS assay. While this aspect was not explored in my thesis, it holds relevance for future work. Site-specific glycan analysis also revealed similar level of complex-type glycans across Env produced by replicon and plasmid DNA. However, there were few sites presenting glycan underoccupancy, but similar underoccupancy has been observed previously within recombinant native-like trimers [3]. Overall, this study reveals the native-like glycan characteristics on Env derived from replicon RNA. However, some proportion observed in NS-EM suggests the presence of non-native like constituents, so to explore that the glycan compositions of Env purified by GNL was compared with an affinity purified Env using quaternary epitope specific antibody, PGT145. To detect the differences in glycan compositions, the glycan processing of oligomannose-type glycans across IMP sites was considered, as these sites are under less structural constraints. Glycan analysis revealed that the GNL purified Env presented more processed oligomannose-glycan structures compared to Env purified by PGT145. This suggests that there is a presence of some material in GNL purified Env which allow the material to be more processed and thus reflect changes in the structural conformations consistent with NS-EM. However,

this potentially suggests that the presence of non-native material is likely due to the absence of post-expression purification and not because of the replicon-based expression.

Furthermore, the glycosylation of material derived from cell-types that likely produce immunogens near the site of intramuscular RNA injection was explored. This study reveals that replicon-transformed dendritic and muscle cell lines generate immunogens displaying similar oligomannose-type glycan content, whereas sites presenting complex-type glycosylation differed substantially in the levels of glycan processing. Overall, the control of the immunogen assembly by protein engineering is sufficient to drive native-like glycosylation at the majority of glycosylation sites independent of producer cells. Furthermore, there were few sites with low glycan occupancy at several sites on Env expressed by replicon RNA and plasmid DNA. To enhance glycan occupancy, the nucleotide-editing approach previously employed on recombinant HIV-1 immunogens was utilized to engineer the RNA. [268]. The strategy is based on the rationale of oligosaccharyltransferase having more affinity towards NXT than NXS. Threonine was substituted with serine at sites characterized by low occupancy, resulting in an observed elevation of glycan occupancy on Env generated by the edited version of replicon RNA.

In conclusion, the mRNA expressed Env exhibits majorly native trimeric structure and glycan compositions with conserved IMP and TAMP sites and near native-like complex-type glycans. However, there were few N-glycan sites showing low glycan occupancy on replicon expressed Env, which similarly have been observed on recombinantly expressed Env. This study is limited to the analysis of soluble Env, it would be important to look at the glycan analysis of full-length membrane-bound HIV-1 Env, currently favored for nucleotide-based vaccines, is less likely to exhibit glycan holes [2,3]. Additionaly, the immunogenicity and binding assays would be quite relevant to understand the efficacy of these vaccine candidates in eliciting neutralizing antibodies.

Chapter 6 Concluding remarks and perspectives

In this thesis, a quantitative site-specific mass spectrometric methodology was employed to understand the glycan shields of immunogens expressed recombinantly and via RNA delivery systems.

This thesis reveals the influence of proline mutations on glycan composition analysis of recombinant SARS-CoV-2 spike protein (S2P and HexaPro) and compares it with the native viral glycosylation. Overall, the glycan compositions are highly similar, with few variations occurring at the site-specific level. This has been supported by the MD-simulation studies which reveals similar protein conformations between S2P and HexaPro. Furthermore, serological studies of HexaPro and 2P suggests highly similar binding with sera isolated from COVID-19 infected patients. Interestingly, the S protein treated with kifunensine aim to elicits all oligomannose-type glycans, demonstrates similar sera binding as untreated S protein. This underscores that the fine processing in glycan composition within SARS-CoV-2 S are unlikely to substantially impact the immune response when compared to the heavily glycosylated viruses like HIV-1, where numerous glycans serve as an epitope for antibodies.

This thesis compares the glycan processing in various SARS-CoV-2 S protein variants which are designed to stabilize the S protein in prefusion conformation. As this conformation is known to contain the epitopes for neutralizing antibodies and thus holds more potential as a vaccine immunogen. Additionally, these recombinant proteins also serve as a valuable resource for diagnostics and research purposes, therefore different variants have been developed to enhance protein stability and expression. This thesis reveals the glycosylation of SARS-CoV-2 S2P, HexaPro, and Closed-S, differes in RBD configurations. Overall, these variants present highly similar glycan composition, with slight variations at site-specific level. Notably, the N234 site which is inherent to the structural property of the S protein was conserved across all these variants and on native virion. However, the complex-type glycan processing, including sialic acids addition, differs between native virions and recombinant S proteins. Despite these changes in complex-type glycan processing, the vaccine is able to provide the protection. One such example is of Novavax vaccine candidate

produced in insect cells, which displays distinct complex-type processing compared to mammalian-cell produced S protein, has demonstrated efficacy of 85% for the B.1.1.7 SARS-CoV-2 variant [117,473].

The development of SARS-CoV-2 vaccines has demonstrated the efficacy of RNA-based platform for the rapid design and testing of new vaccines. However, one potential limitation of RNA-based vaccines is that the viral immunogens must be entirely encoded by the nucleotide sequence which precludes the opportunity for post-expression purification commonly employed in the production of recombinant-based immunogens. Therefore, it is important to understand how the protein is assembled *in situ* and is particularly important in case of HIV-1 vaccine, where immunogen assembly and conservation of glycan epitopes are important to elicit broadly neutralizing antibody response and thus required to mimic native-like assembly and post-translational modifications.

This thesis demonstrates the assembly and glycan composition of single-chain soluble trimeric HIV-1 Env derived from RNA replicons. Native-like glycan processing and assembly is demonstrated in majority of the resulting material, as well as cell-specific glycosylation of muscle and dendritic cells that predominate at the site of RNA injection. Replicon-transformed dendritic and muscle cell lines generate immunogens displaying similar oligomannose-type glycan content, whereas sites presenting complex-type glycosylation differed substantially in the levels of glycan processing. This is consistent with previous studies, as the composition of complex-type glycans is dependent on the repertoire of glycosyltransferases of the producing cells [23]. Overall, the control of the immunogen assembly by protein engineering is sufficient to drive native-like glycosylation at the majority of glycosylation sites independent of producer cells. Furthermore, this study shows that the nucleotide editing can enhance the glycan occupancy of RNA replicon-derived immunogens. Controlling immunogen assembly at the nucleotide level offers a route to enhanced RNA-based immunogens.

While glycan occupancy and native-like glycosylation are an important aspect to consider in vaccine design, merely mimicking of native-like viral spike may not be sufficient. For instance, highly engineered immunogen, eOD-gt8 60-mer, germline-targeting strategy has revealed the elicitation of both B and T cell responses against the HIV-1 immunogen in both clinical and pre-

clinical studies [288,289]. However, this approach also involves the immunization with a final 'polishing immunogen' displaying native-like features for elicitation of matured antibodies. Therefore, it is important to explore strategies to stabilize the Env structure and filling the artificial glycan holes. Furthermore, in this thesis, I have analyzed the glycan analysis of soluble glycoproteins, it will be important to extend these observations to the membrane-associated Env which would potentially be more suitable vaccine candidates for RNA-based platforms. RNA encoding transmembrane proteins and virus-like particles (VLPs) are currently being utilized in preclinical studies. The expression and antigenicity data of VLPs are described in Appendix C, which is part of future work. In addition, cell-specific glycosylation experiments are carried out in in vitro settings, it will be important to know about the glycosylation compositions of vaccine candidates expressed in in vivo settings.

Overall, this thesis highlights the pivotal interplay between protein architecture, production routes, and glycan processing which will illuminate how viral immunogens work and aid in their optimization.

Appendix A Molecular dynamics simulation data of HexaPro and 2P S

A.1 Comparison of ASA values between two-RBD-up HexaPro and 2P S protein

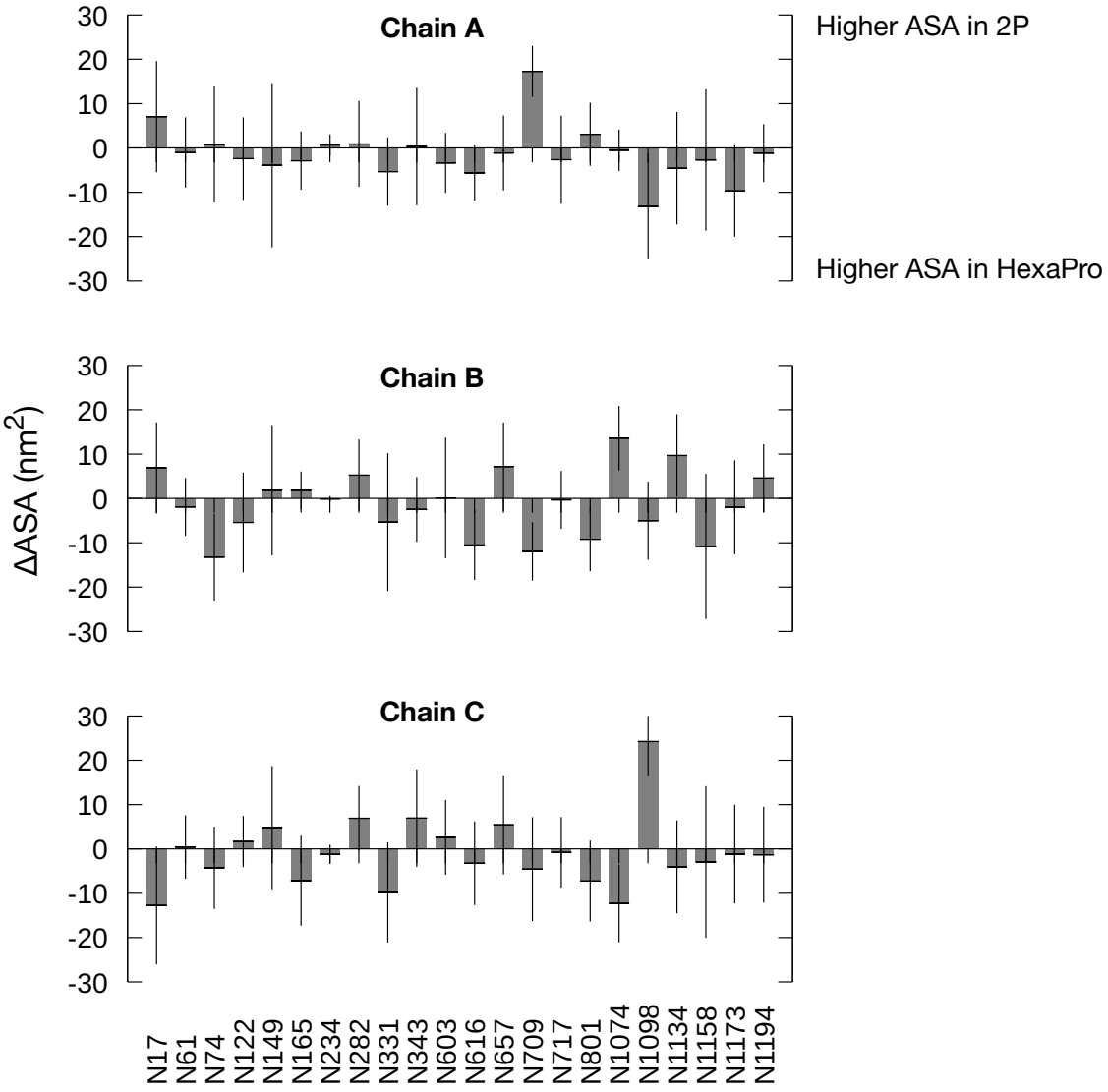


Figure A.1. ΔASA is calculated as the arithmetic difference between the ASA values (ASA_{2P} — $ASA_{HexaPro}$) averaged over the last 50 ns of the simulations and across three replicate simulations. A positive value thus represents a lower accessibility in HexaPro, therefore a potential increase in oligomannose-type glycans, and vice versa. The error bars show standard deviations along the trajectories and across repeat simulation. The simulations were performed by the laboratory of Prof. Bond (A*STAR Bioinformatics Institute, Singapore).

A.2 Comparison of glycan N165 from simulations of two-RBD-up HexaPro and 2P S protein variants

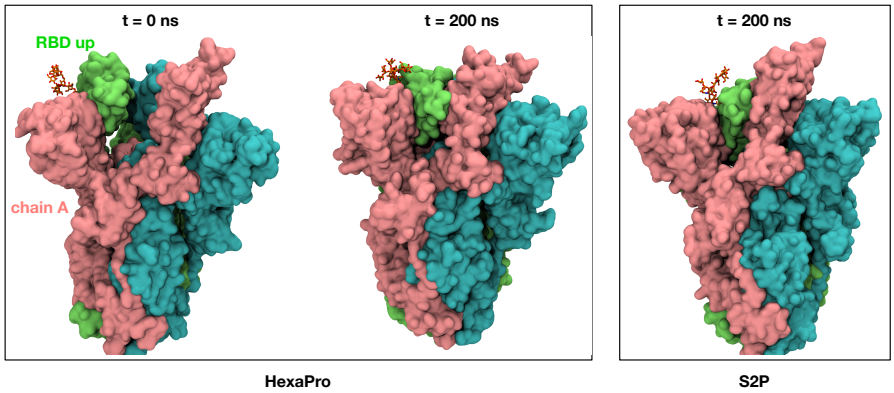


Figure A.2. Snapshot at the beginning (left) and at the end (middle) of the HexaPro simulation, highlighting N165 glycan from chain A. Snapshot at the end of the 2P simulation (right) showing the same glycans. Protein is shown in surface representation and coloured pink (chain A), cyan (chain B), or green (chain C), whilst the glycan is shown in stick representation and coloured in orange. The simulations were performed by the laboratory of Prof. Bond (A*STAR Bioinformatics Institute, Singapore).

A.3 Comparison of protein dynamics between HexaPro and 2P

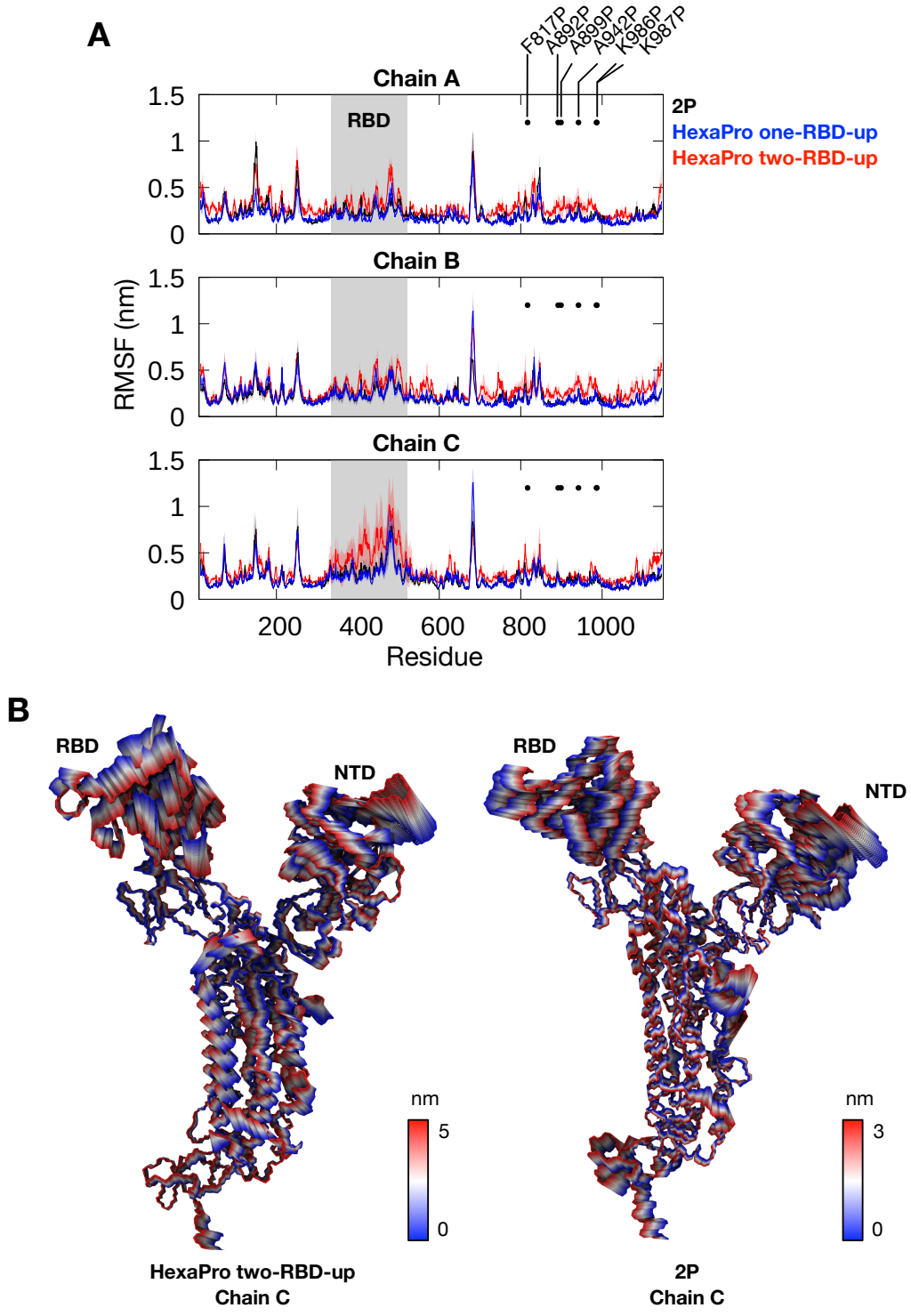


Figure A.3. (A) Root mean square fluctuation (RMSF) of each residue in the ectodomain (ECD) of HexaPro and 2P from 200 ns simulations. The position of the RBD is highlighted in grey and the positions of six proline mutations in HexaPro are marked with black circles and labelled. Thick lines represent average values from three repeat simulations and the shaded areas indicate standard deviation. (B) The first principal motion of all backbone atoms on the ECD of chain C of HexaPro two-RBD-up (left) and 2P (right) as determined by principal component analysis. The simulations were performed by the laboratory of Prof. Bond (A*STAR Bioinformatics Institute, Singapore).

A.4 Comparison of ASA values from replicate simulations

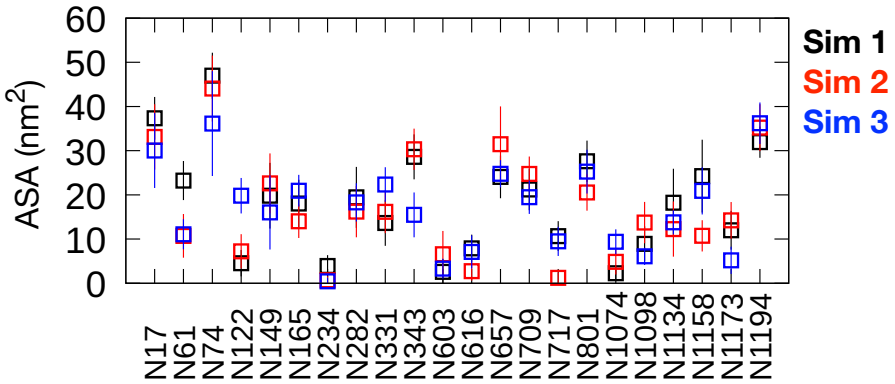


Figure A.4. Example ASA values calculated from three independent 200 ns simulations. Data taken from the last 50 ns of HexaPro two-RBD-up simulations from S protein chain A. Average values are shown with standard deviations along the trajectories depicted as error bars. The simulations were performed by the laboratory of Prof. Bond (A*STAR Bioinformatics Institute, Singapore).

Appendix B Cell-specific glycan composition of Env derived from RNA replicon

B.1 Glycan composition analysis of replicon expressed Env in C2C12 cells

	N88	N133	N142	N156	N160	N185e	N185h	N197	N234	N262	N276	N295	N301	N332	N339	N355	N363	N386	N392	N398	N406	N411	N448	N462	N611	N618	N625	N637	Total
M9Glc	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
M9	0	16	0	66	14	0	0	15	66	48	2	81	0	90	54	0	54	81	46	0		0	42	0	0	0	0	3	25
M8	0	22	0	34	38	0	0	17	29	25	22	13	0	10	17	10	8	9	44	0		42	30	0	0	0	19	23	15
M7	0	5	0	0	45	1	0	19	2	3	34	1	0	0	3	28	3	2	3	0		11	11	1	1	0	17	31	8
M6	0	6	0	0	3	2	0	13	0	5	8	0	0	0	4	34	1	1	2	0		45	9	13	27	0	63	30	10
M5	12	5	0	0	0	0	0	6	1	7	15	0	3	0	1	14	1	2	3	0		1	6	9	7	0	0	2	4
M4	0	0	0	0	0	0	0	0	0	1	4	0	0	0	1	1	0	0	2	0		0	0	0	0	0	0	0	0
M3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0		0	1	0	0	0	0	0	0
FM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
Hybrid	48	25	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0		0	0	0	0	0	0	3	3
Fhybrid	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
HexNAc(3)(x)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
HexNAc(3)(F)(x)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
HexNAc(4)(x)	36	0	77	0	0	0	1	0	0	0	0	0	97	0	0	0	0	0	0	0		0	0	0	1	1	0	0	8
HexNAc(4)(F)(x)	3	0	0	0	0	1	0	6	2	3	3	0	0	0	0	9	2	0	0	99		0	1	37	59	99	0	0	12
HexNAc(5)(x)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
HexNAc(5)(F)(x)	0	20	0	0	0	0	0	0	0	0	0	6	0	0	0	1	0	0	0	1		0	0	31	6	0	0	0	2
HexNAc(6+)(x)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
HexNAc(6+)(F)(x)	0	0	23	0	0	0	99	0	0	0	0	0	0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	5
UNOCC	0	0	0	0	0	96	0	22	0	3	8	0	0	0	19	0	30	5	0	0		0	0	0	0	0	0	7	7
core	0	0	0	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	0	0		0	0	9	0	0	0	0	1

%	N88	N133	N142	N156	N160	N185e	N185h	N197	N234	N262	N276	N295	N301	N332	N339	N355	N363	N386	N392	N398	N406	N411	N448	N462	N611	N618	N625	N637	Total
Mannose	12	55	0	100	100	3	0	69	98	94	85	94	3	100	80	87	67	95	100	0	0	99	98	23	34	0	100	90	60
Hybrid	48	25	0	0	0	0	0	2	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	3	3
Complex	40	20	100	0	0	1	100	7	2	4	3	6	97	0	0	11	2	0	0	100	0	0	1	68	66	100	0	0	26
Unoccupied	0	0	0	0	0	96	0	22	0	3	8	0	0	0	19	0	30	5	0	0	0	1	1	0	0	0	0	7	7
Fucosylation	3	20	23	0	0	1	99	7	2	4	3	6	0	0	0	11	2	0	0	100	0	0	2	68	65	99	0	0	18
Sialylation	73	45	23	0	0	1	0	2	0	4	3	0	97	0	0	10	1	0	0	100	0	0	0	8	62	99	0	0	19
core	0	0	0	0	0	0	0	0	0	0	4	0	0	0	1	0	0	0	0	0	0	0	0	0	9	0	0	0	0

B.2 Glycan composition analysis of replicon expressed Env in DC2.4 cells

	N88	N133	N142	N156	N160	N185e	N185h	N197	N234	N262	N276	N295	N301	N332	N339	N355	N363	N386	N392	N398	N406	N411	N448	N462	N611	N618	N625	N637
M9Glc	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M9	0		0	0	61	0	0	2	47	15	0	26	0	69	41	0	64	92	51	0	0	3	38	2	0	0	0	0
M8	0		0	0	0	0	0	4	20	57	59	7	0	10	18	3	16	5	44	0	0	7	26	0	0	0	0	2
M7	1		0	0	0	1	0	4	2	1	9	0	0	2	3	4	2	1	0	0	0	10	3	5	0	0	0	6
M6	4		13	13	39	0	0	2	1	0	11	22	0	0	1	34	0	1	0	0	0	11	1	0	0	0	0	5
M5	15		0	87	0	1	0	4	1	20	0	35	0	16	1	14	2	1	1	0	100	53	10	14	2	0	0	4
M4	0		0	0	0	0	0	0	0	0	0	0	0	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0
M3	7		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FM	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hybrid	32		0	0	0	0	0	1	0	0	1	0	0	0	0	8	0	0	4	0	0	2	0	0	0	0	0	16
Fhybrid	0	N.D.	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
HexNAc(3)(x)	0		0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HexNAc(3)(F)(x)	0		0	0	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0	0	0	0	12
HexNAc(4)(x)	23		0	0	0	0	0	0	2	0	0	0	67	0	0	7	1	0	0	0	0	0	0	0	0	4	0	0
HexNAc(4)(F)(x)	5		0	0	0	0	0	2	26	0	20	0	0	0	0	15	0	0	0	99	0	0	0	62	0	21	0	1
HexNAc(5)(x)	0		50	0	0	0	0	0	1	0	0	0	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HexNAc(5)(F)(x)	0		0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	14	0	25	0	0	18
HexNAc(6+)(x)	0		3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HexNAc(6+)(F)(x)	0		34	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	75	0	4
UNOCC	11		0	0	0	94	0	80	0	7	0	0	0	0	0	34	5	15	0	0	0	0	14	8	0	72	0	100	33
core	0		0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	0	0

Total	
M9Glc	0
M9	19
M8	10
M7	2
M6	6
M5	14
M4	0
M3	0
FM	0
Hybrid	2
Fhybrid	0
HexNAc(3)(x)	0
HexNAc(3)(F)(x)	1
HexNAc(4)(x)	4
HexNAc(4)(F)(x)	9
HexNAc(5)(x)	3
HexNAc(5)(F)(x)	2
HexNAc(6+)(x)	0
HexNAc(6+)(F)(x)	8
Unoccupied	18
core	1

%	N88	N133	N142	N156	N160	N185e	N185h	N197	N234	N262	N276	N295	N301	N332	N339	N355	N363	N386	N392	N398	N406	N411	N448	N462	N611	N618	N625	N637
Mannose	28		13	100	100	2	0	17	71	93	79	90	0	100	65	56	85	100	96	0	100	84	77	21	3	0	0	16
Hybrid	32		0	0	0	0	0	1	0	0	1	0	0	0	0	8	0	0	4	0	0	2	0	0	0	0	0	16
Complex	29		87	0	0	0	100	2	29	0	20	10	100	0	0	31	1	0	0	100	0	0	14	62	25	100	0	35
Unoccupied	11	N.D.	0	0	0	93	0	80	0	7	0	0	0	0	34	5	15	0	0	0	0	14	8	0	72	0	100	33
Fucosylation	5		34	0	0	0	100	2	26	0	20	0	0	0	0	25	0	0	0	100	0	0	14	62	25	96	0	35
Sialylation	60		37	0	0	0	0	0	15	0	11	0	66	0	0	22	1	0	4	50	0	2	0	19	25	0	0	34
core	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0

Total	
Mannose	52
Hybrid	2
Complex	28
Unoccupied	17
Fucosylation	20
Sialylation	13
core	0

B.3 Glycan composition analysis of transiently expressed Env via plasmid DNA in HEK 293F cells

	N88	N133	N142	N156	N160	N185e	N185h	N197	N234	N262	N276	N295	N301	N332	N339	N355	N363	N386	N392	N398	N406	N411	N448	N462	N611	N618	N625	N637
M9Glc	0	0			0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M9	0	0			30	0	0	16	45	36	0	41		71	42	0	54	47	25	0	0	0	27	0	0	0	0	0
M8	0	0			17	0	0	11	25	32	16	17		19	27	0	11	19	43	0	0	4	25	0	0	0	0	6
M7	0	0			7	0	0	5	13	7	15	10		5	10	1	9	9	9	0	0	10	10	0	0	0	0	19
M6	1	0			3	0	0	5	5	9	8	23		2	5	1	6	8	8	0	0	13	5	0	0	0	0	18
M5	17	0			30	0	0	16	5	5	25	9		2	6	24	12	13	12	0	100	57	16	1	6	0	0	25
M4	0	0			2	0	0	0	0	1	9	0		0	2	1	1	1	1	0	0	1	3	0	0	0	0	2
M3	0	0			0	0	0	0	0	0	1	0		0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
FM	0	0			0	0	0	0	0	0	0	0		0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Hybrid	1	0			1	0	0	0	0	0	14	0		0	0	7	2	1	0	0	0	8	3	0	0	0	5	11
Fhybrid	0	0			0	0	0	0	0	0	0	0		0	0	10	1	0	0	1	0	1	1	0	0	0	0	2
HexNAc(3)(x)	1	0			0	0	0	0	0	0	1	0		0	0	1	0	0	0	0	0	2	0	0	0	0	1	0
HexNAc(3)(F)(x)	1	37			0	1	0	3	0	0	0	0		0	0	14	2	0	0	6	0	1	1	1	0	0	0	1
HexNAc(4)(x)	11	0			0	0	100	0	0	0	2	0		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HexNAc(4)(F)(x)	15	63			3	3	0	10	2	6	5	0		0	1	20	1	0	0	28	0	1	6	45	4	6	0	7
HexNAc(5)(x)	24	0			0	0	0	0	0	0	1	0		0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
HexNAc(5)(F)(x)	20	0			0	5	0	3	0	3	1	0		0	0	16	1	0	0	37	0	0	2	42	33	2	0	2
HexNAc(6+)(x)	1	0			0	0	0	0	0	0	0	0		0	0	0	0	0	0	29	0	0	0	0	0	0	0	0
HexNAc(6+)(F)(x)	1	0			0	1	0	0	0	0	2	0		0	0	2	0	0	0	0	0	0	0	9	31	0	0	2
UNOCC	3	0			5	88	0	30	2	0	0	0		0	5	0	0	0	0	0	0	0	0	0	24	92	92	4
core	6	0			1	0	0	1	1	1	1	0		0	0	2	0	0	0	0	0	1	0	2	1	0	0	1

Total	
M9Glc	0
M9	17
M8	11
M7	6
M6	5
M5	15
M4	1
M3	0
FM	0
Hybrid	2
Fhybrid	1
HexNAc(3)(x)	0
HexNAc(3)(F)(x)	3
HexNAc(4)(x)	5
HexNAc(4)(F)(x)	9
HexNAc(5)(x)	1
HexNAc(5)(F)(x)	7
HexNAc(6+)(x)	1
HexNAc(6+)(F)(x)	2
UNOCC	14
core	1

%	N88	N133	N142	N156	N160	N185e	N185h	N197	N234	N262	N276	N295	N301	N332	N339	N355	N363	N386	N392	N398	N406	N411	N448	N462	N611	N618	N625	N637
Mannose	18	0			90	1	0	53	94	89	74	100		99	93	27	93	98	99	0	100	86	87	1	6	0	0	70
Hybrid	1	0			1	0	0	0	0	0	14	0		0	0	18	2	1	0	1	0	9	4	0	0	0	5	13
Complex	72	100			3	10	100	17	3	10	11	0		1	1	54	4	1	0	99	0	4	9	97	69	8	3	13
Unoccupied	3	0			5	88	0	30	2	0	0	0		0	5	0	0	0	0	0	0	0	0	0	24	92	92	4
Fucosylation	36	100			3	10	0	17	3	10	8	0		1	1	63	4	1	0	71	0	3	10	97	68	8	1	14
Sialylation	26	25			2	5	0	1	1	5	12	0		0	0	40	0	0	0	56	0	5	4	57	53	0	4	6
core	6	0			1	0	0	1	1	1	1	0		0	0	2	0	0	0	0	0	1	0	2	1	0	0	1

Total	
Mannose	49
Hybrid	3
Complex	25
Unoccupied	12
Fucosylation	19
Sialylation	11
core	1

B.4 Expanded glycan composition analysis of Env produced by replicons in HEK 293F cells

Average	88	133	142	156	160	185e	185h	197	234	262	276	295	301	332	339	355	363	386	392	398	406	411	448	462	611	618	625	637
HexNAc(1)	1%	0%	0%	0%	2%	1%	0%	0%	3%	0%	0%	0%	0%	0%	1%	0%	1%	2%	0%	0%	0%	1%	1%	1%	3%	0%	0%	1%
HexNAc(1)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	1%	0%	0%	0%	0%
HexNAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(1)	0%	0%	0%	0%	0%	0%	3%	1%	0%	2%	1%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	1%	3%	0%	0%	0%	0%
HexNAc(2)Hex(1)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(2)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	7%	0%
HexNAc(2)Hex(3)	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	1%	0%	1%	1%	0%	0%	2%	3%	0%	0%	0%	0%	3%
HexNAc(2)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	2%	0%
HexNAc(2)Hex(3)Fuc(2)	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(4)	0%	14%	0%	0%	1%	0%	0%	1%	0%	2%	3%	0%	0%	0%	2%	2%	1%	2%	2%	1%	10%	4%	3%	2%	1%	0%	0%	2%
HexNAc(2)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(5)	11%	21%	0%	0%	9%	2%	1%	10%	9%	15%	29%	6%	0%	1%	6%	42%	7%	10%	11%	3%	34%	23%	9%	6%	14%	10%	0%	17%
HexNAc(2)Hex(5)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(6)	2%	6%	0%	0%	9%	0%	0%	4%	5%	6%	7%	0%	0%	2%	6%	1%	3%	9%	7%	0%	6%	13%	5%	1%	1%	0%	0%	11%
HexNAc(2)Hex(7)	2%	0%	0%	22%	10%	0%	0%	3%	11%	9%	15%	0%	0%	3%	12%	1%	7%	12%	10%	0%	5%	19%	8%	1%	2%	0%	0%	21%
HexNAc(2)Hex(8)	0%	7%	0%	0%	24%	0%	0%	10%	25%	22%	14%	8%	0%	22%	31%	0%	18%	22%	34%	0%	5%	7%	20%	2%	1%	0%	0%	3%
HexNAc(2)Hex(9)	0%	0%	0%	78%	24%	0%	0%	13%	26%	39%	0%	86%	0%	72%	14%	0%	38%	42%	28%	0%	1%	1%	30%	0%	0%	0%	0%	1%
HexNAc(2)Hex(10)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	2%	0%	0%	0%
HexNAc(3)Hex(3)Fuc(1)	0%	11%	0%	0%	0%	1%	0%	1%	0%	0%	1%	0%	0%	0%	0%	1%	0%	0%	0%	6%	2%	0%	0%	4%	1%	0%	0%	1%
HexNac(3)Hex(3)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(3)Fuc(2)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)	1%	0%	1%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	1%	0%	1%	0%
HexNAc(3)Hex(4)Fuc(1)	1%	17%	0%	0%	2%	0%	0%	1%	0%	0%	1%	0%	7%	0%	0%	1%	0%	0%	0%	5%	3%	2%	0%	1%	1%	0%	0%	1%
HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(4)Fuc(1)NeuAc(1)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)NeuAc(1)	0%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)	0%	0%	3%	0%	0%	0%	0%	1%	0%	0%	2%	0%	0%	0%	0%	2%	1%	0%	0%	0%	2%	9%	1%	0%	0%	0%	1%	2%
HexNAc(3)Hex(5)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	2%	1%	0%	0%	1%	2%	0%	0%	0%	1%	0%	0%	2%
HexNac(3)Hex(5)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)NeuAc(1)	0%	0%	28%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)	1%	0%	0%	0%	1%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	7%	0%	1%	0%	0%	3%	0%	2%	0%	0%	0%	0%	3%
HexNAc(3)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)NeuAc(1)	0%	0%	35%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	1%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(6)NeuAc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)	7%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)Fuc(1)	6%	0%	0%	0%	0%	1%	0%	3%	1%	1%	3%	0%	0%	0%	0%	2%	1%	0%	0%	10%	0%	0%	3%	15%	2%	12%	0%	3%
HexNAc(4)Hex(3)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%
HexNac(4)Hex(3)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(3)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)	3%	0%	0%	0%	2%	0%	0%	2%	0%	1%	0%	0%	0%	0%	0%	3%	1%	0%	0%	0%	1%	0%	0%	3%	0%	0%	0%	1%
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	1%	0%	0%	0%	0%	0%	0%	1%	1%	0%	1%	0%	0%	0%	0%	2%	0%	0%	0%	2%	1%	0%	0%	5%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)NeuAc(1)	1%	0%	9%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(1)	8%	0%	0%	0%	0%	0%	0%	2%	3%	1%	0%	0%	0%	0%	0%	3%	1%	0%	2%	0%	5%	0%	3%	4%	2%	9%	1%	3%
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	7%	0%	0%	0%	0%	1%	0%	3%	3%	3%	3%	0%	0%	0%	0%	4%	0%	0%	0%	38%	6%	0%	1%	16%	1%	4%	0%	3%
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	2%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(2)	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(5)Fuc(2)NeuAc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	4%	0%	0%	1%
HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	14%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)NeuAc(1)	4%	0%	14%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	12%	1%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(5)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	77%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(7)	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(7)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)	1%	7%	0%	0%	0%	1%	0%	1%	0%	0%	1%	0%	0%	0%	0%	0%	1%	0%	0%	13%	1%	0%	1%	9%	11%	7%	0%	1%
HexNAc(5)Hex(3)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(5)Hex(3)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)	0%	0%	0%	0%	0%	0%	72%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)	1%	0%	0%	0%	0%	1%	0%	1%	0%	1%	0%	0%	0%	0%	1%	11%	1%	0%	0%	3%	0%	0%	1%	3%	5%	0%	0%	1%
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(5)Hex(4)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	4%	1%	0%	0%	1%
HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)NeuAc(1)	0%	0%	0%	0%	0%	0%	10%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)	0%	0%	0%	0%	2%	0%	1%	0%	0%	0%	1%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	2%	2%	2%	0%	1%
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	2%	0%	1%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	5%	3%	0%	0%	1%
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	2%	1%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(1)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(7)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(7)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(9)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(10)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(11)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	2%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(6)Hex(3)Fuc(3)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
HexNAc(6)Hex(4)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	2%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	1%	1%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%
HexNAc(6)Hex(6)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)	2%	11%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	1%
HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	0%	8%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	5%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(4)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(8)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(9)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	7%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(4)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(7)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(7)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(8)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(8)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(8)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(8)Hex(9)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
UNOCCUPIED	9%	0%	0%	0%	15%	89%	0%	34%	12%	0%	0%	0%	0%	0%	21%	4%	14%	0%	0%	0%	0%	17%	1%	0%	25%	50%	87%	14%

The average abundance of glycan compositions across three biological replicates is shown, with a white to red colour scale used to highlight higher abundances of particular glycan compositions.

B.5 Expanded glycan composition analysis of Env produced by replicons in C2C12 cells

AVERAGE_VEEVNFL_C2C12	88	133	142	156	160	185e	185h	197	234	262	276	295	301	332	339	355	363	386	392	398	406	411	448	462	611	618	625	637
HexNAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	1%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(1)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	9%	0%	0%	0%	0%
HexNAc(2)Hex(1)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(2)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	1%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	4%	0%	0%	0%	1%	1%	0%	0%	2%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(5)	12%	5%	0%	0%	0%	0%	0%	6%	1%	7%	15%	0%	3%	0%	1%	14%	1%	2%	3%	0%		1%	6%	9%	7%	0%	0%	2%
HexNAc(2)Hex(5)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(6)	0%	6%	0%	0%	2%	2%	0%	13%	0%	5%	8%	0%	0%	0%	4%	34%	1%	1%	2%	0%		45%	9%	13%	27%	0%	63%	30%
HexNAc(2)Hex(7)	0%	5%	0%	0%	23%	1%	0%	19%	2%	3%	34%	1%	0%	0%	3%	28%	3%	2%	3%	0%		11%	11%	1%	1%	0%	17%	31%
HexNAc(2)Hex(8)	0%	22%	0%	34%	45%	0%	0%	17%	29%	25%	22%	13%	0%	10%	17%	10%	8%	9%	44%	0%		42%	30%	0%	0%	0%	19%	23%
HexNAc(2)Hex(9)	0%	16%	0%	66%	30%	0%	0%	15%	66%	48%	2%	81%	0%	90%	54%	0%	54%	81%	46%	0%		0%	42%	0%	0%	0%	0%	3%
HexNAc(2)Hex(10)	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(3)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(3)Fuc(2)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(4)Fuc(1)NeuAc(1)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)NeuAc(1)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	1%
HexNAc(3)Hex(5)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(5)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)NeuAc(1)	14%	7%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	2%
HexNAc(3)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)NeuAc(1)	34%	19%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(6)NeuAc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)	0%	0%	77%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(3)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(3)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	1%	0%	0%
HexNAc(4)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%		0%	0%	3%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	99%	0%	0%
HexNAc(4)Hex(4)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)NeuAc(1)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)	15%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	3%	2%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%		0%	1%	13%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	2%	0%	0%	0%	0%	1%	0%	2%	0%	2%	3%	0%	0%	0%	0%	8%	1%	0%	0%	49%		0%	0%	8%	47%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	4%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(5)Fuc(2)NeuAc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	12%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	2%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)NeuAc(1)	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	97%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	1%	0%	0%	0%
HexNAc(4)Hex(5)NeuAc(2)	15%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(5)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	1%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	50%		0%	0%	0%	2%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	3%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(7)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(7)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(5)Hex(3)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	4%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(5)Hex(4)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	28%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	2%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	2%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	1%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(7)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(7)Fuc(1)NeuAc(2)	0%	20%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(9)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(10)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(11)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)NeuAc(1)	0%	0%	23%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)	0%	0%	0%	0%	0%	0%	99%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)Sulfo(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNac(6)Hex(3)Fuc(3)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Sulfo(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(4)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(4)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(8)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(9)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(4)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(4)Fuc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(6)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(7)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(7)Fuc(1)NeuAc(2)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(8)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(8)Fuc(1)NeuAc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(8)Hex(3)Fuc(1)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
HexNAc(8)Hex(9)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%
UNOCCUPIED	0%	0%	0%	0%	0%	96%	0%	22%	0%	3%	8%	0%	0%	0%	19%	0%	30%	5%	0%	0%		0%	0%	0%	0%	0%	0%	7%

The average abundance of glycan compositions across three biological replicates is shown, with a white to red colour scale used to highlight higher abundances of particular glycan compositions.

B.6 Expanded glycan composition analysis of Env produced by replicons in DC2.4 cells

Average_VEEVNFL_DC2.4	88	133	142	156	160	185e	185h	197	234	262	276	295	301	332	339	355	363	386	392	398	406	411	448	462	611	618	625	637
HexNAc(1)	0%		0%	0%	0%	4%		1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	15%	0%	0%	0%	0%
HexNAc(1)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%
HexNAc(2)Hex(1)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(2)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)	7%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(4)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	4%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(4)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(5)	15%		0%	87%	0%	1%		4%	1%	20%	0%	26%	0%	16%	1%	14%	2%	1%	1%	0%	100%	53%	10%	14%	2%	0%	0%	4%
HexNAc(2)Hex(5)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(6)	4%		13%	13%	39%	0%		2%	1%	0%	11%	16%	0%	0%	1%	34%	0%	1%	0%	0%	0%	11%	1%	0%	0%	0%	0%	5%
HexNAc(2)Hex(7)	1%		0%	0%	0%	1%		4%	2%	1%	9%	0%	0%	2%	3%	4%	2%	1%	0%	0%	0%	10%	3%	5%	0%	0%	0%	6%
HexNAc(2)Hex(8)	0%		0%	0%	0%	0%		4%	20%	57%	59%	5%	0%	10%	18%	3%	16%	5%	44%	0%	0%	7%	26%	0%	0%	0%	0%	2%
HexNAc(2)Hex(9)	0%		0%	0%	61%	0%		2%	47%	15%	0%	20%	0%	69%	41%	0%	64%	92%	51%	0%	0%	3%	38%	2%	0%	0%	0%	0%
HexNAc(2)Hex(10)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(3)Fuc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	25%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	12%
HexNac(3)Hex(3)Fuc(2)Sulfo(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)	0%		0%	0%	0%	0%		0%	0%	0%	0%	8%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)	0%		0%	0%	0%	0%		0%	1%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(4)Fuc(1)NeuAc(1)Sulfo(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(5)Fuc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)NeuAc(1)	12%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	15%
HexNAc(3)Hex(6)	1%		0%	0%	0%	0%		1%	0%	0%	1%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%
HexNAc(3)Hex(6)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)NeuAc(1)	19%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	6%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(6)NeuAc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)Fuc(1)	0%		0%	0%	0%	0%		1%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	3%	0%	14%	0%	0%
HexNAc(4)Hex(3)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	33%	0%	0%
HexNac(4)Hex(3)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(3)Sulfo(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)Sulfo(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	50%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%
HexNAc(4)Hex(5)Fuc(1)	0%		0%	0%	0%	0%		1%	0%	0%	9%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	10%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	1%		0%	0%	0%	0%		0%	12%	0%	11%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	19%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	1%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(5)Fuc(2)NeuAc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	24%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(4)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)NeuAc(1)	4%		0%	0%	0%	0%		0%	2%	0%	0%	0%	66%	0%	0%	7%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)NeuAc(2)	19%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(5)Sulfo(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(1)	0%		0%	0%	0%	0%		0%	13%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	3%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	8%	0%	0%	0%	49%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(7)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(7)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	7%	0%	0%	0%	0%
HexNAc(5)Hex(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%
HexNAc(5)Hex(3)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(5)Hex(3)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	33%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	6%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(5)Hex(4)Sulfo(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	14%
HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)	0%		50%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	7%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	25%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(3)	0%		0%	0%	0%	0%		0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(7)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%
HexNAc(5)Hex(7)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(9)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(10)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(11)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)Sulfo(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)Sulfo(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(6)Hex(3)Fuc(3)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Sulfo(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	47%	0%	0%
HexNAc(6)Hex(6)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%
HexNAc(6)Hex(6)Fuc(1)NeuAc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(4)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(4)NeuAc(1)	0%		2%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(4)	0%		3%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(8)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(9)Fuc(1)NeuAc(2)	0%		32%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(3)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(4)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(4)Fuc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%
HexNAc(7)Hex(6)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(7)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(7)Fuc(1)NeuAc(2)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(8)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(8)Fuc(1)NeuAc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(8)Hex(3)Fuc(1)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(8)Hex(9)	0%		0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
UNOCCUPIED	11%		0%	0%	0%	94%		80%	0%	7%	0%	0%	0%	0%	34%	5%	15%	0%	0%	0%	0%	14%	8%	0%	72%	0%	100%	33%

The average abundance of glycan compositions across three biological replicates is shown, with a white to red colour scale used to highlight higher abundances of particular glycan compositions.

B.7 Expanded glycan composition analysis of Env produced by plasmid DNA in HEK293F cells

Average	88	133	142	156	160	185e	185h	197	234	262	276	295	301	332	339	355	363	386	392	398	406	411	448	462	611	618	625	637
HexNAc(1)	3%	0%			0%	0%	0%	1%	1%	0%	1%	0%		0%	0%	1%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	0%
HexNAc(1)Fuc(1)	3%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%
HexNAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%
HexNAc(2)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(1)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(2)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)	0%	0%			0%	0%	0%	0%	0%	0%	1%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(3)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(4)	0%	0%			2%	0%	0%	0%	0%	1%	9%	0%		0%	2%	1%	1%	1%	1%	0%	0%	1%	3%	0%	0%	0%	0%	2%
HexNAc(2)Hex(4)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(5)	17%	0%			30%	0%	0%	16%	5%	5%	25%	9%		2%	6%	24%	12%	13%	12%	0%	100%	57%	16%	1%	6%	0%	0%	25%
HexNAc(2)Hex(5)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(2)Hex(6)	1%	0%			3%	0%	0%	5%	5%	9%	8%	23%		2%	5%	1%	6%	8%	8%	0%	0%	13%	5%	0%	0%	0%	0%	18%
HexNAc(2)Hex(7)	0%	0%			7%	0%	0%	5%	13%	7%	15%	10%		5%	10%	1%	9%	9%	9%	0%	0%	10%	10%	0%	0%	0%	0%	19%
HexNAc(2)Hex(8)	0%	0%			17%	0%	0%	11%	25%	32%	16%	17%		19%	27%	0%	11%	19%	43%	0%	0%	4%	25%	0%	0%	0%	0%	6%
HexNAc(2)Hex(9)	0%	0%			30%	0%	0%	16%	45%	36%	0%	41%		71%	42%	0%	54%	47%	25%	0%	0%	0%	27%	0%	0%	0%	0%	0%
HexNAc(2)Hex(10)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)Fuc(1)	0%	12%			0%	0%	0%	2%	0%	0%	0%	0%		0%	0%	1%	1%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(3)Fuc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(3)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(3)Fuc(2)Sulfo(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)	0%	0%			0%	0%	0%	0%	0%	0%	1%	0%		0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	4%	1%	0%	0%	2%	0%	1%	0%	0%	0%	0%	0%	1%
HexNAc(3)Hex(4)Fuc(1)NeuAc(1)	0%	25%			0%	0%	0%	1%	0%	0%	0%	0%		0%	0%	9%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(3)Hex(4)Fuc(1)NeuAc(1)Sulfo(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(4)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	1%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)	0%	0%			0%	0%	0%	0%	0%	0%	10%	0%		0%	0%	1%	0%	0%	0%	0%	0%	3%	1%	0%	0%	0%	2%	5%
HexNAc(3)Hex(5)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	4%	1%	0%	0%	1%	0%	1%	1%	0%	0%	0%	0%	1%
HexNac(3)Hex(5)Fuc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(5)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	2%	0%		0%	0%	1%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	2%	0%
HexNAc(3)Hex(6)	1%	0%			0%	0%	0%	0%	0%	0%	1%	0%		0%	0%	2%	1%	0%	0%	0%	0%	2%	1%	0%	0%	0%	0%	5%
HexNAc(3)Hex(6)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(3)Hex(6)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	1%	0%		0%	0%	3%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	1%	1%
HexNac(3)Hex(6)NeuAc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)	2%	0%			0%	0%	100%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)Fuc(1)	3%	13%			0%	1%	0%	6%	0%	0%	0%	0%		0%	0%	1%	0%	0%	0%	4%	0%	0%	1%	3%	0%	0%	0%	0%
HexNAc(4)Hex(3)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(3)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(3)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(3)Sulfo(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)	3%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	2%	0%	0%	0%	2%	0%	0%	1%	3%	0%	0%	0%	1%
HexNAc(4)Hex(4)Fuc(1)NeuAc(1)	1%	0%			0%	1%	0%	0%	0%	0%	0%	0%		0%	0%	1%	0%	0%	0%	8%	0%	0%	0%	5%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(1)Sulfo(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(4)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)	4%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(1)	1%	0%			1%	0%	0%	5%	2%	2%	0%	0%		0%	0%	3%	0%	0%	0%	1%	0%	0%	1%	2%	3%	6%	0%	2%
HexNAc(4)Hex(5)Fuc(1)NeuAc(1)	5%	0%			2%	1%	0%	0%	1%	4%	5%	0%		0%	0%	12%	0%	0%	0%	13%	0%	0%	2%	23%	1%	0%	0%	3%
HexNAc(4)Hex(5)Fuc(1)NeuAc(2)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	1%
HexNAc(4)Hex(5)Fuc(2)	0%	50%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(2)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(5)Fuc(2)NeuAc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(3)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%
HexNAc(4)Hex(5)Fuc(4)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)NeuAc(1)	3%	0%			0%	0%	0%	0%	0%	0%	1%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(5)NeuAc(2)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(4)Hex(5)Sulfo(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)Fuc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(6)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(7)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(4)Hex(7)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)	2%	0%			0%	0%	0%	0%	0%	0%	1%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)	2%	0%			0%	1%	0%	1%	0%	0%	0%	0%		0%	0%	2%	0%	0%	0%	17%	0%	0%	0%	8%	2%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(3)Fuc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(5)Hex(3)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)	2%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)	1%	0%			0%	1%	0%	2%	0%	1%	0%	0%		0%	0%	3%	0%	0%	0%	13%	0%	0%	1%	11%	1%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)	0%	0%			0%	1%	0%	0%	0%	0%	1%	0%		0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(1)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)Fuc(2)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(4)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(5)Hex(4)Sulfo(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)	4%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	2%	0%	0%	0%	2%	0%	0%	0%	4%	1%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)NeuAc(1)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	1%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(2)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)Fuc(3)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(5)NeuAc(1)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)	8%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)	9%	0%			0%	0%	0%	0%	0%	1%	0%	0%		0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	2%	2%	2%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(1)	2%	0%			0%	0%	0%	0%	0%	1%	0%	0%		0%	0%	4%	0%	0%	0%	5%	0%	0%	0%	5%	15%	0%	0%	1%
HexNAc(5)Hex(6)Fuc(1)NeuAc(2)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	8%	10%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(1)NeuAc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(2)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)Fuc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%
HexNAc(5)Hex(6)Fuc(3)NeuAc(1)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(1)	6%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(2)	1%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(6)NeuAc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(7)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(7)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(5)Hex(9)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(10)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(11)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	2%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(1)Sulfo(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(2)Sulfo(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(3)Fuc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNac(6)Hex(3)Fuc(3)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(4)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	29%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(1)NeuAc(2)Sulfo(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(5)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(1)NeuAc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)Fuc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(6)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	1%
HexNAc(6)Hex(7)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	10%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	2%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6%	10%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(3)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(1)NeuAc(4)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)Fuc(4)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(7)NeuAc(4)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(8)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(6)Hex(9)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(3)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(4)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(4)Fuc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(6)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(7)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(7)Fuc(1)NeuAc(2)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(8)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(7)Hex(8)Fuc(1)NeuAc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(8)Hex(3)Fuc(1)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HexNAc(8)Hex(9)	0%	0%			0%	0%	0%	0%	0%	0%	0%	0%		0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
UNOCCUPIED	3%	0%			5%	88%	0%	30%	2%	0%	0%	0%		0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	24%	92%	90%	4%

The average abundance of glycan compositions across three biological replicates is shown, with a white to red colour scale used to highlight higher abundances of particular glycan compositions.

B.8 Negative stain – electron micrographs of Env

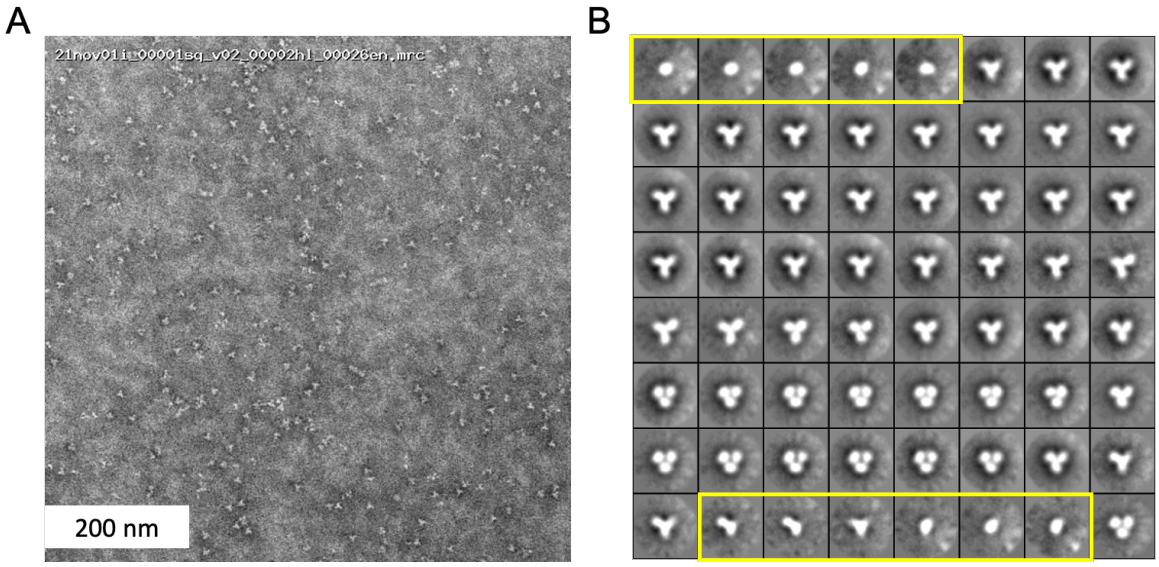


Figure B.8. A Representative of NS-EM image of BG505 NFL.664 Env particles secreted by HEK 293F cells transfected with plasmid DNA. B 2D class averages of the particles. The expressed Env was purified with GNL beads. Amongst all of the particles imaged, 20% fall within class averages corresponding to non-native conformations shown in yellow. The NS-EM was performed by the laboratory of Prof. Andrew Ward (The Scripps Research Institute, USA).

Appendix C Synthesis of RNA encoding Virus like particles and binding assays

Numerous vaccination strategies have been explored in both preclinical and clinical HIV-1 studies, but a successful vaccine remains elusive [474–476]. These approaches encompass various methods, such as HIV-1 Env trimer development based on SOSIP trimers and their variants [201,279,283,477], epitope-based vaccines targeting critical structural vulnerabilities recognized by bnAbs [478,479], as well as sequential vaccine strategies involving germline-engaging antigens [287,453] and prime-boost regimens [480–483]. Nevertheless, the available evidence from these investigations indicates that none of these vaccine approaches have demonstrated complete protection. There are multiple factors which can influence the bnAb-inducing HIV-1 vaccine and will potentially require the combination of efficient B-cell priming, optimization of Env design and presentation and sustained heterologous Env [286].

This section discusses about one of the approaches which utilizes mRNA co-expressing membrane-anchored HIV-1 envelope (Env) and simian immunodeficiency virus (SIV) Gag proteins to generate virus-like particles (VLPs), which closely mimic native viral particles produced by HIV-1 infection [484]. VLPs refers to particle that self-assemble into particles that resemble or mimic the structure, size and symmetry of original viruses. VLPs have been extensively used as prophylactic or therapeutic vaccine platforms for a wide range of diseases. These particles are useful as they present antigens in repetitive manner which allows the efficient crosslinking of B cell receptors [485]. In additions, VLPs can be used to deliver transmembrane Env which will mimic the virus better than soluble couterparts [486].

C.1 Design of constructs for mRNA encoding VLP

The SIVmac239.gag.p55 (Gene accession code: MZ362876) gene was commercially synthesized using IDT. BG505 MD39 20mutF Env gene was provided by Prof. William R Schief (The Scripps Research Institute, USA). modRNA backbone was commercially obtained from Takara. The *gag* and *env* genes were cloned in the modRNA backbone separately. mRNAs were generated by T7 RNA

polymerase-mediated in vitro transcription reactions using linearized DNA templates containing the immunogen open reading frame flanked by 5′ untranslated region (UTR) and 3′ UTR sequences terminated by an encoded polyA tail. CleanCap 5' cap structures (TriLink) were incorporated into the 5′ end co-transcriptionally. Uridine was completely replaced with N1-methyl-pseudouridine to reduce immunogenicity. Purified mRNAs were stored at -80 °C.

C.2 Co-expression of *env* and *gag* mRNAs in mouse muscle cells

The modRNAs encoding HIV-1 env (5 µg) and SIVmac239 Gag (5 µg) were co-transfected into mouse muscle cells (C2C12). The RNA mixture was mixed with 1×10^6 cells per well of 6 well plate and then transfected using neon electroporation. The conditions used for electroporation was 1400mV, 20ms, and 1 pulse, as suggested in manufacturer's protocol. Supernatants were collected 48 hrs after transfection, followed by addition of fresh media and then next harvest after 72 hrs of transfection. The collected supernatant was centrifuged at 1500g for 5mins.

The VLPs were extracted from the supernatant using PEG-it Virus Precipitation Solution (System Biosciences) as recommended in manufacturer's protocol. Briefly, 1 volume of ice-cold PEG-it precipitation solution was added to every 4 volumes of supernatant. Followed by overnight incubation at 4^{o}C and then centrifugation at 1500g for 30 mins at 4^{o}C. After centrifugation, the supernatant was transferred to a fresh tube and the residual PEG-it solution was spun down by centrifugation at 1500g for 30minutes at 4^{o}C. Both the pellets were resuspended in ice cold 1× PBS at 1/50 volume of original volume. For example, if the supernatant in 5ml then in 100µl of PBS.

Subsequently, the size and Z average of expressed proteins, Gag alone, Env alone, and VLP, was measured using dynamic light scattering (Malvern Nano ZS Zetasizer). The Z average of VLP were ~200 nm, relatively higher from Env and Gag alone, suggesting the budding of VLPs from cells (**Figure C.1**)

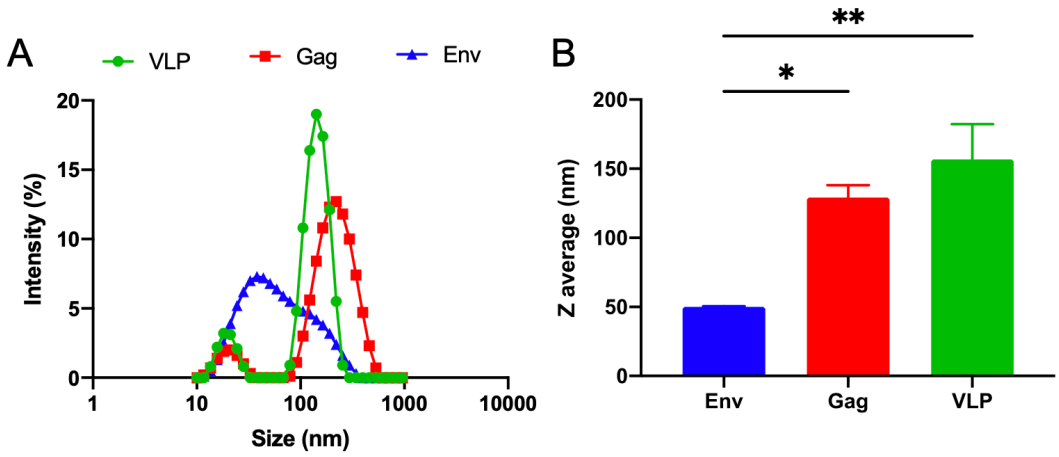


Figure C.1. Dynamic light scattering data for the VLP, Env, Gag proteins produced in the C2C12 cells, representing the intensity of size (A) and Z average of the particle (B). This work has been carried out in the laboratory of Prof. Darrell J. Irvine (Massachusetts Institute of Technology).

To further assess the expression of Gag, Env, and VLPs in mouse muscle cells, we conducted ELISA assays employing antibodies targeting p55 Gag and HIV-1 Env bnAbs. Gag titration involved coating high protein binding ELISA plates with supernatants containing Gag alone and Gag-Env complexes forming VLPs. Both supernatants were coated, both in their unlysed and lysed states using RIPA lysis buffer (Thermo Fischer). SIV Gag p27 antibody was employed to detect Gag expression, and it exhibited significant binding to lysed Gag and VLPs, thus confirming Gag expression (**Figure C.2**).

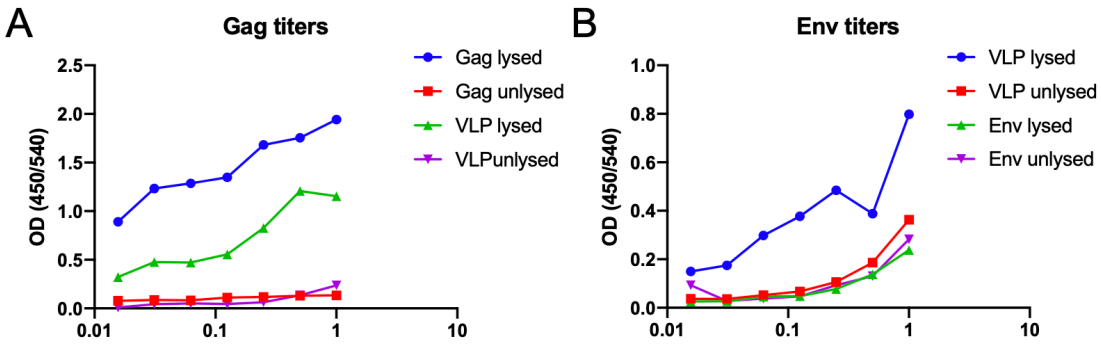


Figure C.2 Binding assay of Gag, Env, VLP expressed by mRNA-based platform. A) Gag titres in supernatant containing Gag and VLP were measured using Anti p27 Gag antibody. B) Env titres were measured using HIV-1 bnAb, VRC01. This work has been carried out in the laboratory of Prof. Darrell J. Irvine (Massachusetts Institute of Technology).

Moreover, to validate VLP budding with Env, Env titration was conducted using a panel of multiple bnAbs (PG9, PCT64, PGT121, PGT145, PGT151, 35022, VRCO1, BG18), as well as non-neutralizing antibody A32. Both lysed and unlysed VLPs displayed binding to HIV-1 antibodies, given that Env is expressed on the surface of Gag. Notably, PGT121 and PGT151 exhibited higher binding to Env compared to the other antibodies, while A32 also showed a modest level of binding to VLPs (**Figure C.3**). These experiments represent preliminary results, and further biological replicates are required to confirm the antigenicity of the VLPs.

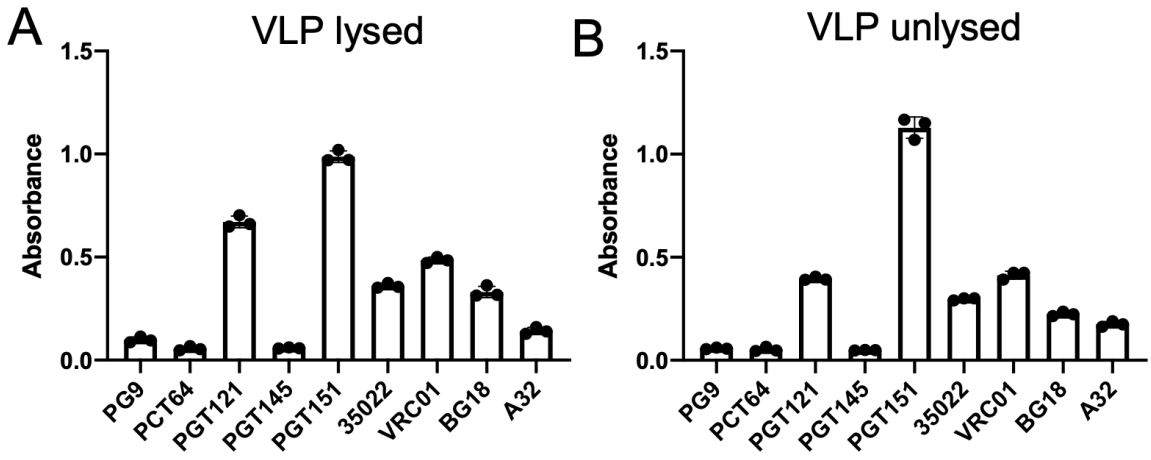


Figure C.3 Binding assay of HIV-1 Env in mRNA-expressed VLPs. This was done using a trimer-specific ELISA with a cocktail of anti-Env human bNAbs (that is, PG9, PCT64, PGT121, PGT145, PGT151, 35022, VRC01, BG18) and non-neutralizing antibody, A32, with VLP lysed using RIPA lysis buffer (A) and unlysed VLPs (B). This work has been carried out in the laboratory of Prof. Darrell J. Irvine (Massachusetts Institute of Technology).

Appendix D Sequences of Env proteins

SARS-CoV-2 S2P

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNV
TWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNAT
NVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQ
GNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHR
SYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL
YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDF
TGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYF
PLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTG
VLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICAS
YQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSV
DCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDF
GGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLT
VLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYEN
QKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDIL
SRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDF
CGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTH
WFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD
VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLI
AIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYTLESGGGSAWSHP
QFEKGGGSGGGSGGSSAWSHPQFEK

HexaPro

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNV
TWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNAT
NVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQ
GNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHR
SYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL
YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDF
TGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYF
PLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTG
VLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICAS
YQTQTNSPGSASSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSV
DCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDF
GGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLT
VLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMAYRFNGIGVTQNVLYENQ
KLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILS
RLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFC
GKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHW
FVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDV
DLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQGSGYIPEAPRDGQAYVRK
DGEWVLLSTFLGRSLEVLFQGPGHHHHHHHHHSAWSHPQFEK

Closed-S

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNV
TWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNAT
NVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQ
GNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHR
SYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSF
TVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL
YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDF
TGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYF
PLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTG
VLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLY
QNVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICAS
YQTQTNSPSRAGSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSV
DCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDF
GGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLT
VLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFAMQMAYRFNGIGVTQNVLYEN
QKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDIL
SRLDKPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDF
CGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTH
WFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPD
VDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQGSGYIPEAPRDGQAYVR
KDGEWVLLSTFLGRSLEVLFQGPGSLPETGGGSDYKDDDDKGGGGSGGGGSGGGGSGGG
GSGGGGSHHHHHH

SARS-CoV-2 RBD

MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARVQPTESIVRFPNITNLCPFGEVFNAT
RFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGD
EVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFE
RDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCG
PGTKHHHHHHGGSGGSGLNDIFEAQKIEWHE

SARS-CoV-1 RBD

MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARVVPSGDVVRFPNITNLCPFGEVFNA
TKFPSVYAWERKKISNCVADYSVLYNSTFFSTFKCYGVSATKLNDLCFSNVYADSFVVKG
DDVRQIAPGQTGVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRP
FERDISNVPFSPDGKPCTPPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVC
GPGTKHHHHHHGGSGGSGLNDIFEAQKIEWHE

MERS RBD

MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRRGARAKPSGSVVEQAEGVECDFSPLLSGT
PPQVYNFKRLVFTNCNYNLTKLLSLFSVNDFTCSQISPAAIASNCYSSLILDYFSYPLSMKSD
LSVSSAGPISQFNYKQSFSNPTCLILATVPHNLTTITKPLKYSYINKCSRLLSDDRTEVPQLV
NANQYSPCVSIVPSTVWEDGDYYRKQLSPLEGGGWLVASGSTVAMTEQLQMGFGITVQY
GTDTNSVCPKGTKHHHHHHGGSGGSGLNDIFEAQKIEWHE

RBD g5.1

NITNLCPFGEVFNATRFASVYAWNRTRISNCVADYSVLYNSASFSTFKCYGVNPTKLNDLC
FTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDNFTGCVIAWNSNNLDSKVGGNYNY
LYRLFRKSNLSPFERDISTEIYQAGSTPCNGTEGFNCYFPLQSYGFQPTNGVGYQPYRVVVL
SFELLHAPATVCGPHHHHHH

RBD g8.2

NITNLCPFGEVFNATRFASVYAWNRTRISNCTADYSVLYNSSSFSTFKCYGVNPTKLNDLC
FTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDNFTGCVIAWNSNNLDSKVGGNYNY
LYRLFRKSNLSPFERDISTEIYQAGSTPCNGTEGFNCYFPLQSYGFQPTNGVGYQPYRVVVL
SFENLSAPATVCGPHHHHHH

RBD g5.1 24mer

LRFGIVASRANHALVGGSGGNITNLCPFGEVFNATRFASVYAWNRTRISNCVADYSVLYN
SASFSTFKCYGVNPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDNFTG
CVIAWNSNNLDSKVGGNYNYLYRLFRKSNLSPFERDISTEIYQAGSTPCNGTEGFNCYFPL
QSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPGGSGGSGGSGGSGGGLSKDIIKLLNE
QVNKEMQSSNLYMSMSSWCYTHSLDGAGLFLFDHAAEEYEHAKKLIIFLNENNVPVQLTS
ISAPEHKFEGLTQIFQKAYEHEQHISESINNIVDHAIKSKDHATFNFLQWYVAEQHEEEVLF
KDILDKIELIGNENHGLYLADQYVKGIAKSRKS

BG505NFL.664

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLEN
VTEEFNMWKNNMVEQMHTDIISLWDQSLKPCVKLTPLCVTLQCTNVTNNITDDMRGELK
NCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPK
VSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEE
VMIRSENITNNAKNILVQFNTPVQINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCNV
SKATWNETLGKVVKQLRKHFGNNTIIRFANSSGGDLEVTTHSFNCGGEFFYCNTSGLFNST
WISNTSVQGSNSTGSNDSITLPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRDG
GSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRAKRRVVGGGGGSGGGGS
AVGIGAVFLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEAQQHLLKLTV
WGIKQLQARVLAVERYLRDQQLLGIWGCSGKLICTTNVPWNSSWSNRNLSEIWDNMTWL
QWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALD

p55 SIV GAG

MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQI
LGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQNKSKKKAQQAAAD
TGHSSQVSQNYPIVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQ
DLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTS
TLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTL
RAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARV
LAEAMSQATNTATIMMQRGNFRNQRKMVKCFNCGKEGHTARNCRAPRKKGCWKCGKE
GHQMKDCTERQANFLGKIWPSYKGRPGNFLQSRPEPTAPPFLQSRPEPTAPPEESFRSGVET
TTPPQKQEPIDKELYPLTSLRSLFGNDPSSQ

BG505 MD 39 20mutF

AENLWVTVYYGVPVWKDAETTLFCASDAKAYETEKHNVWATHACVPTDPNPQEIHLEN
VTEEFNMWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLQCTNYAPKLRSMMRGEIK
NCSFNMTTELRDKKQKVYSLFYRLDVVQINENQGNRSNNSNKEYRLINCNTSAITQACPK
VSFEPIPIHYCAPAGFAILKCKDKKFNGTGPCPSVSTVQCTHGIKPVVSTQLLLNGSLAEEE
VIIRSENITNNAKNILVQLNTPVQINCTRPNNNTVKSIRIGPGQAFYYFGDVLGDVRMAHCN
ISKATWNETLGKVVKQLRKHFGNNTIIRFAQSSGGDLEVTTHSFNCGGEFFYCNTSGLFNS
TWISNTSVQGSNSTGSNDSLILPCRIKQIINMWQRIGQAMYAPPIQGVIRCVSNITGLILTRD
GGSTNSTTETFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTRCKRRVVGSHSGSGGSGS
GGHAAVGIGAVSLGFLGAAGSTMGAASMTLTVQARNLLSGIVQQQSNLLRAPEPQQHLL
KDTHWGIKQLQARVLAVEHYLRDQQLLGIWGCSGKLICCTNVPWNSSWSNRNLSEIWDN
MTWLQWDKEISNYTQIIYGLLEESQNQQEKNEQDLLALDKWASLWNWFDISNWLWYIKI
FIMIVGGLIGLRIVFAVLSVIHRVR

Appendix E Vector maps

E.1 Replicon construct of BG505 NFL.664

Figure E.1. The replicon construct used for the expression of HIV-1 Env. The vector contains VEEV non-structural proteins (nsp's 1—4) as the backbone, with structural proteins replaced by BG505 NFL.664, shown in red. It also included an ampicillin resistance marker (amp marker).

E.2 modRNA backbone

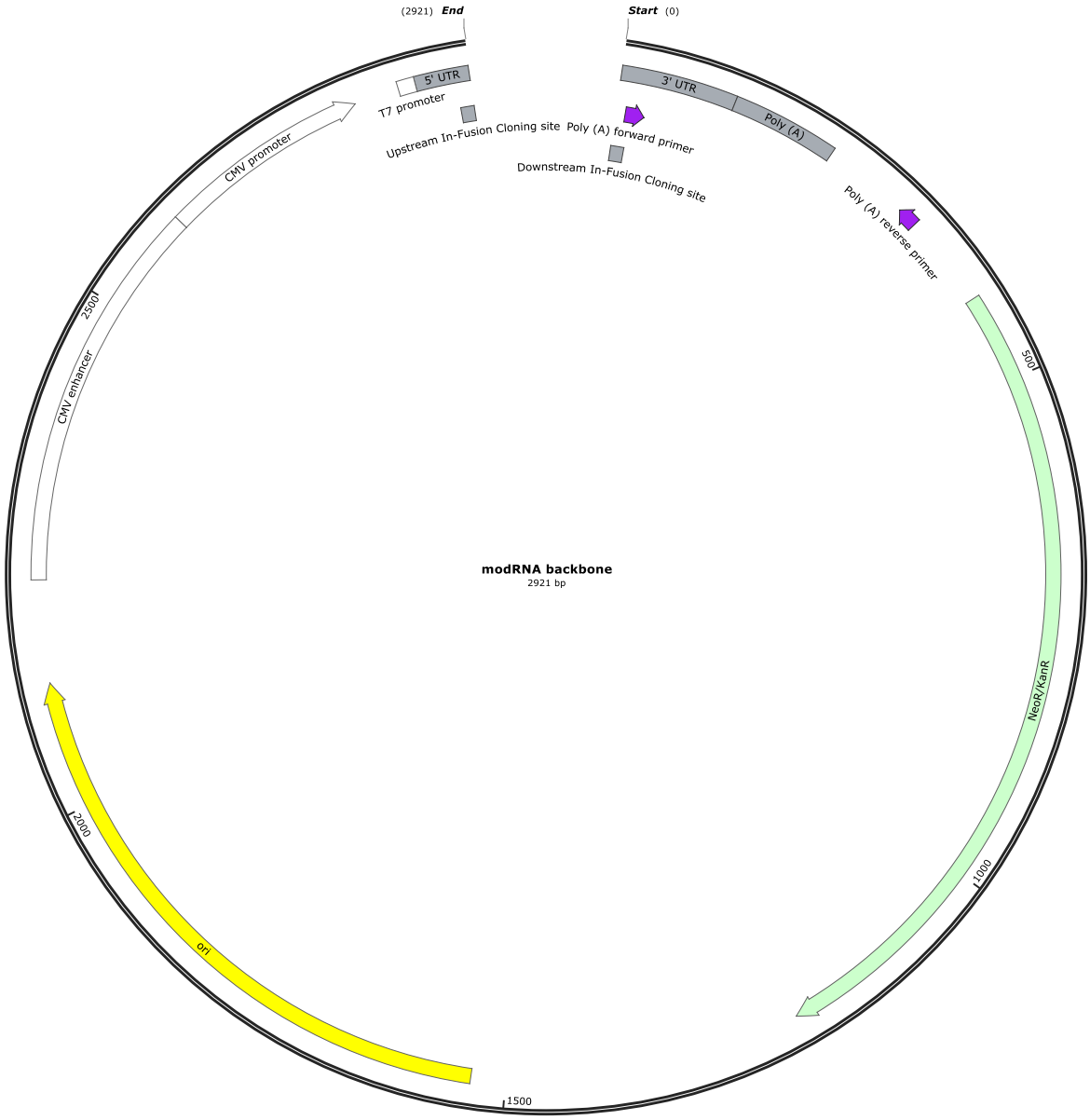


Figure E.2. The modRNA backbone used for the design of RNA construct for VLP expression. The vector contains 5′ and 3′ UTR and poly A tail for insertion of gene of interest for mRNA expression. It includes Neomycin (Neo) and kanamycin (Kan) resistance (R) marker, cytomegalovirus (CMV) enhancer and promoter, and origin of replication (ori).

Appendix F List of publications and conferences

Publications

- Tong T, D'Addabbo A, Xu J, **Chawla H**, Nguyen A, Ochoa P, Crispin M, Binley JM. *Plos Pathogens* 2023. Impact of stabilizing mutations on the antigenic profile and glycosylation of membrane-expressed HIV-1 envelope glycoprotein.

- **Chawla H**, Jossi SE, Faustini SE, Samsudin F, Allen JD, Watanabe Y, Newby ML, Juárez EM, Lamerton RE, McLellan JS, Bond PJ, Richter AG, Cunningham AF, Crispin M. *J Mol Biol.* 2022. Glycosylation and Serological Reactivity of an Expression-enhanced SARS-CoV-2 Viral Spike Mimetic.

- Reiss EIMM, Haaren MMV, Schooten JV, Claireaux MAF, Maisonnasse P, Antanasijevic A, Allen JD, Bontjer I, Torres JL, Lee W-H, Ozorowski G, Vázquez Bernat N, Kaduk M, Aldon Y, Burger JA, **Chawla H**, Aartse A, Tolazzi M, Gao H, Mundsperger P, Crispin M, Montefiori DC, Karlsson Hedestam GB, Scarlatti G, Ward AB, Grand RL, Shattock R, Dereuddre-Bosquet N, Sanders RW & van Gils MJ. *npj Vaccines*. 2022. Fine-mapping the immunodominant antibody epitopes on consensus sequence-based HIV-1 envelope trimer vaccine candidates.

- Konrath KM, Liaw K, Wu Y, Zhu X, Walker SN, Xu Z, Schultheis K, Chokkalingam N, **Chawla H**, Du J, Tursi NJ, Moore A, Adolf-Bryfogle J, Purwar M, Reuschel EL, Frase D, Sullivan M, Fry B, Maricic I, Andrade VM, Iffland C, Crispin M, Broderick KE, Humeau LMPF, Patel A, Smith TRF, Pallesen J, Weiner DB, Kulp DW.*Cell Reports*. 2022. Nucleic acid delivery of immune-focused SARS-CoV-2 nanoparticle drives rapid and potent immunogenicity capable of single dose protection.

- Allen J. D*****, **Chawla, H***, Samsudin, F*****, Zuzic L, Shivgan AT, Watanabe Y, He W-H, Callaghan S, Song G, Yong P, Brouwer PJM, Song Y, Cai Y, Duyvesteyn HME, Malinauskas T, Kint J, Pino P, Wurm MJ, Frank M, Chen B, Stuart DI, Sanders RW, Andrabi R, Burton DR, Li S, Bond PJ, Crispin M. *Biochemistry*. 2021. Site-Specific Steric Control of SARS-CoV-2 Spike Glycosylation. Site-Specific Steric Control of SARS-CoV-2 Spike Glycosylation.

- Watanabe Y, Mendonça L, Allen ER, Howe A, Lee M, Allen JD, **Chawla H**, Pulido D, Donnellan F, Davies H, Ulaszewska M, Belij-Rammerstorfer S, Morris S, Krebs A-S, Dejnirattisai W, Mongkolsapaya J, Supasa P, Screaton GR, Green CM, Lambe T, Zhang P, Gilbert SC, Crispin M. *ACS Cent. Sci.* 2021. Native-like SARS-CoV-2 Spike Glycoprotein Expressed by ChAdOx1 nCoV-19/AZD1222 Vaccine.

- Allen, JD, Watanabe Y, **Chawla H**, Newby ML, Crispin M. *J. Mol. Biol.* 2021. Subtle Influence of ACE2 Glycan Processing on SARS-CoV-2 Recognition.

- Mishra N, Sharma S, Dobhal A, Kumar S, **Chawla H**, Singh R, Makhdoomi MA, Das BK, Lodha R, Kabra SK, Luthra K. *Nature Comms.* 2020. Broadly neutralizing plasma antibodies effective against autologous circulating viruses in infants with multivariant HIV-1 infection.

- Mishra N, Sharma S, Dobhal A, Kumar S, **Chawla H**, Singh R, Das BK, Kabra SK, Lodha R, Luthra K. *Journal of Virology*. 2020. A rare mutation in an infant-derived HIV-1 envelope glycoprotein alters interprotomer stability and susceptibility to broadly neutralizing antibodies targeting the trimer apex.

- Kumar S, Batra H, Singh S, **Chawla H**, Singh R, Katpara S, Hussain AW, Das BK, Lodha R, Kabra SK, Luthra K. *Journal of General Virology*. 2020. Effect of combination antiretroviral

therapy on human immunodeficiency virus 1 specific antibody responses in subtype-C infected children.

- Mishra N, Sharma S, Dobhal A, Kumar S, **Chawla H**, Singh S, Singh R, Das BK, Kabra SK, Lodha R, Luthra K. bioRxiv. 2020. Plasma neutralizing antibodies in an infant with interclade HIV-1 superinfection preferentially neutralize superinfecting HIV-1 strains.

- Kumar S, Panda H, Makhdoomi MA, Mishra N, Safdari HA, **Chawla H,** Aggarwal H, Reddy ES, Lodha R, Kabra SK, Chandele A, Dutta S, Luthra K. *Journal of Virology*. 2019. An HIV-1 broadly neutralizing antibody from a clade C-infected pediatric elite neutralizer potently neutralizes the contemporaneous and autologous evolving viruses.

- Mishra N, Makhdoomi MA, Sharma S, Kumar S, Dobhal A, Kumar D, **Chawla H**, Singh R, Kanga U, Das BK, Lodha R, Kabra SK, Luthra K. *Journal of Virology*. 2019.Viral characteristics associated with maintenance of elite neutralizing activity in chronically HIV-1 clade C-infected monozygotic pediatric twins.

Review

- **Chawla H,** Fadda E, Crispin M. *Curr Opin Struct Biol*. 2022. Principles of SARS-CoV-2 glycosylation.

Submitted, under review

- **Chawla H**, Hayes GM, Silva M, Lee W-H, Ozorowski G, Mckay PF, Shattock RJ, Ward AB, Irvine DJ, Crispin M. *Nature Comms*. 2023. Signatures of native-like glycosylation in RNA replicon-derived HIV-1 immunogens. (Submitted, under review).

*The authors contributed equally

Conferences and talks

05/2020 EAVI weekly meeting - Development in immunogen glycosylation (oral presentation).

10/2020 EAVI weekly meeting - Site-specific glycan analysis of Elite neutralizer SOSIPs (oral presentation).

01/2021 NAC CAVD touchpoint meeting (Annual presentation to the Bill & Melinda Gates Foundation) - Glycan Analysis: mRNA + muscle-cell glycan analysis (oral presentation).

03/2021 CHAVD Principal Investigators Meeting - Glycosylation in protein- and nucleotide-based immunogens (oral presentation).

04/2021 VxPDC CAVD grant meeting - Defining the role and impact of HIV glycosylation on antigenicity, immunogenicity, and vaccine strategies (oral presentation).

07/2021 PhD Shadowing Day- Understanding viral resistance to the immune system (oral presentation).

01/2022 Annual School of Biological Sciences postgraduate symposium, University of Southampton - Glycosylation and serological reactivity of expression enhanced SARS-CoV-2 viral mimetic (Poster presentation).

05/2022 Alumni event, University of Southampton - RNA based vaccines (oral presentation).

02/2023 CHAVD YI Awardee presentation, 4th Annual All-Investigators Retreat, LA JOLLA, CA -Glycosylation of replicon (saRNA) delivered soluble natively flexible linked HIV-1 immunogens (oral presentation).

06/2023 Keystone Symposia Global Health Travel Awardee – Z6 HIV Vaccines, Immunoprophylaxis and Drugs Conference (Poster presentation)

Bibliography

[1] Y. Watanabe, T.A. Bowden, I.A. Wilson, M. Crispin, Exploitation of glycosylation in enveloped virus pathobiology, *Biochim. Biophys. Acta - Gen. Subj.* 1863 (2019) 1480–1497.

[2] L. Cao, J.K. Diedrich, D.W. Kulp, M. Pauthner, L. He, S.-K.K.R. Park, D. Sok, C.Y. Su, C.M. Delahunty, S. Menis, R. Andrabi, J. Guenaga, E. Georgeson, M. Kubitz, Y. Adachi, D.R. Burton, W.R. Schief, J.R. Yates, J.C. Paulson, J.R. Yates III, J.C. Paulson, Global site-specific N-glycosylation analysis of HIV envelope glycoprotein, *Nat. Commun.* 8 (2017) 14954.

[3] W.B. Struwe, E. Chertova, J.D. Allen, G.E. Seabright, Y. Watanabe, D.J. Harvey, M. Medina-Ramirez, J.D. Roser, R. Smith, D. Westcott, B.F. Keele, J.W. Bess, R.W. Sanders, J.D. Lifson, J.P. Moore, M. Crispin, Site-Specific Glycosylation of Virion-Derived HIV-1 Env Is Mimicked by a Soluble Trimeric Immunogen., *Cell Rep.* 24 (2018) 1958-1966.e5.

[4] K. Wanzeck, K.L. Boyd, J.A. McCullers, Glycan Shielding of the Influenza Virus Hemagglutinin Contributes to Immunopathology in Mice, *Am. J. Respir. Crit. Care Med.* 183 (2011) 767–773.

[5] Y. Kobayashi, Y. Suzuki, Evidence for N-Glycan Shielding of Antigenic Sites during Evolution of Human Influenza A Virus Hemagglutinin, *J. Virol.* 86 (2012) 3446–3451.

[6] X. Xiong, M.A. Tortorici, J. Snijder, C. Yoshioka, A.C. Walls, W. Li, A.T. McGuire, F.A. Rey, B.-J. Bosch, D. Veesler, Glycan Shield and Fusion Activation of a Deltacoronavirus Spike Glycoprotein Fine-Tuned for Enteric Infections., *J. Virol.* 92 (2018) e01628-17.

[7] Y. Watanabe, Z.T. Berndsen, J. Raghwani, G.E. Seabright, J.D. Allen, O.G. Pybus, J.S. McLellan, I.A. Wilson, T.A. Bowden, A.B. Ward, M. Crispin, Vulnerabilities in coronavirus glycan shields despite extensive glycosylation, *Nat. Commun.* 11 (2020) 2688.

[8] Y. Watanabe, J.D. Allen, D. Wrapp, J.S. McLellan, M. Crispin, Site-specific glycan analysis of the SARS-CoV-2 spike., *Science*. 369 (2020) 330–333.

[9] G. Ritchie, D.J. Harvey, U. Stroeher, F. Feldmann, H. Feldmann, V. Wahl-Jensen, L. Royle, R.A. Dwek, P.M. Rudd, Identification of N-glycans from Ebola virus glycoproteins by matrix-assisted laser desorption/ionisation time-of-flight and negative ion electrospray tandem mass spectrometry, *Rapid Commun. Mass Spectrom.* 24 (2010) 571.

[10] J.E. Lee, M.L. Fusco, A.J. Hessell, W.B. Oswald, D.R. Burton, E.O. Saphire, Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor, *Nature*. 454 (2008) 177–182.

[11] Y. Watanabe, J. Raghwani, J.D. Allen, G.E. Seabright, S. Li, F. Moser, J.T. Huiskonen, T. Strecker, T.A. Bowden, M. Crispin, Structure of the Lassa virus glycan shield provides a model for immunological resistance, *Proc. Natl. Acad. Sci.* 115 (2018) 7320–7325.

[12] R. Sommerstein, L. Flatz, M.M. Remy, P. Malinge, G. Magistrelli, N. Fischer, M. Sahin, A. Bergthaler, S. Igonet, J. ter Meulen, D. Rigo, P. Meda, N. Rabah, B. Coutard, T.A. Bowden, P.-H. Lambert, C.-A. Siegrist, D.D. Pinschewer, Arenavirus Glycan Shield Promotes Neutralizing Antibody Evasion and Protracted Infection, *PLOS Pathog.* 11 (2015) e1005276.

[13] C.R. Fontes-Garfias, C. Shan, H. Luo, A.E. Muruato, D.B.A. Medeiros, E. Mays, X. Xie, J. Zou, C.M. Roundy, M. Wakamiya, S.L. Rossi, T. Wang, S.C. Weaver, P.Y. Shi, Functional Analysis of Glycosylation of Zika Virus Envelope Protein, *Cell Rep.* 21 (2017) 1180–1190.

[14] J.A. Mondotte, P.-Y. Lozach, A. Amara, A. V. Gamarnik, Essential Role of Dengue Virus Envelope Protein N Glycosylation at Asparagine-67 during Viral Propagation, *J. Virol.* 81 (2007) 7136–7148.

[15] G.S. Mohan, W. Li, L. Ye, R.W. Compans, C. Yang, Antigenic Subversion: A Novel Mechanism of Host Immune Evasion by Ebola Virus, *PLoS Pathog*. 8 (2012) 1003065.

[16] I. Bagdonaite, H.H. Wandall, Global aspects of viral glycosylation, *Glycobiology*. 28 (2018) 443–467.

[17] E. Margolin, M. Crispin, A. Meyers, R. Chapman, E.P. Rybicki, A Roadmap for the Molecular Farming of Viral Glycoprotein Vaccines: Engineering Glycosylation and Glycosylation-Directed Folding., *Front. Plant Sci.* 11 (2020) 609207.

[18] L.K. Pritchard, S. Vasiljevic, G. Ozorowski, G.E. Seabright, A. Cupo, R. Ringe, H.J. Kim, R.W. Sanders, K.J. Doores, D.R. Burton, I.A. Wilson, A.B. Ward, J.P. Moore, M. Crispin, Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers, *Cell Rep.* 11 (2015) 1604–1613.

[19] D.L. Carbaugh, H.M. Lazear, Flavivirus Envelope Protein Glycosylation: Impacts on Viral Infection and Pathogenesis., *J. Virol.* 94 (2020) e00104-20.

[20] R.L. Schnaar, Glycobiology simplified: diverse roles of glycan recognition in inflammation, *J. Leukoc. Biol.* 99 (2016) 825–838.

[21] P.H. Seeberger, Monosaccharide Diversity, Cold Spring Harbor Laboratory Press, 2022.

[22] J. Dawson, Carbohydrate Chemistry: simple sugars, rings and glycosides, (2021) Chapter 13.

[23] C. Reily, T.J. Stewart, M.B. Renfrow, J. Novak, Glycosylation in health and disease, *Nat. Rev. Nephrol.* 15 (2019) 346–366.

[24] D. Shental-Bechor, Y. Levy, Effect of glycosylation on protein folding: A close look at thermodynamic stabilization, *Proc. Natl. Acad. Sci.* 105 (2008) 8256–8261.

[25] G.B.E. Stewart-Jones, C. Soto, T. Lemmin, G.Y. Chuang, A. Druz, R. Kong, P. V. Thomas, K. Wagh, T. Zhou, A.J. Behrens, T. Bylund, C.W. Choi, J.R. Davison, I.S. Georgiev, M.G. Joyce, Y. Do Kwon, M. Pancera, J. Taft, Y. Yang, B. Zhang, S.S. Shivatare, V.S. Shivatare, C.C.D. Lee, C.Y. Wu, C.A. Bewley, D.R. Burton, W.C. Koff, M. Connors, M. Crispin, U. Baxa, B.T. Korber, C.H. Wong, J.R. Mascola, P.D. Kwong, Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G, *Cell*. 165 (2016) 813–826.

[26] S.A. Jeffers, D.A. Sanders, A. Sanchez, Covalent Modifications of the Ebola Virus Glycoprotein, *J. Virol.* 76 (2002) 12463–12472.

[27] M. Nita-Lazar, M. Wacker, B. Schegg, S. Amber, M. Aebi, The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation, *Glycobiology*. 15 (2005) 361–367.

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

[29] R. Kornfeld, S. Kornfeld, ASSEMBLY OF ASPARAGINE-LINKED OLIGOSACCHARIDES, *Annu. Rev. Biochem.* 54 (1985) 631–664.

[30] S. Stertz, M. Reichelt, M. Spiegel, T. Kuri, L. Martínez-Sobrido, A. García-Sastre, F. Weber, G. Kochs, The intracellular sites of early replication and budding of SARS-coronavirus, *Virology*. 361 (2007) 304–315.

[31] A. Varki, R. Cummings, J. Esko, H. Freeze, G. Hart, J. Marth, Glycosyltransferases, *Cold Spring Harb. Lab. Press*. (1999).

[32] F.E. Ware, A. Vassilakos, P.A. Peterson, M.R. Jackson, M.A. Lehrman, D.B. Williams, The Molecular Chaperone Calnexin Binds Glc1Man9GlcNAc2 Oligosaccharide as an Initial Step in Recognizing Unfolded Glycoproteins, *J. Biol. Chem.* 270 (1995) 4697–4704.

[33] F. Helle, G. Vieyres, L. Elkrief, C.-I. Popescu, C. Wychowski, V. Descamps, S. Castelain, P. Roingeard, G. Duverlie, J. Dubuisson, Role of N-Linked Glycans in the Functions of Hepatitis C Virus Envelope Proteins Incorporated into Infectious Virions, *J. Virol.* 84 (2010) 11905–11915.

[34] T.G. Senkevich, C.L. White, E. V Koonin, B. Moss, Complete pathway for protein disulfide bond formation encoded by poxviruses, *Proc. Natl. Acad. Sci.* 99 (2002) 6667–6672.

[35] A. Gallina, T.M. Hanley, R. Mandel, M. Trahey, C.C. Broder, G.A. Viglianti, H.J.P. Ryser, Inhibitors of Protein-Disulfide Isomerase Prevent Cleavage of Disulfide Bonds in Receptor-bound Glycoprotein 120 and Prevent HIV-1 Entry, *J. Biol. Chem.* 277 (2002) 50579–50588.

[36] R.P. Ringe, A. Yasmeen, G. Ozorowski, E.P. Go, L.K. Pritchard, M. Guttman, T.A. Ketas, C.A. Cottrell, I.A. Wilson, R.W. Sanders, A. Cupo, M. Crispin, K.K. Lee, H. Desaire, A.B. Ward, P.J. Klasse, J.P. Moore, Influences on the Design and Purification of Soluble, Recombinant Native-Like HIV-1 Envelope Glycoprotein Trimers., *J. Virol.* 89 (2015) 12189–210.

[37] M. Crispin, D.J. Harvey, V.T. Chang, C. Yu, A.R. Aricescu, E.Y. Jones, S.J. Davis, R.A. Dwek, P.M. Rudd, Inhibition of hybrid- and complex-type glycosylation reveals the presence of the GlcNAc transferase I-independent fucosylation pathway, *Glycobiology*. 16 (2006) 748–756.

[38] J. Rini, J. Esko, A. Varki, Glycosyltransferases and Glycan-processing Enzymes, Cold Spring Harbor Laboratory Press, 2009.

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

[40] P. Stanley, K.W. Moremen, N.E. Lewis, N. Taniguchi, M. Aebi, N-Glycans, *Encycl. Cell Biol. Vol. 1-6, Second Ed.* 2 (2022) 487–494.

[41] G. Vieyres, J. Dubuisson, T. Pietschmann, Incorporation of Hepatitis C Virus E1 and E2 Glycoproteins: The Keystones on a Peculiar Virion, *Viruses*. 6 (2014) 1149–1187.

[42] T.A. Gerken, O. Jamison, C.L. Perrine, J.C. Collette, H. Moinova, L. Ravi, S.D. Markowitz, W. Shen, H. Patel, L.A. Tabak, Emerging Paradigms for the Initiation of Mucin-type Protein O-Glycosylation by the Polypeptide GalNAc Transferase Family of Glycosyltransferases, *J. Biol. Chem.* 286 (2011) 14493–14507.

[43] H. Wilkinson, R. Saldova, Current Methods for the Characterization of O-Glycans, *Cite This J. Proteome Res*. 19 (2020) 3905.

[44] D.T. Tran, K.G. Ten Hagen, Mucin-type O-Glycosylation during Development, *J. Biol. Chem.* 288 (2013) 6921–6929.

[45] I. Brockhausen, G. Möller, G. Merz, K. Adermann, H. Paulsen, Control of mucin synthesis: the peptide portion of synthetic O-glycopeptide substrates influences the activity of O-

glycan core 1 UDPgalactose:N-acetyl-alpha-galactosaminyl-R beta 3-galactosyltransferase., *Biochemistry*. 29 (1990) 10206–12.

[46] M. Dalziel, C. Whitehouse, I. McFarlane, I. Brockhausen, S. Gschmeissner, T. Schwientek, H. Clausen, J.M. Burchell, J. Taylor-Papadimitriou, The Relative Activities of the C2GnT1 and ST3Gal-I Glycosyltransferases Determine O-Glycan Structure and Expression of a Tumor-associated Epitope on MUC1, *J. Biol. Chem.* 276 (2001) 11007–11015.

[47] O. Martinez, L. Tantral, N. Mulherkar, K. Chandran, C.F. Basler, Impact of Ebola Mucin-Like Domain on Antiglycoprotein Antibody Responses Induced by Ebola Virus-Like Particles, *J. Infect. Dis.* 204 (2011) S825–S832.

[48] R. Nordén, A. Halim, K. Nyström, E.P. Bennett, U. Mandel, S. Olofsson, J. Nilsson, G. Larson, O-Linked Glycosylation of the Mucin Domain of the Herpes Simplex Virus Type 1-specific Glycoprotein gC-1 Is Temporally Regulated in a Seed-and-spread Manner, *J. Biol. Chem.* 290 (2015) 5078–5091.

[49] Z.A. Silver, A. Antonopoulos, S.M. Haslam, A. Dell, G.M. Dickinson, M.S. Seaman, R.C. Desrosiers, Discovery of O-Linked Carbohydrate on HIV-1 Envelope and Its Role in Shielding against One Category of Broadly Neutralizing Antibodies, *Cell Rep.* 30 (2020) 1862-1869.e4.

[50] P. Zhao, J.L. Praissman, O.C. Grant, Y. Cai, T. Xiao, K.E. Rosenbalm, K. Aoki, B.P. Kellman, R. Bridger, D.H. Barouch, M.A. Brindley, N.E. Lewis, M. Tiemeyer, B. Chen, R.J. Woods, L. Wells, Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor, *Cell Host Microbe*. 28 (2020) 586-601.e6.

[51] L. Zhang, M. Mann, Z.A. Syed, H.M. Reynolds, E. Tian, N.L. Samara, D.C. Zeldin, L.A. Tabak, K.G. Ten Hagen, Furin cleavage of the SARS-CoV-2 spike is modulated by O-glycosylation, *Proc. Natl. Acad. Sci. U. S. A.* 118 (2021) 2109905118.

[52] E. Gonzalez-Rodriguez, M. Zol-Hanlon, G. Bineva-Todd, A. Marchesi, M. Skehel, K.E. Mahoney, C. Roustan, A. Borg, L. Di Vagno, S. Kjær, A.G. Wrobel, D.J. Benton, P. Nawrath, S.L. Flitsch, D. Joshi, A.M. González-Ramírez, K.A. Wilkinson, R.J. Wilkinson, E.C. Wall, R. Hurtado-Guerrero, S.A. Malaker, B. Schumann, O-Linked Sialoglycans Modulate the Proteolysis of SARS-CoV-2 Spike and Likely Contribute to the Mutational Trajectory in Variants of Concern, *ACS Cent. Sci.* 9 (2023) 393–404.

[53] R. O'Flaherty, I. Trbojević-Akmačić, G. Greville, P.M. Rudd, G. Lauc, The sweet spot for biologics: recent advances in characterization of biotherapeutic glycoproteins, *Expert Rev. Proteomics*. 15 (2018) 13–29.

[54] L.R. Ruhaak, G. Xu, Q. Li, E. Goonatilleke, C.B. Lebrilla, Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses, *Chem. Rev.* 118 (2018) 7886–7930.

[55] K.T. Schjoldager, Y. Narimatsu, H.J. Joshi, H. Clausen, Global view of human protein glycosylation pathways and functions, *Nat. Rev. Mol. Cell Biol.* 21 (2020) 729–749.

[56] H. Shelton, G. Ayora-Talavera, J. Ren, S. Loureiro, R.J. Pickles, W.S. Barclay, I.M. Jones, Receptor Binding Profiles of Avian Influenza Virus Hemagglutinin Subtypes on Human Cells as a Predictor of Pandemic Potential, *J. Virol.* 85 (2011) 1875–1880.

[57] I. Brockhausen, Pathways of O-glycan biosynthesis in cancer cells, *Biochim. Biophys. Acta - Gen. Subj.* 1473 (1999) 67–95.

[58] S.J.D. Neil, Á. McKnight, K. Gustafsson, R.A. Weiss, HIV-1 incorporates ABO histo-blood group antigens that sensitize virions to complement-mediated inactivation, *Blood*. 105 (2005) 4693–4699.

[59] L. Cooling, Blood groups in infection and host susceptibility, *Clin. Microbiol. Rev.* 28 (2015) 801–870.

[60] P. Arend, Position of human blood group O(H) and phenotype-determining enzymes in growth and infectious disease, *Ann. N. Y. Acad. Sci.* 1425 (2018) 5–18.

[61] N.Y. Kim, W.W. Jung, Y.K. Oh, T. Chun, H.Y. Park, H.T. Lee, I.K. Han, J.M. Yang, Y.B. Kim, Natural protection from zoonosis by alpha-gal epitopes on virus particles in xenotransmission, *Xenotransplantation*. 14 (2007) 104–111.

[62] A. Antonopoulos, S.J. North, S.M. Haslam, A. Dell, Glycosylation of mouse and human immune cells: insights emerging from N-glycomics analyses, *Biochem. Soc. Trans.* 39 (2011) 1334–1340.

[63] C. Hilger, J. Fischer, F. Wölbing, T. Biedermann, Role and Mechanism of Galactose-Alpha-1,3-Galactose in the Elicitation of Delayed Anaphylactic Reactions to Red Meat, *Curr. Allergy Asthma Rep.* 19 (2019) 3.

[64] U. Galili, Natural anti-carbohydrate antibodies contributing to evolutionary survival of primates in viral epidemics?, *Glycobiology*. 26 (2016) 1140–1150.

[65] R.M. Seymour, M.J. Allan, A. Pomiankowski, K. Gustafsson, Evolution of the human ABO polymorphism by two complementary selective pressures., *Proceedings. Biol. Sci.* 271 (2004) 1065–72.

[66] J.E. Stencel-Baerenwald, K. Reiss, D.M. Reiter, T. Stehle, T.S. Dermody, The sweet spot: defining virus–sialic acid interactions, *Nat. Rev. Microbiol.* 12 (2014) 739–749.

[67] K. Shinya, M. Ebina, S. Yamada, M. Ono, N. Kasai, Y. Kawaoka, Influenza virus receptors in the human airway, *Nat. 2006 4407083*. 440 (2006) 435–436.

[68] U. Kumlin, S. Olofsson, K. Dimock, N. Arnberg, Sialic acid tissue distribution and influenza virus tropism, *Influenza Other Respi. Viruses*. 2 (2008) 147.

[69] J.M. Nicholls, A.J. Bourne, H. Chen, Y. Guan, J.M. Peiris, Sialic acid receptor detection in the human respiratory tract: Evidence for widespread distribution of potential binding sites for human and avian influenza viruses, *Respir. Res.* 8 (2007) 1–10.

[70] W. Widagdo, N.M.A. Okba, W. Li, A. de Jong, R.L. de Swart, L. Begeman, J.M.A. van den Brand, B.-J. Bosch, B.L. Haagmans, Species-Specific Colocalization of Middle East Respiratory Syndrome Coronavirus Attachment and Entry Receptors., *J. Virol.* 93 (2019) e00107-19.

[71] E. Qing, M. Hantak, S. Perlman, T. Gallagher, Distinct Roles for Sialoside and Protein Receptors in Coronavirus Infection, *MBio*. 11 (2020) e02764-19.

[72] L. Nguyen, K.A. McCord, D.T. Bui, K.M. Bouwman, E.N. Kitova, M. Elaish, D. Kumawat, G.C. Daskhan, I. Tomris, L. Han, P. Chopra, T.-J. Yang, S.D. Willows, A.L. Mason, L.K. Mahal, T.L. Lowary, L.J. West, S.-T.D. Hsu, T. Hobman, S.M. Tompkins, G.-J. Boons, R.P. de Vries, M.S. Macauley, J.S. Klassen, Sialic acid-containing glycolipids mediate binding and viral entry of SARS-CoV-2, *Nat. Chem. Biol.* 18 (2022) 81–90.

[73] E. Song, C. Zhang, B. Israelow, A. Lu-Culligan, A.V. Prado, S. Skriabine, P. Lu, O.-E. Weizman, F. Liu, Y. Dai, K. Szigeti-Buck, Y. Yasumoto, G. Wang, C. Castaldi, J. Heltke, E. Ng, J. Wheeler, M.M. Alfajaro, E. Levavasseur, B. Fontes, N.G. Ravindra, D. Van Dijk, S. Mane, M. Gunel, A. Ring, S.A.J. Kazmi, K. Zhang, C.B. Wilen, T.L. Horvath, I. Plu, S. Haik, J.-L. Thomas, A. Louvi, S.F. Farhadian, A. Huttner, D. Seilhean, N. Renier, K. Bilguvar, A. Iwasaki, Neuroinvasion of SARS-CoV-2 in human and mouse brain, *J. Exp. Med.* 218 (2021) e20202135.

[74] Z. Ye, Y. Mao, H. Clausen, S.Y. Vakhrushev, Glyco-DIA: a method for quantitative O-glycoproteomics with in silico-boosted glycopeptide libraries, *Nat. Methods 2019 169*. 16 (2019) 902–910.

[75] Public health surveillance for COVID-19: interim guidance, (n.d.).

[76] L. van der Hoek, Human coronaviruses: what do they cause?, *Antivir. Ther.* 12 (2007) 651–8.

[77] N. Zhong, B. Zheng, Y. Li, L. Poon, Z. Xie, K. Chan, P. Li, S. Tan, Q. Chang, J. Xie, X. Liu, J. Xu, D. Li, K. Yuen, J. Peiris, Y. Guan, Epidemiology and cause of severe acute

respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003, *Lancet*. 362 (2003) 1353–1358.

[78] A.M. Zaki, S. van Boheemen, T.M. Bestebroer, A.D.M.E. Osterhaus, R.A.M. Fouchier, Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia., *N. Engl. J. Med.* 367 (2012) 1814–20.

[79] F. Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. Song, Y. Hu, Z.-W. Tao, J.-H. Tian, Y.-Y. Pei, M.-L. Yuan, Y.-L. Zhang, F.-H. Dai, Y. Liu, Q.-M. Wang, J.-J. Zheng, L. Xu, E.C. Holmes, Y.-Z. Zhang, A new coronavirus associated with human respiratory disease in China, *Nature*. 579 (2020) 265–269.

[80] A.A.T. Naqvi, K. Fatima, T. Mohammad, U. Fatima, I.K. Singh, A. Singh, S.M. Atif, G. Hariprasad, G.M. Hasan, M.I. Hassan, Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach, *Biochim. Biophys. Acta - Mol. Basis Dis.* 1866 (2020) 165878.

[81] C.B. Jackson, M. Farzan, B. Chen, H. Choe, Mechanisms of SARS-CoV-2 entry into cells, *Nat. Rev. Mol. Cell Biol.* 23 (2022) 3–20.

[82] H. Yang, Z. Rao, Structural biology of SARS-CoV-2 and implications for therapeutic development, *Nat. Rev. Microbiol.* 19 (2021) 685–700.

[83] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, M.A. Müller, C. Drosten, S. Pöhlmann, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor., *Cell*. 181 (2020) 271-280.e8.

[84] B. Malone, N. Urakova, E.J. Snijder, E.A. Campbell, Structures and functions of coronavirus replication–transcription complexes and their relevance for SARS-CoV-2 drug design, *Nat. Rev. Mol. Cell Biol.* 23 (2022) 21.

[85] N.H. Moeller, K. Shi, Ö. Demir, C. Belica, S. Banerjee, L. Yin, C. Durfee, R.E. Amaro, H. Aihara, Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease ExoN., *Proc. Natl. Acad. Sci. U. S. A.* 119 (2022) e2106379119.

[86] X. He, E.H.Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y.C. Lau, J.Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B.J. Cowling, F. Li, G.M. Leung, Temporal dynamics in viral shedding and transmissibility of COVID-19, *Nat. Med. 2020 265*. 26 (2020) 672–675.

[87] W. Fischer, E.E. Giorgi, S. Chakraborty, K. Nguyen, T. Bhattacharya, J. Theiler, P.A. Goloboff, H. Yoon, W. Abfalterer, B.T. Foley, H. Tegally, J.E. San, T. de Oliveira, S.

Gnanakaran, B. Korber, E. Wilkinson, N. Msomi, A. Iranzadeh, V. Fonseca, D. Doolabh, K. Mlisana, A. von Gottberg, S. Walaza, M. Allam, A. Ismail, T. Mohale, A.J. Glass, S. Engelbrecht, G. Van Zyl, W. Preiser, F. Petruccione, A. Sigal, D. Hardie, G. Marais, M. Hsiao, S. Korsman, M.-A. Davies, L. Tyers, I. Mudau, D. York, C. Maslo, D. Goedhals, S. Abrahams, O. Laguda-Akingba, A. Alisoltani-Dehkordi, A. Godzik, C.K. Wibmer, B.T. Sewell, J. Lourenço, S.L.K. Pond, S. Weaver, M. Giovanetti, L.C.J. Alcantara, D. Martin, J.N. Bhiman, C. Williamson, HIV-1 and SARS-CoV-2: Patterns in the evolution of two pandemic pathogens, *Cell Host Microbe*. 29 (2021) 1093–1110.

[88] S. Santopolo, A. Riccio, M.G. Santoro, The biogenesis of SARS-CoV-2 spike glycoprotein: multiple targets for host-directed antiviral therapy, *Biochem. Biophys. Res. Commun.* 538 (2021) 80–87.

[89] J. Shang, G. Ye, K. Shi, Y. Wan, C. Luo, H. Aihara, Q. Geng, A. Auerbach, F. Li, Structural basis of receptor recognition by SARS-CoV-2, *Nature*. 581 (2020) 221–224.

[90] Q. Wang, Y. Zhang, L. Wu, S. Niu, C. Song, Z. Zhang, G. Lu, C. Qiao, Y. Hu, K.Y. Yuen, Q. Wang, H. Zhou, J. Yan, J. Qi, Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2, *Cell*. 181 (2020) 894-904.e9.

[91] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G.F. Gao, W. Tan, A Novel Coronavirus from Patients with Pneumonia in China, 2019, *N. Engl. J. Med.* 382 (2020) 727–733.

[92] B. Coutard, C. Valle, X. de Lamballerie, B. Canard, N.G. Seidah, E. Decroly, The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade., *Antiviral Res.* 176 (2020) 104742.

[93] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor, *Nature*. 581 (2020) 215–220.

[94] J.K. Millet, G.R. Whittaker, Physiological and molecular triggers for SARS-CoV membrane fusion and entry into host cells., *Virology*. 517 (2018) 3–8.

[95] C.F.S. Eldrid, J.D. Allen, M.L. Newby, M. Crispin, Suppression of O-Linked Glycosylation of the SARS-CoV-2 Spike by Quaternary Structural Restraints, *Anal. Chem.* 93 (2021) 14392–14400.

[96] L. Guo, S. Lin, Z. Chen, Y. Cao, B. He, G. Lu, Targetable elements in SARS-CoV-2 S2 subunit for the design of pan-coronavirus fusion inhibitors and vaccines, *Signal Transduct. Target. Ther.* 8 (2023) 197.

[97] J.D. Allen, D.P. Ivory, S.G. Song, W.-T. He, T. Capozzola, P. Yong, D.R. Burton, R. Andrabi, M. Crispin, The diversity of the glycan shield of sarbecoviruses related to SARS-CoV-2, *Cell Rep.* 42 (2023) 112307.

[98] P.J.M.M. Brouwer, T.G. Caniels, K. van der Straten, J.L. Snitselaar, Y. Aldon, S. Bangaru, J.L. Torres, N.M.A.A. Okba, M. Claireaux, G. Kerster, A.E.H.H. Bentlage, M.M. van Haaren, D. Guerra, J.A. Burger, E.E. Schermer, K.D. Verheul, N. van der Velde, A. van der Kooi, J. van Schooten, M.J. van Breemen, T.P.L.L. Bijl, K. Sliepen, A. Aartse, R. Derking, I. Bontjer, N.A. Kootstra, W.J. Wiersinga, G. Vidarsson, B.L. Haagmans, A.B. Ward, G.J. de Bree, R.W. Sanders, M.J. van Gils, Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability., *Science*. 369 (2020) 643–650.

[99] B. Ju, Q. Zhang, J. Ge, R. Wang, J. Sun, X. Ge, J. Yu, S. Shan, B. Zhou, S. Song, X. Tang, J. Yu, J. Lan, J. Yuan, H. Wang, J. Zhao, S. Zhang, Y. Wang, X. Shi, L. Liu, J. Zhao, X. Wang, Z. Zhang, L. Zhang, Human neutralizing antibodies elicited by SARS-CoV-2 infection, *Nature*. 584 (2020) 115–119.

[100] A.B. Vogel, I. Kanevsky, Y. Che, K.A. Swanson, A. Muik, M. Vormehr, L.M. Kranz, K.C. Walzer, S. Hein, A. Güler, J. Loschko, M.S. Maddur, A. Ota-Setlik, K. Tompkins, J. Cole, B.G. Lui, T. Ziegenhals, A. Plaschke, D. Eisel, S.C. Dany, S. Fesser, S. Erbar, F. Bates, D. Schneider, B. Jesionek, B. Sänger, A.-K.K. Wallisch, Y. Feuchter, H. Junginger, S.A. Krumm, A.P. Heinen, P. Adams-Quack, J. Schlereth, S. Schille, C. Kröner, R. de la Caridad Güimil Garcia, T. Hiller, L. Fischer, R.S. Sellers, S. Choudhary, O. Gonzalez, F. Vascotto, M.R. Gutman, J.A. Fontenot, S. Hall-Ursone, K. Brasky, M.C. Griffor, S. Han, A.A.H.H. Su, J.A. Lees, N.L. Nedoma, E.H. Mashalidis, P. V. Sahasrabudhe, C.Y. Tan, D. Pavliakova, G. Singh, C. Fontes-Garfias, M. Pride, I.L. Scully, T. Ciolino, J. Obregon, M. Gazi, R. Carrion, K.J. Alfson, W. V. Kalina, D. Kaushal, P.-Y.Y. Shi, T. Klamp, C. Rosenbaum, A.N. Kuhn, Ö. Türeci, P.R. Dormitzer, K.U. Jansen, U. Sahin, BNT162b vaccines protect rhesus macaques from SARS-CoV-2, *Nature*. 592 (2021) 283–289.

[101] K.S. Corbett, B. Flynn, K.E. Foulds, J.R. Francica, S. Boyoglu-Barnum, A.P. Werner, B. Flach, S. O'Connell, K.W. Bock, M. Minai, B.M. Nagata, H. Andersen, D.R. Martinez, A.T. Noe, N. Douek, M.M. Donaldson, N.N. Nji, G.S. Alvarado, D.K. Edwards, D.R. Flebbe, E. Lamb, N.A. Doria-Rose, B.C. Lin, M.K. Louder, S. O'Dell, S.D. Schmidt, E. Phung, L.A. Chang, C. Yap, J.-P.M. Todd, L. Pessaint, A. Van Ry, S. Browne, J. Greenhouse, T. Putman-Taylor, A. Strasbaugh, T.-A. Campbell, A. Cook, A. Dodson, K. Steingrebe, W. Shi, Y. Zhang, O.M. Abiona, L. Wang, A. Pegu, E.S. Yang, K. Leung, T. Zhou, I.-T. Teng, A. Widge, I. Gordon, L. Novik, R.A. Gillespie, R.J. Loomis, J.I. Moliva, G. Stewart-Jones, S. Himansu, W.-P. Kong, M.C. Nason, K.M. Morabito, T.J. Ruckwardt, J.E. Ledgerwood, M.R. Gaudinski, P.D. Kwong, J.R. Mascola, A. Carfi, M.G. Lewis, R.S.

Baric, A. McDermott, I.N. Moore, N.J. Sullivan, M. Roederer, R.A. Seder, B.S. Graham, Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates, *N. Engl. J. Med.* 383 (2020) 1544–1555.

[102] Q. Gao, L. Bao, H. Mao, L. Wang, K. Xu, M. Yang, Y. Li, L. Zhu, N. Wang, Z. Lv, H. Gao, X. Ge, B. Kan, Y. Hu, J. Liu, F. Cai, D. Jiang, Y. Yin, C. Qin, J. Li, X. Gong, X. Lou, W. Shi, D. Wu, H. Zhang, L. Zhu, W. Deng, Y. Li, J. Lu, C. Li, X. Wang, W. Yin, Y. Zhang, C. Qin, Development of an inactivated vaccine candidate for SARS-CoV-2, *Science*. 369 (2020) 77–81.

[103] P.M. Folegatti, K.J. Ewer, P.K. Aley, B. Angus, S. Becker, S. Belij-Rammerstorfer, D. Bellamy, S. Bibi, M. Bittaye, E.A. Clutterbuck, C. Dold, S.N. Faust, A. Finn, A.L. Flaxman, B. Hallis, P. Heath, D. Jenkin, R. Lazarus, R. Makinson, A.M. Minassian, K.M. Pollock, M. Ramasamy, H. Robinson, M. Snape, R. Tarrant, M. Voysey, C. Green, A.D. Douglas, A.V.S. Hill, T. Lambe, S.C. Gilbert, A.J. Pollard, J. Aboagye, K. Adams, A. Ali, E. Allen, J.L. Allison, R. Anslow, E.H. Arbe-Barnes, G. Babbage, K. Baillie, M. Baker, N. Baker, P. Baker, I. Baleanu, J. Ballaminut, E. Barnes, J. Barrett, L. Bates, A. Batten, K. Beadon, R. Beckley, E. Berrie, L. Berry, A. Beveridge, K.R. Bewley, E.M. Bijker, T. Bingham, L. Blackwell, C.L. Blundell, E. Bolam, E. Boland, N. Borthwick, T. Bower, A. Boyd, T. Brenner, P.D. Bright, C. Brown-O'Sullivan, E. Brunt, J. Burbage, S. Burge, K.R. Buttigieg, N. Byard, I. Cabera Puig, A. Calvert, S. Camara, M. Cao, F. Cappuccini, M. Carr, M.W. Carroll, V. Carter, K. Cathie, R.J. Challis, S. Charlton, I. Chelysheva, J.-S. Cho, P. Cicconi, L. Cifuentes, H. Clark, E. Clark, T. Cole, R. Colin-Jones, C.P. Conlon, A. Cook, N.S. Coombes, R. Cooper, C.A. Cosgrove, K. Coy, W.E.M. Crocker, C.J. Cunningham, B.E. Damratoski, L. Dando, M.S. Datoo, H. Davies, H. De Graaf, T. Demissie, C. Di Maso, I. Dietrich, T. Dong, F.R. Donnellan, N. Douglas, C. Downing, J. Drake, R. Drake-Brockman, R.E. Drury, S.J. Dunachie, N.J. Edwards, F.D.L. Edwards, C.J. Edwards, S.C. Elias, M.J. Elmore, K.R.W. Emary, M.R. English, S. Fagerbrink, S. Felle, S. Feng, S. Field, C. Fixmer, C. Fletcher, K.J. Ford, J. Fowler, P. Fox, E. Francis, J. Frater, J. Furze, M. Fuskova, E. Galiza, D. Gbesemete, C. Gilbride, K. Godwin, G. Gorini, L. Goulston, C. Grabau, L. Gracie, Z. Gray, L.B. Guthrie, M. Hackett, S. Halwe, E. Hamilton, J. Hamlyn, B. Hanumunthadu, I. Harding, S.A. Harris, A. Harris, D. Harrison, C. Harrison, T.C. Hart, L. Haskell, S. Hawkins, I. Head, J.A. Henry, J. Hill, S.H.C. Hodgson, M.M. Hou, E. Howe, N. Howell, C. Hutlin, S. Ikram, C. Isitt, P. Iveson, S. Jackson, F. Jackson, S.W. James, M. Jenkins, E. Jones, K. Jones, C.E. Jones, B. Jones, R. Kailath, K. Karampatsas, J. Keen, S. Kelly, D. Kelly, D. Kerr, S. Kerridge, L. Khan, U. Khan, A. Killen, J. Kinch, T.B. King, L. King, J. King, L. Kingham-Page, P. Klenerman, F. Knapper, J.C. Knight, D. Knott, S. Koleva, A. Kupke, C.W. Larkworthy, J.P.J. Larwood, A. Laskey, A.M. Lawrie, A. Lee, K.Y. Ngan Lee, E.A. Lees, H. Legge, A. Lelliott, N.-M. Lemm, A.M. Lias, A. Linder, S.

Lipworth, X. Liu, S. Liu, R. Lopez Ramon, M. Lwin, F. Mabesa, M. Madhavan, G. Mallett, K. Mansatta, I. Marcal, S. Marinou, E. Marlow, J.L. Marshall, J. Martin, J. McEwan, L. McInroy, G. Meddaugh, A.J. Mentzer, N. Mirtorabi, M. Moore, E. Moran, E. Morey, V. Morgan, S.J. Morris, H. Morrison, G. Morshead, R. Morter, Y.F. Mujadidi, J. Muller, T. Munera-Huertas, C. Munro, A. Munro, S. Murphy, V.J. Munster, P. Mweu, A. Noé, F.L. Nugent, E. Nuthall, K. O'Brien, D. O'Connor, B. Oguti, J.L. Oliver, C. Oliveira, P.J. O'Reilly, M. Osborn, P. Osborne, C. Owen, D. Owens, N. Owino, M. Pacurar, K. Parker, H. Parracho, M. Patrick-Smith, V. Payne, J. Pearce, Y. Peng, M.P. Peralta Alvarez, J. Perring, K. Pfafferott, D. Pipini, E. Plested, H. Pluess-Hall, K. Pollock, I. Poulton, L. Presland, S. Provstgaard-Morys, D. Pulido, K. Radia, F. Ramos Lopez, J. Rand, H. Ratcliffe, T. Rawlinson, S. Rhead, A. Riddell, A.J. Ritchie, H. Roberts, J. Robson, S. Roche, C. Rohde, C.S. Rollier, R. Romani, I. Rudiansyah, S. Saich, S. Sajjad, S. Salvador, L. Sanchez Riera, H. Sanders, K. Sanders, S. Sapaun, C. Sayce, E. Schofield, G. Screaton, B. Selby, C. Semple, H.R. Sharpe, I. Shaik, A. Shea, H. Shelton, S. Silk, L. Silva-Reyes, D.T. Skelly, H. Smee, C.C. Smith, D.J. Smith, R. Song, A.J. Spencer, E. Stafford, A. Steele, E. Stefanova, L. Stockdale, A. Szigeti, A. Tahiri-Alaoui, M. Tait, H. Talbot, R. Tanner, I.J. Taylor, V. Taylor, R. Te Water Naude, N. Thakur, Y. Themistocleous, A. Themistocleous, M. Thomas, T.M. Thomas, A. Thompson, S. Thomson-Hill, J. Tomlins, S. Tonks, J. Towner, N. Tran, J.A. Tree, A. Truby, K. Turkentine, C. Turner, N. Turner, S. Turner, T. Tuthill, M. Ulaszewska, R. Varughese, N. Van Doremalen, K. Veighey, M.K. Verheul, I. Vichos, E. Vitale, L. Walker, M.E.E. Watson, B. Welham, J. Wheat, C. White, R. White, A.T. Worth, D. Wright, S. Wright, X.L. Yao, Y. Yau, Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial, *Lancet*. 396 (2020) 467–478.

[104] N.B. Mercado, R. Zahn, F. Wegmann, C. Loos, A. Chandrashekar, J. Yu, J. Liu, L. Peter, K. McMahan, L.H. Tostanoski, X. He, D.R. Martinez, L. Rutten, R. Bos, D. van Manen, J. Vellinga, J. Custers, J.P. Langedijk, T. Kwaks, M.J.G. Bakkers, D. Zuijdgeest, S.K. Rosendahl Huber, C. Atyeo, S. Fischinger, J.S. Burke, J. Feldman, B.M. Hauser, T.M. Caradonna, E.A. Bondzie, G. Dagotto, M.S. Gebre, E. Hoffman, C. Jacob-Dolan, M. Kirilova, Z. Li, Z. Lin, S.H. Mahrokhian, L.F. Maxfield, F. Nampanya, R. Nityanandam, J.P. Nkolola, S. Patel, J.D. Ventura, K. Verrington, H. Wan, L. Pessaint, A. Van Ry, K. Blade, A. Strasbaugh, M. Cabus, R. Brown, A. Cook, S. Zouantchangadou, E. Teow, H. Andersen, M.G. Lewis, Y. Cai, B. Chen, A.G. Schmidt, R.K. Reeves, R.S. Baric, D.A. Lauffenburger, G. Alter, P. Stoffels, M. Mammen, J. Van Hoof, H. Schuitemaker, D.H. Barouch, Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques, *Nature*. 586 (2020) 583–588.

[105] R. Bos, L. Rutten, J.E.M. van der Lubbe, M.J.G. Bakkers, G. Hardenberg, F. Wegmann, D.

Zuijdgeest, A.H. de Wilde, A. Koornneef, A. Verwilligen, D. van Manen, T. Kwaks, R. Vogels, T.J. Dalebout, S.K. Myeni, M. Kikkert, E.J. Snijder, Z. Li, D.H. Barouch, J. Vellinga, J.P.M. Langedijk, R.C. Zahn, J. Custers, H. Schuitemaker, Ad26 vector-based COVID-19 vaccine encoding a prefusion-stabilized SARS-CoV-2 Spike immunogen induces potent humoral and cellular immune responses, *Npj Vaccines*. 5 (2020) 91.

[106] F. Amanat, D. Stadlbauer, S. Strohmeier, T.H.O.O. Nguyen, V. Chromikova, M. McMahon, K. Jiang, G.A. Arunkumar, D. Jurczyszak, J. Polanco, M. Bermudez-Gonzalez, G. Kleiner, T. Aydillo, L. Miorin, D.S. Fierer, L.A. Lugo, E.M. Kojic, J. Stoever, S.T.H.H. Liu, C. Cunningham-Rundles, P.L. Felgner, T. Moran, A. García-Sastre, D. Caplivski, A.C. Cheng, K. Kedzierska, O. Vapalahti, J.M. Hepojoki, V. Simon, F. Krammer, A serological assay to detect SARS-CoV-2 seroconversion in humans., *Nat. Med.* 26 (2020) 1033–1036.

[107] S.E. Faustini, S.E. Jossi, M. Perez-Toledo, A.M. Shields, J.D. Allen, Y. Watanabe, M.L. Newby, A. Cook, C.R. Willcox, M. Salim, M. Goodall, J.L. Heaney, E. Marcial-Juarez, G.L. Morley, B. Torlinska, D.C. Wraith, T. V. Veenith, S. Harding, S. Jolles, M.J. Ponsford, T. Plant, A. Huissoon, M.K. O'Shea, B.E. Willcox, M.T. Drayson, M. Crispin, A.F. Cunningham, A.G. Richter, Development of a high-sensitivity ELISA detecting IgG, IgA and IgM antibodies to the SARS-CoV-2 spike glycoprotein in serum and saliva, *Immunology*. 164 (2021) 135–147.

[108] N. van Doremalen, T. Lambe, A. Spencer, S. Belij-Rammerstorfer, J.N. Purushotham, J.R. Port, V.A. Avanzato, T. Bushmaker, A. Flaxman, M. Ulaszewska, F. Feldmann, E.R. Allen, H. Sharpe, J. Schulz, M. Holbrook, A. Okumura, K. Meade-White, L. Pérez-Pérez, N.J. Edwards, D. Wright, C. Bissett, C. Gilbride, B.N. Williamson, R. Rosenke, D. Long, A. Ishwarbhai, R. Kailath, L. Rose, S. Morris, C. Powers, J. Lovaglio, P.W. Hanley, D. Scott, G. Saturday, E. de Wit, S.C. Gilbert, V.J. Munster, ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques, *Nature*. 586 (2020) 578–582.

[109] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B.S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation, *Science*. 367 (2020) 1260–1263.

[110] C.-L. Hsieh, J.A. Goldsmith, J.M. Schaub, A.M. DiVenere, H.-C. Kuo, K. Javanmardi, K.C. Le, D. Wrapp, A.G. Lee, Y. Liu, C.-W. Chou, P.O. Byrne, C.K. Hjorth, N. V. Johnson, J. Ludes-Meyers, A.W. Nguyen, J. Park, N. Wang, D. Amengor, J.J. Lavinder, G.C. Ippolito, J.A. Maynard, I.J. Finkelstein, J.S. McLellan, Structure-based design of prefusion-stabilized SARS-CoV-2 spikes, *Science*. 369 (2020) 1501–1505.

[111] J. Juraszek, L. Rutten, S. Blokland, P. Bouchier, R. Voorzaat, T. Ritschel, M.J.G. Bakkers, L.L.R. Renault, J.P.M. Langedijk, Stabilizing the closed SARS-CoV-2 spike trimer, *Nat.*

Commun. 12 (2021) 244.

[112] R.W. Sanders, J.P. Moore, Virus vaccines: proteins prefer prolines, *Cell Host Microbe*. 29 (2021) 327–333.

[113] Y. Valdes-Balbin, D. Santana-Mederos, F. Paquet, S. Fernandez, Y. Climent, F. Chiodo, L. Rodríguez, B. Sanchez Ramirez, K. Leon, T. Hernandez, L. Castellanos-Serra, R. Garrido, G.-W. Chen, D. Garcia-Rivera, D.G. Rivera, V. Verez-Bencomo, Molecular Aspects Concerning the Use of the SARS-CoV-2 Receptor Binding Domain as a Target for Preventive Vaccines, *ACS Cent. Sci.* 7 (2021) 757–767.

[114] M.E. Toledo-Romaní, M. García-Carmenate, C. Valenzuela-Silva, W. Baldoquín-Rodríguez, M. Martínez-Pérez, M. Rodríguez-González, B. Paredes-Moreno, I. Mendoza-Hernández, R. González-Mujica Romero, O. Samón-Tabio, P. Velazco-Villares, J.P. Bacallao-Castillo, E. Licea-Martín, M. Rodríguez-Ortega, N. Herrera-Marrero, E. Caballero-González, L. Egües-Torres, R. Duartes-González, S. García-Blanco, S. Pérez-Cabrera, S. Huete-Ferreira, K. Idalmis-Cisnero, O. Fonte-Galindo, D. Meliá-Pérez, I. Rojas-Remedios, D. Doroud, M.M. Gouya, A. Biglari, S. Fernández-Castillo, Y. Climent-Ruiz, Y. Valdes-Balbín, D. García-Rivera, P. Van der Stuyft, V. Verez-Bencomo, Safety and efficacy of the two doses conjugated protein-based SOBERANA-02 COVID-19 vaccine and of a heterologous three-dose combination with SOBERANA-Plus: a double-blind, randomised, placebo-controlled phase 3 clinical trial, *Lancet Reg Heal. Am.* 18 (2023) 100423.

[115] T. Li, Q. Zheng, H. Yu, D. Wu, W. Xue, H. Xiong, X. Huang, M. Nie, M. Yue, R. Rong, S. Zhang, Y. Zhang, Y. Wu, S. Wang, Z. Zha, T. Chen, T. Deng, Y. Wang, T. Zhang, Y. Chen, Q. Yuan, Q. Zhao, J. Zhang, Y. Gu, S. Li, N. Xia, SARS-CoV-2 spike produced in insect cells elicits high neutralization titres in non-human primates, *Emerg. Microbes Infect.* 9 (2020) 2076–2090.

[116] V. Shinde, S. Bhikha, Z. Hoosain, M. Archary, Q. Bhorat, L. Fairlie, U. Lalloo, M.S.L. Masilela, D. Moodley, S. Hanley, L. Fouche, C. Louw, M. Tameris, N. Singh, A. Goga, K. Dheda, C. Grobbelaar, G. Kruger, N. Carrim-Ganey, V. Baillie, T. de Oliveira, A. Lombard Koen, J.J. Lombaard, R. Mngqibisa, A.E. Bhorat, G. Benadé, N. Lalloo, A. Pitsi, P.-L. Vollgraaff, A. Luabeya, A. Esmail, F.G. Petrick, A. Oommen-Jose, S. Foulkes, K. Ahmed, A. Thombrayil, L. Fries, S. Cloney-Clark, M. Zhu, C. Bennett, G. Albert, E. Faust, J.S. Plested, A. Robertson, S. Neal, I. Cho, G.M. Glenn, F. Dubovsky, S.A. Madhi, Efficacy of NVX-CoV2373 Covid-19 Vaccine against the B.1.351 Variant, *N. Engl. J. Med.* 384 (2021) 1899–1909.

[117] D. Wang, J. Baudys, J.L. Bundy, M. Solano, T. Keppel, J.R. Barr, Comprehensive Analysis

of the Glycan Complement of SARS-CoV-2 Spike Proteins Using Signature Ions-Triggered Electron-Transfer/Higher-Energy Collisional Dissociation (EThcD) Mass Spectrometry, *Anal. Chem.* 92 (2020) 14730–14739.

[118] B.J. Ward, P. Gobeil, A. Séguin, J. Atkins, I. Boulay, P.-Y. Charbonneau, M. Couture, M.-A. D'Aoust, J. Dhaliwall, C. Finkle, K. Hager, A. Mahmood, A. Makarkov, M.P. Cheng, S. Pillet, P. Schimke, S. St-Martin, S. Trépanier, N. Landry, Phase 1 randomized trial of a plant-derived virus-like particle vaccine for COVID-19, *Nat. Med.* 27 (2021) 1071–1078.

[119] M.S. Makatsa, M.B. Tincho, J.M. Wendoh, S.D. Ismail, R. Nesamari, F. Pera, S. de Beer, A. David, S. Jugwanth, M.P. Gededzha, N. Mampeule, I. Sanne, W. Stevens, L. Scott, J. Blackburn, E.S. Mayne, R.S. Keeton, W.A. Burgers, SARS-CoV-2 Antigens Expressed in Plants Detect Antibody Responses in COVID-19 Patients, *Front. Plant Sci.* 12 (2021) 550.

[120] R.L. Willey, R. Shibata, E.O. Freed, M.W. Cho, M.A. Martin, Differential glycosylation, virion incorporation, and sensitivity to neutralizing antibodies of human immunodeficiency virus type 1 envelope produced from infected primary T-lymphocyte and macrophage cultures, *J. Virol.* 70 (1996) 6431–6436.

[121] L. Casalino, Z. Gaieb, J.A. Goldsmith, C.K. Hjorth, A.C. Dommer, A.M. Harbison, C.A. Fogarty, E.P. Barros, B.C. Taylor, J.S. McLellan, E. Fadda, R.E. Amaro, Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein, *ACS Cent. Sci.* 6 (2020) 1722–1734.

[122] T. Tokatlian, B.J. Read, C.A. Jones, D.W. Kulp, S. Menis, J.Y.H. Chang, J.M. Steichen, S. Kumari, J.D. Allen, E.L. Dane, A. Liguori, M. Sangesland, D. Lingwood, M. Crispin, W.R. Schief, D.J. Irvine, Innate immune recognition of glycans targets HIV nanoparticle immunogens to germinal centers., *Science*. 363 (2019) 649–654.

[123] L.E. McCoy, E. Falkowska, K.J. Doores, K. Le, D. Sok, M.J. van Gils, Z. Euler, J.A. Burger, M.S. Seaman, R.W. Sanders, H. Schuitemaker, P. Poignard, T. Wrin, D.R. Burton, Incomplete Neutralization and Deviation from Sigmoidal Neutralization Curves for HIV Broadly Neutralizing Monoclonal Antibodies, *PLOS Pathog.* 11 (2015) e1005110.

[124] H. Yao, Y. Song, Y. Chen, N. Wu, J. Xu, C. Sun, J. Zhang, T. Weng, Z. Zhang, Z. Wu, L. Cheng, D. Shi, X. Lu, J. Lei, M. Crispin, Y. Shi, L. Li, S. Li, Molecular Architecture of the SARS-CoV-2 Virus, *Cell*. 183 (2020) 730-738.e13.

[125] J.D. Allen, H. Chawla, F. Samsudin, L. Zuzic, A.T. Shivgan, Y. Watanabe, W. He, S. Callaghan, G. Song, P. Yong, P.J.M. Brouwer, Y. Song, Y. Cai, H.M.E. Duyvesteyn, T. Malinauskas, J. Kint, P. Pino, M.J. Wurm, M. Frank, B. Chen, D.I. Stuart, R.W. Sanders, R. Andrabi, D.R. Burton, S. Li, P.J. Bond, M. Crispin, Site-Specific Steric Control of SARS-

CoV-2 Spike Glycosylation, *Biochemistry*. 60 (2021) 2153–2169.

[126] J. Brun, S. Vasiljevic, B. Gangadharan, M. Hensen, A. V Chandran, M.L. Hill, J.L. Kiappes, R.A. Dwek, D.S. Alonzi, W.B. Struwe, N. Zitzmann, Assessing Antigen Structural Integrity through Glycosylation Analysis of the SARS-CoV-2 Viral Spike., *ACS Cent. Sci.* 7 (2021) 586–593.

[127] A. Casas-Sanchez, A. Romero-Ramirez, E. Hargreaves, C.C. Ellis, B.I. Grajeda, I.L. Estevao, E.I. Patterson, G.L. Hughes, I.C. Almeida, T. Zech, Á. Acosta-Serrano, Inhibition of Protein N- Glycosylation Blocks SARS-CoV-2 Infection, *MBio*. 13 (2022) e03718-21.

[128] T. Sztain, S.-H. Ahn, A.T. Bogetti, L. Casalino, J.A. Goldsmith, E. Seitz, R.S. McCool, F.L. Kearns, F. Acosta-Reyes, S. Maji, G. Mashayekhi, J.A. McCammon, A. Ourmazd, J. Frank, J.S. McLellan, L.T. Chong, R.E. Amaro, A glycan gate controls opening of the SARS-CoV-2 spike protein, *Nat. Chem.* 13 (2021) 963–968.

[129] B. Li, L. Wang, H. Ge, X. Zhang, P. Ren, Y. Guo, W. Chen, J. Li, W. Zhu, W. Chen, L. Zhu, F. Bai, Identification of Potential Binding Sites of Sialic Acids on the RBD Domain of SARS-CoV-2 Spike Protein, *Front. Chem.* 9 (2021) 659764.

[130] D. Hoffmann, S. Mereiter, Y. Jin Oh, V. Monteil, E. Elder, R. Zhu, D. Canena, L. Hain, E. Laurent, C. Grünwald-Gruber, M. Klausberger, G. Jonsson, M.J. Kellner, M. Novatchkova, M. Ticevic, A. Chabloz, G. Wirnsberger, A. Hagelkruys, F. Altmann, L. Mach, J. Stadlmann, C. Oostenbrink, A. Mirazimi, P. Hinterdorfer, J.M. Penninger, Identification of lectin receptors for conserved SARS-CoV-2 glycosylation sites., *EMBO J.* 40 (2021) e108375.

[131] H.-Y. Huang, H.-Y. Liao, X. Chen, S.-W. Wang, C.-W. Cheng, M. Shahed-Al-Mahmud, Y.-M. Liu, A. Mohapatra, T.-H. Chen, J.M. Lo, Y.-M. Wu, H.-H. Ma, Y.-H. Chang, H.-Y. Tsai, Y.-C. Chou, Y.-P. Hsueh, C.-Y. Tsai, P.-Y. Huang, S.-Y. Chang, T.-L. Chao, H.-C. Kao, Y.-M. Tsai, Y.-H. Chen, C.-Y. Wu, J.-T. Jan, T.-J.R. Cheng, K.-I. Lin, C. Ma, C.-H. Wong, Vaccination with SARS-CoV-2 spike protein lacking glycan shields elicits enhanced protective responses in animal models, *Sci. Transl. Med.* 14 (2022) eabm0899.

[132] A.M. Harbison, C.A. Fogarty, T.K. Phung, A. Satheesan, B.L. Schulz, E. Fadda, Fine-tuning the spike: role of the nature and topology of the glycan shield in the structure and dynamics of the SARS-CoV-2 S, *Chem. Sci.* 13 (2022) 386–395.

[133] L. Kang, G. He, A.K. Sharp, X. Wang, A.M. Brown, P. Michalak, J. Weger-Lucarelli, A selective sweep in the Spike gene has driven SARS-CoV-2 human adaptation, *Cell*. 184 (2021) 4392-4400.e4.

[134] S. Zhang, Q. Liang, X. He, C. Zhao, W. Ren, Z. Yang, Z. Wang, Q. Ding, H. Deng, T. Wang, L. Zhang, X. Wang, Loss of Spike N370 glycosylation as an important evolutionary event for the enhanced infectivity of SARS-CoV-2, *Cell Res. 2022*. (2022) 1–4.

[135] L.F. Zacchi, B.L. Schulz, N-glycoprotein macroheterogeneity: biological implications and proteomic characterization, *Glycoconj. J.* 33 (2016) 359–376.

[136] C. Mathew, R.G. Weiß, C. Giese, C.W. Lin, M.E. Losfeld, R. Glockshuber, S. Riniker, M. Aebi, Glycan–protein interactions determine kinetics of N-glycan remodeling, *RSC Chem. Biol.* 2 (2021) 917–931.

[137] R. Henderson, R.J. Edwards, K. Mansouri, K. Janowska, V. Stalls, M. Kopp, B.F. Haynes, P. Acharya, Glycans on the SARS-CoV-2 Spike Control the Receptor Binding Domain Conformation., *BioRxiv Prepr. Serv. Biol.* (2020).

[138] I. Hang, C. Lin, O.C. Grant, S. Fleurkens, T.K. Villiger, M. Soos, M. Morbidelli, R.J. Woods, R. Gauss, M. Aebi, Analysis of site-specific N -glycan remodeling in the endoplasmic reticulum and the Golgi, *Glycobiology*. 25 (2015) 1335–1349.

[139] W. Li, M.J. Moore, N. Vasilieva, J. Sui, S.K. Wong, M.A. Berne, M. Somasundaran, J.L. Sullivan, K. Luzuriaga, T.C. Greenough, H. Choe, M. Farzan, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus., *Nature*. 426 (2003) 450–4.

[140] J.D. Allen, Y. Watanabe, H. Chawla, M.L. Newby, M. Crispin, Subtle Influence of ACE2 Glycan Processing on SARS-CoV-2 Recognition, *J. Mol. Biol.* 433 (2021) 166762.

[141] Q. Yang, T.A. Hughes, A. Kelkar, X. Yu, K. Cheng, S.J. Park, W.C. Huang, J.F. Lovell, S. Neelamegham, Inhibition of SARS-CoV-2 viral entry upon blocking N-and O-glycan elaboration, *Elife*. 9 (2020) 1–44.

[142] H.C. Huang, Y.J. Lai, C.C. Liao, W.F. Yang, K. Bin Huang, I.J. Lee, W.C. Chou, S.H. Wang, L.H. Wang, J.M. Hsu, C.P. Sun, C.T. Kuo, J. Wang, T.C. Hsiao, P.J. Yang, T.A. Lee, W. Huang, F.A. Li, C.Y. Shen, Y.L. Lin, M.H. Tao, C.W. Li, Targeting conserved N-glycosylation blocks SARS-CoV-2 variant infection in vitro, *EBioMedicine*. 74 (2021) 103712.

[143] J. Shang, Y. Wan, C. Luo, G. Ye, Q. Geng, A. Auerbach, F. Li, Cell entry mechanisms of SARS-CoV-2, *Proc. Natl. Acad. Sci. U. S. A.* 117 (2020) 11727–11734.

[144] D. Niemeyer, S. Stenzel, T. Veith, S. Schroeder, K. Friedmann, F. Weege, J. Trimpert, J. Heinze, A. Richter, J. Jansen, J. Emanuel, J. Kazmierski, F. Pott, L.M. Jeworowski, R. Olmer, M.-C. Jaboreck, B. Tenner, J. Papies, F. Walper, M.L. Schmidt, N. Heinemann, E. Möncke-Buchner, M. Baumgardt, K. Hoffmann, M. Widera, T.T.N. Thao, A. Balázs, J.

Schulze, C. Mache, T.C. Jones, M. Morkel, S. Ciesek, L.G. Hanitsch, M.A. Mall, A.C. Hocke, V. Thiel, K. Osterrieder, T. Wolff, U. Martin, V.M. Corman, M.A. Müller, C. Goffinet, C. Drosten, SARS-CoV-2 variant Alpha has a spike-dependent replication advantage over the ancestral B.1 strain in human cells with low ACE2 expression, *PLOS Biol.* 20 (2022) e3001871.

[145] M. Hoffmann, A. Sidarovich, P. Arora, N. Krüger, I. Nehlmeier, A. Kempf, L. Graichen, M.S. Winkler, D. Niemeyer, C. Goffinet, C. Drosten, S. Schulz, H.M. Jäck, S. Pöhlmann, Evidence for an ACE2-Independent Entry Pathway That Can Protect from Neutralization by an Antibody Used for COVID-19 Therapy, *MBio*. 13 (2022).

[146] M. Puray-Chavez, K.M. LaPak, T.P. Schrank, J.L. Elliott, D.P. Bhatt, M.J. Agajanian, R. Jasuja, D.Q. Lawson, K. Davis, P.W. Rothlauf, Z. Liu, H. Jo, N. Lee, K. Tenneti, J.E. Eschbach, C. Shema Mugisha, E.M. Cousins, E.W. Cloer, H.R. Vuong, L.A. VanBlargan, A.L. Bailey, P. Gilchuk, J.E. Crowe, M.S. Diamond, D.N. Hayes, S.P.J. Whelan, A. Horani, S.L. Brody, D. Goldfarb, M. Ben Major, S.B. Kutluay, Systematic analysis of SARS-CoV-2 infection of an ACE2-negative human airway cell, *Cell Rep.* 36 (2021) 109364.

[147] J. Baggen, M. Jacquemyn, L. Persoons, E. Vanstreels, V.E. Pye, A.G. Wrobel, V. Calvaresi, S.R. Martin, C. Roustan, N.B. Cronin, E. Reading, H.J. Thibaut, T. Vercruysse, P. Maes, F. De Smet, A. Yee, T. Nivitchanyong, M. Roell, N. Franco-Hernandez, H. Rhinn, A.A. Mamchak, M. Ah Young-Chapon, E. Brown, P. Cherepanov, D. Daelemans, TMEM106B is a receptor mediating ACE2-independent SARS-CoV-2 cell entry, *Cell*. 186 (2023) 3427-3442.e22.

[148] I. Hamming, W. Timens, M.L.C. Bulthuis, A.T. Lely, G.J. Navis, H. van Goor, Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis, *J. Pathol.* 203 (2004) 631–637.

[149] T.M. Clausen, D.R. Sandoval, C.B. Spliid, J. Pihl, H.R. Perrett, C.D. Painter, A. Narayanan, S.A. Majowicz, E.M. Kwong, R.N. McVicar, B.E. Thacker, C.A. Glass, Z. Yang, J.L. Torres, G.J. Golden, P.L. Bartels, R.N. Porell, A.F. Garretson, L. Laubach, J. Feldman, X. Yin, Y. Pu, B.M. Hauser, T.M. Caradonna, B.P. Kellman, C. Martino, P.L.S.M. Gordts, S.K. Chanda, A.G. Schmidt, K. Godula, S.L. Leibel, J. Jose, K.D. Corbett, A.B. Ward, A.F. Carlin, J.D. Esko, SARS-CoV-2 Infection Depends on Cellular Heparan Sulfate and ACE2, *Cell*. 183 (2020) 1043-1057.e15.

[150] C.J. Mycroft-West, D. Su, I. Pagani, T.R. Rudd, S. Elli, N.S. Gandhi, S.E. Guimond, G.J. Miller, M.C.Z. Meneghetti, H.B. Nader, Y. Li, Q.M. Nunes, P. Procter, N. Mancini, M. Clementi, A. Bisio, N.R. Forsyth, V. Ferro, J.E. Turnbull, M. Guerrini, D.G. Fernig, E. Vicenzi, E.A. Yates, M.A. Lima, M.A. Skidmore, Heparin Inhibits Cellular Invasion by

SARS-CoV-2: Structural Dependence of the Interaction of the Spike S1 Receptor-Binding Domain with Heparin, *Thromb. Haemost.* 120 (2020) 1700–1715.

[151] P.S. Kwon, H. Oh, S.-J. Kwon, W. Jin, F. Zhang, K. Fraser, J.J. Hong, R.J. Linhardt, J.S. Dordick, Sulfated polysaccharides effectively inhibit SARS-CoV-2 in vitro, *Cell Discov.* 6 (2020) 50.

[152] L. Nguyen, K.A. McCord, D.T. Bui, K.M. Bouwman, E.N. Kitova, M. Elaish, D. Kumawat, G.C. Daskhan, I. Tomris, L. Han, P. Chopra, T.-J. Yang, S.D. Willows, A.L. Mason, L.K. Mahal, T.L. Lowary, L.J. West, S.-T.D. Hsu, T. Hobman, S.M. Tompkins, G.-J. Boons, R.P. de Vries, M.S. Macauley, J.S. Klassen, Sialic acid-containing glycolipids mediate binding and viral entry of SARS-CoV-2., *Nat. Chem. Biol.* 18 (2022) 81–90.

[153] A.N. Baker, S.J. Richards, C.S. Guy, T.R. Congdon, M. Hasan, A.J. Zwetsloot, A. Gallo, J.R. Lewandowski, P.J. Stansfeld, A. Straube, M. Walker, S. Chessa, G. Pergolizzi, S. Dedola, R.A. Field, M.I. Gibson, The SARS-COV-2 Spike Protein Binds Sialic Acids and Enables Rapid Detection in a Lateral Flow Point of Care Diagnostic Device, *ACS Cent. Sci.* 6 (2020) 2046–2052.

[154] M.P. Lenza, I. Oyenarte, T. Diercks, J.I. Quintana, A. Gimeno, H. Coelho, A. Diniz, F. Peccati, S. Delgado, A. Bosch, M. Valle, O. Millet, N.G.A. Abrescia, A. Palazón, F. Marcelo, G. Jiménez-Osés, J. Jiménez-Barbero, A. Ardá, J. Ereño-Orbea, Structural Characterization of N-Linked Glycans in the Receptor Binding Domain of the SARS-CoV-2 Spike Protein and their Interactions with Human Lectins, *Angew. Chemie Int. Ed.* 59 (2020) 23763–23771.

[155] M. Stravalaci, I. Pagani, E.M. Paraboschi, M. Pedotti, A. Doni, F. Scavello, S.N. Mapelli, M. Sironi, C. Perucchini, L. Varani, M. Matkovic, A. Cavalli, D. Cesana, P. Gallina, N. Pedemonte, V. Capurro, N. Clementi, N. Mancini, P. Invernizzi, R. Bayarri-Olmos, P. Garred, R. Rappuoli, S. Duga, B. Bottazzi, M. Uguccioni, R. Asselta, E. Vicenzi, A. Mantovani, C. Garlanda, Recognition and inhibition of SARS-CoV-2 by humoral innate immunity pattern recognition molecules, *Nat. Immunol.* 23 (2022) 275–286.

[156] G. Paiardi, S. Richter, P. Oreste, C. Urbinati, M. Rusnati, R.C. Wade, The binding of heparin to spike glycoprotein inhibits SARS-CoV-2 infection by three mechanisms, *J. Biol. Chem.* 298 (2022) 101507.

[157] H. Qi, B. Liu, X. Wang, L. Zhang, The humoral response and antibodies against SARS-CoV-2 infection, *Nat. Immunol.* 23 (2022) 1008–1020.

[158] K. Röltgen, S.D. Boyd, Antibody and B cell responses to SARS-CoV-2 infection and vaccination, *Cell Host Microbe*. 29 (2021) 1063–1075.

[159] M.C. Woodruff, R.P. Ramonell, D.C. Nguyen, K.S. Cashman, A.S. Saini, N.S. Haddad, A.M. Ley, S. Kyu, J.C. Howell, T. Ozturk, S. Lee, N. Suryadevara, J.B. Case, R. Bugrovsky, W. Chen, J. Estrada, A. Morrison-Porter, A. Derrico, F.A. Anam, M. Sharma, H.M. Wu, S.N. Le, S.A. Jenks, C.M. Tipton, B. Staitieh, J.L. Daiss, E. Ghosn, M.S. Diamond, R.H. Carnahan, J.E. Crowe, W.T. Hu, F.E.-H. Lee, I. Sanz, Extrafollicular B cell responses correlate with neutralizing antibodies and morbidity in COVID-19, *Nat. Immunol.* 21 (2020) 1506–1516.

[160] R.A. Elsner, M.J. Shlomchik, Germinal Center and Extrafollicular B Cell Responses in Vaccination, Immunity, and Autoimmunity, *Immunity*. 53 (2020) 1136–1150.

[161] D. Lapuente, T.H. Winkler, M. Tenbusch, B-cell and antibody responses to SARS-CoV-2: infection, vaccination, and hybrid immunity., *Cell. Mol. Immunol.* (2023).

[162] A.C. Walls, M.C. Miranda, A. Schäfer, M.N. Pham, A. Greaney, P.S. Arunachalam, M.-J. Navarro, M.A. Tortorici, K. Rogers, M.A. O'Connor, L. Shirreff, D.E. Ferrell, J. Bowen, N. Brunette, E. Kepl, S.K. Zepeda, T. Starr, C.-L. Hsieh, B. Fiala, S. Wrenn, D. Pettie, C. Sydeman, K.R. Sprouse, M. Johnson, A. Blackstone, R. Ravichandran, C. Ogohara, L. Carter, S.W. Tilles, R. Rappuoli, S.R. Leist, D.R. Martinez, M. Clark, R. Tisch, D.T. O'Hagan, R. Van Der Most, W.C. Van Voorhis, D. Corti, J.S. McLellan, H. Kleanthous, T.P. Sheahan, K.D. Smith, D.H. Fuller, F. Villinger, J. Bloom, B. Pulendran, R.S. Baric, N.P. King, D. Veesler, Elicitation of broadly protective sarbecovirus immunity by receptor-binding domain nanoparticle vaccines, *Cell*. 184 (2021) 5432-5447.e16.

[163] G. Cerutti, Y. Guo, T. Zhou, J. Gorman, M. Lee, M. Rapp, E.R. Reddem, J. Yu, F. Bahna, J. Bimela, Y. Huang, P.S. Katsamba, L. Liu, M.S. Nair, R. Rawi, A.S. Olia, P. Wang, B. Zhang, G.Y. Chuang, D.D. Ho, Z. Sheng, P.D. Kwong, L. Shapiro, Potent SARS-CoV-2 neutralizing antibodies directed against spike N-terminal domain target a single supersite, *Cell Host Microbe*. 29 (2021) 819-833.e7.

[164] Z. Chen, E. John Wherry, T cell responses in patients with COVID-19, *Nat. Rev. Immunol.* 20 (2020) 529–536.

[165] A. Chandrashekar, J. Liu, A.J. Martinot, K. Mcmahan, N.B. Mercado, L. Peter, L.H. Tostanoski, J. Yu, Z. Maliga, M. Nekorchuk, K. Busman-Sahay, M. Terry, L.M. Wrijil, S. Ducat, D.R. Martinez, C. Atyeo, S. Fischinger, J.S. Burke, M.D. Slein, L. Pessaint, A. Van Ry, J. Greenhouse, T. Taylor, K. Blade, A. Cook, B. Finneyfrock, R. Brown, E. Teow, J. Velasco, R. Zahn, F. Wegmann, P. Abbink, E.A. Bondzie, G. Dagotto, M.S. Gebre, X. He, C. Jacob-Dolan, N. Kordana, Z. Li, M.A. Lifton, S.H. Mahrokhian, L.F. Maxfield, R. Nityanandam, J.P. Nkolola, A.G. Schmidt, A.D. Miller, R.S. Baric, G. Alter, P.K. Sorger, J.D. Estes, H. Andersen, M.G. Lewis, D.H. Barouch, SARS-CoV-2 infection protects

against rechallenge in rhesus macaques, (n.d.).

[166] D. Pinto, M.M. Sauer, N. Czudnochowski, J.S. Low, M. Alejandra Tortorici, M.P. Housley, J. Noack, A.C. Walls, J.E. Bowen, B. Guarino, L.E. Rosen, J. di Iulio, J. Jerak, H. Kaiser, S. Islam, S. Jaconi, N. Sprugasci, K. Culap, R. Abdelnabi, C. Foo, L. Coelmont, I. Bartha, S. Bianchi, C. Silacci-Fregni, J. Bassi, R. Marzi, E. Vetti, A. Cassotta, A. Ceschi, P. Ferrari, P.E. Cippà, O. Giannini, S. Ceruti, C. Garzoni, A. Riva, F. Benigni, E. Cameroni, L. Piccoli, M.S. Pizzuto, M. Smithey, D. Hong, A. Telenti, F.A. Lempp, J. Neyts, C. Havenar-Daughton, A. Lanzavecchia, F. Sallusto, G. Snell, H.W. Virgin, M. Beltramello, D. Corti, D. Veesler, Broad betacoronavirus neutralization by a stem helix–specific human antibody, *Science*. 373 (2021) 1109–1116.

[167] M. Ghorbani, B.R. Brooks, J.B. Klauda, Exploring dynamics and network analysis of spike glycoprotein of SARS-COV-2, *Biophys. J.* 120 (2021) 2902–2913.

[168] X. Chi, R. Yan, J.J.J.J.J. Zhang, G. Zhang, Y. Zhang, M. Hao, Z. Zhang, P. Fan, Y. Dong, Y. Yang, Z. Chen, Y. Guo, J.J.J.J.J. Zhang, Y. Li, X. Song, Y. Chen, L. Xia, L. Fu, L. Hou, J. Xu, C. Yu, J. Li, Q. Zhou, W. Chen, A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2., *Science*. 369 (2020) 650–655.

[169] M.A. Tortorici, M. Beltramello, F.A. Lempp, D. Pinto, H. V. Dang, L.E. Rosen, M. McCallum, J. Bowen, A. Minola, S. Jaconi, F. Zatta, A. De Marco, B. Guarino, S. Bianchi, E.J. Lauron, H. Tucker, J. Zhou, A. Peter, C. Havenar-Daughton, J.A. Wojcechowskyj, J.B. Case, R.E. Chen, H. Kaiser, M. Montiel-Ruiz, M. Meury, N. Czudnochowski, R. Spreafico, J. Dillen, C. Ng, N. Sprugasci, K. Culap, F. Benigni, R. Abdelnabi, S.Y.C. Foo, M.A. Schmid, E. Cameroni, A. Riva, A. Gabrieli, M. Galli, M.S. Pizzuto, J. Neyts, M.S. Diamond, H.W. Virgin, G. Snell, D. Corti, K. Fink, D. Veesler, Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms, *Science*. 370 (2020) 950–957.

[170] C.O. Barnes, C.A. Jette, M.E. Abernathy, K.-M.A. Dam, S.R. Esswein, H.B. Gristick, A.G. Malyutin, N.G. Sharaf, K.E. Huey-Tubman, Y.E. Lee, D.F. Robbiani, M.C. Nussenzweig, A.P. West, P.J. Bjorkman, SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies, *Nature*. 588 (2020) 682–687.

[171] L. Liu, P. Wang, M.S. Nair, J. Yu, M. Rapp, Q. Wang, Y. Luo, J.F.W. Chan, V. Sahi, A. Figueroa, X. V. Guo, G. Cerutti, J. Bimela, J. Gorman, T. Zhou, Z. Chen, K.-Y. Yuen, P.D. Kwong, J.G. Sodroski, M.T. Yin, Z. Sheng, Y. Huang, L. Shapiro, D.D. Ho, Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike, *Nature*. 584 (2020) 450–456.

[172] T. Lemmin, C. Soto, J. Stuckey, P.D. Kwong, Microsecond Dynamics and Network Analysis of the HIV-1 SOSIP Env Trimer Reveal Collective Behavior and Conserved Microdomains of the Glycan Shield, *Structure*. 25 (2017) 1631-1639.e2.

[173] Z.T. Berndsen, S. Chakraborty, X. Wang, C.A. Cottrell, J.L. Torres, J.K. Diedrich, C.A. López, J.R. Yates, M.J. van Gils, J.C. Paulson, S. Gnanakaran, A.B. Ward, Visualization of the HIV-1 Env glycan shield across scales, *Proc. Natl. Acad. Sci. U. S. A.* 117 (2020) 28014–28025.

[174] Y. Watanabe, L. Mendonça, E.R. Allen, A. Howe, M. Lee, J.D. Allen, H. Chawla, D. Pulido, F. Donnellan, H. Davies, M. Ulaszewska, S. Belij-Rammerstorfer, S. Morris, A.-S. Krebs, W. Dejnirattisai, J. Mongkolsapaya, P. Supasa, G.R. Screaton, C.M. Green, T. Lambe, P. Zhang, S.C. Gilbert, M. Crispin, Native-like SARS-CoV-2 Spike Glycoprotein Expressed by ChAdOx1 nCoV-19/AZD1222 Vaccine, *ACS Cent. Sci.* 7 (2021) 594–602.

[175] R. Thai, G. Moine, M. Desmadril, D. Servent, J.L. Tarride, A. Ménez, M. Léonetti, Antigen stability controls antigen presentation, *J. Biol. Chem.* 279 (2004) 50257–50266.

[176] H.A. Doyle, M.J. Mamula, Post-translational protein modifications in antigen recognition and autoimmunity, *Trends Immunol.* 22 (2001) 443–449.

[177] K. Hymes, J. Greene, A. Marcus, D. William, T. Cheung, N. Prose, H. Ballard, L. Laubenstein, KAPOSI'S SARCOMA IN HOMOSEXUAL MEN—A REPORT OF EIGHT CASES, *Lancet*. 318 (1981) 598–600.

[178] M.S. Gottlieb, R. Schroff, H.M. Schanker, J.D. Weisman, P.T. Fan, R.A. Wolf, A. Saxon, Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men, *N. Engl. J. Med.* 305 (1981) 1425–1431.

[179] F. Barré-Sinoussi, J.C. Chermann, F. Rey, M.T. Nugeyre, S. Chamaret, J. Gruest, C. Dauguet, C. Axler-Blin, F. Vézinet-Brun, C. Rouzioux, W. Rozenbaum, L. Montagnier, Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS), *Science*. 220 (1983) 868–871.

[180] R.C. Gallo, P.S. Sarin, E.P. Gelmann, M. Robert-Guroff, E. Richardson, V.S. Kalyanaraman, D. Mann, G.D. Sidhu, R.E. Stahl, S. Zolla-Pazner, J. Leibowitch, M. Popovic, Isolation of Human T-Cell Leukemia Virus in Acquired Immune Deficiency Syndrome (AIDS), *Science*. 220 (1983) 865–867.

[181] A.M. Munis, Gene Therapy Applications of Non-Human Lentiviral Vectors, *Viruses*. 12 (2020) 1106.

[182] R.A. Weiss, How Does HIV Cause AIDS?, *Science*. 260 (1993) 1273–1279.

[183] M.M. Santoro, C.F. Perno, HIV-1 Genetic Variability and Clinical Implications, *ISRN Microbiol.* 2013 (2013) 1–20.

[184] F.E. McCutchan, Global epidemiology of HIV, *J. Med. Virol.* 78 (2006) S7–S12.

[185] N. Bbosa, P. Kaleebu, D. Ssemwanga, HIV subtype diversity worldwide, *Curr. Opin. HIV AIDS.* 14 (2019) 153–160.

[186] A.D. Frankel, J.A.T. Young, HIV-1: Fifteen Proteins and an RNA, *Annu. Rev. Biochem.* 67 (1998) 1–25.

[187] S. Scarlata, C. Carter, Role of HIV-1 Gag domains in viral assembly, *Biochim. Biophys. Acta - Biomembr.* 1614 (2003) 62–72.

[188] M.A. Checkley, B.G. Luttge, E.O. Freed, HIV-1 Envelope Glycoprotein Biosynthesis, Trafficking, and Incorporation, *J. Mol. Biol.* 410 (2011) 582–608.

[189] S. Philpott, HIV-1 Coreceptor Usage, Transmission, and Disease Progression, *Curr. HIV Res.* 1 (2003) 217–227.

[190] R.K. Gupta, D. Peppa, A.L. Hill, C. Gálvez, M. Salgado, M. Pace, L.E. McCoy, S.A. Griffith, J. Thornhill, A. Alrubayyi, L.E.P. Huyveneers, E. Nastouli, P. Grant, S.G. Edwards, A.J. Innes, J. Frater, M. Nijhuis, A.M.J. Wensing, J. Martinez-Picado, E. Olavarria, Evidence for HIV-1 cure after CCR5Δ32/Δ32 allogeneic haemopoietic stem-cell transplantation 30 months post analytical treatment interruption: a case report, *Lancet HIV*. 7 (2020) e340–e347.

[191] B. Chen, Molecular Mechanism of HIV-1 Entry, *Trends Microbiol.* 27 (2019) 878.

[192] K. Bebenek, J. Abbotts, J.D. Roberts, S.H. Wilson, T.A. Kunkel, Specificity and Mechanism of Error-prone Replication by Human Immunodeficiency Virus-1 Reverse Transcriptase, *J. Biol. Chem.* 264 (1989) 16948–16956.

[193] T. Chiu, D. Davies, Structure and Function of HIV-1 Integrase, *Curr. Top. Med. Chem.* 4 (2005) 965–977.

[194] A. Battistini, M. Sgarbanti, HIV-1 Latency: An Update of Molecular Mechanisms and Therapeutic Strategies, *Viruses*. 6 (2014) 1715–1758.

[195] F. Barré-Sinoussi, A.L. Ross, J.-F. Delfraissy, Past, present and future: 30 years of HIV research, *Nat. Rev. Microbiol.* 11 (2013) 877–883.

[196] M. Lu, X. Ma, L.R. Castillo-Menendez, J. Gorman, N. Alsahafi, U. Ermel, D.S. Terry, M. Chambers, D. Peng, B. Zhang, T. Zhou, N. Reichard, K. Wang, J.R. Grover, B.P. Carman,

M.R. Gardner, I. Nikić-Spiegel, A. Sugawara, J. Arthos, E.A. Lemke, A.B. Smith, M. Farzan, C. Abrams, J.B. Munro, A.B. McDermott, A. Finzi, P.D. Kwong, S.C. Blanchard, J.G. Sodroski, W. Mothes, Associating HIV-1 envelope glycoprotein structures with states on the virus observed by smFRET, *Nature*. 568 (2019) 415–419.

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

[198] T. Zhou, L. Xu, B. Dey, A.J. Hessell, D. Van Ryk, S.-H. Xiang, X. Yang, M.-Y. Zhang, M.B. Zwick, J. Arthos, D.R. Burton, D.S. Dimitrov, J. Sodroski, R. Wyatt, G.J. Nabel, P.D. Kwong, Structural definition of a conserved neutralization epitope on HIV-1 gp120, *Nature*. 445 (2007) 732–737.

[199] M. Pancera, S. Majeed, Y.E.A. Ban, L. Chen, C.C. Huang, L. Kong, Y. Do Kwon, J. Stuckey, T. Zhou, J.E. Robinson, W.R. Schief, J. Sodroski, R. Wyatt, P.D. Kwong, 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 (2010) 1166–1171.

[200] C. Huang, M. Tang, M.-Y. Zhang, S. Majeed, E. Montabana, R.L. Stanfield, D.S. Dimitrov, B. Korber, J. Sodroski, I.A. Wilson, R. Wyatt, P.D. Kwong, Structure of a V3-Containing HIV-1 gp120 Core, *Science*. 310 (2005) 1025–1028.

[201] R.W. Sanders, R. Derking, A. Cupo, J.-P.P. Julien, A. Yasmeen, N. de Val, H.J. Kim, C. Blattner, A.T. de la Peña, J. Korzun, M. Golabek, K. de los Reyes, T.J. Ketas, M.J. van Gils, C.R. King, I.A. Wilson, A.B. Ward, P.J. Klasse, J.P. Moore, 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 (2013) e1003618.

[202] J.M. Binley, R.W. Sanders, B. Clas, N. Schuelke, A. Master, Y. Guo, F. Kajumo, D.J. Anselma, P.J. Maddon, W.C. Olson, J.P. Moore, 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 (2000) 627–643.

[203] R.W. Sanders, M. Vesanen, N. Schuelke, A. Master, L. Schiffner, R. Kalyanaraman, M. Paluch, B. Berkhout, P.J. Maddon, W.C. Olson, M. Lu, J.P. Moore, Stabilization of the Soluble, Cleaved, Trimeric Form of the Envelope Glycoprotein Complex of Human Immunodeficiency Virus Type 1, *J. Virol.* 76 (2002) 8875–8889.

[204] X. Wu, A.B. Parast, B.A. Richardson, R. Nduati, G. John-Stewart, D. Mbori-Ngacha, S.M.J. Rainwater, J. Overbaugh, Neutralization Escape Variants of Human Immunodeficiency Virus Type 1 Are Transmitted from Mother to Infant, *J. Virol.* 80 (2006) 835–844.

[205] J.-P.P. Julien, A. Cupo, D. Sok, R.L. Stanfield, D. Lyumkis, M.C. Deller, P.-J.J. Klasse, D.R. Burton, R.W. Sanders, J.P. Moore, A.B. Ward, I.A. Wilson, Crystal Structure of a Soluble Cleaved HIV-1 Envelope Trimer, *Science*. 342 (2013) 1477–1483.

[206] L. Kong, A. Torrents de la Peña, M.C. Deller, F. Garces, K. Sliepen, Y. Hua, R.L. Stanfield, R.W. Sanders, I.A. Wilson, Complete epitopes for vaccine design derived from a crystal structure of the broadly neutralizing antibodies PGT128 and 8ANC195 in complex with an HIV-1 Env trimer, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 71 (2015) 2099–2108.

[207] M. Pancera, T. Zhou, A. Druz, I.S. Georgiev, C. Soto, J. Gorman, J. Huang, P. Acharya, G.-Y. Chuang, G. Ofek, G.B.E. Stewart-Jones, J. Stuckey, R.T. Bailer, M.G. Joyce, M.K. Louder, N. Tumba, Y. Yang, B. Zhang, M.S. Cohen, B.F. Haynes, J.R. Mascola, L. Morris, J.B. Munro, S.C. Blanchard, W. Mothes, M. Connors, P.D. Kwong, Structure and immune recognition of trimeric pre-fusion HIV-1 Env, *Nature*. 514 (2014) 455–461.

[208] M. Vey, W. Schäfer, S. Berghöfer, H.D. Klenk, W. Garten, Maturation of the trans-Golgi network protease furin: compartmentalization of propeptide removal, substrate cleavage, and COOH-terminal truncation., *J. Cell Biol.* 127 (1994) 1829–42.

[209] S.S. Molloy, P.A. Bresnahan, S.H. Leppla, K.R. Klimpel, G. Thomas, Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen., *J. Biol. Chem.* 267 (1992) 16396–402.

[210] L.K. Pritchard, D.J. Harvey, C. Bonomelli, M. Crispin, K.J. Doores, Cell- and Protein-Directed Glycosylation of Native Cleaved HIV-1 Envelope., *J. Virol.* 89 (2015) 8932–44.

[211] R.P. Ringe, R.W. Sanders, A. Yasmeen, H.J. Kim, J.H. Lee, A. Cupo, J. Korzun, R. Derking, T. Van Montfort, J.P. Julien, I.A. Wilson, P.J. Klasse, A.B. Ward, J.P. Moore, Cleavage strongly influences whether soluble HIV-1 envelope glycoprotein trimers adopt a native-like conformation, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 18256–18261.

[212] A.K. Dey, K.B. David, M. Lu, J.P. Moore, Biochemical and biophysical comparison of cleaved and uncleaved soluble, trimeric HIV-1 envelope glycoproteins, *Virology*. 385 (2009) 275–281.

[213] A. Yasmeen, R. Ringe, R. Derking, A. Cupo, J.-P. Julien, D.R. Burton, A.B. Ward, I.A. Wilson, R.W. Sanders, J.P. Moore, P.J. Klasse, Differential binding of neutralizing and non-neutralizing antibodies to native-like soluble HIV-1 Env trimers, uncleaved Env proteins,

and monomeric subunits., *Retrovirology*. 11 (2014) 41.

[214] A.-J. Behrens, D.J. Harvey, E. Milne, A. Cupo, A. Kumar, N. Zitzmann, W.B. Struwe, J.P. Moore, M. Crispin, Molecular Architecture of the Cleavage-Dependent Mannose Patch on a Soluble HIV-1 Envelope Glycoprotein Trimer., *J. Virol.* 91 (2017) e01894-16.

[215] L.M. Walker, M. Huber, K.J. Doores, E. Falkowska, R. Pejchal, J.-P. Julien, S.-K. Wang, A. Ramos, P.-Y. Chan-Hui, M. Moyle, J.L. Mitcham, P.W. Hammond, O.A. Olsen, P. Phung, S. Fling, C.-H. Wong, S. Phogat, T. Wrin, M.D. Simek, P.G. Principal Investigators, W.C. Koff, I.A. Wilson, D.R. Burton, P. Poignard, Broad neutralization coverage of HIV by multiple highly potent antibodies, *Nature*. 477 (2011) 466–470.

[216] E. Falkowska, K.M. Le, A. Ramos, K.J. Doores, J.H. Lee, C. Blattner, A. Ramirez, R. Derking, M.J. van Gils, C.-H. Liang, R. Mcbride, B. von Bredow, S.S. Shivatare, C.-Y. Wu, P.-Y. Chan-Hui, Y. Liu, T. Feizi, M.B. Zwick, W.C. Koff, M.S. Seaman, K. Swiderek, J.P. Moore, D. Evans, J.C. Paulson, C.-H. Wong, A.B. Ward, I.A. Wilson, R.W. Sanders, P. Poignard, D.R. Burton, Broadly Neutralizing HIV Antibodies Define a Glycan-Dependent Epitope on the Prefusion Conformation of gp41 on Cleaved Envelope Trimers, *Immunity*. 40 (2014) 657–668.

[217] A. Sarkar, S. Bale, A.-J. Behrens, S. Kumar, S.K. Sharma, N. de Val, J. Pallesen, A. Irimia, D.C. Diwanji, R.L. Stanfield, A.B. Ward, M. Crispin, R.T. Wyatt, I.A. Wilson, Structure of a cleavage-independent HIV Env recapitulates the glycoprotein architecture of the native cleaved trimer, *Nat. Commun.* 9 (2018) 1956.

[218] S.K. Sharma, N. de Val, S. Bale, J. Guenaga, K. Tran, Y. Feng, V. Dubrovskaya, A.B. Ward, R.T. Wyatt, Cleavage-Independent HIV-1 Env Trimers Engineered as Soluble Native Spike Mimetics for Vaccine Design, *Cell Rep.* 11 (2015) 539–550.

[219] I.S. Georgiev, M.G. Joyce, Y. Yang, M. Sastry, B. Zhang, U. Baxa, R.E. Chen, A. Druz, C.R. Lees, S. Narpala, A. Schön, J. Van Galen, G.-Y. Chuang, J. Gorman, A. Harned, M. Pancera, G.B.E. Stewart-Jones, C. Cheng, E. Freire, A.B. McDermott, J.R. Mascola, P.D. Kwong, Single-Chain Soluble BG505.SOSIP gp140 Trimers as Structural and Antigenic Mimics of Mature Closed HIV-1 Env, *J. Virol.* 89 (2015) 5318–5329.

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

[221] X. Wei, J.M. Decker, S. Wang, H. Hui, J.C. Kappes, X. Wu, J.F. Salazar-Gonzalez, M.G. Salazar, J.M. Kilby, M.S. Saag, N.L. Komarova, M.A. Nowak, B.H. Hahn, P.D. Kwong, G.M. Shaw, Antibody neutralization and escape by HIV-1, *Nature*. 422 (2003) 307–312.

[222] C. Bonomelli, K.J. Doores, D.C. Dunlop, V. Thaney, R.A. Dwek, D.R. Burton, M. Crispin, C.N. Scanlan, The Glycan Shield of HIV Is Predominantly Oligomannose Independently of Production System or Viral Clade, *PLoS One*. 6 (2011) e23521.

[223] C.K. Leonard, M.W. Spellman, L. Riddle, R.J. Harris, J.N. Thomas, T.J. Gregory, 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 (1990) 10373–10382.

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

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

[226] K.J. Doores, C. Bonomelli, D.J. Harvey, S. Vasiljevic, R.A. Dwek, D.R. Burton, M. Crispin, C.N. Scanlan, Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens, *Proc. Natl. Acad. Sci.* 107 (2010) 13800–13805.

[227] K.P. Coss, S. Vasiljevic, L.K. Pritchard, S.A. Krumm, M. Glaze, S. Madzorera, P.L. Moore, M. Crispin, K.J. Doores, HIV-1 Glycan Density Drives the Persistence of the Mannose Patch within an Infected Individual, *J. Virol.* 90 (2016) 11132–11144.

[228] L. Kong, I.A. Wilson, P.D. Kwong, Crystal structure of a fully glycosylated HIV-1 gp120 core reveals a stabilizing role for the glycan at Asn262, *Proteins Struct. Funct. Bioinforma.* 83 (2015) 590–596.

[229] J.H. Lee, G. Ozorowski, A.B. Ward, Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer, *Science*. 351 (2016) 1043–1048.

[230] E.P. Go, A. Herschhorn, C. Gu, L. Castillo-Menendez, S. Zhang, Y. Mao, H. Chen, H. Ding, J.K. Wakefield, D. Hua, H.-X. Liao, J.C. Kappes, J. Sodroski, H. Desaire, Comparative Analysis of the Glycosylation Profiles of Membrane-Anchored HIV-1 Envelope Glycoprotein Trimers and Soluble gp140., *J. Virol.* 89 (2015) 8245–57.

[231] M. Crispin, K.J. Doores, Targeting host-derived glycans on enveloped viruses for antibody-based vaccine design., *Curr. Opin. Virol.* 11 (2015) 63–9.

[232] A.-J.J. Behrens, M. Crispin, Structural principles controlling HIV envelope glycosylation, *Curr. Opin. Struct. Biol.* 44 (2017) 125–133.

[233] J.P. Moore, Y. Cao, D.D. Ho, R.A. Koup, Development of the anti-gp120 antibody response during seroconversion to human immunodeficiency virus type 1, *J. Virol.* 68 (1994) 5142–5155.

[234] I. Mikell, D.N. Sather, S.A. Kalams, M. Altfeld, G. Alter, L. Stamatatos, Characteristics of the Earliest Cross-Neutralizing Antibody Response to HIV-1, *PLOS Pathog.* 7 (2011) e1001251.

[235] P.L. Moore, E.T. Crooks, L. Porter, P. Zhu, C.S. Cayanan, H. Grise, P. Corcoran, M.B. Zwick, M. Franti, L. Morris, K.H. Roux, D.R. Burton, J.M. Binley, Nature of Nonfunctional Envelope Proteins on the Surface of Human Immunodeficiency Virus Type 1, *J. Virol.* 80 (2006) 2515.

[236] D.N. Sather, J. Armann, L.K. Ching, A. Mavrantoni, G. Sellhorn, Z. Caldwell, X. Yu, B. Wood, S. Self, S. Kalams, L. Stamatatos, Factors Associated with the Development of Cross-Reactive Neutralizing Antibodies during Human Immunodeficiency Virus Type 1 Infection, *J. Virol.* 83 (2009) 757–769.

[237] P.L. Moore, E.S. Gray, C.K. Wibmer, J.N. Bhiman, M. Nonyane, D.J. Sheward, T. Hermanus, S. Bajimaya, N.L. Tumba, M.-R. Abrahams, B.E. Lambson, N. Ranchobe, L. Ping, N. Ngandu, Q.A. Karim, S.S.A. Karim, R.I. Swanstrom, M.S. Seaman, C. Williamson, L. Morris, Evolution of an HIV glycan–dependent broadly neutralizing antibody epitope through immune escape, *Nat. Med.* 18 (2012) 1688–1692.

[238] M.S. Seaman, H. Janes, N. Hawkins, L.E. Grandpre, C. Devoy, A. Giri, R.T. Coffey, L. Harris, B. Wood, M.G. Daniels, T. Bhattacharya, A. Lapedes, V.R. Polonis, F.E. McCutchan, P.B. Gilbert, S.G. Self, B.T. Korber, D.C. Montefiori, J.R. Mascola, Tiered Categorization of a Diverse Panel of HIV-1 Env Pseudoviruses for Assessment of Neutralizing Antibodies, *J. Virol.* 84 (2010) 1439.

[239] I. del Moral-Sánchez, K. Sliepen, Strategies for inducing effective neutralizing antibody responses against HIV-1, *Expert Rev. Vaccines*. 18 (2019) 1127.

[240] E. Landais, P.L. Moore, Development of broadly neutralizing antibodies in HIV-1 infected elite neutralizers, *Retrovirology 2018 151*. 15 (2018) 1–14.

[241] C.A. Derdeyn, P.L. Moore, L. Morris, Development of Broadly Neutralizing Antibodies from Autologous Neutralizing Antibody Responses, *Curr. Opin. HIV AIDS*. 9 (2014) 210.

[242] J.H. Lee, R. Andrabi, C.-Y.Y. Su, A. Yasmeen, J.-P.P. Julien, L. Kong, N.C. Wu, R. McBride, D. Sok, M. Pauthner, C.A. Cottrell, T. Nieusma, C. Blattner, J.C. Paulson, P.J. Klasse, I.A. Wilson, D.R. Burton, A.B. Ward, A Broadly Neutralizing Antibody Targets the

Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic β-Hairpin Structure, *Immunity*. 46 (2017) 690–702.

[243] J.M. Steichen, D.W. Kulp, T. Tokatlian, A. Escolano, P. Dosenovic, R.L. Stanfield, L.E. McCoy, G. Ozorowski, X. Hu, O. Kalyuzhniy, B. Briney, T. Schiffner, F. Garces, N.T. Freund, A.D. Gitlin, S. Menis, E. Georgeson, M. Kubitz, Y. Adachi, M. Jones, A.A. Mutafyan, D.S. Yun, C.T. Mayer, A.B. Ward, D.R. Burton, I.A. Wilson, D.J. Irvine, M.C. Nussenzweig, W.R. Schief, HIV Vaccine Design to Target Germline Precursors of Glycan-Dependent Broadly Neutralizing Antibodies, *Immunity*. 45 (2016) 483–496.

[244] P. Dosenovic, L. von Boehmer, A. Escolano, J. Jardine, N.T. Freund, A.D. Gitlin, A.T. McGuire, D.W. Kulp, T. Oliveira, L. Scharf, J. Pietzsch, M.D. Gray, A. Cupo, M.J. van Gils, K.-H. Yao, C. Liu, A. Gazumyan, M.S. Seaman, P.J. Björkman, R.W. Sanders, J.P. Moore, L. Stamatatos, W.R. Schief, M.C. Nussenzweig, Immunization for HIV-1 Broadly Neutralizing Antibodies in Human Ig Knockin Mice, *Cell*. 161 (2015) 1505–1515.

[245] A.B. Ward, I.A. Wilson, Insights into the trimeric HIV-1 envelope glycoprotein structure, *Trends Biochem. Sci.* 40 (2015) 101–107.

[246] D.R. Burton, L. Hangartner, Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design, *Annu. Rev. Immunol.* 34 (2016) 635–659.

[247] B.F. Haynes, New approaches to HIV vaccine development, *Curr. Opin. Immunol.* 35 (2015) 39–47.

[248] J.H. Lee, N. de Val, D. Lyumkis, A.B. Ward, Model Building and Refinement of a Natively Glycosylated HIV-1 Env Protein by High-Resolution Cryoelectron Microscopy, *Structure*. 23 (2015) 1943–1951.

[249] M. Crispin, A.B. Ward, I.A. Wilson, Structure and Immune Recognition of the HIV Glycan Shield, *Annu. Rev. Biophys.* 47 (2018) 499–523.

[250] J.S. McLellan, M. Pancera, C. Carrico, J. Gorman, J.-P. Julien, R. Khayat, R. Louder, R. Pejchal, M. Sastry, K. Dai, S. O'Dell, N. Patel, S. Shahzad-ul-Hussan, Y. Yang, B. Zhang, T. Zhou, J. Zhu, J.C. Boyington, G.-Y. Chuang, D. Diwanji, I. Georgiev, Y. Do Kwon, D. Lee, M.K. Louder, S. Moquin, S.D. Schmidt, Z.-Y. Yang, M. Bonsignori, J.A. Crump, S.H. Kapiga, N.E. Sam, B.F. Haynes, D.R. Burton, W.C. Koff, L.M. Walker, S. Phogat, R. Wyatt, J. Orwenyo, L.-X. Wang, J. Arthos, C.A. Bewley, J.R. Mascola, G.J. Nabel, W.R. Schief, A.B. Ward, I.A. Wilson, P.D. Kwong, Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9, *Nature*. 480 (2011) 336–343.

[251] M. Pancera, S. Shahzad-Ul-Hussan, N.A. Doria-Rose, J.S. McLellan, R.T. Bailer, K. Dai, S.

Loesgen, M.K. Louder, R.P. Staupe, Y. Yang, B. Zhang, R. Parks, J. Eudailey, K.E. Lloyd, J. Blinn, S.M. Alam, B.F. Haynes, M.N. Amin, L.-X. Wang, D.R. Burton, W.C. Koff, G.J. Nabel, J.R. Mascola, C.A. Bewley, P.D. Kwong, Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16., *Nat. Struct. Mol. Biol.* 20 (2013) 804–13.

[252] D. Sok, M.J.V. Gils, M. Pauthner, J.P. Julien, K.L. Saye-Francisco, J. Hsueh, B. Briney, J.H. Lee, K.M. Le, P.S. Lee, Y. Hua, M.S. Seaman, J.P. Moore, A.B. Ward, I.A. Wilson, R.W. Sanders, D.R. Burton, Recombinant HIV envelope trimer selects for quaternary-dependent antibodies targeting the trimer apex, *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) 17624–17629.

[253] N.A. Doria-Rose, J.N. Bhiman, R.S. Roark, C.A. Schramm, J. Gorman, G.-Y. Chuang, M. Pancera, E.M. Cale, M.J. Ernandes, M.K. Louder, M. Asokan, R.T. Bailer, A. Druz, I.R. Fraschilla, N.J. Garrett, M. Jarosinski, R.M. Lynch, K. McKee, S. O'Dell, A. Pegu, S.D. Schmidt, R.P. Staupe, M.S. Sutton, K. Wang, C.K. Wibmer, B.F. Haynes, S. Abdool-Karim, L. Shapiro, P.D. Kwong, P.L. Moore, L. Morris, J.R. Mascola, New Member of the V1V2-Directed CAP256-VRC26 Lineage That Shows Increased Breadth and Exceptional Potency, *J. Virol.* 90 (2016) 76–91.

[254] A. Cupo, V.M. Cruz Portillo, P. Gelfand, A. Yasmeen, P.J. Klasse, J.P. Moore, Optimizing the production and affinity purification of HIV-1 envelope glycoprotein SOSIP trimers from transiently transfected CHO cells, *PLoS One*. 14 (2019) e0215106.

[255] P.J.M. Brouwer, A. Antanasijevic, M. de Gast, J.D. Allen, T.P.L. Bijl, A. Yasmeen, R. Ravichandran, J.A. Burger, G. Ozorowski, J.L. Torres, C. LaBranche, D.C. Montefiori, R.P. Ringe, M.J. van Gils, J.P. Moore, P.J. Klasse, M. Crispin, N.P. King, A.B. Ward, R.W. Sanders, Immunofocusing and enhancing autologous Tier-2 HIV-1 neutralization by displaying Env trimers on two-component protein nanoparticles, *Npj Vaccines*. 6 (2021) 24.

[256] S.W. de Taeye, J.P. Moore, R.W. Sanders, HIV-1 Envelope Trimer Design and Immunization Strategies To Induce Broadly Neutralizing Antibodies, *Trends Immunol.* 37 (2016) 221–232.

[257] K. Rantalainen, Z.T. Berndsen, S. Murrell, L. Cao, O. Omorodion, J.L. Torres, M. Wu, J. Umotoy, J. Copps, P. Poignard, E. Landais, J.C. Paulson, I.A. Wilson, A.B. Ward, Co-evolution of HIV Envelope and Apex-Targeting Neutralizing Antibody Lineage Provides Benchmarks for Vaccine Design, *Cell Rep.* 23 (2018) 3249–3261.

[258] L. Cao, M. Pauthner, R. Andrabi, K. Rantalainen, Z. Berndsen, J.K. Diedrich, S. Menis, D. Sok, R. Bastidas, S.-K.R. Park, C.M. Delahunty, L. He, J. Guenaga, R.T. Wyatt, W.R.

Schief, A.B. Ward, J.R. Yates, D.R. Burton, J.C. Paulson, Differential processing of HIV envelope glycans on the virus and soluble recombinant trimer, *Nat. Commun.* 9 (2018) 3693.

[259] C.A. Cottrell, J. van Schooten, C.A. Bowman, M. Yuan, D. Oyen, M. Shin, R. Morpurgo, P. van der Woude, M. van Breemen, J.L. Torres, R. Patel, J. Gross, L.M. Sewall, J. Copps, G. Ozorowski, B. Nogal, D. Sok, E.G. Rakasz, C. Labranche, V. Vigdorovich, S. Christley, D.G. Carnathan, D.N. Sather, D. Montefiori, G. Silvestri, D.R. Burton, J.P. Moore, I.A. Wilson, R.W. Sanders, A.B. Ward, M.J. van Gils, Mapping the immunogenic landscape of near-native HIV-1 envelope trimers in non-human primates, *PLOS Pathog.* 16 (2020) e1008753.

[260] R.W. Sanders, M.J. van Gils, R. Derking, D. Sok, T.J. Ketas, J.A. Burger, G. Ozorowski, A. Cupo, C. Simonich, L. Goo, H. Arendt, H.J. Kim, J.H. Lee, P. Pugach, M. Williams, G. Debnath, B. Moldt, M.J. Van Breemen, G. Isik, M. Medina-Ramírez, J.W. Back, W.C. Koff, J.-P.P. Julien, E.G. Rakasz, M.S. Seaman, M. Guttman, K.K. Lee, P.J. Klasse, C. LaBranche, W.R. Schief, I.A. Wilson, J. Overbaugh, D.R. Burton, A.B. Ward, D.C. Montefiori, H. Dean, J.P. Moore, HIV-1 neutralizing antibodies induced by native-like envelope trimers, *Science*. 349 (2015) aac4223.

[261] L.E. McCoy, M.J. van Gils, G. Ozorowski, T. Messmer, B. Briney, J.E. Voss, D.W. Kulp, M.S. Macauley, D. Sok, M. Pauthner, S. Menis, C.A. Cottrell, J.L. Torres, J. Hsueh, W.R. Schief, I.A. Wilson, A.B. Ward, R.W. Sanders, D.R. Burton, Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies, *Cell Rep.* 16 (2016) 2327–2338.

[262] P.J. Klasse, T.J. Ketas, C.A. Cottrell, G. Ozorowski, G. Debnath, D. Camara, E. Francomano, P. Pugach, R.P. Ringe, C.C. LaBranche, M.J. van Gils, C.A. Bricault, D.H. Barouch, S. Crotty, G. Silvestri, S. Kasturi, B. Pulendran, I.A. Wilson, D.C. Montefiori, R.W. Sanders, A.B. Ward, J.P. Moore, Epitopes for neutralizing antibodies induced by HIV-1 envelope glycoprotein BG505 SOSIP trimers in rabbits and macaques, *PLOS Pathog.* 14 (2018) e1006913.

[263] R.P. Ringe, P. Pugach, C.A. Cottrell, C.C. LaBranche, G.E. Seabright, T.J. Ketas, G. Ozorowski, S. Kumar, A. Schorcht, M.J. van Gils, M. Crispin, D.C. Montefiori, I.A. Wilson, A.B. Ward, R.W. Sanders, P.J. Klasse, J.P. Moore, Closing and Opening Holes in the Glycan Shield of HIV-1 Envelope Glycoprotein SOSIP Trimers Can Redirect the Neutralizing Antibody Response to the Newly Unmasked Epitopes, *J. Virol.* 93 (2019) 1656–1674.

[264] H. Duan, X. Chen, J.C. Boyington, C. Cheng, Y. Zhang, A.J. Jafari, T. Stephens, Y.

Tsybovsky, O. Kalyuzhniy, P. Zhao, S. Menis, M.C. Nason, E. Normandin, M. Mukhamedova, B.J. DeKosky, L. Wells, W.R. Schief, M. Tian, F.W. Alt, P.D. Kwong, J.R. Mascola, Glycan Masking Focuses Immune Responses to the HIV-1 CD4-Binding Site and Enhances Elicitation of VRC01-Class Precursor Antibodies, *Immunity*. 49 (2018) 301-311.e5.

[265] J. Jardine, J.-P. Julien, S. Menis, T. Ota, O. Kalyuzhniy, A. McGuire, D. Sok, P.-S. Huang, S. MacPherson, M. Jones, T. Nieusma, J. Mathison, D. Baker, A.B. Ward, D.R. Burton, L. Stamatatos, D. Nemazee, I.A. Wilson, W.R. Schief, Rational HIV Immunogen Design to Target Specific Germline B Cell Receptors, *Science*. 340 (2013) 711–716.

[266] K. Wagh, E.F. Kreider, Y. Li, H.J. Barbian, G.H. Learn, E. Giorgi, P.T. Hraber, T.G. Decker, A.G. Smith, M. V. Gondim, L. Gillis, J. Wandzilak, G.-Y. Chuang, R. Rawi, F. Cai, P. Pellegrino, I. Williams, J. Overbaugh, F. Gao, P.D. Kwong, B.F. Haynes, G.M. Shaw, P. Borrow, M.S. Seaman, B.H. Hahn, B. Korber, Completeness of HIV-1 Envelope Glycan Shield at Transmission Determines Neutralization Breadth, *Cell Rep.* 25 (2018) 893-908.e7.

[267] Y.-W. Huang, H.-I. Yang, Y.-T. Wu, T.-L. Hsu, T.-W. Lin, J.W. Kelly, C.-H. Wong, Residues Comprising the Enhanced Aromatic Sequon Influence Protein N-Glycosylation Efficiency, *J. Am. Chem. Soc*. 139 (2017) 8.

[268] R. Derking, J.D. Allen, C.A. Cottrell, K. Sliepen, G.E. Seabright, W.-H. Lee, Y. Aldon, K. Rantalainen, A. Antanasijevic, J. Copps, A. Yasmeen, A. Cupo, V.M. Cruz Portillo, M. Poniman, N. Bol, P. van der Woude, S.W. de Taeye, T.L.G.M. van den Kerkhof, P.J. Klasse, G. Ozorowski, M.J. van Gils, J.P. Moore, A.B. Ward, M. Crispin, R.W. Sanders, Enhancing glycan occupancy of soluble HIV-1 envelope trimers to mimic the native viral spike, *Cell Rep.* 35 (2021) 108933.

[269] S. Gerber, C. Lizak, G. Michaud, M. Bucher, T. Darbre, M. Aebi, J.-L. Reymond, K.P. Locher, Mechanism of Bacterial Oligosaccharyltransferase, *J. Biol. Chem.* 288 (2013) 8849–8861.

[270] E. Bause, Model studies on N-glycosylation of proteins, *Biochem. Soc. Trans.* 12 (1984) 514–517.

[271] L. Kasturi, J.R. Eshleman, W.H. Wunner, S.H. Shakin-Eshleman, The Hydroxy Amino Acid in an Asn-X-Ser/Thr Sequon Can Influence N-Linked Core Glycosylation Efficiency and the Level of Expression of a Cell Surface Glycoprotein, *J. Biol. Chem.* 270 (1995) 14756–14761.

[272] L. Kong, N.C. Sheppard, G.B.E.E. Stewart-Jones, C.L. Robson, H. Chen, X. Xu, G. Krashias, C. Bonomelli, C.N. Scanlan, P.D. Kwong, S.A. Jeffs, I.M. Jones, Q.J. Sattentau,

Expression-system-dependent modulation of HIV-1 envelope glycoprotein antigenicity and immunogenicity., *J. Mol. Biol.* 403 (2010) 131–147.

[273] P.R. Crocker, J.C. Paulson, A. Varki, Siglecs and their roles in the immune system, *Nat. Rev. Immunol.* 7 (2007) 255–266.

[274] A.K. Dey, A. Cupo, G. Ozorowski, V.K. Sharma, A.J. Behrens, E.P. Go, T.J. Ketas, A. Yasmeen, P.J. Klasse, E. Sayeed, H. Desaire, M. Crispin, I.A. Wilson, R.W. Sanders, T. Hassell, A.B. Ward, J.P. Moore, cGMP production and analysis of BG505 SOSIP.664, an extensively glycosylated, trimeric HIV-1 envelope glycoprotein vaccine candidate, *Biotechnol. Bioeng.* 115 (2018) 885.

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

[276] E. Falkowska, K.M. Le, A. Ramos, K.J. Doores, J.H. Lee, C. Blattner, A. Ramirez, R. Derking, M.J. van Gils, C.-H. Liang, R. Mcbride, B. von Bredow, S.S. Shivatare, C.-Y. Wu, P.-Y. Chan-Hui, Y. Liu, T. Feizi, M.B. Zwick, W.C. Koff, M.S. Seaman, K. Swiderek, J.P. Moore, D. Evans, J.C. Paulson, C.-H. Wong, A.B. Ward, I.A. Wilson, R.W. Sanders, P. Poignard, D.R. Burton, Broadly Neutralizing HIV Antibodies Define a Glycan-Dependent Epitope on the Prefusion Conformation of gp41 on Cleaved Envelope Trimers, *Immunity*. 40 (2014) 657–668.

[277] V.S. Shivatare, S.S. Shivatare, C.-C.D. Lee, C.-H. Liang, K.-S. Liao, Y.-Y. Cheng, G. Saidachary, C.-Y. Wu, N.-H. Lin, P.D. Kwong, D.R. Burton, C.-Y. Wu, C.-H. Wong, Unprecedented Role of Hybrid N-Glycans as Ligands for HIV-1 Broadly Neutralizing Antibodies., *J. Am. Chem. Soc.* 140 (2018) 5202–5210.

[278] Z. Zou, A. Chastain, S. Moir, J. Ford, K. Trandem, E. Martinelli, C. Cicala, P. Crocker, J. Arthos, P.D. Sun, Siglecs Facilitate HIV-1 Infection of Macrophages through Adhesion with Viral Sialic Acids, *PLoS One*. 6 (2011) e24559.

[279] S.W. De Taeye, G. Ozorowski, A. Torrents De La Peña, M. Guttman, J.-P.P. Julien, T.L.G.M. Van Den Kerkhof, J.A. Burger, L.K. Pritchard, P. Pugach, A. Yasmeen, J. Crampton, J. Hu, I. Bontjer, J.L. Torres, H. Arendt, J. DeStefano, W.C. Koff, H. Schuitemaker, D. Eggink, B. Berkhout, H. Dean, C. LaBranche, S. Crotty, M. Crispin, D.C. Montefiori, P.J.J. Klasse, K.K. Lee, J.P. Moore, I.A. Wilson, A.B. Ward, R.W. Sanders, Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes, *Cell*. 163 (2015) 1702–1715.

[280] I. del Moral-Sánchez, R.A. Russell, E.E. Schermer, C.A. Cottrell, J.D. Allen, A. Torrents de la Peña, C.C. LaBranche, S. Kumar, M. Crispin, A.B. Ward, D.C. Montefiori, Q.J. Sattentau, K. Sliepen, R.W. Sanders, High thermostability improves neutralizing antibody responses induced by native-like HIV-1 envelope trimers, *Npj Vaccines*. 7 (2022) 27.

[281] P.J. Klasse, C.C. LaBranche, T.J. Ketas, G. Ozorowski, A. Cupo, P. Pugach, R.P. Ringe, M. Golabek, M.J. van Gils, M. Guttman, K.K. Lee, I.A. Wilson, S.T. Butera, A.B. Ward, D.C. Montefiori, R.W. Sanders, J.P. Moore, Sequential and Simultaneous Immunization of Rabbits with HIV-1 Envelope Glycoprotein SOSIP.664 Trimers from Clades A, B and C, *PLOS Pathog.* 12 (2016) e1005864.

[282] Y. Aldon, P.F. McKay, J. Allen, G. Ozorowski, R. Felfödiné Lévai, M. Tolazzi, P. Rogers, L. He, N. de Val, K. Fábián, G. Scarlatti, J. Zhu, A.B. Ward, M. Crispin, R.J. Shattock, Rational Design of DNA-Expressed Stabilized Native-Like HIV-1 Envelope Trimers, *Cell Rep.* 24 (2018) 3324-3338.e5.

[283] A. Torrents de la Peña, J.-P. Julien, S.W. de Taeye, F. Garces, M. Guttman, G. Ozorowski, L.K. Pritchard, A.-J. Behrens, E.P. Go, J.A. Burger, E.E. Schermer, K. Sliepen, T.J. Ketas, P. Pugach, A. Yasmeen, C.A. Cottrell, J.L. Torres, C.D. Vavourakis, M.J. van Gils, C. LaBranche, D.C. Montefiori, H. Desaire, M. Crispin, P.J. Klasse, K.K. Lee, J.P. Moore, A.B. Ward, I.A. Wilson, R.W. Sanders, Improving the Immunogenicity of Native-like HIV-1 Envelope Trimers by Hyperstabilization, *Cell Rep.* 20 (2017) 1805–1817.

[284] D. Sok, K.M. Le, M. Vadnais, K.L. Saye-Francisco, J.G. Jardine, J.L. Torres, Z.T. Berndsen, L. Kong, R. Stanfield, J. Ruiz, A. Ramos, C.-H. Liang, P.L. Chen, M.F. Criscitiello, W. Mwangi, I.A. Wilson, A.B. Ward, V. V Smider, D.R. Burton, Rapid elicitation of broadly neutralizing antibodies to HIV by immunization in cows, *Nature*. 548 (2017) 108–111.

[285] B. Heydarchi, D.S. Fong, H. Gao, N.A. Salazar-Quiroz, J.M. Edwards, C.A. Gonelli, S. Grimley, T.E. Aktepe, C. Mackenzie, W.J. Wales, M.J. van Gils, A. Cupo, I. Rouiller, P.R. Gooley, J.P. Moore, R.W. Sanders, D. Montefiori, A. Sethi, D.F.J. Purcell, Broad and ultra-potent cross-clade neutralization of HIV-1 by a vaccine-induced CD4 binding site bovine antibody, *Cell Reports Med.* 3 (2022) 100635.

[286] B.F. Haynes, K. Wiehe, P. Borrow, K.O. Saunders, B. Korber, K. Wagh, A.J. McMichael, G. Kelsoe, B.H. Hahn, F. Alt, G.M. Shaw, Strategies for HIV-1 vaccines that induce broadly neutralizing antibodies, *Nat. Rev. Immunol.* 23 (2023) 142–158.

[287] J.M. Steichen, Y.-C. Lin, C. Havenar-Daughton, S. Pecetta, G. Ozorowski, J.R. Willis, L. Toy, D. Sok, A. Liguori, S. Kratochvil, J.L. Torres, O. Kalyuzhniy, E. Melzi, D.W. Kulp, S.

Raemisch, X. Hu, S.M. Bernard, E. Georgeson, N. Phelps, Y. Adachi, M. Kubitz, E. Landais, J. Umotoy, A. Robinson, B. Briney, I.A. Wilson, D.R. Burton, A.B. Ward, S. Crotty, F.D. Batista, W.R. Schief, A generalized HIV vaccine design strategy for priming of broadly neutralizing antibody responses., *Science*. 366 (2019) eaax4380.

[288] D.J. Leggat, K.W. Cohen, J.R. Willis, W.J. Fulp, A.C. DeCamp, O. Kalyuzhniy, C.A. Cottrell, S. Menis, G. Finak, L. Ballweber-Fleming, A. Srikanth, J.R. Plyler, T. Schiffner, A. Liguori, F. Rahaman, A. Lombardo, V. Philiponis, R.E. Whaley, A. Seese, J. Brand, A.M. Ruppel, W. Hoyland, N.L. Yates, L.D. Williams, K. Greene, H. Gao, C.R. Mahoney, M.M. Corcoran, A. Cagigi, A. Taylor, D.M. Brown, D.R. Ambrozak, T. Sincomb, X. Hu, R. Tingle, E. Georgeson, S. Eskandarzadeh, N. Alavi, D. Lu, T.-M. Mullen, M. Kubitz, B. Groschel, J. Maenza, O. Kolokythas, N. Khati, J. Bethony, S. Crotty, M. Roederer, G.B. Karlsson Hedestam, G.D. Tomaras, D. Montefiori, D. Diemert, R.A. Koup, D.S. Laufer, M.J. McElrath, A.B. McDermott, W.R. Schief, Vaccination induces HIV broadly neutralizing antibody precursors in humans., *Science*. 378 (2022) eadd6502.

[289] Study Record | ClinicalTrials.gov, (n.d.). https://clinicaltrials.gov/study/NCT05414786.

[290] K.W. Cohen, S.C. De Rosa, W.J. Fulp, A.C. deCamp, A. Fiore-Gartland, C.R. Mahoney, S. Furth, J. Donahue, R.E. Whaley, L. Ballweber-Fleming, A. Seese, K. Schwedhelm, D. Geraghty, G. Finak, S. Menis, D.J. Leggat, F. Rahaman, A. Lombardo, B.R. Borate, V. Philiponis, J. Maenza, D. Diemert, O. Kolokythas, N. Khati, J. Bethony, O. Hyrien, D.S. Laufer, R.A. Koup, A.B. McDermott, W.R. Schief, M.J. McElrath, A first-in-human germline-targeting HIV nanoparticle vaccine induced broad and publicly targeted helper T cell responses, *Sci. Transl. Med.* 15 (2023) eadf3309.

[291] C. Sun, T. Zuo, Z. Wen, First clinical study of germline-targeting strategy: One step closer to a successful bnAb-based HIV vaccine, *Innov.* 4 (2023) 100374.

[292] P.T. Wingfield, Overview of the Purification of Recombinant Proteins, *Curr. Protoc. Protein Sci.* 80 (2015) 6.1.1–6.1.35.

[293] D. Weissman, mRNA transcript therapy, *Expert Rev. Vaccines*. 14 (2015) 265–281.

[294] N. Nafissi, S. Alqawlaq, E.A. Lee, M. Foldvari, P.A. Spagnuolo, R.A. Slavcev, DNA Ministrings: Highly Safe and Effective Gene Delivery Vectors, *Mol. Ther. - Nucleic Acids*. 3 (2014) e165.

[295] D. Hobernik, M. Bros, DNA Vaccines—How Far From Clinical Use?, *Int. J. Mol. Sci.* 19 (2018) 3605.

[296] N. Pardi, M.J. Hogan, F.W. Porter, D. Weissman, mRNA vaccines — a new era in

vaccinology, *Nat. Rev. Drug Discov.* 17 (2018) 261–279.

[297] N. Chaudhary, D. Weissman, K.A. Whitehead, mRNA vaccines for infectious diseases: principles, delivery and clinical translation, *Nat. Rev. Drug Discov.* 20 (2021) 817–838.

[298] J.A. Wolff, R.W. Malone, P. Williams, W. Chong, G. Acsadi, A. Jani, P.L. Felgner, Direct gene transfer into mouse muscle in vivo, *Science*. 247 (1990) 1465–1468.

[299] L.R. Baden, H.M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S.A. Spector, N. Rouphael, C.B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L.L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B.S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, T. Zaks, Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine, *N. Engl. J. Med.* 384 (2021) 403–416.

[300] F.P. Polack, S.J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J.L. Perez, G. Pérez Marc, E.D. Moreira, C. Zerbini, R. Bailey, K.A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R.W. Frenck, L.L. Hammitt, Ö. Türeci, H. Nell, A. Schaefer, S. Ünal, D.B. Tresnan, S. Mather, P.R. Dormitzer, U. Şahin, K.U. Jansen, W.C. Gruber, Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine, *N. Engl. J. Med.* 383 (2020) 2603–2615.

[301] M.N. Uddin, M.A. Roni, Challenges of Storage and Stability of mRNA-Based COVID-19 Vaccines, *Vaccines*. 9 (2021) 1033.

[302] M.N. Fleeton, M. Chen, P. Berglund, G. Rhodes, S.E. Parker, M. Murphy, G.J. Atkins, P. Liljeström, Self-replicative RNA vaccines elicit protection against influenza A virus, respiratory syncytial virus, and a tickborne encephalitis virus., *J. Infect. Dis.* 183 (2001) 1395–8.

[303] I.T. Lee, R. Nachbagauer, D. Ensz, H. Schwartz, L. Carmona, K. Schaefers, A. Avanesov, D. Stadlbauer, C. Henry, R. Chen, W. Huang, D.R. Schrempp, J. Ananworanich, R. Paris, Safety and immunogenicity of a phase 1/2 randomized clinical trial of a quadrivalent, mRNA-based seasonal influenza vaccine (mRNA-1010) in healthy adults: interim analysis, *Nat. Commun.* 14 (2023) 3631.

[304] P.F. McKay, K. Hu, A.K. Blakney, K. Samnuan, J.C. Brown, R. Penn, J. Zhou, C.R. Bouton, P. Rogers, K. Polra, P.J.C. Lin, C. Barbosa, Y.K. Tam, W.S. Barclay, R.J. Shattock, Self-amplifying RNA SARS-CoV-2 lipid nanoparticle vaccine candidate induces high neutralizing antibody titers in mice, *Nat. Commun.* 11 (2020) 3523.

[305] K. Bloom, F. van den Berg, P. Arbuthnot, Self-amplifying RNA vaccines for infectious diseases, *Gene Ther.* 28 (2021) 117–129.

[306] V. Krähling, S. Erbar, A. Kupke, S.S. Nogueira, K.C. Walzer, H. Berger, E. Dietzel, S. Halwe, C. Rohde, L. Sauerhering, L. Aragão-Santiago, J. Moreno Herrero, S. Witzel, H. Haas, S. Becker, U. Sahin, Self-amplifying RNA vaccine protects mice against lethal Ebola virus infection, *Mol. Ther.* 31 (2023) 374–386.

[307] A.J. Geall, C.W. Mandl, J.B. Ulmer, RNA: The new revolution in nucleic acid vaccines, *Semin. Immunol.* 25 (2013) 152–159.

[308] K. Karikó, H. Muramatsu, F.A. Welsh, J. Ludwig, H. Kato, S. Akira, D. Weissman, Incorporation of Pseudouridine Into mRNA Yields Superior Nonimmunogenic Vector With Increased Translational Capacity and Biological Stability, *Mol. Ther.* 16 (2008) 1833–1840.

[309] A.S. Espeseth, P.J. Cejas, M.P. Citron, D. Wang, D.J. DiStefano, C. Callahan, G.O. Donnell, J.D. Galli, R. Swoyer, S. Touch, Z. Wen, J. Antonello, L. Zhang, J.A. Flynn, K.S. Cox, D.C. Freed, K.A. Vora, K. Bahl, A.H. Latham, J.S. Smith, M.E. Gindy, G. Ciaramella, D. Hazuda, C.A. Shaw, A.J. Bett, Modified mRNA/lipid nanoparticle-based vaccines expressing respiratory syncytial virus F protein variants are immunogenic and protective in rodent models of RSV infection, *Npj Vaccines*. 5 (2020) 16.

[310] S. Chang Kim, S. Singh Sekhon, W.-R. Shin, G. Ahn, B.-K. Cho, J.-Y. Ahn, Y.-H. Kim, Modifications of mRNA vaccine structural elements for improving mRNA stability and translation efficiency, 18 (2022) 1–8.

[311] J.B. Sandbrink, R.J. Shattock, RNA Vaccines: A Suitable Platform for Tackling Emerging Pandemics?, *Front. Immunol.* 11 (2020) 608460.

[312] A.B. Vogel, L. Lambert, E. Kinnear, D. Busse, S. Erbar, K.C. Reuter, L. Wicke, M. Perkovic, T. Beissert, H. Haas, S.T. Reece, U. Sahin, J.S. Tregoning, Self-Amplifying RNA Vaccines Give Equivalent Protection against Influenza to mRNA Vaccines but at Much Lower Doses, *Mol. Ther.* 26 (2018) 446–455.

[313] M. Melo, E. Porter, Y. Zhang, M. Silva, N. Li, B. Dobosh, A. Liguori, P. Skog, E. Landais, S. Menis, D. Sok, D. Nemazee, W.R. Schief, R. Weiss, D.J. Irvine, Immunogenicity of RNA Replicons Encoding HIV Env Immunogens Designed for Self-Assembly into Nanoparticles, *Mol. Ther.* 27 (2019) 2080–2090.

[314] A. Utt, T. Quirin, S. Saul, K. Hellström, T. Ahola, A. Merits, Versatile Trans-Replication Systems for Chikungunya Virus Allow Functional Analysis and Tagging of Every Replicase Protein, *PLoS One*. 11 (2016) e0151616.

[315] A. Olshefsky, C. Richardson, S.H. Pun, N.P. King, Engineering Self-Assembling Protein Nanoparticles for Therapeutic Delivery, *Bioconjug. Chem.* 33 (2022) 2018–2034.

[316] J.P. Bost, H. Barriga, M.N. Holme, A. Gallud, M. Maugeri, D. Gupta, T. Lehto, H. Valadi, E.K. Esbjörner, M.M. Stevens, S. El-Andaloussi, Delivery of Oligonucleotide Therapeutics: Chemical Modifications, Lipid Nanoparticles, and Extracellular Vesicles, *ACS Nano*. 15 (2021) 13993–14021.

[317] X. Hou, T. Zaks, R. Langer, Y. Dong, Lipid nanoparticles for mRNA delivery, *Nat. Rev. Mater.* 6 (2021) 1078–1094.

[318] L. Sissoëff, M. Mousli, P. England, C. Tuffereau, Stable trimerization of recombinant rabies virus glycoprotein ectodomain is required for interaction with the p75NTR receptor, *J. Gen. Virol.* 86 (2005) 2543–2552.

[319] Y. Tao, S. V Strelkov, V. V Mesyanzhinov, M.G. Rossmann, Structure of bacteriophage T4 fibritin: a segmented coiled coil and the role of the C-terminal domain, (n.d.). http://biomednet.com/elecref/0969212600500789 (accessed November 12, 2023).

[320] X. Yang, J. Lee, E.M. Mahony, P.D. Kwong, R. Wyatt, J. Sodroski, Highly Stable Trimers Formed by Human Immunodeficiency Virus Type 1 Envelope Glycoproteins Fused with the Trimeric Motif of T4 Bacteriophage Fibritin, *J. Virol.* 76 (2002) 4634–4642.

[321] J.M. Binley, R.W. Sanders, A. Master, C.S. Cayanan, C.L. Wiley, L. Schiffner, B. Travis, S. Kuhmann, D.R. Burton, S.-L. Hu, W.C. Olson, J.P. Moore, Enhancing the Proteolytic Maturation of Human Immunodeficiency Virus Type 1 Envelope Glycoproteins, *J. Virol.* 76 (2002) 2606–2616.

[322] K.S. Corbett, D.K. Edwards, S.R. Leist, O.M. Abiona, S. Boyoglu-Barnum, R.A. Gillespie, S. Himansu, A. Schäfer, C.T. Ziwawo, A.T. DiPiazza, K.H. Dinnon, S.M. Elbashir, C.A. Shaw, A. Woods, E.J. Fritch, D.R. Martinez, K.W. Bock, M. Minai, B.M. Nagata, G.B. Hutchinson, K. Wu, C. Henry, K. Bahl, D. Garcia-Dominguez, L.Z. Ma, I. Renzi, W.P. Kong, S.D. Schmidt, L. Wang, Y. Zhang, E. Phung, L.A. Chang, R.J. Loomis, N.E. Altaras, E. Narayanan, M. Metkar, V. Presnyak, C. Liu, M.K. Louder, W. Shi, K. Leung, E.S. Yang, A. West, K.L. Gully, L.J. Stevens, N. Wang, D. Wrapp, N.A. Doria-Rose, G. Stewart-Jones, H. Bennett, G.S. Alvarado, M.C. Nason, T.J. Ruckwardt, J.S. McLellan, M.R. Denison, J.D. Chappell, I.N. Moore, K.M. Morabito, J.R. Mascola, R.S. Baric, A. Carfi, B.S. Graham, SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness, *Nature*. 586 (2020) 567–571.

[323] M. Sousa De Almeida, E. Susnik, B. Drasler, P. Taladriz-Blanco, A. Petri-Fink, B. Rothen-Rutishauser, Understanding nanoparticle endocytosis to improve targeting strategies in

nanomedicine, *Chem. Soc. Rev.* 50 (2021) 5397.

[324] M.L. Guevara, F. Persano, S. Persano, Advances in Lipid Nanoparticles for mRNA-Based Cancer Immunotherapy, *Front. Chem.* 8 (2020).

[325] C. Hald Albertsen, J.A. Kulkarni, D. Witzigmann, M. Lind, K. Petersson, J.B. Simonsen, The role of lipid components in lipid nanoparticles for vaccines and gene therapy., *Adv. Drug Deliv. Rev.* 188 (2022) 114416.

[326] K.A. Hajj, K.A. Whitehead, Tools for translation: non-viral materials for therapeutic mRNA delivery, *Nat. Rev. Mater.* 2 (2017) 17056.

[327] Q. Xiong, G.Y. Lee, J. Ding, W. Li, J. Shi, Biomedical applications of mRNA nanomedicine, *Nano Res.* 11 (2018) 5281.

[328] S. Guan, J. Rosenecker, Nanotechnologies in delivery of mRNA therapeutics using nonviral vector-based delivery systems, *Gene Ther.* 24 (2017) 133–143.

[329] P.S. Kowalski, A. Rudra, L. Miao, D.G. Anderson, Delivering the Messenger: Advances in Technologies for Therapeutic mRNA Delivery, *Mol. Ther.* 27 (2019) 710–728.

[330] M.S. Gebre, L.A. Brito, L.H. Tostanoski, D.K. Edwards, A. Carfi, D.H. Barouch, Novel approaches for vaccine development, *Cell*. 184 (2021) 1589–1603.

[331] T.P. Peacock, D.H. Goldhill, J. Zhou, L. Baillon, R. Frise, O.C. Swann, R. Kugathasan, R. Penn, J.C. Brown, R.Y. Sanchez-David, L. Braga, M.K. Williamson, J.A. Hassard, E. Staller, B. Hanley, M. Osborn, M. Giacca, A.D. Davidson, D.A. Matthews, W.S. Barclay, The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets, *Nat. Microbiol.* 6 (2021) 899–909.

[332] S. Jiang, C. Hillyer, L. Du, Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses, *Trends Immunol.* 41 (2020) 355–359.

[333] L. Rutten, M.S.A. Gilman, S. Blokland, J. Juraszek, J.S. McLellan, J.P.M. Langedijk, Structure-Based Design of Prefusion-Stabilized Filovirus Glycoprotein Trimers, *Cell Rep.* 30 (2020) 4540-4550.e3.

[334] S. Perri, C.E. Greer, K. Thudium, B. Doe, H. Legg, H. Liu, R.E. Romero, Z. Tang, Q. Bin, T.W. Dubensky, M. Vajdy, G.R. Otten, J.M. Polo, An alphavirus replicon particle chimera derived from venezuelan equine encephalitis and sindbis viruses is a potent gene-based vaccine delivery vector, *J. Virol.* 77 (2003) 10394–10403.

[335] A.-J. Behrens, S. Vasiljevic, L.K. Pritchard, D.J. Harvey, R.S. Andev, S.A. Krumm, W.B. Struwe, A. Cupo, A. Kumar, N. Zitzmann, G.E. Seabright, H.B. Kramer, D.I.R. Spencer, L.

Royle, J.H. Lee, P.J. Klasse, D.R. Burton, I.A. Wilson, A.B. Ward, R.W. Sanders, J.P. Moore, K.J. Doores, M. Crispin, Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein, *Cell Rep.* 14 (2016) 2695–2706.

[336] Q. Hu, R.J. Noll, H. Li, A. Makarov, M. Hardman, R.G. Cooks, The Orbitrap: a new mass spectrometer, *J. Mass Spectrom.* 40 (2005) 430–443.

[337] M. Wilm, Principles of Electrospray Ionization, *Mol. Cell. Proteomics*. 10 (2011).

[338] J. V. Olsen, B. Macek, O. Lange, A. Makarov, S. Horning, M. Mann, Higher-energy C-trap dissociation for peptide modification analysis, *Nat. Methods 2007 49*. 4 (2007) 709–712.

[339] L.R. Heil, P.M. Remes, M.J. Maccoss, Comparison of Unit Resolution Versus High-Resolution Accurate Mass for Parallel Reaction Monitoring, *J. Proteome Res.* 20 (2021) 4435–4442.

[340] N.M. Riley, S.A. Malaker, M.D. Driessen, C.R. Bertozzi, Optimal Dissociation Methods Differ for N- A nd O-Glycopeptides, *J. Proteome Res.* 19 (2020) 3286–3301.

[341] O.J. Hale, H.J. Cooper, In situ mass spectrometry analysis of intact proteins and protein complexes from biological substrates, *Biochem. Soc. Trans.* 48 (2020) 317–326.

[342] A. Šali, T.L. Blundell, Comparative Protein Modelling by Satisfaction of Spatial Restraints, *J. Mol. Biol.* 234 (1993) 779–815.

[343] L.G. Trabuco, E. Villa, K. Mitra, J. Frank, K. Schulten, Flexible Fitting of Atomic Structures into Electron Microscopy Maps Using Molecular Dynamics, *Structure*. 16 (2008) 673–683.

[344] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics, *J. Mol. Graph.* 14 (1996) 33–38.

[345] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD, *J. Comput. Chem.* 26 (2005) 1781–1802.

[346] J. Huang, A.D. MacKerell, CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data, *J. Comput. Chem.* 34 (2013) 2135–2145.

[347] S. Hakansson-McReynolds, S. Jiang, L. Rong, M. Caffrey, Solution Structure of the Severe Acute Respiratory Syndrome-Coronavirus Heptad Repeat 2 Domain in the Prefusion State, *J. Biol. Chem.* 281 (2006) 11965–11971.

[348] J. Dev, D. Park, Q. Fu, J. Chen, H.J. Ha, F. Ghantous, T. Herrmann, W. Chang, Z. Liu, G.

Frey, M.S. Seaman, B. Chen, J.J. Chou, Structural basis for membrane anchoring of HIV-1 envelope spike, *Science*. 353 (2016) 172–175.

[349] Y. Cai, J. Zhang, T. Xiao, H. Peng, S.M. Sterling, R.M. Walsh, S. Rawson, S. Rits-Volloch, B. Chen, Distinct conformational states of SARS-CoV-2 spike protein, *Science*. 21 (2020) 1586–1592.

[350] L. Zuzic, F. Samsudin, A.T. Shivgan, P. V. Raghuvamsi, J.K. Marzinek, A. Boags, C. Pedebos, N.K. Tulsian, J. Warwicker, P. MacAry, M. Crispin, S. Khalid, G.S. Anand, P.J. Bond, Uncovering cryptic pockets in the SARS-CoV-2 spike glycoprotein, *Structure*. 30 (2022) 1062-1074.e4.

[351] S.J. Park, J. Lee, Y. Qi, N.R. Kern, H.S. Lee, S. Jo, I. Joung, K. Joo, J. Lee, W. Im, CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates, *Glycobiology*. 29 (2019) 320–331.

[352] G. van Meer, D.R. Voelker, G.W. Feigenson, G. V Meer, Membrane lipids: where they are and how they behave., *Nat. Rev. Mol. Cell Biol.* 9 (2008) 112–124.

[353] J. Lee, D.S. Patel, J. Ståhle, S.-J. Park, N.R. Kern, S. Kim, J. Lee, X. Cheng, M.A. Valvano, O. Holst, Y.A. Knirel, Y. Qi, S. Jo, J.B. Klauda, G. Widmalm, W. Im, CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans, *J. Chem. Theory Comput.* 15 (2019) 775–786.

[354] J. Lee, X. Cheng, J.M. Swails, M.S. Yeom, P.K. Eastman, J.A. Lemkul, S. Wei, J. Buckner, J.C. Jeong, Y. Qi, S. Jo, V.S. Pande, D.A. Case, C.L. Brooks, A.D. MacKerell, J.B. Klauda, W. Im, CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field, *J. Chem. Theory Comput.* 12 (2016) 405–413.

[355] W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, *Phys. Rev. A*. 31 (1985) 1695–1697.

[356] S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, *Mol. Phys.* 52 (1984) 255–268.

[357] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.* 52 (1981) 7182–7190.

[358] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.* 103 (1995) 8577–8593.

[359] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver

for molecular simulations, *J. Comp. Chem.* 18 (1997) 1463–1472.

[360] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX*. 1 (2015) 19–25.

[361] C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, S. Stagg, C.S. Potter, B. Carragher, Automated molecular microscopy: the new Leginon system, *J. Struct. Biol.* 151 (2005) 41–60.

[362] N.R. Voss, C.K. Yoshioka, M. Radermacher, C.S. Potter, B. Carragher, DoG Picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy, *J. Struct. Biol.* 166 (2009) 205–213.

[363] T. Ogura, K. Iwasaki, C. Sato, Topology representing network enables highly accurate classification of protein images taken by cryo electron-microscope without masking, *J. Struct. Biol.* 143 (2003) 185–200.

[364] C.C.D. Joe, S. Chatterjee, G. Lovrecz, T.E. Adams, M. Thaysen-Andersen, R. Walsh, S.A. Locarnini, P. Smooker, H.J. Netter, Glycoengineered hepatitis B virus-like particles with enhanced immunogenicity., *Vaccine*. 38 (2020) 3892–3901.

[365] M. Hyakumura, R. Walsh, M. Thaysen-Andersen, N.J. Kingston, M. La, L. Lu, G. Lovrecz, N.H. Packer, S. Locarnini, H.J. Netter, Modification of Asparagine-Linked Glycan Density for the Design of Hepatitis B Virus Virus-Like Particles with Enhanced Immunogenicity, *J. Virol.* 89 (2015) 11312–11322.

[366] L.K. Pritchard, D.I.R. Spencer, L. Royle, S. Vasiljevic, S.A. Krumm, K.J. Doores, M. Crispin, Glycan Microheterogeneity at the PGT135 Antibody Recognition Site on HIV-1 gp120 Reveals a Molecular Mechanism for Neutralization Resistance, *J. Virol.* 89 (2015) 6952–6959.

[367] R. Pejchal, K.J. Doores, L.M. Walker, R. Khayat, P.-S. Huang, S.-K. Wang, R.L. Stanfield, J.-P. Julien, A. Ramos, M. Crispin, R. Depetris, U. Katpally, A. Marozsan, A. Cupo, S. Maloveste, Y. Liu, R. McBride, Y. Ito, R.W. Sanders, C. Ogohara, J.C. Paulson, T. Feizi, C.N. Scanlan, C.-H. Wong, J.P. Moore, W.C. Olson, A.B. Ward, P. Poignard, W.R. Schief, D.R. Burton, I.A. Wilson, A Potent and Broad Neutralizing Antibody Recognizes and Penetrates the HIV Glycan Shield, *Science*. 334 (2011) 1097–1103.

[368] Y. Huang, S.O. Owino, C.J. Crevar, D.M. Carter, T.M. Ross, N-Linked Glycans and K147 Residue on Hemagglutinin Synergize To Elicit Broadly Reactive H1N1 Influenza Virus Antibodies, *J. Virol.* 94 (2020) e01432-19.

[369] G.E. Seabright, C.A. Cottrell, M.J. van Gils, A. D'addabbo, D.J. Harvey, A.-J. Behrens, J.D. Allen, Y. Watanabe, N. Scaringi, T.M. Polveroni, A. Maker, S. Vasiljevic, N. de Val, R.W. Sanders, A.B. Ward, M. Crispin, Networks of HIV-1 Envelope Glycans Maintain Antibody Epitopes in the Face of Glycan Additions and Deletions, *Structure*. 28 (2020) 897-909.e6.

[370] M. Butler, Optimisation of the Cellular Metabolism of Glycosylation for Recombinant Proteins Produced by Mammalian Cell Systems, *Cytotechnology*. 50 (2006) 57–76.

[371] S.C.L. Ho, E.Y.C. Koh, M. van Beers, M. Mueller, C. Wan, G. Teo, Z. Song, Y.W. Tong, M. Bardor, Y. Yang, Control of IgG LC:HC ratio in stably transfected CHO cells and study of the impact on expression, aggregation, glycosylation and conformational stability., *J. Biotechnol.* 165 (2013) 157–66.

[372] M. Butler, M. Spearman, The choice of mammalian cell host and possibilities for glycosylation engineering, *Curr. Opin. Biotechnol.* 30 (2014) 107–112.

[373] J. Yolitz, C. Schwing, J. Chang, D. Van Ryk, F. Nawaz, D. Wei, C. Cicala, J. Arthos, A.S. Fauci, Signal peptide of HIV envelope protein impacts glycosylation and antigenicity of gp120, *Proc. Natl. Acad. Sci.* 115 (2018) 2443–2448.

[374] A.-J.J. Behrens, W.B. Struwe, M. Crispin, Glycosylation profiling to evaluate glycoprotein immunogens against HIV-1, *Expert Rev. Proteomics*. 14 (2017) 881–890.

[375] E.P. Go, H. Ding, S. Zhang, R.P. Ringe, N. Nicely, D. Hua, R.T. Steinbock, M. Golabek, J. Alin, S.M. Alam, A. Cupo, B.F. Haynes, J.C. Kappes, J.P. Moore, J.G. Sodroski, H. Desaire, Glycosylation Benchmark Profile for HIV-1 Envelope Glycoprotein Production Based on Eleven Env Trimers, *J. Virol.* 91 (2017) 2428–2444.

[376] S. Alexander, J. Elder, Carbohydrate dramatically influences immune reactivity of antisera to viral glycoprotein antigens, *Science*. 226 (1984) 1328–1330.

[377] X. Yang, Y. Yu, J. Xu, H. Shu, J. Xia, H. Liu, Y. Wu, L. Zhang, Z. Yu, M. Fang, T. Yu, Y. Wang, S. Pan, X. Zou, S. Yuan, Y. Shang, Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study, *Lancet Respir. Med.* 8 (2020) 475–481.

[378] M.L. Holshue, C. DeBolt, S. Lindquist, K.H. Lofy, J. Wiesman, H. Bruce, C. Spitters, K. Ericson, S. Wilkerson, A. Tural, G. Diaz, A. Cohn, L. Fox, A. Patel, S.I. Gerber, L. Kim, S. Tong, X. Lu, S. Lindstrom, M.A. Pallansch, W.C. Weldon, H.M. Biggs, T.M. Uyeki, S.K. Pillai, First Case of 2019 Novel Coronavirus in the United States, *N. Engl. J. Med.* 382 (2020) 929–936.

[379] B.J. Bosch, R. van der Zee, C.A.M. de Haan, P.J.M. Rottier, The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex, *J. Virol.* 77 (2003) 8801–8811.

[380] M.A. Tortorici, D. Veesler, Structural insights into coronavirus entry., *Adv. Virus Res.* 105 (2019) 93–116.

[381] M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses, *Nat. Microbiol.* 5 (2020) 562–569.

[382] T.F. Rogers, F. Zhao, D. Huang, N. Beutler, A. Burns, W. He, O. Limbo, C. Smith, G. Song, J. Woehl, L. Yang, R.K. Abbott, S. Callaghan, E. Garcia, J. Hurtado, M. Parren, L. Peng, S. Ramirez, J. Ricketts, M.J. Ricciardi, S.A. Rawlings, N.C. Wu, M. Yuan, D.M. Smith, D. Nemazee, J.R. Teijaro, J.E. Voss, I.A. Wilson, R. Andrabi, B. Briney, E. Landais, D. Sok, J.G. Jardine, D.R. Burton, Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model, *Science*. 369 (2020) 956–963.

[383] M. Yuan, H. Liu, N.C. Wu, C.-C.D. Lee, X. Zhu, F. Zhao, D. Huang, W. Yu, Y. Hua, H. Tien, T.F. Rogers, E. Landais, D. Sok, J.G. Jardine, D.R. Burton, I.A. Wilson, Structural basis of a shared antibody response to SARS-CoV-2, *Science*. 369 (2020) 1119–1123.

[384] Y. Wu, F. Wang, C. Shen, W. Peng, D. Li, C. Zhao, Z. Li, S. Li, Y. Bi, Y. Yang, Y. Gong, H. Xiao, Z. Fan, S. Tan, G. Wu, W. Tan, X. Lu, C. Fan, Q. Wang, Y. Liu, C. Zhang, J. Qi, G.F. Gao, F. Gao, L. Liu, A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2., *Science*. 368 (2020) 1274–1278.

[385] C.O. Barnes, A.P. West, K.E. Huey-Tubman, M.A.G. Hoffmann, N.G. Sharaf, P.R. Hoffman, N. Koranda, H.B. Gristick, C. Gaebler, F. Muecksch, J.C.C. Lorenzi, S. Finkin, T. Hägglöf, A. Hurley, K.G. Millard, Y. Weisblum, F. Schmidt, T. Hatziioannou, P.D. Bieniasz, M. Caskey, D.F. Robbiani, M.C. Nussenzweig, P.J. Bjorkman, Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies, *Cell*. 182 (2020) 828-842.e16.

[386] W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, Y. Zhou, L. Du, Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine, *Cell. Mol. Immunol.* 17 (2020) 613–620.

[387] V. Indenbaum, R. Koren, S. Katz-Likvornik, M. Yitzchaki, O. Halpern, G. Regev-Yochay, C. Cohen, A. Biber, T. Feferman, N. Cohen Saban, R. Dhan, T. Levin, Y. Gozlan, M. Weil, O. Mor, M. Mandelboim, D. Sofer, E. Mendelson, Y. Lustig, Testing IgG antibodies against the RBD of SARS-CoV-2 is sufficient and necessary for COVID-19 diagnosis, *PLoS One*.

15 (2020) e0241164.

[388] L. Premkumar, B. Segovia-Chumbez, R. Jadi, D.R. Martinez, R. Raut, A.J. Markmann, C. Cornaby, L. Bartelt, S. Weiss, Y. Park, C.E. Edwards, E. Weimer, E.M. Scherer, N. Rouphael, S. Edupuganti, D. Weiskopf, L. V. Tse, Y.J. Hou, D. Margolis, A. Sette, M.H. Collins, J. Schmitz, R.S. Baric, A.M. de Silva, A.M. de Silva, The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients, *Sci. Immunol.* 5 (2020) eabc8413.

[389] D. Stadlbauer, F. Amanat, V. Chromikova, K. Jiang, S. Strohmeier, G.A. Arunkumar, J. Tan, D. Bhavsar, C. Capuano, E. Kirkpatrick, P. Meade, R.N. Brito, C. Teo, M. McMahon, V. Simon, F. Krammer, SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup., *Curr. Protoc. Microbiol.* 57 (2020) e100.

[390] A. Bryan, G. Pepper, M.H. Wener, S.L. Fink, C. Morishima, A. Chaudhary, K.R. Jerome, P.C. Mathias, A.L. Greninger, Performance Characteristics of the Abbott Architect SARS-CoV-2 IgG Assay and Seroprevalence in Boise, Idaho., *J. Clin. Microbiol.* 58 (2020) e00941-20.

[391] P. Escribano, A. Álvarez-Uría, R. Alonso, P. Catalán, L. Alcalá, P. Muñoz, J. Guinea, Detection of SARS-CoV-2 antibodies is insufficient for the diagnosis of active or cured COVID-19, *Sci. Rep.* 10 (2020) 19893.

[392] A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein, *Cell*. 181 (2020) 281-292.e6.

[393] A. Petherick, Developing antibody tests for SARS-CoV-2, *Lancet*. 395 (2020) 1101–1102.

[394] J.S. Turner, W. Kim, E. Kalaidina, C.W. Goss, A.M. Rauseo, A.J. Schmitz, L. Hansen, A. Haile, M.K. Klebert, I. Pusic, J.A. O'Halloran, R.M. Presti, A.H. Ellebedy, SARS-CoV-2 infection induces long-lived bone marrow plasma cells in humans, *Nature*. 595 (2021) 421–425.

[395] A. Wajnberg, F. Amanat, A. Firpo, D.R. Altman, M.J. Bailey, M. Mansour, M. McMahon, P. Meade, D.R. Mendu, K. Muellers, D. Stadlbauer, K. Stone, S. Strohmeier, V. Simon, J. Aberg, D.L. Reich, F. Krammer, C. Cordon-Cardo, Robust neutralizing antibodies to SARS-CoV-2 infection persist for months, *Science*. 370 (2020) 1227–1230.

[396] D.F. Robbiani, C. Gaebler, F. Muecksch, J.C.C. Lorenzi, Z. Wang, A. Cho, M. Agudelo, C.O. Barnes, A. Gazumyan, S. Finkin, T. Hägglöf, T.Y. Oliveira, C. Viant, A. Hurley, H.H.

Hoffmann, K.G. Millard, R.G. Kost, M. Cipolla, K. Gordon, F. Bianchini, S.T. Chen, V. Ramos, R. Patel, J. Dizon, I. Shimeliovich, P. Mendoza, H. Hartweger, L. Nogueira, M. Pack, J. Horowitz, F. Schmidt, Y. Weisblum, E. Michailidis, A.W. Ashbrook, E. Waltari, J.E. Pak, K.E. Huey-Tubman, N. Koranda, P.R. Hoffman, A.P. West, C.M. Rice, T. Hatziioannou, P.J. Bjorkman, P.D. Bieniasz, M. Caskey, M.C. Nussenzweig, Convergent antibody responses to SARS-CoV-2 in convalescent individuals, *Nature*. 584 (2020) 437–442.

[397] J. Pallesen, N. Wang, K.S. Corbett, D. Wrapp, R.N. Kirchdoerfer, H.L. Turner, C.A. Cottrell, M.M. Becker, L. Wang, W. Shi, W.-P. Kong, E.L. Andres, A.N. Kettenbach, M.R. Denison, J.D. Chappell, B.S. Graham, A.B. Ward, J.S. McLellan, Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen, *Proc. Natl. Acad. Sci.* 114 (2017) E7348–E7357.

[398] J.E. Robinson, K.M. Hastie, R.W. Cross, R.E. Yenni, D.H. Elliott, J.A. Rouelle, C.B. Kannadka, A.A. Smira, C.E. Garry, B.T. Bradley, H. Yu, J.G. Shaffer, M.L. Boisen, J.N. Hartnett, M.A. Zandonatti, M.M. Rowland, M.L. Heinrich, L. Martínez-Sobrido, B. Cheng, J.C. de la Torre, K.G. Andersen, A. Goba, M. Momoh, M. Fullah, M. Gbakie, L. Kanneh, V.J. Koroma, R. Fonnie, S.C. Jalloh, B. Kargbo, M.A. Vandi, M. Gbetuwa, O. Ikponmwosa, D.A. Asogun, P.O. Okokhere, O.A. Follarin, J.S. Schieffelin, K.R. Pitts, J.B. Geisbert, P.C. Kulakoski, R.B. Wilson, C.T. Happi, P.C. Sabeti, S.M. Gevao, S.H. Khan, D.S. Grant, T.W. Geisbert, E.O. Saphire, L.M. Branco, R.F. Garry, Most neutralizing human monoclonal antibodies target novel epitopes requiring both Lassa virus glycoprotein subunits., *Nat. Commun.* 7 (2016) 11544.

[399] A. Shields, S.E. Faustini, M. Perez-Toledo, S. Jossi, E. Aldera, J.D. Allen, S. Al-Taei, C. Backhouse, A. Bosworth, L.A. Dunbar, D. Ebanks, B. Emmanuel, M. Garvey, J. Gray, I.M. Kidd, G. McGinnell, D.E. McLoughlin, G. Morley, J. O'Neill, D. Papakonstantinou, O. Pickles, C. Poxon, M. Richter, E.M. Walker, K. Wanigasooriya, Y. Watanabe, C. Whalley, A.E. Zielinska, M. Crispin, D.C. Wraith, A.D. Beggs, A.F. Cunningham, M.T. Drayson, A.G. Richter, SARS-CoV-2 seroprevalence and asymptomatic viral carriage in healthcare workers: a cross-sectional study, *Thorax*. 75 (2020) 1089–1094.

[400] M. Perez-Toledo, S.E. Faustini, S.E. Jossi, A.M. Shields, E. Marcial-Juarez, H.K. Kanthimathinathan, J.D. Allen, Y. Watanabe, M. Goodall, B.E. Willcox, C.R. Willcox, M. Salim, D.C. Wraith, T. V. Veenith, E. Syrimi, M.T. Drayson, D. Jyothish, E. Al-Abadi, A. Chikermane, S.B. Welch, K. Masilamani, S. Hackett, M. Crispin, B.R. Scholefield, A.F. Cunningham, A.G. Richter, SARS-CoV-2-specific IgG1/IgG3 but not IgM in children with Pediatric Inflammatory Multi-System Syndrome, *Pediatr. Allergy Immunol.* 32 (2021) 1125–1129.

[401] C. Seephetdee, N. Buasri, K. Bhukhai, K. Srisanga, S. Manopwisedjaroen, S. Lertjintanakit, N. Phueakphud, C. Pakiranay, N. Kangwanrangsan, S. Srichatrapimuk, S. Kirdlarp, S. Sungkanuparph, S. Chutipongtanate, A. Thitithanyanont, S. Hongeng, P. Wongtrakoongate, Mice Immunized with the Vaccine Candidate HexaPro Spike Produce Neutralizing Antibodies against SARS-CoV-2., *Vaccines*. 9 (2021) 498.

[402] C.L.D. McMillan, J.J.Y. Choo, A. Idris, A. Supramaniam, N. Modhiran, A.A. Amarilla, A. Isaacs, S.T.M. Cheung, B. Liang, H. Bielefeldt-Ohmann, A. Azuar, D. Acharya, G. Kelly, G.J.P. Fernando, M.J. Landsberg, A.A. Khromykh, D. Watterson, P.R. Young, N.A.J. McMillan, D.A. Muller, Complete protection by a single-dose skin patch-delivered SARS-CoV-2 spike vaccine., *Sci. Adv.* 7 (2021) eabj8065.

[403] D. Pinto, Y.-J.J. Park, M. Beltramello, A.C. Walls, M.A. Tortorici, S. Bianchi, S. Jaconi, K. Culap, F. Zatta, A. De Marco, A. Peter, B. Guarino, R. Spreafico, E. Cameroni, J.B. Case, R.E. Chen, C. Havenar-Daughton, G. Snell, A. Telenti, H.W. Virgin, A. Lanzavecchia, M.S. Diamond, K. Fink, D. Veesler, D. Corti, Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody, *Nature*. 583 (2020) 290–295.

[404] C.-C. Wang, J.-R. Chen, Y.-C. Tseng, C.-H. Hsu, Y.-F. Hung, S.-W. Chen, C.-M. Chen, K.-H. Khoo, T.-J. Cheng, Y.-S.E. Cheng, J.-T. Jan, C.-Y. Wu, C. Ma, C.-H. Wong, Glycans on influenza hemagglutinin affect receptor binding and immune response, *Proc. Natl. Acad. Sci.* 106 (2009) 18137–18142.

[405] Y.-M. She, X. Li, T.D. Cyr, Remarkable Structural Diversity of N -Glycan Sulfation on Influenza Vaccines, *Anal. Chem.* 91 (2019) 5083–5090.

[406] J.A. Klein, J. Zaia, Assignment of coronavirus spike protein site-specific glycosylation using GlycReSoft, *Biorxiv*. (2020).

[407] H. Chawla, S.E. Jossi, S.E. Faustini, F. Samsudin, J.D. Allen, Y. Watanabe, M.L. Newby, E. Marcial-Juárez, R.E. Lamerton, J.S. McLellan, P.J. Bond, A.G. Richter, A.F. Cunningham, M. Crispin, Glycosylation and Serological Reactivity of an Expression-enhanced SARS-CoV-2 Viral Spike Mimetic., *J. Mol. Biol.* 434 (2022) 167332.

[408] A.S. Kim, D.P. Leaman, M.B. Zwick, Antibody to gp41 MPER Alters Functional Properties of HIV-1 Env without Complete Neutralization, *PLoS Pathog*. 10 (2014) 1004271.

[409] P.S. Arunachalam, A.C. Walls, N. Golden, C. Atyeo, S. Fischinger, C. Li, P. Aye, M.J. Navarro, L. Lai, V.V. Edara, K. Röltgen, K. Rogers, L. Shirreff, D.E. Ferrell, S. Wrenn, D. Pettie, J.C. Kraft, M.C. Miranda, E. Kepl, C. Sydeman, N. Brunette, M. Murphy, B. Fiala, L. Carter, A.G. White, M. Trisal, C.-L. Hsieh, K. Russell-Lodrigue, C. Monjure, J. Dufour, S. Spencer, L. Doyle-Meyers, R.P. Bohm, N.J. Maness, C. Roy, J.A. Plante, K.S. Plante, A.

Zhu, M.J. Gorman, S. Shin, X. Shen, J. Fontenot, S. Gupta, D.T. O'Hagan, R. Van Der Most, R. Rappuoli, R.L. Coffman, D. Novack, J.S. McLellan, S. Subramaniam, D. Montefiori, S.D. Boyd, J.L. Flynn, G. Alter, F. Villinger, H. Kleanthous, J. Rappaport, M.S. Suthar, N.P. King, D. Veesler, B. Pulendran, Adjuvanting a subunit COVID-19 vaccine to induce protective immunity, *Nature*. 594 (2021) 253–258.

[410] J.R. Francica, B.J. Flynn, K.E. Foulds, A.T. Noe, A.P. Werner, I.N. Moore, M. Gagne, T.S. Johnston, C. Tucker, R.L. Davis, B. Flach, S. O'Connell, S.F. Andrew, E. Lamb, D.R. Flebbe, S.T. Nurmukhambetova, M.M. Donaldson, J.P.M. Todd, A.L. Zhu, C. Atyeo, S. Fischinger, M.J. Gorman, S. Shin, V.V. Edara, K. Floyd, L. Lai, S. Boyoglu-Barnum, R. Van De Wetering, A. Tylor, E. McCarthy, V. Lecouturier, S. Ruiz, C. Berry, T. Tibbitts, H. Andersen, A. Cook, A. Dodson, L. Pessaint, A. Van Ry, M. Koutsoukos, C. Gutzeit, I.T. Teng, T. Zhou, D. Li, B.F. Haynes, P.D. Kwong, A. McDermott, M.G. Lewis, T.M. Fu, R. Chicz, R. Van der Most, K.S. Corbett, M.S. Suthar, G. Alter, M. Roederer, N.J. Sullivan, D.C. Douek, B.S. Graham, D. Casimiro, R.A. Seder, Protective antibodies elicited by SARS-CoV-2 spike protein vaccination are boosted in the lung after challenge in nonhuman primates, *Sci. Transl. Med.* 13 (2021) 4547.

[411] T.-Y. Kuo, M.-Y. Lin, R.L. Coffman, J.D. Campbell, P. Traquina, Y.-J. Lin, L.T.-C. Liu, J. Cheng, Y.-C. Wu, C.-C. Wu, W.-H. Tang, C.-G. Huang, K.-C. Tsao, C. Chen, Development of CpG-adjuvanted stable prefusion SARS-CoV-2 spike antigen as a subunit vaccine against COVID-19, *Sci. Rep.* 10 (2020) 20085.

[412] P.I. Kontou, G.G. Braliou, N.L. Dimou, G. Nikolopoulos, P.G. Bagos, Antibody Tests in Detecting SARS-CoV-2 Infection: A Meta-Analysis, *Diagnostics*. 10 (2020) 319.

[413] J.T. Ladner, S.N. Henson, A.S. Boyle, A.L. Engelbrektson, Z.W. Fink, F. Rahee, J. D'ambrozio, K.E. Schaecher, M. Stone, W. Dong, S. Dadwal, J. Yu, M.A. Caligiuri, P. Cieplak, M. Bjørås, M.H. Fenstad, S.A. Nordbø, D.E. Kainov, N. Muranaka, M.S. Chee, S.A. Shiryaev, J.A. Altin, Epitope-resolved profiling of the SARS-CoV-2 antibody response identifies cross-reactivity with endemic human coronaviruses, *Cell Reports Med.* 2 (2021) 100189.

[414] K.M. Konrath, K. Liaw, Y. Wu, X. Zhu, S.N. Walker, Z. Xu, K. Schultheis, N. Chokkalingam, H. Chawla, J. Du, N.J. Tursi, A. Moore, J. Adolf-Bryfogle, M. Purwar, E.L. Reuschel, D. Frase, M. Sullivan, B. Fry, I. Maricic, V.M. Andrade, C. Iffland, M. Crispin, K.E. Broderick, L.M.P.F. Humeau, A. Patel, T.R.F. Smith, J. Pallesen, D.B. Weiner, D.W. Kulp, Nucleic acid delivery of immune-focused SARS-CoV-2 nanoparticles drives rapid and potent immunogenicity capable of single-dose protection, *Cell Rep.* 38 (2022) 110318.

[415] D. Pinto, Y.J. Park, M. Beltramello, A.C. Walls, M.A. Tortorici, S. Bianchi, S. Jaconi, K.

Culap, F. Zatta, A. De Marco, A. Peter, B. Guarino, R. Spreafico, E. Cameroni, J.B. Case, R.E. Chen, C. Havenar-Daughton, G. Snell, A. Telenti, H.W. Virgin, A. Lanzavecchia, M.S. Diamond, K. Fink, D. Veesler, D. Corti, Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody, *Nature*. 583 (2020) 290–295.

[416] M.N. Ramasamy, A.M. Minassian, K.J. Ewer, A.L. Flaxman, P.M. Folegatti, D.R. Owens, M. Voysey, P.K. Aley, B. Angus, G. Babbage, S. Belij-Rammerstorfer, L. Berry, S. Bibi, M. Bittaye, K. Cathie, H. Chappell, S. Charlton, P. Cicconi, E.A. Clutterbuck, R. Colin-Jones, C. Dold, K.R.W. Emary, S. Fedosyuk, M. Fuskova, D. Gbesemete, C. Green, B. Hallis, M.M. Hou, D. Jenkin, C.C.D. Joe, E.J. Kelly, S. Kerridge, A.M. Lawrie, A. Lelliott, M.N. Lwin, R. Makinson, N.G. Marchevsky, Y. Mujadidi, A.P.S. Munro, M. Pacurar, E. Plested, J. Rand, T. Rawlinson, S. Rhead, H. Robinson, A.J. Ritchie, A.L. Ross-Russell, S. Saich, N. Singh, C.C. Smith, M.D. Snape, R. Song, R. Tarrant, Y. Themistocleous, K.M. Thomas, T.L. Villafana, S.C. Warren, M.E.E. Watson, A.D. Douglas, A.V.S. Hill, T. Lambe, S.C. Gilbert, S.N. Faust, A.J. Pollard, J. Aboagye, K. Adams, A. Ali, E.R. Allen, L. Allen, J.L. Allison, F. Andritsou, R. Anslow, E.H. Arbe-Barnes, M. Baker, N. Baker, P. Baker, I. Baleanu, D. Barker, E. Barnes, J.R. Barrett, K. Barrett, L. Bates, A. Batten, K. Beadon, R. Beckley, D. Bellamy, A. Berg, L. Bermejo, E. Berrie, A. Beveridge, K. Bewley, E.M. Bijker, G. Birch, L. Blackwell, H. Bletchly, C.L. Blundell, S.R. Blundell, E. Bolam, E. Boland, D. Bormans, N. Borthwick, K. Boukas, T. Bower, F. Bowring, A. Boyd, T. Brenner, P. Brown, C. Brown-O'Sullivan, S. Bruce, E. Brunt, J. Burbage, J. Burgoyne, K.R. Buttigieg, N. Byard, I. Cabera Puig, S. Camara, M. Cao, F. Cappuccini, M. Carr, M.W. Carroll, P. Cashen, A. Cavey, J. Chadwick, R. Challis, D. Chapman, D. Charles, I. Chelysheva, J.-S. Cho, L. Cifuentes, E. Clark, S. Collins, C.P. Conlon, N.S. Coombes, R. Cooper, C. Cooper, W.E.M. Crocker, S. Crosbie, D. Cullen, C. Cunningham, F. Cuthbertson, B.E. Datoo, L. Dando, M.S. Datoo, C. Datta, H. Davies, S. Davies, E.J. Davis, J. Davis, D. Dearlove, T. Demissie, S. Di Marco, C. Di Maso, D. DiTirro, C. Docksey, T. Dong, F.R. Donnellan, N. Douglas, C. Downing, J. Drake, R. Drake-Brockman, R.E. Drury, S.J. Dunachie, C.J. Edwards, N.J. Edwards, O. El Muhanna, S.C. Elias, R.S. Elliott, M.J. Elmore, M.R. English, S. Felle, S. Feng, C. Ferreira Da Silva, S. Field, R. Fisher, C. Fixmer, K.J. Ford, J. Fowler, E. Francis, J. Frater, J. Furze, P. Galian-Rubio, C. Galloway, H. Garlant, M. Gavrila, F. Gibbons, K. Gibbons, C. Gilbride, H. Gill, K. Godwin, K. Gordon-Quayle, G. Gorini, L. Goulston, C. Grabau, L. Gracie, N. Graham, N. Greenwood, O. Griffiths, G. Gupta, E. Hamilton, B. Hanumunthadu, S.A. Harris, T. Harris, D. Harrison, T.C. Hart, B. Hartnell, L. Haskell, S. Hawkins, J.A. Henry, M. Hermosin Herrera, D. Hill, J. Hill, G. Hodges, S.H.C. Hodgson, K.L. Horton, E. Howe, N. Howell, J. Howes, B. Huang, J. Humphreys, H.E. Humphries, P. Iveson, F. Jackson, S. Jackson, S. Jauregui, H. Jeffers, B. Jones, C.E. Jones, E. Jones, K. Jones, A. Joshi, R. Kailath, J. Keen, D.M. Kelly, S. Kelly, D. Kelly, D. Kerr, L. Khan, B. Khozoee, A. Killen, J. Kinch, L.D.W. King, T.B. King, L.

Kingham, P. Klenerman, J.C. Knight, D. Knott, S. Koleva, G. Lang, C.W. Larkworthy, J.P.J. Larwood, R. Law, A. Lee, K.Y.N. Lee, E.A. Lees, S. Leung, Y. Li, A.M. Lias, A. Linder, S. Lipworth, S. Liu, X. Liu, S. Lloyd, L. Loew, R. Lopez Ramon, M. Madhavan, D.O. Mainwaring, G. Mallett, K. Mansatta, S. Marinou, P. Marius, E. Marlow, P. Marriott, J.L. Marshall, J. Martin, S. Masters, J. McEwan, J.L. McGlashan, L. McInroy, N. McRobert, C. Megson, A.J. Mentzer, N. Mirtorabi, C. Mitton, M. Moore, M. Moran, E. Morey, R. Morgans, S.J. Morris, H.M. Morrison, G. Morshead, R. Morter, N.A. Moya, E. Mukhopadhyay, J. Muller, C. Munro, S. Murphy, P. Mweu, A. Noé, F.L. Nugent, K. O'Brien, D. O'Connor, B. Oguti, V. Olchawski, C. Oliveira, P.J. O'Reilly, P. Osborne, L. Owen, N. Owino, P. Papageorgiou, H. Parracho, K. Parsons, B. Patel, M. Patrick-Smith, Y. Peng, E.J. Penn, M.P. Peralta-Alvarez, J. Perring, C. Petropoulos, D.J. Phillips, D. Pipini, S. Pollard, I. Poulton, D. Pratt, L. Presland, P.C. Proud, S. Provstgaard-Morys, S. Pueschel, D. Pulido, R. Rabara, K. Radia, D. Rajapaska, F. Ramos Lopez, H. Ratcliffe, S. Rayhan, B. Rees, E. Reyes Pabon, H. Roberts, I. Robertson, S. Roche, C.S. Rollier, R. Romani, Z. Rose, I. Rudiansyah, S. Sabheha, S. Salvador, H. Sanders, K. Sanders, I. Satti, C. Sayce, A.B. Schmid, E. Schofield, G. Screaton, C. Sedik, S. Seddiqi, R.R. Segireddy, B. Selby, I. Shaik, H.R. Sharpe, R. Shaw, A. Shea, S. Silk, L. Silva-Reyes, D.T. Skelly, D.J. Smith, D.C. Smith, N. Smith, A.J. Spencer, L. Spoors, E. Stafford, I. Stamford, L. Stockdale, D. Stockley, L. V. Stockwell, M. Stokes, L.H. Strickland, A. Stuart, S. Sulaiman, E. Summerton, Z. Swash, A. Szigeti, A. Tahiri-Alaoui, R. Tanner, I. Taylor, K. Taylor, U. Taylor, R. te Water Naude, A. Themistocleous, M. Thomas, T.M. Thomas, A. Thompson, K. Thompson, V. Thornton-Jones, L. Tinh, A. Tomic, S. Tonks, J. Towner, N. Tran, J.A. Tree, A. Truby, C. Turner, R. Turner, M. Ulaszewska, R. Varughese, D. Verbart, M.K. Verheul, I. Vichos, L. Walker, M.E. Wand, B. Watkins, J. Welch, A.J. West, C. White, R. White, P. Williams, M. Woodyer, A.T. Worth, D. Wright, T. Wrin, X.L. Yao, D.-A. Zbarcea, D. Zizi, Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial, *Lancet*. 396 (2020) 1979–1993.

[417] W.-H. Chen, L. Du, S.M. Chag, C. Ma, N. Tricoche, X. Tao, C.A. Seid, E.M. Hudspeth, S. Lustigman, C.-T.K. Tseng, M.E. Bottazzi, P.J. Hotez, B. Zhan, S. Jiang, Yeast-expressed recombinant protein of the receptor-binding domain in SARS-CoV spike protein with deglycosylated forms as a SARS vaccine candidate, *Hum. Vaccin. Immunother.* 10 (2014) 648–658.

[418] S. Kumar, V.K. Maurya, A.K. Prasad, M.L.B. Bhatt, S.K. Saxena, Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV), *VirusDisease*. 31 (2020) 13–21.

[419] F.R. Carbone, P.A. Gleeson, Carbohydrates and antigen recognition by T cells,

Glycobiology. 7 (1997) 725–730.

[420] H. Chu, B. Hu, X. Huang, Y. Chai, Y. Wang, H. Shuai, D. Yang, Y. Hou, X. Zhang, T.T.-T. Yuen, J.-P. Cai, A.J. Zhang, J. Zhou, S. Yuan, K.K.-W. To, I.H.-Y. Chan, K.-Y. Sit, D.C.-C. Foo, I.Y.-H. Wong, A.T.-L. Ng, T.T. Cheung, S.Y.-K. Law, W.-K. Au, K.-H. Kok, J.F.-W. Chan, K.-Y. Yuen, Host and viral determinants for efficient SARS-CoV-2 infection of the human lung, *Nat. Commun.* 12 (2020) 134.

[421] C. Xu, D.T.W. Ng, Glycosylation-directed quality control of protein folding, *Nat. Rev. Mol. Cell Biol.* 16 (2015) 742–752.

[422] L.A. Jackson, E.J. Anderson, N.G. Rouphael, P.C. Roberts, M. Makhene, R.N. Coler, M.P. McCullough, J.D. Chappell, M.R. Denison, L.J. Stevens, A.J. Pruijssers, A. McDermott, B. Flach, N.A. Doria-Rose, K.S. Corbett, K.M. Morabito, S. O'Dell, S.D. Schmidt, P.A. Swanson, M. Padilla, J.R. Mascola, K.M. Neuzil, H. Bennett, W. Sun, E. Peters, M. Makowski, J. Albert, K. Cross, W. Buchanan, R. Pikaart-Tautges, J.E. Ledgerwood, B.S. Graham, J.H. Beigel, An mRNA Vaccine against SARS-CoV-2 — Preliminary Report, *N. Engl. J. Med.* 383 (2020) 1920–1931.

[423] A. Sternberg, C. Naujokat, Structural features of coronavirus SARS-CoV-2 spike protein: Targets for vaccination., *Life Sci.* 257 (2020) 118056.

[424] J.F. Cipollo, L.M. Parsons, Glycomics and glycoproteomics of viruses: Mass spectrometry applications and insights toward structure–function relationships, *Mass Spectrom. Rev.* 39 (2020) 371–409.

[425] S. Bangaru, G. Ozorowski, H.L. Turner, A. Antanasijevic, D. Huang, X. Wang, J.L. Torres, J.K. Diedrich, J.-H. Tian, A.D. Portnoff, N. Patel, M.J. Massare, J.R. Yates, D. Nemazee, J.C. Paulson, G. Glenn, G. Smith, A.B. Ward, Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate, *Science*. 370 (2020) 1089–1094.

[426] Y. Watanabe, L. Mendonça, E.R.E.R. Allen, A. Howe, M. Lee, J.D.J.D. Allen, H. Chawla, D. Pulido, F. Donnellan, H. Davies, M. Ulaszewska, S. Belij-Rammerstorfer, S. Morris, A.-S.S.A.-S. Krebs, W. Dejnirattisai, J. Mongkolsapaya, P. Supasa, G.R.G.R. Screaton, C.M.C.M.C.M. Green, T. Lambe, P. Zhang, S.C.S.C. Gilbert, M. Crispin, Native-like SARS-CoV-2 Spike Glycoprotein Expressed by ChAdOx1 nCoV-19/AZD1222 Vaccine, *ACS Cent. Sci.* 7 (2021) 594–602.

[427] C. Gstöttner, T. Zhang, A. Resemann, S. Ruben, S. Pengelley, D. Suckau, T. Welsink, M. Wuhrer, E. Domínguez-Vega, Structural and Functional Characterization of SARS-CoV-2 RBD Domains Produced in Mammalian Cells, *Anal. Chem.* 93 (2021) 6839–6847.

[428] A. Antonopoulos, S. Broome, V. Sharov, C. Ziegenfuss, R.L. Easton, M. Panico, A. Dell, H.R. Morris, S.M. Haslam, Site-specific characterization of SARS-CoV-2 spike glycoprotein receptor-binding domain, *Glycobiology*. 31 (2021) 181–187.

[429] G.L. Morley, S. Taylor, S. Jossi, M. Perez-Toledo, S.E. Faustini, E. Marcial-Juarez, A.M. Shields, M. Goodall, J.D. Allen, Y. Watanabe, M.L. Newby, M. Crispin, M.T. Drayson, A.F. Cunningham, A.G. Richter, M.K. O'Shea, Sensitive detection of SARS-CoV-2-Specific antibodies in dried blood spot samples, *Emerg. Infect. Dis.* 26 (2020) 2970–2973.

[430] S.E. Faustini, S.E. Jossi, M. Perez-Toledo, A.M. Shields, J.D. Allen, Y. Watanabe, M.L. Newby, A. Cook, C.R. Willcox, M. Salim, M. Goodall, J.L. Heaney, E. Marcial-Juarez, G.L. Morley, B. Torlinska, D.C. Wraith, T. V. Veenith, S. Harding, S. Jolles, M.J. Ponsford, T. Plant, A. Huissoon, M.K. O'Shea, B.E. Willcox, M.T. Drayson, M. Crispin, A.F. Cunningham, A.G. Richter, Development of a high sensitivity ELISA detecting IgG, A & M antibodies to the SARS-CoV-2 spike glycoprotein in serum and saliva, *Immunology*. 00 (2021) 1–13.

[431] A.M. Cook, S.E. Faustini, L.J. Williams, A.F. Cunningham, M.T. Drayson, A.M. Shields, D. Kay, L. Taylor, T. Plant, A. Huissoon, G. Wallis, S. Beck, S.E. Jossi, M. Perez-Toledo, M.L. Newby, J.D. Allen, M. Crispin, S. Harding, A.G. Richter, Validation of a combined ELISA to detect IgG, IgA and IgM antibody responses to SARS-CoV-2 in mild or moderate non-hospitalised patients, *J. Immunol. Methods*. 494 (2021) 113046.

[432] A.-J. Behrens, D.J. Harvey, E. Milne, A. Cupo, A. Kumar, N. Zitzmann, W.B. Struwe, J.P. Moore, M. Crispin, Molecular Architecture of the Cleavage-Dependent Mannose Patch on a Soluble HIV-1 Envelope Glycoprotein Trimer, *J. Virol.* 91 (2017) e01894-16.

[433] Y.T. Pang, A. Acharya, D.L. Lynch, A. Pavlova, J.C. Gumbart, SARS-CoV-2 spike opening dynamics and energetics reveal the individual roles of glycans and their collective impact, *Commun. Biol.* 5 (2022) 1170.

[434] F. Higel, T. Sandl, C.-Y.Y. Kao, N. Pechinger, F. Sörgel, W. Friess, F. Wolschin, A. Seidl, N-glycans of complex glycosylated biopharmaceuticals and their impact on protein clearance., *Eur. J. Pharm. Biopharm.* 139 (2019) 123–131.

[435] S. Chia, S.J. Tay, Z. Song, Y. Yang, I. Walsh, K.T. Pang, Enhancing pharmacokinetic and pharmacodynamic properties of recombinant therapeutic proteins by manipulation of sialic acid content, *Biomed. Pharmacother.* 163 (2023) 114757.

[436] B. Boson, V. Legros, B. Zhou, E. Siret, C. Mathieu, F.-L. Cosset, D. Lavillette, S. Denolly, The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles, *J. Biol. Chem.* 296 (2021)

100111.

[437] N. Li, Y. Qi, F.Y. Zhang, X.H. Yu, Y.G. Wu, Y. Chen, C.L. Jiang, W. Kong, Overexpression of α-2,6 sialyltransferase stimulates propagation of human influenza viruses in Vero cells., *Acta Virol.* 55 (2011) 147–53.

[438] Y. Cai, J. Zhang, T. Xiao, H. Peng, S.M. Sterling, R.M. Walsh, S. Rawson, S. Rits-Volloch, B. Chen, Distinct conformational states of SARS-CoV-2 spike protein., *Science*. 369 (2020) 1586–1592.

[439] Z. Ke, J. Oton, K. Qu, M. Cortese, V. Zila, L. McKeane, T. Nakane, J. Zivanov, C.J. Neufeldt, B. Cerikan, J.M. Lu, J. Peukes, X. Xiong, H.-G. Kräusslich, S.H.W. Scheres, R. Bartenschlager, J.A.G. Briggs, Structures and distributions of SARS-CoV-2 spike proteins on intact virions, *Nature*. 588 (2020) 498–502.

[440] L. Liu, P. Wang, M.S. Nair, J. Yu, M. Rapp, Q. Wang, Y. Luo, J.F.-W. Chan, V. Sahi, A. Figueroa, X. V Guo, G. Cerutti, J. Bimela, J. Gorman, T. Zhou, Z. Chen, K.-Y. Yuen, P.D. Kwong, J.G. Sodroski, M.T. Yin, Z. Sheng, Y. Huang, L. Shapiro, D.D. Ho, Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike, *Nature*. 584 (2020) 450–456.

[441] L. Zhang, C.B. Jackson, H. Mou, A. Ojha, H. Peng, B.D. Quinlan, E.S. Rangarajan, A. Pan, A. Vanderheiden, M.S. Suthar, W. Li, T. Izard, C. Rader, M. Farzan, H. Choe, SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity, *Nat. Commun.* 11 (2020) 6013.

[442] Z. Xu, N. Chokkalingam, E. Tello-Ruiz, M.C. Wise, M.A. Bah, S. Walker, N.J. Tursi, P.D. Fisher, K. Schultheis, K.E. Broderick, L. Humeau, D.W. Kulp, D.B. Weiner, A DNA-Launched Nanoparticle Vaccine Elicits CD8+ T-cell Immunity to Promote In Vivo Tumor Control, *Cancer Immunol. Res.* 8 (2020) 1354–1364.

[443] S.A. Mendonça, R. Lorincz, P. Boucher, D.T. Curiel, Adenoviral vector vaccine platforms in the SARS-CoV-2 pandemic, *Npj Vaccines*. 6 (2021) 97.

[444] E. Dolgin, How COVID unlocked the power of RNA vaccines, *Nature*. 589 (2021) 189–191.

[445] J.H. Kim, P. Hotez, C. Batista, O. Ergonul, J.P. Figueroa, S. Gilbert, M. Gursel, M. Hassanain, G. Kang, B. Lall, H. Larson, D. Naniche, T. Sheahan, S. Shoham, A. Wilder-Smith, N. Strub-Wourgaft, P. Yadav, M.E. Bottazzi, Operation Warp Speed: implications for global vaccine security, *Lancet Glob. Heal.* 9 (2021) e1017–e1021.

[446] M. Zhao, M. Vandersluis, J. Stout, U. Haupts, M. Sanders, R. Jacquemart, Affinity

chromatography for vaccines manufacturing: Finally ready for prime time?, *Vaccine*. 37 (2019) 5491–5503.

[447] R.W. Sanders, J.P. Moore, Native-like Env trimers as a platform for HIV-1 vaccine design, *Immunol. Rev.* 275 (2017) 161–182.

[448] M. Medina-Ramírez, F. Garces, A. Escolano, P. Skog, S.W. de Taeye, I. Del Moral-Sanchez, A.T. McGuire, A. Yasmeen, A.J. Behrens, G. Ozorowski, T.L.G.M. van den Kerkhof, N.T. Freund, P. Dosenovic, Y. Hua, A.D. Gitlin, A. Cupo, P. van der Woude, M. Golabek, K. Sliepen, T. Blane, N. Kootstra, M.J. van Breemen, L.K. Pritchard, R.L. Stanfield, M. Crispin, A.B. Ward, L. Stamatatos, P.J. Klasse, J.P. Moore, D. Nemazee, M.C. Nussenzweig, I.A. Wilson, R.W. Sanders, Design and crystal structure of a native-like HIV-1 envelope trimer that engages multiple broadly neutralizing antibody precursors in vivo, *J. Exp. Med.* 214 (2017) 2573–2590.

[449] J.M. Kovacs, J.P. Nkolola, H. Peng, A. Cheung, J. Perry, C.A. Miller, M.S. Seaman, D.H. Barouch, B. Chen, HIV-1 envelope trimer elicits more potent neutralizing antibody responses than monomeric gp120., *Proc. Natl. Acad. Sci. U. S. A.* 109 (2012) 12111–6.

[450] G.E. Seabright, K.J. Doores, D.R. Burton, M. Crispin, Protein and Glycan Mimicry in HIV Vaccine Design, *J. Mol. Biol.* 431 (2019) 2223–2247.

[451] M. Guttman, A. Cupo, J.-P. Julien, R.W. Sanders, I.A. Wilson, J.P. Moore, K.K. Lee, Antibody potency relates to the ability to recognize the closed, pre-fusion form of HIV Env, *Nat. Commun.* 6 (2015) 6144.

[452] L. He, S. Kumar, J.D. Allen, D. Huang, X. Lin, C.J. Mann, K.L. Saye-Francisco, J. Copps, A. Sarkar, G.S. Blizard, G. Ozorowski, D. Sok, M. Crispin, A.B. Ward, D. Nemazee, D.R. Burton, I.A. Wilson, J. Zhu, HIV-1 vaccine design through minimizing envelope metastability., *Sci. Adv.* 4 (2018) eaau6769.

[453] J.G. Jardine, T. Ota, D. Sok, M. Pauthner, D.W. Kulp, O. Kalyuzhniy, P.D. Skog, T.C. Thinnes, D. Bhullar, B. Briney, S. Menis, M. Jones, M. Kubitz, S. Spencer, Y. Adachi, D.R. Burton, W.R. Schief, D. Nemazee, Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen, *Science*. 349 (2015) 156–161.

[454] R. Derking, R.W. Sanders, Structure-guided envelope trimer design in HIV-1 vaccine development: a narrative review., *J. Int. AIDS Soc.* 24 Suppl 7 (2021) e25797.

[455] A. Torrents de la Peña, R.W. Sanders, Stabilizing HIV-1 envelope glycoprotein trimers to induce neutralizing antibodies., *Retrovirology*. 15 (2018) 63.

[456] J.T. Sullivan, C. Sulli, A. Nilo, A. Yasmeen, G. Ozorowski, R.W. Sanders, A.B. Ward, P.J.

Klasse, J.P. Moore, B.J. Doranz, High-Throughput Protein Engineering Improves the Antigenicity and Stability of Soluble HIV-1 Envelope Glycoprotein SOSIP Trimers, *J. Virol.* 91 (2017).

[457] R. Rawi, L. Rutten, Y.-T. Lai, A.S. Olia, S. Blokland, J. Juraszek, C.-H. Shen, Y. Tsybovsky, R. Verardi, Y. Yang, B. Zhang, T. Zhou, G.-Y. Chuang, P.D. Kwong, J.P.M. Langedijk, Automated Design by Structure-Based Stabilization and Consensus Repair to Achieve Prefusion-Closed Envelope Trimers in a Wide Variety of HIV Strains, *Cell Rep.* 33 (2020) 108432.

[458] L. Rutten, Y.-T. Lai, S. Blokland, D. Truan, I.J.M. Bisschop, N.M. Strokappe, A. Koornneef, D. van Manen, G.-Y. Chuang, S.K. Farney, H. Schuitemaker, P.D. Kwong, J.P.M. Langedijk, A Universal Approach to Optimize the Folding and Stability of Prefusion-Closed HIV-1 Envelope Trimers, *Cell Rep.* 23 (2018) 584–595.

[459] M. Moulard, E. Decroly, Maturation of HIV envelope glycoprotein precursors by cellular endoproteases, *Biochim. Biophys. Acta - Rev. Biomembr.* 1469 (2000) 121–132.

[460] A.A. Hargett, M.B. Renfrow, Glycosylation of viral surface proteins probed by mass spectrometry., *Curr. Opin. Virol.* 36 (2019) 56–66.

[461] M. Zhang, Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin, *Glycobiology*. 14 (2004) 1229–1246.

[462] I. Frolov, T.A. Hoffman, B.M. Prágai, S.A. Dryga, H. V. Huang, S. Schlesinger, C.M. Rice, Alphavirus-based expression vectors: strategies and applications., *Proc. Natl. Acad. Sci.* 93 (1996) 11371–11377.

[463] K.M. Pollock, H.M. Cheeseman, A.J. Szubert, V. Libri, M. Boffito, D. Owen, H. Bern, L.R. McFarlane, J. O'Hara, N.-M. Lemm, P. McKay, T. Rampling, Y.T.N. Yim, A. Milinkovic, C. Kingsley, T. Cole, S. Fagerbrink, M. Aban, M. Tanaka, S. Mehdipour, A. Robbins, W. Budd, S. Faust, H. Hassanin, C.A. Cosgrove, A. Winston, S. Fidler, D. Dunn, S. McCormack, R.J. Shattock, K. Adams, F. Amini, N.B. Atako, A. Bakri, W. Barclay, E. Brodnicki, J.C. Brown, R. Byrne, R. Chilvers, S. Coelho, S. Day, M. Desai, E. Dorman, T. Elliott, K.E. Flight, J. Fletcher, J. Galang, J. Gohil, A. Gupta, C. Harlow, K. Hu, M. Kalyan, D. Lagrue, E. Liscano, C. Njenga, K. Polra, D.A. Powlette, P. Randell, M. Rauchenberger, I. Redknap, M. Ricamara, P. Rogers, H. Sallah, K. Samnuan, M. Schumacher, Z. Shah, R. Shaw, T. Shaw, S. Sivapatham, S. Slater, K. Sorley, R. Storch, E. Tan, T. Tan, L. Thielemans, S. Whitely, C. Valentine, J. Varghese, A. Vikraman, M. Wilkins, Safety and immunogenicity of a self-amplifying RNA vaccine against COVID-19: COVAC1, a phase

I, dose-ranging trial, *EClinicalMedicine*. 44 (2022) 101262.

[464] A. Schorcht, C.A. Cottrell, P. Pugach, R.P. Ringe, A.X. Han, J.D. Allen, T.L.G.M. van den Kerkhof, G.E. Seabright, E.E. Schermer, T.J. Ketas, J.A. Burger, J. van Schooten, C.C. LaBranche, G. Ozorowski, N. de Val, D.L. V. Bader, H. Schuitemaker, C.A. Russell, D.C. Montefiori, M.J. van Gils, M. Crispin, P.J. Klasse, A.B. Ward, J.P. Moore, R.W. Sanders, The Glycan Hole Area of HIV-1 Envelope Trimers Contributes Prominently to the Induction of Autologous Neutralization., *J. Virol.* 96 (2022) e0155221.

[465] B. Nogal, M. Bianchi, C.A. Cottrell, R.N. Kirchdoerfer, L.M. Sewall, H.L. Turner, F. Zhao, D. Sok, D.R. Burton, L. Hangartner, A.B. Ward, Mapping Polyclonal Antibody Responses in Non-human Primates Vaccinated with HIV Env Trimer Subunit Vaccines, *Cell Rep.* 30 (2020) 3755-3765.e7.

[466] P.A. Blundell, D. Lu, A. Dell, S. Haslam, R.J. Pleass, Choice of Host Cell Line Is Essential for the Functional Glycosylation of the Fc Region of Human IgG1 Inhibitors of Influenza B Viruses., *J. Immunol.* 204 (2020) 1022–1034.

[467] J.H. Nam, F. Zhang, M. Ermonval, R.J. Linhardt, S.T. Sharfstein, The effects of culture conditions on the glycosylation of secreted human placental alkaline phosphatase produced in Chinese hamster ovary cells, *Biotechnol. Bioeng.* 100 (2008) 1178–1192.

[468] J. Dumont, D. Euwart, B. Mei, S. Estes, R. Kshirsagar, Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives, *Crit. Rev. Biotechnol.* 36 (2016) 1110–1122.

[469] E. Böhm, B.K. Seyfried, M. Dockal, M. Graninger, M. Hasslacher, M. Neurath, C. Konetschny, P. Matthiessen, A. Mitterer, F. Scheiflinger, Differences in N-glycosylation of recombinant human coagulation factor VII derived from BHK, CHO, and HEK293 cells., *BMC Biotechnol.* 15 (2015) 87.

[470] M.S. Macauley, P.R. Crocker, J.C. Paulson, Siglec-mediated regulation of immune cell function in disease, *Nat. Rev. Immunol.* 14 (2014) 653–666.

[471] M. Aebi, N-linked protein glycosylation in the ER, *Biochim. Biophys. Acta - Mol. Cell Res.* 1833 (2013) 2430–2437.

[472] F. Schwarz, M. Aebi, Mechanisms and principles of N-linked protein glycosylation, *Curr. Opin. Struct. Biol.* 21 (2011) 576–582.

[473] P.T. Heath, E.P. Galiza, D.N. Baxter, M. Boffito, D. Browne, F. Burns, D.R. Chadwick, R. Clark, C. Cosgrove, J. Galloway, A.L. Goodman, A. Heer, A. Higham, S. Iyengar, A. Jamal, C. Jeanes, P.A. Kalra, C. Kyriakidou, D.F. McAuley, A. Meyrick, A.M. Minassian,

J. Minton, P. Moore, I. Munsoor, H. Nicholls, O. Osanlou, J. Packham, C.H. Pretswell, A. San Francisco Ramos, D. Saralaya, R.P. Sheridan, R. Smith, R.L. Soiza, P.A. Swift, E.C. Thomson, J. Turner, M.E. Viljoen, G. Albert, I. Cho, F. Dubovsky, G. Glenn, J. Rivers, A. Robertson, K. Smith, S. Toback, Safety and Efficacy of NVX-CoV2373 Covid-19 Vaccine, *N. Engl. J. Med.* 385 (2021) 1172–1183.

[474] L. Corey, P.B. Gilbert, G.D. Tomaras, B.F. Haynes, G. Pantaleo, A.S. Fauci, Immune correlates of vaccine protection against HIV-1 acquisition., *Sci. Transl. Med.* 7 (2015) 310rv7.

[475] A.J. Hessell, D.C. Malherbe, N.L. Haigwood, Passive and active antibody studies in primates to inform HIV vaccines, *Expert Rev. Vaccines*. 17 (2018) 127.

[476] T. Day, J. Kublin, Lessons Learned from HIV Vaccine Clinical Efficacy Trials, *Curr. HIV Res.* 11 (2014) 441–449.

[477] Y. Do Kwon, M. Pancera, P. Acharya, I.S. Georgiev, E.T. Crooks, J. Gorman, M.G. Joyce, M. Guttman, X. Ma, S. Narpala, C. Soto, D.S. Terry, Y. Yang, T. Zhou, G. Ahlsen, R.T. Bailer, M. Chambers, G.-Y. Chuang, N.A. Doria-Rose, A. Druz, M.A. Hallen, A. Harned, T. Kirys, M.K. Louder, S. O'Dell, G. Ofek, K. Osawa, M. Prabhakaran, M. Sastry, G.B.E. Stewart-Jones, J. Stuckey, P. V. Thomas, T. Tittley, C. Williams, B. Zhang, H. Zhao, Z. Zhou, B.R. Donald, L.K. Lee, S. Zolla-Pazner, U. Baxa, A. Schön, E. Freire, L. Shapiro, K.K. Lee, J. Arthos, J.B. Munro, S.C. Blanchard, W. Mothes, J.M. Binley, A.B. McDermott, J.R. Mascola, P.D. Kwong, Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env, *Nat. Struct. Mol. Biol.* 22 (2015) 522–531.

[478] P.D. Kwong, J.R. Mascola, HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure, *Immunity*. 48 (2018) 855–871.

[479] P.D. Kwong, B.J. DeKosky, J.B. Ulmer, Antibody-guided structure-based vaccines., *Semin. Immunol.* 50 (2020) 101428.

[480] T. Bradley, J. Pollara, S. Santra, N. Vandergrift, S. Pittala, C. Bailey-Kellogg, X. Shen, R. Parks, D. Goodman, A. Eaton, H. Balachandran, L. V Mach, K.O. Saunders, J.A. Weiner, R. Scearce, L.L. Sutherland, S. Phogat, J. Tartaglia, S.G. Reed, S.-L. Hu, J.F. Theis, A. Pinter, D.C. Montefiori, T.B. Kepler, K.K. Peachman, M. Rao, N.L. Michael, T.J. Suscovich, G. Alter, M.E. Ackerman, M.A. Moody, H.-X. Liao, G. Tomaras, G. Ferrari, B.T. Korber, B.F. Haynes, Pentavalent HIV-1 vaccine protects against simian-human immunodeficiency virus challenge, *Nat. Commun.* 8 (2017) 15711.

[481] D.H. Barouch, F.L. Tomaka, F. Wegmann, D.J. Stieh, G. Alter, M.L. Robb, N.L. Michael,

L. Peter, J.P. Nkolola, E.N. Borducchi, A. Chandrashekar, D. Jetton, K.E. Stephenson, W. Li, B. Korber, G.D. Tomaras, D.C. Montefiori, G. Gray, N. Frahm, M.J. McElrath, L. Baden, J. Johnson, J. Hutter, E. Swann, E. Karita, H. Kibuuka, J. Mpendo, N. Garrett, K. Mngadi, K. Chinyenze, F. Priddy, E. Lazarus, F. Laher, S. Nitayapan, P. Pitisuttithum, S. Bart, T. Campbell, R. Feldman, G. Lucksinger, C. Borremans, K. Callewaert, R. Roten, J. Sadoff, L. Scheppler, M. Weijtens, K. Feddes-de Boer, D. van Manen, J. Vreugdenhil, R. Zahn, L. Lavreys, S. Nijs, J. Tolboom, J. Hendriks, Z. Euler, M.G. Pau, H. Schuitemaker, Evaluation of a mosaic HIV-1 vaccine in a multicentre, randomised, double-blind, placebo-controlled, phase 1/2a clinical trial (APPROACH) and in rhesus monkeys (NHP 13-19), *Lancet*. 392 (2018) 232–243.

[482] K. Om, D. Paquin-Proulx, M. Montero, K. Peachman, X. Shen, L. Wieczorek, Z. Beck, J.A. Weiner, D. Kim, Y. Li, T. Mdluli, Z. Shubin, C. Bryant, V. Sharma, A. Tokarev, P. Dawson, Y. White, O. Appelbe, N.R. Klatt, S. Tovanabutra, J.D. Estes, G.R. Matyas, G. Ferrari, C.R. Alving, G.D. Tomaras, M.E. Ackerman, N.L. Michael, M.L. Robb, V. Polonis, M. Rolland, M.A. Eller, M. Rao, D.L. Bolton, Adjuvanted HIV-1 vaccine promotes antibody-dependent phagocytic responses and protects against heterologous SHIV challenge, *PLOS Pathog.* 16 (2020) e1008764.

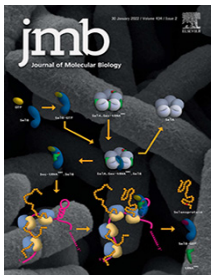
[483] B.K. Felber, Z. Lu, X. Hu, A. Valentin, M. Rosati, C.A.L. Remmel, J.A. Weiner, M.C. Carpenter, K. Faircloth, S. Stanfield-Oakley, W.B. Williams, X. Shen, G.D. Tomaras, C.C. LaBranche, D. Montefiori, H. V Trinh, M. Rao, M.S. Alam, N.A. Vandergrift, K.O. Saunders, Y. Wang, W. Rountree, J. Das, G. Alter, S.G. Reed, P.P. Aye, F. Schiro, B. Pahar, J.P. Dufour, R.S. Veazey, P.A. Marx, D.J. Venzon, G.M. Shaw, G. Ferrari, M.E. Ackerman, B.F. Haynes, G.N. Pavlakis, Co-immunization of DNA and Protein in the Same Anatomical Sites Induces Superior Protective Immune Responses against SHIV Challenge, *Cell Rep.* 31 (2020) 107624.

[484] P. Zhang, E. Narayanan, Q. Liu, Y. Tsybovsky, K. Boswell, S. Ding, Z. Hu, D. Follmann, Y. Lin, H. Miao, H. Schmeisser, D. Rogers, S. Falcone, S.M. Elbashir, V. Presnyak, K. Bahl, M. Prabhakaran, X. Chen, E.K. Sarfo, D.R. Ambrozak, R. Gautam, M.A. Martin, J. Swerczek, R. Herbert, D. Weiss, J. Misamore, G. Ciaramella, S. Himansu, G. Stewart-Jones, A. McDermott, R.A. Koup, J.R. Mascola, A. Finzi, A. Carfi, A.S. Fauci, P. Lusso, A multiclade env–gag VLP mRNA vaccine elicits tier-2 HIV-1-neutralizing antibodies and reduces the risk of heterologous SHIV infection in macaques, *Nat. Med.* 27 (2021) 2234–2245.

[485] F. Zabel, D. Mohanan, J. Bessa, A. Link, A. Fettelschoss, P. Saudan, T.M. K€, M.F. Bachmann, Viral Particles Drive Rapid Differentiation of Memory B Cells into Secondary Plasma Cells Producing Increased Levels of Antibodies, *J. Immunol.* 192 (2014) 5499–

5508.

[486] C. Ludwig, R. Wagner, Virus-like particles-universal molecular toolboxes, *Curr. Opin. Biotechnol.* 18 (2007) 537–545.



Glycosylation and Serological Reactivity of an Expression-enhanced SARS-CoV-2 Viral Spike Mimetic

Himanshi Chawla [1], Sian E. Jossi [2], Sian E. Faustini [2], Firdaus Samsudin [3], Joel D. Allen [1], Yasunori Watanabe [1,4], Maddy L. Newby [1], Edith Marcial-Juárez [2], Rachel E. Lamerton [2], Jason S. McLellan [5], Peter J. Bond [3,6], Alex G. Richter [2], Adam F. Cunningham [2] and Max Crispin [1]*

*1 - **School of Biological Sciences,** University of Southampton, Southampton SO17 1BJ, UK*
*2 - **Institute of Immunology and Immunotherapy,** University of Birmingham, Birmingham B15 2TT, UK*
*3 - **Bioinformatics Institute,** Agency for Science, Technology and Research (A*STAR), Singapore 138671, Singapore*
*4 - **Oxford Glycobiology Institute,** Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK*
*5 - **Department of Molecular Biosciences,** The University of Texas at Austin, Austin, TX 78712, USA*
*6 - **Department of Biological Sciences,** National University of Singapore, Singapore 117543, Singapore*

***Correspondence to Max Crispin:** max.crispin@soton.ac.uk (M. Crispin)*
https://doi.org/10.1016/j.jmb.2021.167332
Edited by M.F. Summers

Abstract

Extensive glycosylation of viral glycoproteins is a key feature of the antigenic surface of viruses and yet glycan processing can also be influenced by the manner of their recombinant production. The low yields of the soluble form of the trimeric spike (S) glycoprotein from SARS-CoV-2 has prompted advances in protein engineering that have greatly enhanced the stability and yields of the glycoprotein. The latest expression-enhanced version of the spike incorporates six proline substitutions to stabilize the prefusion conformation (termed SARS-CoV-2 S HexaPro). Although the substitutions greatly enhanced expression whilst not compromising protein structure, the influence of these substitutions on glycan processing has not been explored. Here, we show that the site-specific N-linked glycosylation of the expression-enhanced HexaPro resembles that of an earlier version containing two proline substitutions (2P), and that both capture features of native viral glycosylation. However, there are site-specific differences in glycosylation of HexaPro when compared to 2P. Despite these discrepancies, analysis of the serological reactivity of clinical samples from infected individuals confirmed that both HexaPro and 2P protein are equally able to detect IgG, IgA, and IgM responses in all sera analysed. Moreover, we extend this observation to include an analysis of glycan engineered S protein, whereby all N-linked glycans were converted to oligomannose-type and conclude that serological activity is not impacted by large scale changes in glycosylation. These observations suggest that variations in glycan processing will not impact the serological assessments currently being performed across the globe.

Introduction

Recombinant viral glycoproteins are an important resource for vaccine development, diagnostics and as research reagents. Viral glycoprotein glycosylation can influence an extensive range of properties including immunogen trafficking,[1] immunogenicity,[2,3] antigenicity[4,5] and serum clearance rates[6]. Importantly, recombinant viral spike glycosylation can be influenced both by features

of the glycoprotein sequence, such as glycan density and protein architecture[7–9] and an array of expression conditions, such as producer cell type and expression conditions.[10–13] It is therefore important to define the glycosylation of recombinant viral glycoproteins and monitor changes that may occur during target optimization and the development of manufacturing procedures.[14,15] As carbohydrates on viral proteins can influence the immune response, it is important to look at the binding of sera antibodies to antigens possessing distinct glycoforms.[16] Here, we investigate the antigenic properties of glycoprotein reagents developed in response to the coronavirus infectious disease 2019 (COVID-19) pandemic, focused on the viral S glycoprotein.[17,18]

The causative agent of COVID-19, Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), is a positive-sense single-stranded RNA virus that has caused significant morbidity and mortality throughout the world.[19,20] Like other coronaviruses, SARS-CoV-2 utilizes the S glycoprotein for recognition of the host cell entry receptor and subsequent membrane fusion, which is mediated by the S1 and S2 subunits, respectively.[21] The S protein is a trimeric class I fusion protein and is post-translationally cleaved into S1 and S2 subunits using the host cellular protease, furin.[22] The S1 subunit possesses an N-terminal domain (NTD) and a receptor-binding domain (RBD), also known as domain A and B, respectively.[23]

The exposed position of the RBD enables binding to the angiotensin converting enzyme 2 (ACE2) receptor[24] and, as a result, it is the main target of anti-SARS antibodies during infection.[25–30] Due to this phenomenon, combined with its high recombinant protein yields, several antibody tests have been developed using RBD as a tool to test for previous SARS-CoV-2 infection.[31–34] One disadvantage of using RBD as an antigenic bait for testing is that it may not capture the entire antibody response to the S protein as it lacks the full trimeric structure.[35] In addition to RBD, other antibody tests use the nucleoprotein (N protein) as antigenic bait to detect prior SARS-CoV-2 infection, such as the Abbott test.[36,37] Similarly to the RBD, N protein will not capture the complete antigenic surface and therefore may not reveal the full immune response to SARS-CoV-2 infection. As the S protein is the prime target of neutralizing antibodies, the native-like trimeric S glycoprotein may facilitate the presentation of a more complete range of epitopes for antibody testing.[38,39] Serological testing requires that the protein is both stable and that production is readily scalable for widespread use. There has been significant development in design of S protein constructs to facilitate increased recombinant production and protein stability.

Prefusion stabilization strategies have been employed for class I fusion proteins to increase the recombinant expression of viral glycoproteins.[40–44] A common strategy is the introduction of proline substitutions which impedes the switch to the post fusion conformation.[45] This is crucial as neutralizing antibody epitopes are predominantly presented on the prefusion conformation.[41,46–48] For SARS-CoV-2, the expression of a stable, soluble form of the S-protein was originally achieved by truncation at the transmembrane domain and the incorporation of two proline residues (K986P and V987P)[17] (Figure 1; SARS-CoV-2 2P S, henceforth termed “2P”). Despite the utility of the 2P construct for structural analysis[17] and serological testing,[35,49,50] the low expression levels prompted the development of an expression enhanced version containing four additional prolines (Figure 1; SARS-CoV-2 HexaPro, henceforth termed “HexaPro”).[18] HexaPro exhibits native-like protein architecture, antigenic properties, and contains the twenty-two N-linked glycosylation sites of the native viral spike.[18,51,52] Additionally, HexaPro has shown promising results as a vaccine candidate in mice immunization, resulting in high-titre neutralizing antibodies.[53–55]

Here, we performed Liquid chromatography-mass spectrometry (LC-MS) experiments to establish whether the additional stabilizing mutations impact the presentation of glycans on the surface of the S protein. Further, to explore whether the observed differences in glycans impacted the HexaPro protein function, we compared the binding of HexaPro and 2P protein to ACE2 using surface plasmon resonance (SPR) and demonstrated comparable binding between the two. Molecular dynamics (MD) simulations revealed a similar pattern of surface accessibility of N-linked glycan sites between the 2P and HexaPro S proteins, and support the overall conserved nature of glycosylation. In addition, we explored whether the modest changes in glycosylation impacted the detection of immune responses in patient sera previously infected with COVID-19. Both HexaPro and 2P protein were successful at detecting an immune response towards SARS-CoV-2 in both hospitalized and non-hospitalized patients. Furthermore, we compared the binding of S protein possessing oligomannose-type glycans at all potential N-linked glycosylation sites (PNGS), which was achieved through co-transfection with the ER α mannosidase I inhibitor, kifunensine. Both S protein glycan variants revealed highly similar binding to sera from patients with prior-SARS-CoV-2 infection. These studies further support the use of hyperstabilization using additional proline mutations and demonstrate its utility in serological testing.

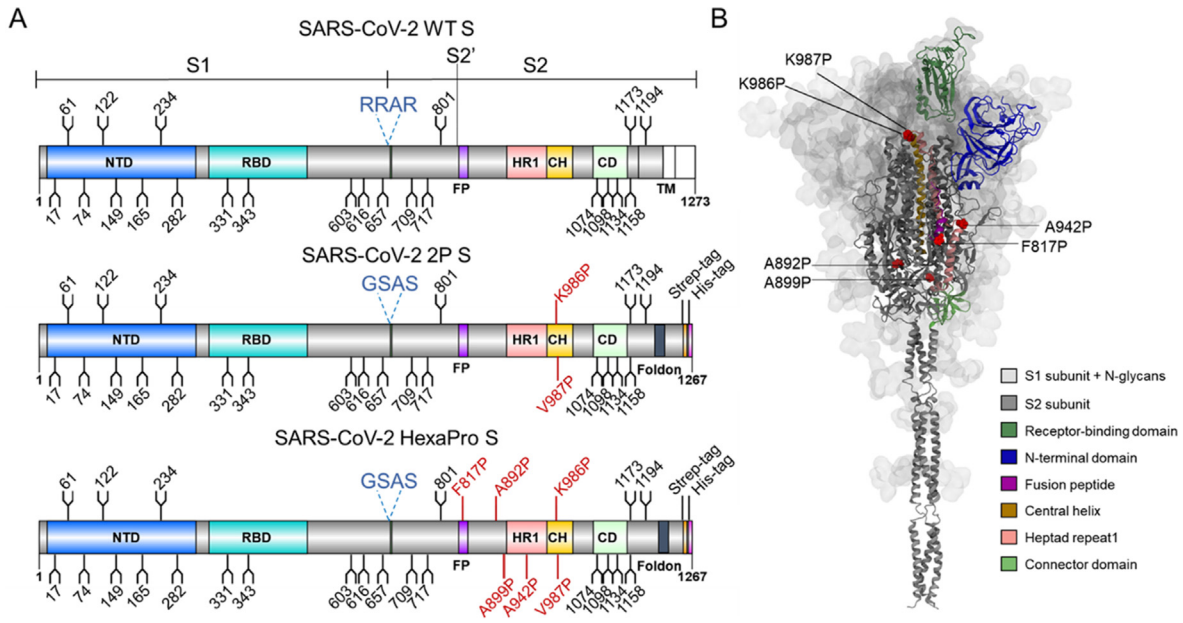


Figure 1. Representation and characterization of recombinant SARS-CoV-2 spike protein. (A) The protein domains are represented as N-terminal domain (NTD), receptor-binding domain (RBD), fusion peptide (FP), heptad repeat 1 (HR1), central helix (CH), connector domain (CD), and transmembrane domain (TM). The fusion cleavage site is illustrated as dashed lines (blue). N-linked glycosylation sequons (N-X-S/T, where X ≠ P) are shown as branches. SARS-CoV-2 WT presents S1 and S2 domain with furin cleavage site (RRAR) and transmembrane domain at C-terminal end. SARS-CoV-2 2P prefusion stabilized protein with proline substitutions at residues 986 and 987 and, a "GSAS" mutation at furin cleavage site (residues 682–685). HexaPro prefusion stabilized protein of SARS-CoV-2 with a "GSAS" mutation at furin cleavage site and six proline substitutions, highlighted in red. (B) Structural representation of HexaPro S protein illustrating six proline substitutions (red spheres) in SARS-CoV-2 ectodomain (PDB ID: 6XKL). The S1 subunit along with N-glycans are shown as transparent molecular surface. The S2 subunit is shown in dark grey. Different domains present in S1/S2 subunits are highlighted in respective colors in ribbon diagram of only one protomer.

Results and discussion

Characterization of prefusion stabilized SARS-CoV-2 HexaPro S protein

For characterization of prefusion stabilized SARS-CoV-2 S protein, we transiently transfected plasmid encoding SARS-CoV-2 S protein containing the HexaPro stabilizing mutations in Human embryonic kidney (HEK) 293F cells. To ensure the analysis of only native-like trimeric protein, the supernatant was first purified using nickel-affinity chromatography followed by size exclusion chromatography (SEC) (Figure 2(A)). To functionally characterize the binding of the expressed protein with ACE-2, the binding affinity (expressed here using the dissociation constant, K_D) of HexaPro with a soluble recombinant ACE2 was determined using surface plasmon resonance (SPR) which was repeated three times to determine the average K_D (Figure 2(B)). A residual plot of the data revealed minimal deviation between observed values and calculated values using a 1:1 binding model between SARS-CoV-2 HexaPro S protein and ACE2 (Figure 2 (C)). The K_D values are comparable to that previously reported for 2P[56] (Figure 2(D)).

Determination of impact of proline mutations on spike glycosylation

To determine the effect of stabilizing mutations on glycosylation, we determined the site-specific glycan compositions of SARS-CoV-2 HexaPro. In preparation for glycopeptide analysis, we independently expressed and purified three replicates of the soluble ectodomain of the SARS-CoV-2 protein truncated at the transmembrane domain. We note that this construct was used to determine the high resolution cryo-electron microscopy (cryo-EM) structure.[18] To generate glycopeptide samples derived from these batches of protein, we employed three different protease enzymes, trypsin, chymotrypsin, and alpha-lytic protease. These proteases cleave the protein chain at specific amino acids and were selected to generate glycopeptides that contain a single glycosylation sequon (N-X-S/T-X, where X-any amino acid except proline). Using liquid chromatography-electrospray ionization mass spectrometry (LC-ESI MS), we were able to quantify the

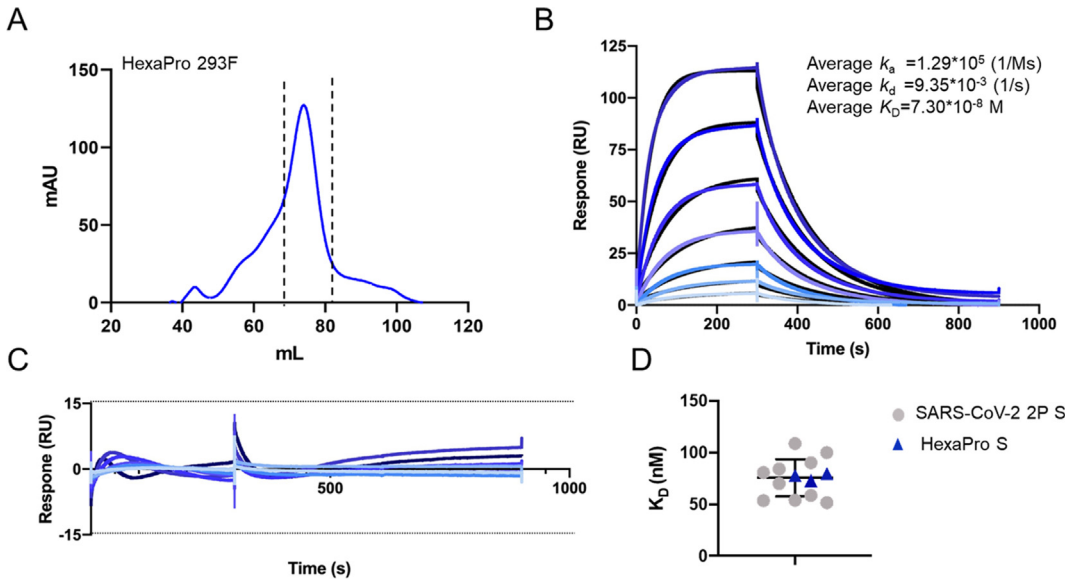


Figure 2. Characterization of recombinant SARS-COV-2 S protein, HexaPro. (A) SEC of affinity-purified recombinantly expressed S protein. Dashed lines indicate fractions collected for subsequent use. (B) SPR of HexaPro (ligand) with soluble ACE2 receptor (analyte). Dark blue to light blue lines represent the serial dilutions of ACE2 protein from 200 nM to 3.125 nM, respectively. Black lines are fitted values of the respective concentration to illustrate the best fit to a 1:1 binding model. Three repeats were performed and averaged to determine the k_a, k_d and K_D values. (C) Residual plot illustrating the deviation of the fitted data to the raw values of the experimental data at different concentrations. (D) Representation of K_D values determined using various repeats, grey dots illustrate the binding of 2P with ACE2 (values reproduced from 56) and the binding of HexaPro with ACE2 are shown as blue triangles. The mean of K_D values of 2P is plotted as a black line and the error bars represent ±standard deviation calculated using GraphPad Prism.

oligomannose-type glycans, complex-type glycans and the proportion of unoccupied potential N-glycosylation sites (PNGS) across all 22 N-linked glycan sites on the HexaPro protomers (Figure 3 (A and B)).

To determine the potential impact of more extensive modifications on the prefusion stabilized S protein, we compared the glycan profile of SARS-CoV-2 S protein of 2P with HexaPro. Interestingly, the overall processing states of the recombinant S protein were conserved across both the versions, with few variations at the site-specific level (Figures 3(B) and 4(A)). Site-specific analysis of these recombinant proteins suggests that the levels of oligomannose-type glycans are consistent with native S protein on infectious virus and also with other coronaviruses.[51,52,57–59]

In addition to the information obtained from studying the populations of oligomannose-type glycans at individual glycosylation sites, understanding the processing of complex-type glycans is also informative when considering immunogen design, reagents for serological studies, and for understanding the extent to which recombinant glycoprotein can be used as mimics of the functional viral spike. For example, the epitope of neutralizing antibody S309, which targets the S protein of SARS-CoV-1 and SARS-CoV-2, contains a fucosylated glycans at N343.[60] HexaPro is 99% fucosylated at the N343 site, with almost all the glycans bearing fucose residues (Supplementary Table S1). Furthermore, sulfated N-linked glycans have been detected on viral glycoproteins and they could potentially play role in immune regulation, as in influenza.[61,62] We detected sulfation at several N-glycosylation sites on HexaPro (N74, N149, N1194) and observed a similar abundance of sulfation across both S proteins (2P and HexaPro), which is in accordance with analysis of other SARS-CoV-2 S proteins[63,64] (Supplementary Figure S1).

In all formats of SARS-CoV-2 S expressed in mammalian cells, a higher proportion of complex-type glycans compared to oligomannose-type glycans were observed.[51,52,59,61,63,65] Moreover, the complex-type glycans somewhat obscure immunogenic surfaces and constitute a shield to evade the immune system,[66–68] albeit not at a level observed in many other viral envelopes.[21,57] Out of 22 N-linked glycan sites on each protomer, HexaPro S presents more than 50% highly processed complex-type glycans on 15 N-linked sites which is comparable to the 2P expressed in different laboratories.[51,52] The underoccupancy at the glycosylation sequon at N657 is present on both HexaPro (Figure 3(A) and Supplementary Table S1) and 2P (Supplementary Table S2). The glycan site at the C terminus, N1194, is fully occupied in the case of

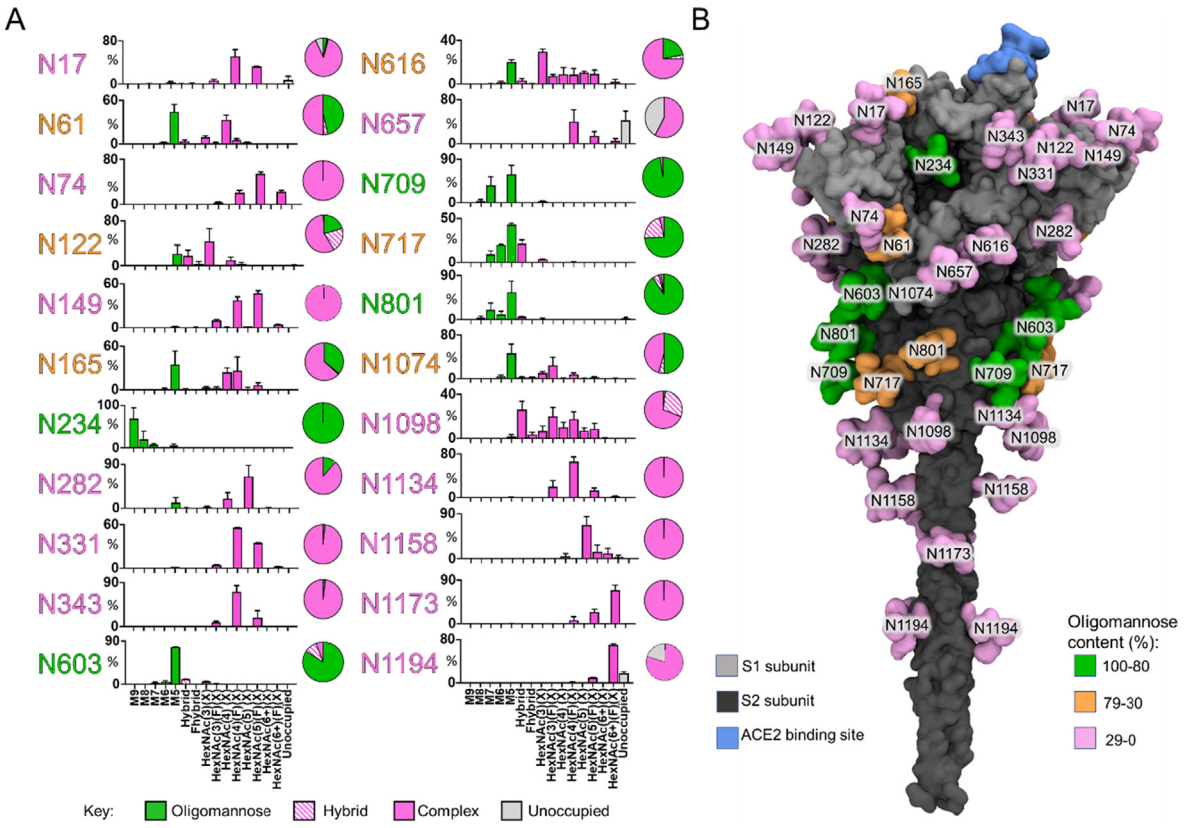


Figure 3. Site-specific glycosylation of expression-enhanced recombinant trimer of SARS-CoV-2 S protein (HexaPro). (A) Relative quantification of the N-linked glycosylation sites of trimeric S protein, produced in HEK293F cells. The bar graph represents the mean of three independently expressed replicates with error bars representing the standard error of the mean. The color codes in the schematic illustrates the processing state of glycans from least processed to most processed, oligomannose (green), hybrid (dashed pink), and complex glycans (pink). The proportion of unoccupied N-linked glycan sites are represented in grey. The pie charts summarize the quantification of these categories. The N-linked glycan site labels are colored based on the oligomannose-type glycan content, green (80–100%), orange (30–79%) and magenta (0–29%). (B) The model was constructed using the prefusion structure of trimeric SARS-CoV-2 S glycoprotein as described in Materials and Methods. The S1 and S2 subunits are shown as light and dark grey, respectively. The glycans sites are categorized as high-mannose type glycans (green), hybrid glycans (orange), and complex-type glycans (pink). The ACE2 receptor binding site is shown in blue.

2P (Supplementary Table S2) however, there is an elevation of unoccupied PNGS (18%) at N1194 on the HexaPro S protein (Supplementary Table S1).

The oligomannose-type glycan content of the glycans of the HexaPro protein (29%) (Supplementary Table S1) is lower when compared to other viral glycoproteins including HIV-1 Env (60%) and LASV GPC (49%).[69,70] This is consistent with earlier observations using 2P protein which indicated that SARS-CoV-2 S is less shielded, which may be beneficial for the elicitation of neutralizing antibodies.[51] Overall, there is a similar level of oligomannose-type glycans across both 2P and HexaPro (Supplementary Table S3).

Thus we have shown here that the site-specific glycosylation of the expression enhanced version of SARS-CoV-2, HexaPro, is highly similar to the glycosylation of 2P and native S protein as presented on the virus.[51,52,58] Also, we confirm earlier observations[18] that both forms of the recombinant protein have indistinguishable binding properties to the receptor, ACE2, indicating the functional form of protein is intact. However, we also detected some difference in glycosylation which could suggest differences in the conformational properties between the HexaPro and 2P. This motivated us to extend the analysis of protein conformational flexibility by performing MD simulations, and to characterize its functional behaviour by performing serological testing. This builds upon previous observations of cryo-EM and a small-scale serological evaluation.[18]

Differences in oligomannose content between 2P and HexaPro

The glycan at a structural site, N234, which has been shown to stabilize the RBD up conformation in the trimeric structure of the S protein,[71] is principally oligomannose-type, and is conserved across

both the constructs of recombinant S protein (Figure 3(A and B)). Also, the oligomannose-type glycans of N234 likely arises from steric clashes with the protein component, which in turn limits the ability of glycan processing enzymes to act, as it is sandwiched between N-terminal and receptor-binding domains[71] (Figure 4(B)). Overall, the oligomannose content is highly similar across both 2P and HexaPro (Supplementary Table S3). However, the $Man_9GlcNAc_2$ content is higher in the case of HexaPro compared to 2P, indicating the reduced accessibility to glycan processing enzymes[51] (Supplementary Table S3). This could possibly be explained by the two RBD "up" conformation observed in HexaPro.[18] Also, at the site-specific level there are several other sites which showed changes in oligomannose content across 2P and HexaPro (Figure 4(A and B)). Differences in glycan processing states were observed at glycan sites N61, N122, N165, N603, N616 and N801. The major differences were observed at N165 and N122, which are in close proximity to the RBD (a roughly 50 percentage point difference in both cases). The N165 site possesses a higher abundance of oligomannose-type glycans on HexaPro shown in dark blue whereas the N122 site possesses a lower abundance of oligomannose-type glycans on HexaPro shown in red (Figure 4(B)).

To investigate the molecular basis of the observed changes in oligomannose-type glycan content across both the S proteins, we performed a series of triplicate 200 ns MD simulations of the HexaPro two-RBD-up, HexaPro one-RBD-up and 2P one-RBD-up constructs. The HexaPro two RBD-up model was generated by fitting the structure of HexaPro one-RBD-up to the electron density map of HexaPro two-RBD-up using molecular dynamics flexible fitting (MDFF) (details in Methods section). All potential N-linked glycosylation sites were glycosylated with $Man_9GlcNAc_2$ glycans, as these represent the primary substrate for glycan processing enzymes and hence may be used to predict glycan processing as previously described.[52]

First, we investigated the correlation between accessible surface area (ASA) for all glycans between HexaPro and 2P. There is a moderate correlation ($r^2 = 0.5$) between both one-RBD-up and two-RBD-up HexaPro when compared to the 2P, suggesting there is some degree of difference in glycan accessibility between HexaPro (two-RBD-up & one-RBD-up) and 2P (one-RBD-up) (Figure 4(C)). To understand whether this difference is due to the protein architecture, or due to the stochastic nature of the simulations, we then compared the ASA for each of the glycosylation site between HexaPro and 2P. Furthermore, we calculated the arithmetic difference between the ASA values of HexaPro (two RBD-up) and 2P (one RBD-up) at each glycan site in all three chains (Supplementary Figure S2). The error bars represent standard deviations throughout the trajectories which capture the variation caused by sampling. A positive value represents lower accessibility in HexaPro, which could correlate to reduced glycan processing and a higher abundance of oligomannose-type glycans. The difference at most of the sites was small and with substantial standard deviations, suggesting that most of the differences arise from the stochastic sampling of the glycans.

Then, we further explored sites which showed changes in oligomannose content in the site-specific glycan analysis. For example, on HexaPro, N165 displays an increase of almost 50 percentage point in oligomannose-type glycan compositions when compared with 2P (Figure 4(A)). As the N165 site is in close proximity to the RBD region, its glycan processing state may be influenced by the orientation of the RBD. When comparing the model generated with two RBD domains in the up configuration, to one-RBD-up, the steric environment of the N165 and N122 sites were expected to change. However, the simulations comparing HexaPro two-RBD-up versus one-RBD-up showed little changes in solvent accessibility (Figure 4(D)). This is likely because in the HexaPro (two-RBD-up) simulation, the additional RBD in the initial up configuration tends to revert to a down state, and the final snapshot in all simulation replicas is similar to the 2P (one-RBD-up) simulation (Supplementary Figure S3). It is noteworthy that in our initial model, derived from the experimental electron density map of the HexaPro two RBD-up conformation, the RBD in chain C resembles an intermediate state between the up and down states, rather than a fully open conformation. The changes in ASA values we observed are only significant at a few sites and minor with variations across the three chains, indicating that the differences are due to sampling.

Finally, we performed protein-structural analyses to compare the dynamics of 2P and HexaPro. We observed similar root-mean square fluctuations (RMSF) profiles between the HexaPro (two-RBD-up and one-RBD-up) and 2P (one-RBD-up) simulations, except for the RBD in chain C, due to the up to down conformational changes described above (Supplementary Figure S4(A)). The principal motion measured during the simulations also reveals similar dynamics between these two S proteins (Supplementary Figure S4(B)). Hence, the simulations suggest that despite local glycan perturbation, the two S protein versions have very similar dynamics and our computational analysis does not provide evidence to support a steric explanation for the differences in glycosylation.

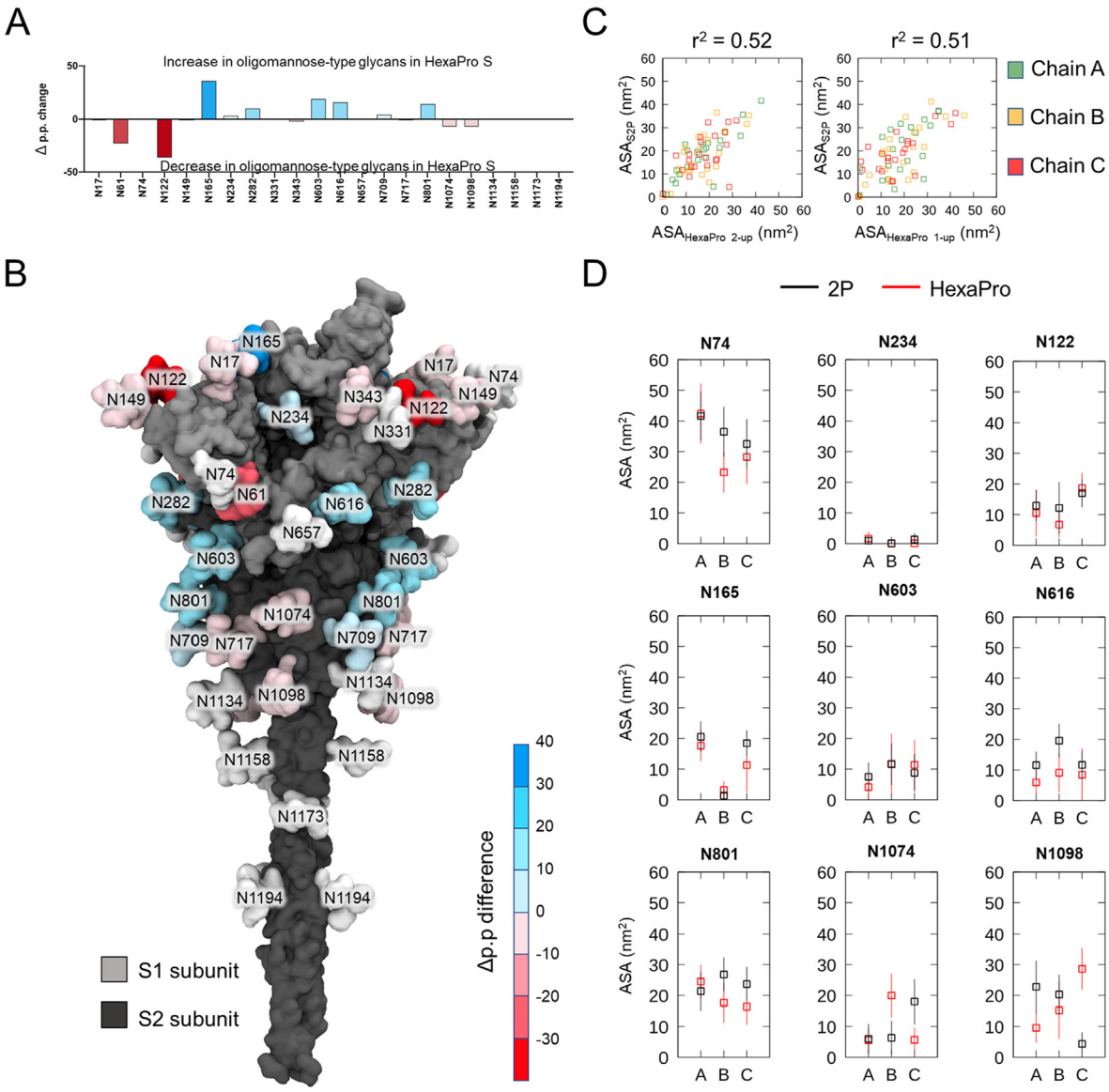


Figure 4. Comparison of glycan composition across prefusion-stabilized SARS-CoV-2 S protein. (A) The percentage point change in oligomannose-type glycan content between SARS-CoV-2 S protein, HexaPro and 2P. The percentage point (p.p.) difference on the y-axis is the arithmetic difference between the percentiles of oligomannose-type glycans between the two populations (here defined as: p.p. = % HexaPro – % 2P). Positive values (blue) indicate a higher abundance of oligomannose-type glycans in HexaPro relative to 2P. Negative values (red) indicate a lower abundance of oligomannose-type glycans in HexaPro relative to 2P. (B) A full length model of SARS-CoV-2 S protein with N-glycans colored based on the percentage point change values. The scale represents the differences in oligomannose-type glycans observed in HexaPro when compared to 2P protein. Colors correspond to p.p. values in Panel A. The model was constructed using prefusion structure of trimeric SARS-CoV-2 S glycoprotein, as detailed in Materials and Methods. (C) Correlation of ASA values between HexaPro and 2P S protein. The average ASA values of all glycans from three replica simulations of two-RBD-up HexaPro (left side) and one-RBD-up (right side) structures plotted against the average ASA values from simulations of the respective 2P structures. (D) The average ASA values for 2P (black) and HexaPro two-RBD-up (red) with error bars showing standard deviations along the trajectories and across three repeat simulations. The displayed sites are those with changes in the oligomannose content across both versions. N74 (high ASA values) and N234 (lowest ASA values) were used as a reference. The chain (A, B, or C) of the trimeric S protein is indicated along the x-axis.

Conservation of serological reactivity across recombinant SARS-CoV-2 S protein

To compare the serological reactivity of the recombinant 2P and HexaPro S protein, we tested the binding of different immunoglobulin isotypes in sera from subjects with or without a prior SARS-CoV-2 infection to these viral antigens. This extends the observations presented by Hsieh et al. by using a larger and geographically distinct donor group and by examining a range of antibody isotypes.[18] Sera from three groups of subjects from the United Kingdom were analyzed: hospitalized subjects (HS) which included individuals that were admitted to hospital and had RT-PCR confirmed SARS-CoV-2 infection; non-hospitalized convalescent (NHC) subjects, who were tested positive by clinically validated antibody test[35] and were not hospitalized and a negative control group, from whom sera was taken before 2019 (Pre-19). As expected, strong IgG, IgA and IgM responses were detected to both S glycoproteins in all hospitalized subjects with severe disease (Figure 5(A) and Supplementary Figure S5(A)). In contrast to the strong responses observed in severe cases, IgG, IgA and IgM responses were observed in the NHC subjects, and in some instances these responses were not above those of control sera (Figure 5(A)). There was minimal binding of IgG to S glycoprotein by Pre19 sera. Both 2P and HexaPro showed comparable serological reactivity, with a slightly increased level of binding of patient, but not control, sera to HexaPro in the NHC sera (Figure 5(A)). Overall, the signal: noise ratio was superior for HexaPro compared to 2P, particularly as sera were diluted (Supplementary Figure S5(B)), but overall, a key conclusion is that the HexaPro was not inferior to the 2P glycoprotein.

Although the small difference in glycosylation between 2P and HexaPro did not result in diminished antibody binding to HexaPro, it is conceivable that wider batch-to-batch variations could impact consistency of serological reactivity. We therefore tested sera for antibody binding to glycoprotein with substantially engineered glycosylation. We expressed HexaPro in the presence of kifunensine (Kif) which results in oligomannose-type glycans and we confirmed the altered processing state using hydrophilic interaction ultra-performance liquid chromatography (HILIC-UPLC) analysis of fluorescently labelled pool of N-linked glycans (Figure 5(D)). The WT HexaPro chromatogram shows the presence of diverse glycans in contrast to Kif-treated HexaPro, where $Man_9GlcNAc_2$ (M9) glycans were predominant (Figure 5(D)). We compared IgG, IgA, IgM and combined IgGAM antibody binding to HexaPro with no glycan engineering (which we term 'wildtype') with Kif-treated HexaPro. The serological response detected using SARS-CoV-2 positive sera were similar whether kif-treated HexaPro or WT HexaPro was used in the assay (Figure 5(B)). Negligible binding was observed with pre-19 sera (shown in black) with both kif-treated and WT HexaPro. Area Under the Curve (AUC) calculations confirmed that both WT S and Kif-treated S protein were bound similarly by IgG, IgA, IgM and IgGAM (Figure 5(C)). This suggests that the immune response elicited following SARS-CoV-2 infection, with respect to immunoglobulin binding, is not dictated by the glycan processing state of the S protein, as converting the glycans at every site from their native-like compositions, does not impact the detection sensitivity.

Perspectives

In this study we aimed to investigate a range of biophysical, glycan composition and serological binding properties of HexaPro, an expression-enhanced version of SARS-CoV-2 S protein. This version contains six proline mutations which lead to its high expression and stabilization, and promising immunogenic properties.[18,53–55,72] We reveal the comparable affinity of HexaPro to the ACE2 receptor with the earlier version, 2P which has been used in several vaccine studies,[73–76] and extensively as a serological reagent.[35,49,77] Furthermore, we determined the impact of these additional proline mutations on glycan composition using LC-ESI MS and compared it with the 2P. It was interesting to note that overall, the glycosylation was highly similar between both versions, except at a few sites. To further explore phenomena that could be directing these changes we performed MD simulations to decipher the conformational properties of both versions.

The structural site, N234 showed fully oligomannose-type glycans in both the versions suggesting that the central integrity of the protein architecture is unperturbed.[52,58,71] Overall, the oligomannose content was similar in both the S proteins, however, the $Man_9GlcNAc_2$ content was higher in HexaPro, suggesting less accessibility to glycan processing enzymes. This motivated us to look at the accessible surface area (ASA) of oligomannose-type glycans in 2P and HexaPro using MD simulations. We did not see much difference in ASA values at N-glycan sites which showed changes in oligomannose-content determined using LC-MS. Also, we observed highly similar protein dynamics in both versions. This suggests that both of the proteins have very similar conformational properties. However, the simulations sample a short timescale while glycans are processed over a much longer timescale and these simulations do not provide evidence for the steric changes predicted from the differential glycosylation observed in the LC-MS data. It is conceivable that the changes in oligomannose-type glycan content could instead be due to high expression of HexaPro resulting in changes in productive enzyme:sub-

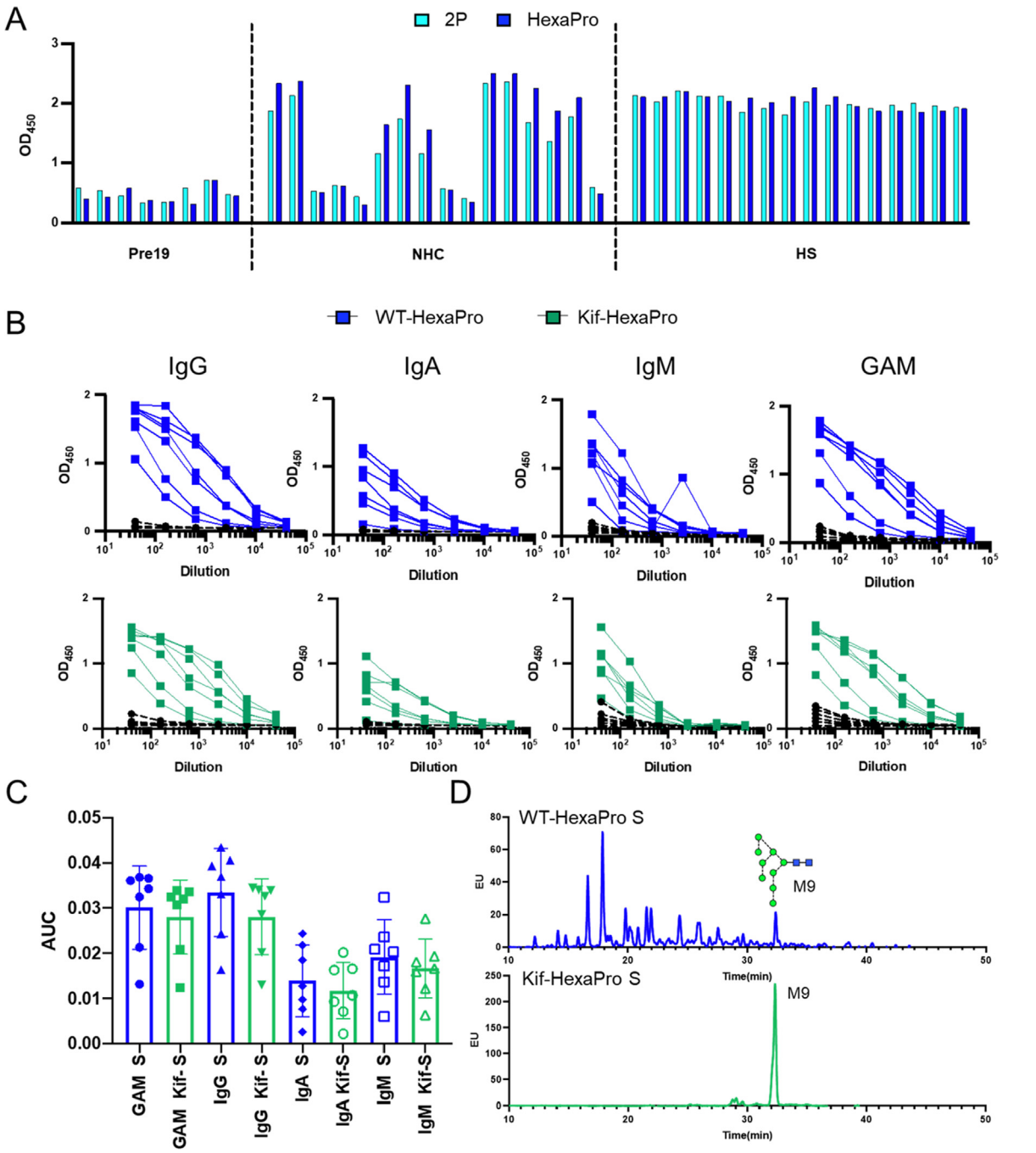


Figure 5. Antibody binding to spike glycoproteins. Individual serological responses from pre-2019 donors (Pre19, $n = 8$), non-hospitalized convalescent donors (NHC, $n = 16$) or PCR + hospitalized subjects (HS, $n = 16$) as determined by ELISA using HRP-labelled combined anti-IgG, IgA and IgM. A) Absorbance values of sera serially diluted from a starting dilution of 1:40 against 0.1 μg 2P (cyan bars) or HexaPro (blue bars). B) Absorbance values of sera serially diluted from a starting dilution of 1:40 from Pre-19 (black circle, dashed lines), NHC of HexaPro protein (blue squares) and Kifunensine-treated HexaPro (green squares) as determined by ELISA using HRP-labelled combined anti-IgG, IgA, IgM, and GAM. C) Area Under the Curve (AUC) of responses shown in figure B. The blue bars representing the AUC of HexaPro with IgG, IgA, IgM and GAM. The green bars representing the AUC of kif-treated HexaPro with different immunoglobulins. The mean ± standard deviation from the mean (SD) is plotted. D) HILIC-UPLC profile of N-linked glycans from WT (wildtype) HexaPro and Kifunensine-treated HexaPro produced in HEK 293F cells and purified by Ni^{+2} column followed by SEC. The blue peaks representing glycan spectra of WT-HexaPro. The green peaks representing glycan spectra of kifunensine-treated HexaPro showing only $Man_9GlcNAc_2$ (M9) glycans.

strate recognition events. Overall, however, these results suggest that there are limited structural and dynamic differences between 2P and HexaPro.

S proteins are being deployed in serological testing and have proven to be effective in confirming prior infection of SARS-CoV-2 in infected patients.[35,78,79] Due to its high expression, HexaPro could widen the availability of S protein for serological testing. In this study, we aimed to investigate the binding of HexaPro with sera IgG, IgA, and IgM to better understand its interaction with antibodies induced in COVID patients, in order to maximize the potential of HexaPro in these applications. We also see highly similar antibody binding with both 2P and HexaPro, suggesting that incorporation of mutations in HexaPro does not seem to affect immune recognition. Our results also reveal highly similar reactivity with glycoengineered HexaPro possessing all oligomannose-type glycans at all PNGS, indicating that sera binding is not readily impacted by the fine processing of the glycans of the S protein. Moreover, since the level of antibody binding was not significantly reduced after glycan engineering of HexaPro, the data may be interpreted as indicating that after natural infection most antibodies do not target epitopes that can be influenced by variations in glycan processing. If so, since antibodies from infected individuals can neutralise infection in vitro, it suggests that protection from infection is not associated with the targeting the terminal region of glycans. We note however that the predominant glycan at N234 is not changed by kifunensine. Overall, these observations suggest that variations in the S protein glycosylation of SARS-CoV-2 will not impact the serological assessments currently being performed across the globe.

Materials and Methods

Protein expression and purification

For expression of prefusion S ectodomain of SARS-CoV-2 HexaPro construct, the base construct of SARS-CoV-2 S 2P (GenBank: MN908947) having proline substitutions at residues 986 and 987, a "GSAS" substitution at furin cleavage site (residues 682–685) and C-terminal foldon trimerization motif, HRV3C protease recognition site, Twin-Strep-tag and octa-histidine tag cloned in mammalian vector pαH was used. The HexaPro construct was created by addition of four proline substitution (residues 817, 892, 899, 942) in 2P base construct as described previously.[18] Plasmid encoding S protein was used to transiently transfect FreeStyle 293-F cells (Thermo Fisher) using polyethylenimine (PEI). Cells were maintained at a density of 0.2–3.0 $\times$ 10^6 cells/mL at 37 °C, 8% CO_2 and 125 rpm shaking in FreeStyle 293F media (Fisher Scientific). Transfection mix was prepared in Opti-MEM (Fisher Scientific) using two solutions, DNA (310 μg/l) and PEI max reagent (1 mg/mL, pH 7) in a ratio of 1:3 in 25 mL of Opti-MEM respectively, followed by addition of DNA solution to the PEI mix and incubated for 30 minutes at room temperature. Cells were transfected at a density of 1 $\times$ 10^6 cells/mL and incubated at 37 °C, 8% CO_2 and 125 rpm shaking. To elicit oligomannose-type glycans on S glycoprotein, 20 μM kifunensine was added at the time of transfection. Culture was harvested after 7 days post transfection and the media was separated from the cells by centrifugation at 4,000 rpm for 30 minutes.

The supernatant was filtered using 0.22 μm pore size filter (Merck) followed by S protein purification using 5 mL His Trap FF column connected to Akta Pure system (GE Healthcare). Prior to loading the sample, the column was washed with washing buffer (50 mM Na_2PO_4, 300 mM NaCl) at pH 7. The sample was loaded onto the column at a speed of 2 mL/min. The column was washed with washing buffer (10 column volumes) containing 50 mM imidazole and eluted in 3 column volumes of elution buffer (300 mM imidazole in washing buffer). The elution was concentrated by a Vivaspin column (100 kDa cut-off) to a volume of 1 mL and buffer exchanged to phosphate buffered saline (PBS). Further, purification of target S protein fraction was carried out using size-exclusion chromatography using a Superdex 200 16 600 column (GE healthcare). The target fraction was concentrated in 100 kDa vivaspin (GE healthcare) to a volume of 1 mL.

Glycopeptide analysis by LC-MS

To perform the glycopeptide analysis using three protease enzymes, three 50 μg aliquots of SARS-COV-2 HexaPro were denatured for 1 h in 50 mM Tris/HCl, pH 8.0 containing 6 M of urea and 5 mM of dithiothreitol (DTT). Next, the S proteins were reduced and alkylated by adding 20 mM iodoacetamide (IAA) and incubated for 1 hr in the dark, followed by incubation with DTT to get rid of any residual IAA. The alkylated S proteins were buffer exchanged into 50 mM Tris/HCl, pH 8.0 using Vivaspin columns (3 kDa) and digested separately overnight using trypsin, chymotrypsin or alpha lytic protease (Mass Spectrometry Grade, Promega) at a ratio of 1:30 (w/w). The next day, the peptides were dried and extracted using C18 Zip-tip (MerckMilipore). The peptides were dried again, re-suspended in 0.1% formic acid and analyzed by nanoLC-ESI MS with an Easy-nLC 1200 (Thermo Fisher Scientific) system coupled to a Fusion mass spectrometer (Thermo Fisher Scientific) using higher energy collision-induced dissociation (HCD) fragmentation. Peptides were separated using an EasySpray PepMap RSLC C18 column (75 μm $\times$ 75 cm). A trapping column (PepMap 100 C18 3 μm (particle size), 75 μm $\times$ 2 cm) was used in line with the LC prior to separation with the analytical column. The LC conditions were as follows: 275 min linear gradient

consisting of 0–32% acetonitrile in 0.1% formic acid over 240 min followed by 35 minutes of 80% acetonitrile in 0.1% formic acid. The flow rate was set to 300 nL/min. The spray voltage was set to 2.7 kV and the temperature of the heated capillary was set to 40 °C. The ion transfer tube temperature was set to 275 °C. The scan range was 400–1600 *m/z*. The HCD collision energy was set to 50%, appropriate for fragmentation of glycopeptide ions. Precursor and fragment detection were performed using an Orbitrap at a resolution MS1 = 100,000. MS2 = 30,000. The Automatic gain control (AGC) target for MS1 = $4e^5$ and MS2 = $5e^4$ and injection time: MS1 = 50 ms MS2 = 54 ms.

Glycopeptide fragmentation data were extracted from the raw file using Byonic™ and Byologic™ software (Version 3.5; Protein Metrics Inc.). The glycopeptide fragmentation data were evaluated manually for each glycopeptide; the peptide was scored as true-positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the glycan identified. The MS data was searched using the Protein Metrics N-glycan library, along with filtering of sulfated type glycans using wildcard search for sulfation (0, 1, 2, 3). Then the N309 mammalian glycan library was modified to include the sulfated glycans identified in previous searches which was then further used for determining glycan composition of 2P and HexaPro S protein. The relative amounts of each glycan at each site as well as the unoccupied proportion were determined by comparing the extracted chromatographic areas for different glycotypes with an identical peptide sequence. All charge states for a single glycopeptide were summed. The precursor mass tolerance was set at 4 ppm and 10 ppm for fragments. A 1% false discovery rate (FDR) was applied. The relative amounts of each glycan at each site as well as the unoccupied proportion were determined by comparing the extracted ion chromatographic areas for different glycopeptides with an identical peptide sequence. Glycans were categorized according to the composition detected. HexNAc(2)Hex(9–5) was classified as M9 to M5. HexNAc(3)Hex(5–6)Neu5Ac(0–4) was classified as Hybrid with HexNAc(3)Hex(5–6)Fuc(1)Neu5Ac(0–4) classified as Fhybrid. Complex-type glycans were classified according to the number of processed antenna and fucosylation. Complex glycans are categorized as HexNAc(3)(X), HexNAc(3)(F)(X), HexNAc(4)(X), HexNAc(4)(F)(X), HexNAc(5)(X), HexNAc(5)(F)(X), HexNAc(6+)(X) and HexNAc(6+)(F)(X). Any glycan containing at least one sialic acid was counted as sialylated. The quantifications of glycan compositions were represented as the mean of three biological replicates ± standard error of the mean.

Determination of affinity using surface plasmon resonance (SPR)

Analysis of SARS-CoV-2 HexaPro binding with ACE2 protein was analyzed using a Biacore T200 (Cytiva/GE Healthcare). The proteins were buffer exchanged in the running buffer used for the SPR, HBS P+ (0.01 M HEPES pH 7.4, 0.15 M NaCl, 0.005% v/v Surfactant P20). Prior to injection of $NiCl_2$, metallic contaminants were removed via a pulse of 350 mM ethylenediaminetetraacetic acid (EDTA) at a flow rate of 30 μL/min for 1 min. Followed by loading of Ni^{2+} at a flow rate of 10 μL/min for 1 min. SARS-CoV-2 S protein (50 μg/mL), ligand was injected at 10 μL/min for 240 s. Flow cell 2–1 was used in which one of the cells was used as a control for determination of non-specific binding to the chip. Control cycles were performed by flowing the analyte (ACE2 protein) over the control cell having absence of ligand (S protein); negligible binding was indicated. The analyte was serially diluted ranging from 200 nM to 3.125 nM in triplicated along with HBS P+ buffer only as a control and were injected at a flow rate of 50 μL/min. Association and dissociation time was set as 300 s and 600 s respectively. After each cycle, the chip was regenerated with a pulse EDTA (350 mM) for 1 min at a flow rate of 30 μL/min. The 1:1 binding model was used for fitting the resulting data using Biacore Evaluation Software (GE Healthcare) and subsequently fitted curves were used to calculate K_D.

Patient sample collection and ethical approval

Serological responses to 2P and HexaPro forms of SARS-CoV-2 S glycoprotein were analyzed in samples from acutely unwell intensive treatment unit (ITU) patients with SARS CoV-2, convalescent individuals who have had mild disease and normal control sera from pre-2019. We have previously shown that severity of disease affects the quantitation of antibody[35] and so a spectrum of samples were used under ethics gained to aid assay development (NRES Committee West Midlands - South Birmingham 2002/201 Amendment Number 4, 24 April 2013) and from a Convalescent health care worker cohort (London - Camden & Kings Cross Research Ethics Committee 20/HRA/1817). Hospitalized subjects also provided nasopharyngeal swabs which were guanidine isothiocyanate inactivated, then analyzed by revers-transcriptase PCR directed against the SARS-CoV-2 ORF1ab and N genes (Viasure, CerTest Biotec). Pre-2019 negative control sera were obtained as part of a University of Birmingham study, reference ERN_16-178. All study participants gave written, informed consent and samples were fully anonymized.

Serum ELISA methodology

All sera were obtained by centrifugation of whole blood at 3500 RPM for 5 mins, then stored at –20 °C until use. Antibodies to S glycoprotein were detected using an in-house developed, high-sensitivity ELISA, as previously described.[35] In short, Nunc 96-well plates (ThermoFisher) or high binding plates (Corning) were coated with 2 μg/mL 2P or HexaPro or kifunensine HexaPro and blocked with 2% bovine serum albumin (Sigma Aldritch) (w/v) in PBS-T (Oxoid phosphate buffered saline with 0.1% Tween-20, Sigma Aldritch). Serum was initially diluted 1:40 in PBS-T and then serially diluted. Secondary antibodies (combined horse radish peroxidase-conjugated mouse anti-human IgG, A and M monoclonal antibodies) were diluted in PBS-T (anti–IgG R-10 1:8,000, anti–IgA MG4.156 1:4,000, and anti-IgM AF6 1:2,000; Abingdon Health). Signal was developed using TMB-Core (Bio-Rad) for between 6 and 12 minutes then stopped with 0.2 M H_2SO_4 (Sigma-Aldrich). Optical density (OD) at 450 nm was detected using the Dynex DSX automated liquid handler (Dynex Technologies, USA). Signal:noise ratio (S:N ratio) was calculated by dividing the individual OD values from PCR+ serum samples (signal) by the average OD from the pre-2019 negative controls (noise). Statistical significance was assessed using a RM 2-way ANOVA with Geisser-Greenhouse correction, followed by Sidak's multiple comparisons test, using Graphpad Prism version 8.

Integrative modelling and molecular dynamics simulation

Three S protein models were built using Modeller version 9.21[80]: (i) 2P with one RBD in the "up" conformation, (ii) HexaPro with one RBD in the "up" conformation, and iii) HexaPro with two RBDs in the "up" conformation. For 2P, the S protein ectodomain (ECD) was modelled using the cryo-EM structure of SARS-CoV-2 2P S ECD in the open state (PDB: 6VSB).[17] The ECD of the one-RBD-up HexaPro was modelled using the cryo-EM structure of SARS-CoV-2 HexaPro S with one RBD up (PDB: 6XKL).[18] For the two-RBD-up ECD, we performed molecular dynamics flexible fitting (MDFF),[81] whereby the atomic coordinates of the HexaPro one-RBD-up structure was fitted to the electron density map of HexaPro two-RBD-up (EMDB: EMD-22222).[18] The initial structure was prepared in VMD[82] and MDFF was performed in vacuum using NAMD version 2.11[83] with the CHARMM36 force field.[84] The MDFF simulation was performed until convergence using a range of scaling factors from 0.3 to 40, with secondary structural and domain restraints applied to the protein. The stalk and the transmembrane (TM) domain of all three models were built using the NMR structure of SARS-CoV HR2 domain (PDB: 2FXP)[85] and the NMR structure of HIV-1 gp-41 TM domain (PDB: 5JYN),[43] respectively, while missing loops in the NTD and the C terminus of the ECD were modelled based on a higher resolution cryo-EM structure of S protein in the closed state (PDB: 6XR8).[86] The same modelling protocol previously described to build a full-length model of the wild-type SARS-CoV-2 S protein was used.[87] The aim of the modelling and MD simulation study was to measure the accessible surface area (ASA) for each unprocessed glycosylation sites, in order to ascertain likely accessibility to glycan processing enzymes, as previously described in Allen et al.[52] Man-9 represents the primary substrate for glycan processing enzymes; as such Man-9 glycans were added to all 22 glycosylation sites using CHARMM-GUI Glycan Reader and Modeller.[88] The glycosylated S protein models were then embedded into a pre-equilibrated model of the endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membrane[89] built using CHARMM-GUI Membrane Builder.[90] The system was solvated with TIP3P water molecules and neutralized with 0.15 M NaCl salt. Stepwise energy minimization and equilibration simulations with decreasing amount of positional and dihedral restraints were conducted following the standard CHARMM-GUI protocols.[91] For each S protein model, three replicates of 200 ns production simulation were performed. The Nosé-Hoover thermostat was used to maintain the temperature at 310 K,[92,93] while a semi-isotropic coupling to the Parrinello-Rahman barostat was used to maintain the pressure at 1 atm.[94] The electrostatic interactions were calculated using the smooth particle mesh Ewald method with a real-space cut-off of 1.2 nm,[95] and the van der Waals interactions were truncated at 1.2 nm with a force switch smoothing applied between 1.0 and 1.2 nm. The simulations employed a 2 fs integration time step with the LINCS algorithm applied on all covalent bonds involving hydrogen atoms.[96] All simulations were run using GROMACS 2018[97] and the CHARMM36 force field.[84] ASA calculation was performed using the GROMACS tool *gmx sasa*, based on the last 50 ns of each trajectory. Comparison of ASA values between replica simulations showed overlapping values within error bars at most sites (Figure S6), and hence an average value was calculated to represent each glycosylation site.

Glycan analysis by HILIC-UPLC

Gel bands corresponding to the HexaPro S and glycan engineered HexaPro S protein (kifunensine-treated) were excised and washed three times with alternating 1 ml acetonitrile and water, incubating and shaking for 5 minutes following addition of each wash solution. All liquid was removed following the final wash stages and N-linked glycans were released in-gel using PNGaseF, (2 μg enzyme in 100 μL H_2O) (New England Biolabs) at 37 °C overnight. Following digestion, the liquid was removed from the gel

bands and placed into a separate Eppendorf. The gel bands were then washed twice with 100 μl MilliQ H_2O and this was pooled with the original solution. The extracted glycans were then dried completely in a speed vac at 30 °C.

The released glycans were subsequently fluorescently labelled with procainamide using 110 mg/ml procainamide and 60 mg/ml sodium cyanoborohydride in a buffer consisting of 70% DMSO, 30% acetic acid. For each sample, 100 μl of labelling mixture was added. Labelling was performed at 60 °C for 2 hours. Excess label and PNGaseF were removed using Spe-ed Amide-2 cartridges (Applied Separations). First, the cartridges were equilibrated sequentially with 1 ml acetonitrile, water and acetonitrile again. Then 1 ml of 95% acetonitrile was added to the procainamide-released glycan mixture and applied to the cartridge, allowing the cartridge to drain by gravity flow. After the mixture has emptied the cartridge, two washes using 97% acetonitrile were performed. To elute the labelled glycans 1 ml HPLC grade water was added to the cartridges and the elution collected. The elution was then dried completely using a speed vac, before resuspending in 24 μl of 50 mM ammonium formate.

A 6 μl aliquot of the resuspended glycans were mixed with 24 μl of acetonitrile and analysed on a Waters Acquity H-Class UPLC instrument with a Glycan BEH Amide column (2.1 mm × 150 mm, 1.7 μM, Waters), with an injection volume of 10 μl. A gradient of two buffers; 50 mM ammonium formate (buffer A) and acetonitrile (buffer B) was used for optimal separation. Gradient conditions were as follows: initial conditions, 0.5 ml/min 22% buffer A, increasing buffer A concentration to 44.1% over 57.75 minutes. Following this the concentration of buffer A was increase to 100% at 59.25 minutes and held there until 66.75 minutes and the flow rate was dropped to 0.25 ml/min, to fully elute from the column. Finally, the %A was reduced to 20% to prepare for subsequent runs. Wavelengths used for detection of the procainamide label were: excitation 310 nm, emission 370 nm. Data were processed using Empower 3 software (Waters, Manchester, UK). The relative abundance of oligomannose-type glycans was measured by digestion with Endoglycosidase H (per sample in 20 μl volume) (Endo H; New England Biolabs). A 6 μl aliquot of labelled glycans was combined with 1 μg endoH to a final volume of 20 μl. Digestion was performed overnight at 37 °C.

Digested glycans were cleaned using a 96-well PVDF protein-binding membrane (Millipore) attached to a vacuum manifold. Prior to application to the membrane, 100 μl HPLC-grade H_2O was added to each sample. Following equilibration with 150 μl ethanol, and 2 × 150 μl HPLC-grade H_2O, the sample was added to the 96-well plate and the flow-through was collected in a 96-well collection plate. Each well was then washed twice with HPLC-grade H_2O to a final elution volume of 300 μl. The elution was then dried completely at 30 °C. Prior to analysis the sample was resuspended in 6 μl ammonium formate and 24 μl acetonitrile and analysed as detailed above.

CRediT authorship contribution statement

Himanshi Chawla: Conceptualization, Investigation, Formal analysis, Writing – original draft, Writing – original draft, Writing – review & editing. **Sian E. Jossi:** Investigation, Formal analysis, Writing – review & editing. **Sian E. Faustini:** Investigation, Formal analysis, Writing – review & editing. **Firdaus Samsudin:** Investigation, Formal analysis, Writing – review & editing. **Joel D. Allen:** Investigation, Writing – review & editing. **Yasunori Watanabe:** Investigation, Writing – review & editing. **Maddy L. Newby:** Investigation, Writing – review & editing. **Edith Marcial-Juárez:** Investigation, Writing – review & editing. **Rachel E. Lamerton:** Investigation, Writing – review & editing. **Jason S. McLellan:** Conceptualization, Supervision, Funding acquisition, Writing – review & editing. **Peter J. Bond:** Conceptualization, Supervision, Funding acquisition, Writing – review & editing. **Alex G. Richter:** Conceptualization, Supervision, Funding acquisition, Writing – review & editing. **Adam F. Cunningham:** Conceptualization, Supervision, Funding acquisition, Writing – review & editing. **Max Crispin:** Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Acknowledgements

This work was supported by the International AIDS Vaccine Initiative (IAVI) through grant INV-008352/OPP1153692 funded by the Bill and Melinda Gates Foundation (M.C.). We also gratefully acknowledge support from the University of Southampton Coronavirus Response Fund (M. C.), a donation from the Bright Future Trust (M. C.), and the gift of laboratory consumables by Binding Site Ltd. P.J.B. and F.S. gratefully acknowledge support from BII core and grant FY21_CF_HTPO SEED_ID_BII_C211418001 funded by A*STAR. Simulations were performed on petascale computer cluster ASPIRE-1 at the National Supercomputing Centre of Singapore (NSCC), the A*STAR Computational Resource Centre (A*CRC), and the supercomputer Fugaku provided by RIKEN through the HPCI System Research Project (Project ID: hp200303). This work was funded in part by Welch Foundation

grant number F-0003-19620604 (J.S.M.). We would like to thank the University of Birmingham Clinical Immunology Service for their invaluable support in sample collection and processing. AFC and AGR are grateful for funding from the Medical Research Council, the Global Challenges Research Fund (GCRF) and the Institute for Global Innovation (IG1 Project, 3107) of the University of Birmingham. We would also like to thank the Wellcome Trust Mechanisms of Inflammatory Disease PhD Programme, 108871/Z/15/Z (R.E.L.). Y.W. has taken up a position at AstraZeneca; all experimental work was performed prior to this development.

Data availability

Mass spectrometry raw files have been deposited in the MassIVE proteomics databsase, MSV000087945.

Declaration of interests

J.S.M. is an inventor on the following U.S. patent applications: no. 62/412,703 ("Prefusion Coronavirus Spike Proteins and Their Use"); no. 62/972,886 ("2019-nCoV Vaccine"); no. 63/032,502 ("Engineered Coronavirus Spike (S) Protein and Methods of Use Thereof").

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2021.167332.

Received 9 August 2021;
Accepted 21 October 2021;
Available online 27 October 2021

Keywords:
SARS-CoV-2;
glycosylation;
serology;
glycoprotein;
antibody

References

1. Tokatlian, T., Read, B.J., Jones, C.A., Kulp, D.W., Menis, S., Chang, J.Y.H., Steichen, J.M., Kumari, S., et al., (2019). Innate immune recognition of glycans targets HIV nanoparticle immunogens to germinal centers. *Science* **363**, 649–654. https://doi.org/10.1126/SCIENCE.AAT9120.
2. Joe, C.C.D., Chatterjee, S., Lovrecz, G., Adams, T.E., Thaysen-Andersen, M., Walsh, R., Locarnini, S.A., Smooker, P., et al., (2020). Glycoengineered hepatitis B virus-like particles with enhanced immunogenicity. *Vaccine* **38**, 3892–3901. https://doi.org/10.1016/j.vaccine.2020.03.007.
3. Hyakumura, M., Walsh, R., Thaysen-Andersen, M., Kingston, N.J., La, M., Lu, L., Lovrecz, G., Packer, N.H., et al., (2015). Modification of Asparagine-Linked Glycan Density for the Design of Hepatitis B Virus Virus-Like Particles with Enhanced Immunogenicity. *J. Virol.* **89**, 11312–11322. https://doi.org/10.1128/JVI.01123-15.
4. Pritchard, L.K., Spencer, D.I.R., Royle, L., Vasiljevic, S., Krumm, S.A., Doores, K.J., Crispin, M., (2015). Glycan Microheterogeneity at the PGT135 Antibody Recognition Site on HIV-1 gp120 Reveals a Molecular Mechanism for Neutralization Resistance. *J. Virol.* **89**, 6952–6959. https://doi.org/10.1128/JVI.00230-15.
5. Pejchal, R., Doores, K.J., Walker, L.M., Khayat, R., Huang, P.-S., Wang, S.-K., Stanfield, R.L., Julien, J.-P., et al., (2011). A Potent and Broad Neutralizing Antibody Recognizes and Penetrates the HIV Glycan Shield. *Science* **334**, 1097–1103. https://doi.org/10.1126/science.1213256.
6. Huang, Y., Owino, S.O., Crevar, C.J., Carter, D.M., Ross, T.M., (2020). N-Linked Glycans and K147 Residue on Hemagglutinin Synergize To Elicit Broadly Reactive H1N1 Influenza Virus Antibodies. *J. Virol.* **94** https://doi.org/10.1128/JVI.01432-19.
7. Go, E.P., Herschhorn, A., Gu, C., Castillo-Menendez, L., Zhang, S., Mao, Y., Chen, H., Ding, H., et al., (2015). Comparative Analysis of the Glycosylation Profiles of Membrane-Anchored HIV-1 Envelope Glycoprotein Trimers and Soluble gp140. *J. Virol.* **89**, 8245–8257. https://doi.org/10.1128/JVI.00628-15.
8. Behrens, A.-J., Crispin, M., (2017). Structural principles controlling HIV envelope glycosylation. *Curr. Opin. Struct. Biol.* **44**, 125–133. https://doi.org/10.1016/j.sbi.2017.03.008.
9. Seabright, G.E., Cottrell, C.A., van Gils, M.J., D'addabbo, A., Harvey, D.J., Behrens, A.-J., Allen, J.D., Watanabe, Y., et al., (2020). Networks of HIV-1 Envelope Glycans Maintain Antibody Epitopes in the Face of Glycan Additions and Deletions. *Structure* **28**, 897–909.e6. https://doi.org/10.1016/j.str.2020.04.022.
10. Butler, M., (2006). Optimisation of the Cellular Metabolism of Glycosylation for Recombinant Proteins Produced by Mammalian Cell Systems. *Cytotechnology* **50**, 57–76. https://doi.org/10.1007/s10616-005-4537-x.
11. Ho, S.C.L., Koh, E.Y.C., van Beers, M., Mueller, M., Wan, C., Teo, G., Song, Z., Tong, Y.W., et al., (2013). Control of IgG LC:HC ratio in stably transfected CHO cells and study of the impact on expression, aggregation, glycosylation and conformational stability. *J. Biotechnol.* **165**, 157–166. https://doi.org/10.1016/j.jbiotec.2013.03.019.
12. Butler, M., Spearman, M., (2014). The choice of mammalian cell host and possibilities for glycosylation engineering. *Curr. Opin. Biotechnol.* **30**, 107–112. https://doi.org/10.1016/J.COPBIO.2014.06.010.
13. Yolitz, J., Schwing, C., Chang, J., Van Ryk, D., Nawaz, F., Wei, D., Cicala, C., Arthos, J., Fauci, A.S., (2018). Signal peptide of HIV envelope protein impacts glycosylation and antigenicity of gp120. *Proc. Natl. Acad. Sci.* **115**, 2443–2448. https://doi.org/10.1073/PNAS.1722627115.
14. Behrens, A.-J., Struwe, W.B., Crispin, M., (2017). Glycosylation profiling to evaluate glycoprotein immunogens against HIV-1. *Expert Rev. Proteomics* **14**, 881–890. https://doi.org/10.1080/14789450.2017.1376658.

15. Go, E.P., Ding, H., Zhang, S., Ringe, R.P., Nicely, N., Hua, D., Steinbock, R.T., Golabek, M., et al., (2017). Glycosylation Benchmark Profile for HIV-1 Envelope Glycoprotein Production Based on Eleven Env Trimers. *J. Virol.* **91** https://doi.org/10.1128/JVI.02428-16.
16. Alexander, S., Elder, J., (1984). Carbohydrate dramatically influences immune reactivity of antisera to viral glycoprotein antigens. *Science* **226**, 1328–1330. https://doi.org/10.1126/science.6505693.
17. Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.-L., Abiona, O., Graham, B.S., McLellan, J.S., (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263. https://doi.org/10.1126/science.abb2507.
18. Hsieh, C.-L., Goldsmith, J.A., Schaub, J.M., DiVenere, A. M., Kuo, H.-C., Javanmardi, K., Le, K.C., Wrapp, D., et al., (2020). Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505. https://doi.org/10.1126/SCIENCE.ABD0826.
19. Yang, X., Yu, Y., Xu, J., Shu, H., Xia, J., Liu, H., Wu, Y., Zhang, L., et al., (2020). Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir. Med.* **8**, 475–481. https://doi.org/10.1016/S2213-2600(20)30079-5.
20. Holshue, M.L., DeBolt, C., Lindquist, S., Lofy, K.H., Wiesman, J., Bruce, H., Spitters, C., Ericson, K., et al., (2020). First Case of 2019 Novel Coronavirus in the United States. *N. Engl. J. Med.* **382**, 929–936. https://doi.org/10.1056/NEJMoa2001191.
21. Watanabe, Y., Bowden, T.A., Wilson, I.A., Crispin, M., (1863). Exploitation of glycosylation in enveloped virus pathobiology. *Biochim. Biophys. Acta - Gen. Subj.* **2019**, 1480–1497. https://doi.org/10.1016/j.bbagen.2019.05.012.
22. Bosch, B.J., van der Zee, R., de Haan, C.A.M., Rottier, P.J. M., (2003). The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *J. Virol.* **77**, 8801–8811. https://doi.org/10.1128/JVI.77.16.8801-8811.2003.
23. Tortorici, M.A., Veesler, D., (2019). Structural insights into coronavirus entry. *In: Adv. Virus Res.*. Elsevier, pp. 93–116. https://doi.org/10.1016/bs.aivir.2019.08.002.
24. Letko, M., Marzi, A., Munster, V., (2020). Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature Microbiol.* **5**, 562–569. https://doi.org/10.1038/s41564-020-0688-y.
25. Brouwer, P.J.M.M., Caniels, T.G., van der Straten, K., Snitselaar, J.L., Aldon, Y., Bangaru, S., Torres, J.L., Okba, N.M.A.A., et al., (2020). Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643–650. https://doi.org/10.1126/science.abc5902.
26. Ju, B., Zhang, Q., Ge, J., Wang, R., Sun, J., Ge, X., Yu, J., Shan, S., et al., (2020). Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115–119. https://doi.org/10.1038/s41586-020-2380-z.
27. Rogers, T.F., Zhao, F., Huang, D., Beutler, N., Burns, A., He, W., Limbo, O., Smith, C., et al., (2020). Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science* **369**, 956–963. https://doi.org/10.1126/science.abc7520.
28. Yuan, M., Liu, H., Wu, N.C., Lee, C.-C.D., Zhu, X., Zhao, F., Huang, D., Yu, W., et al., (2020). Structural basis of a shared antibody response to SARS-CoV-2. *Science* **369**, 1119–1123. https://doi.org/10.1126/science.abd2321.
29. Wu, Y., Wang, F., Shen, C., Peng, W., Li, D., Zhao, C., Li, Z., Li, S., et al., (2020). A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. *Science* **368**, 1274–1278. https://doi.org/10.1126/science.abc2241.
30. Barnes, C.O., West, A.P., Huey-Tubman, K.E., Hoffmann, M.A.G., Sharaf, N.G., Hoffman, P.R., Koranda, N., Gristick, H.B., et al., (2020). Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **182**, 828–842.e16. https://doi.org/10.1016/j.cell.2020.06.025.
31. Tai, W., He, L., Zhang, X., Pu, J., Voronin, D., Jiang, S., Zhou, Y., Du, L., (2020). Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell. Mol. Immunol.* **17**, 613–620. https://doi.org/10.1038/s41423-020-0400-4.
32. Indenbaum, V., Koren, R., Katz-Likvornik, S., Yitzchaki, M., Halpern, O., Regev-Yochay, G., Cohen, C., Biber, A., et al., (2020). Testing IgG antibodies against the RBD of SARS-CoV-2 is sufficient and necessary for COVID-19 diagnosis. *PLoS One* **15**, https://doi.org/10.1371/journal.pone.0241164 e0241164.
33. Premkumar, L., Segovia-Chumbez, B., Jadi, R., Martinez, D.R., Raut, R., Markmann, A.J., Cornaby, C., Bartelt, L., et al., (2020). The receptor binding domain of the viral spike protein is an immunodominant and highly specific target of antibodies in SARS-CoV-2 patients. *Sci. Immunol.* **5**, eabc8413. https://doi.org/10.1126/sciimmunol.abc8413.
34. Stadlbauer, D., Amanat, F., Chromikova, V., Jiang, K., Strohmeier, S., Arunkumar, G.A., Tan, J., Bhavsar, D., et al., (2020). SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. *Curr. Protoc. Microbiol.* **57**, https://doi.org/10.1002/cpmc.100 e100.
35. Faustini, S.E., Jossi, S.E., Perez-Toledo, M., Shields, A.M., Allen, J.D., Watanabe, Y., Newby, M.L., Cook, A., et al., (2021). Development of a high-sensitivity ELISA detecting IgG, IgA and IgM antibodies to the SARS-CoV-2 spike glycoprotein in serum and saliva. *Immunology* **00**, imm.13349. https://doi.org/10.1111/imm.13349.
36. Bryan, A., Pepper, G., Wener, M.H., Fink, S.L., Morishima, C., Chaudhary, A., Jerome, K.R., Mathias, P.C., et al., (2020). Performance Characteristics of the Abbott Architect SARS-CoV-2 IgG Assay and Seroprevalence in Boise, Idaho. *J. Clin. Microbiol.* **58** https://doi.org/10.1128/JCM.00941-20.
37. Escribano, P., Álvarez-Uría, A., Alonso, R., Catalán, P., Alcalá, L., Muñoz, P., Guinea, J., (2020). Detection of SARS-CoV-2 antibodies is insufficient for the diagnosis of active or cured COVID-19. *Sci. Rep.* **10**, 1–7. https://doi.org/10.1038/s41598-020-76914-5.
38. Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., Veesler, D., (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell.* **181**, 281–292.e6. https://doi.org/10.1016/j.cell.2020.02.058.
39. Petherick, A., (2020). Developing antibody tests for SARS-CoV-2. *Lancet* **395**, 1101–1102. https://doi.org/10.1016/S0140-6736(20)30788-1.

40. Kirchdoerfer, R.N., Cottrell, C.A., Wang, N., Pallesen, J., Yassine, H.M., Turner, H.L., Corbett, K.S., Graham, B.S., et al., (2016). Pre-fusion structure of a human coronavirus spike protein. *Nature* **531**, 118–121. https://doi.org/10.1038/nature17200.
41. Pallesen, J., Wang, N., Corbett, K.S., Wrapp, D., Kirchdoerfer, R.N., Turner, H.L., Cottrell, C.A., Becker, M. M., et al., (2017). Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci.* **114**, E7348–E7357. https://doi.org/10.1073/pnas.1707304114.
42. Battles, M.B., Más, V., Olmedillas, E., Cano, O., Vázquez, M., Rodríguez, L., Melero, J.A., McLellan, J.S., (2017). Structure and immunogenicity of pre-fusion-stabilized human metapneumovirus F glycoprotein. *Nature Commun.* **8**, 1528. https://doi.org/10.1038/s41467-017-01708-9.
43. Dev, J., Park, D., Fu, Q., Chen, J., Ha, H.J., Ghantous, F., Herrmann, T., Chang, W., et al., (2016). Structural basis for membrane anchoring of HIV-1 envelope spike. *Science* **353**, 172–175. https://doi.org/10.1126/science.aaf7066.
44. Sanders, R.W., Moore, J.P., (2021). Virus vaccines: proteins prefer prolines. *Cell Host Microbe* **29**, 327–333. https://doi.org/10.1016/j.chom.2021.02.002.
45. Rutten, L., Gilman, M.S.A., Blokland, S., Juraszek, J., McLellan, J.S., Langedijk, J.P.M., (2020). Structure-Based Design of Prefusion-Stabilized Filovirus Glycoprotein Trimers. *Cell Rep.* **30**, 4540–4550.e3. https://doi.org/10.1016/j.celrep.2020.03.025.
46. Sanders, R.W., Derking, R., Cupo, A., Julien, J.-P., Yasmeen, A., de Val, N., Kim, H.J., Blattner, C., 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. https://doi.org/10.1371/journal.ppat.1003618.
47. Robinson, J.E., Hastie, K.M., Cross, R.W., Yenni, R.E., Elliott, D.H., Rouelle, J.A., Kannadka, C.B., Smira, A.A., et al., (2016). Most neutralizing human monoclonal antibodies target novel epitopes requiring both Lassa virus glycoprotein subunits. *Nature Commun.* **7**, 11544. https://doi.org/10.1038/ncomms11544.
48. Jiang, S., Hillyer, C., Du, L., (2020). Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends Immunol.* **41**, 355–359. https://doi.org/10.1016/j.it.2020.03.007.
49. Shields, A., Faustini, S.E., Perez-Toledo, M., Jossi, S., Aldera, E., Allen, J.D., Al-Taei, S., Backhouse, C., et al., (2020). SARS-CoV-2 seroprevalence and asymptomatic viral carriage in healthcare workers: a cross-sectional study. *Thorax* **75**, 1089–1094. https://doi.org/10.1136/thoraxjnl-2020-215414.
50. Perez-Toledo, M., Faustini, S.E., Jossi, S.E., Shields, A.M., Marcial-Juarez, E., Kanthimathinathan, H.K., Allen, J.D., Watanabe, Y., et al., (2021). SARS-CoV-2-specific IgG1/IgG3 but not IgM in children with Pediatric Inflammatory Multi-System Syndrome. *Pediatr. Allergy Immunol.* **32**, 1125–1129. https://doi.org/10.1111/pai.13504.
51. Watanabe, Y., Allen, J.D., Wrapp, D., McLellan, J.S., Crispin, M., (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330–333. https://doi.org/10.1126/science.abb9983.
52. Allen, J.D., Chawla, H., Samsudin, F., Zuzic, L., Shivgan, A.T., Watanabe, Y., He, W., Callaghan, S., et al., (2021). Site-Specific Steric Control of SARS-CoV-2 Spike Glycosylation. *Biochemistry* **60**, 2153–2169. https://doi.org/10.1021/acs.biochem.1c00279.
53. Seephetdee, C., Buasri, N., Bhukhai, K., Srisanga, K., Manopwisedjaroen, S., Lertjintanakit, S., Phueakphud, N., Pakiranay, C., et al., (2021). Mice Immunized with the Vaccine Candidate HexaPro Spike Produce Neutralizing Antibodies against SARS-CoV-2. *Vaccines* **9**, 498. https://doi.org/10.3390/vaccines9050498.
54. McMillan, C.L.D., Choo, J.J.Y., Idris, A., Supramaniam, A., Modhiran, N., Amarilla, A.A., Isaacs, A., Cheung, S.T.M., et al., (2021). Complete protection by a single dose skin patch delivered SARS-CoV-2 spike vaccine. *BioRxiv*. https://doi.org/10.1101/2021.05.30.446357. 2021.05.30.446357.
55. Walls, A.C., Miranda, M.C., Pham, M.N., Schäfer, A., Greaney, A., Arunachalam, P.S., Navarro, M.-J., Tortorici, M.A., et al., (2021). Elicitation of broadly protective sarbecovirus immunity by receptor-binding domain nanoparticle vaccines. *BioRxiv*. https://doi.org/10.1101/2021.03.15.435528. 2021.03.15.435528.
56. Allen, J.D., Watanabe, Y., Chawla, H., Newby, M.L., Crispin, M., (2021). Subtle Influence of ACE2 Glycan Processing on SARS-CoV-2 Recognition. *J. Mol. Biol.* **433**, https://doi.org/10.1016/j.jmb.2020.166762 166762.
57. Watanabe, Y., Berndsen, Z.T., Raghwani, J., Seabright, G. E., Allen, J.D., Pybus, O.G., McLellan, J.S., Wilson, I.A., et al., (2020). Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nature Commun.* **11**, 2688. https://doi.org/10.1038/s41467-020-16567-0.
58. Yao, H., Song, Y., Chen, Y., Wu, N., Xu, J., Sun, C., Zhang, J., Weng, T., et al., (2020). Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **183**, 730–738.e13. https://doi.org/10.1016/J.CELL.2020.09.018.
59. Brun, J., Vasiljevic, S., Gangadharan, B., Hensen, M., Chandran, A.V., Hill, M.L., Kiappes, J.L., Dwek, R.A., et al., (2021). Assessing Antigen Structural Integrity through Glycosylation Analysis of the SARS-CoV-2 Viral Spike. *ACS Cent. Sci.* **7**, 586–593. https://doi.org/10.1021/acscentsci.1c00058.
60. Pinto, D., Park, Y.-J., Beltramello, M., Walls, A.C., Tortorici, M.A., Bianchi, S., Jaconi, S., Culap, K., et al., (2020). Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295. https://doi.org/10.1038/s41586-020-2349-y.
61. Wang, C.-C., Chen, J.-R., Tseng, Y.-C., Hsu, C.-H., Hung, Y.-F., Chen, S.-W., Chen, C.-M., Khoo, K.-H., et al., (2009). Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proc. Natl. Acad. Sci.* **106**, 18137–18142. https://doi.org/10.1073/pnas.0909696106.
62. She, Y.-M., Li, X., Cyr, T.D., (2019). Remarkable Structural Diversity of N -Glycan Sulfation on Influenza Vaccines. *Anal. Chem.* **91**, 5083–5090. https://doi.org/10.1021/acs.analchem.8b05372.
63. Zhao, P., Praissman, J.L., Grant, O.C., Cai, Y., Xiao, T., Rosenbalm, K.E., Aoki, K., Kellman, B.P., et al., (2020). Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor. *Cell Host Microbe* **28**, 586–601.e6. https://doi.org/10.1016/j.chom.2020.08.004.
64. Klein, J.A., Zaia, J., (2020). Assignment of coronavirus spike protein site-specific glycosylation using GlycReSoft. *Biorxiv*. https://doi.org/10.1101/2020.05.31.125302.
65. McCoy, L.E., van Gils, M.J., Ozorowski, G., Messmer, T., Briney, B., Voss, J.E., Kulp, D.W., Macauley, M.S., et al.,

(2016). Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies. *Cell Rep.* **16**, 2327–2338. https://doi.org/10.1016/j.celrep.2016.07.074.
66. Doores, K.J., Bonomelli, C., Harvey, D.J., Vasiljevic, S., Dwek, R.A., Burton, D.R., Crispin, M., Scanlan, C.N., (2010). Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc. Natl. Acad. Sci.* **107**, 13800–13805. https://doi.org/10.1073/pnas.1006498107.
67. Crispin, M., Ward, A.B., Wilson, I.A., (2018). Structure and Immune Recognition of the HIV Glycan Shield. *Annu. Rev. Biophys.* **47**, 499–523. https://doi.org/10.1146/annurev-biophys-060414-034156.
68. H.-Y. Huang, H.-Y. Liao, X. Chen, S.-W. Wang, C.-W. Cheng, M. Shahed-Al-Mahmud, T.-H. Chen, J.M. Lo, et al., Impact of glycosylation on a broad-spectrum vaccine against SARS-CoV-2, (n.d.). https://doi.org/10.1101/2021.05.25.445523.
69. Watanabe, Y., Raghwani, J., Allen, J.D., Seabright, G.E., Li, S., Moser, F., Huiskonen, J.T., Strecker, T., et al., (2018). Structure of the Lassa virus glycan shield provides a model for immunological resistance. *Proc. Natl. Acad. Sci.* **115**, 7320–7325. https://doi.org/10.1073/pnas.1803990115.
70. Behrens, A.-J., Vasiljevic, S., Pritchard, L.K., Harvey, D.J., Andev, R.S., Krumm, S.A., Struwe, W.B., Cupo, A., et al., (2016). Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein. *Cell Rep.* **14**, 2695–2706. https://doi.org/10.1016/j.celrep.2016.02.058.
71. Casalino, L., Gaieb, Z., Goldsmith, J.A., Hjorth, C.K., Dommer, A.C., Harbison, A.M., Fogarty, C.A., Barros, E.P., et al., (2020). Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent. Sci.* **6**, 1722–1734. https://doi.org/10.1021/acscentsci.0c01056.
72. Arunachalam, P.S., Walls, A.C., Golden, N., Atyeo, C., Fischinger, S., Li, C., Aye, P., Navarro, M.J., et al., (2021). Adjuvanting a subunit COVID-19 vaccine to induce protective immunity. *Nature* **594**, 253–258. https://doi.org/10.1038/s41586-021-03530-2.
73. Francica, J.R., Flynn, B.J., Foulds, K.E., Noe, A.T., Werner, A.P., Moore, I.N., Gagne, M., Johnston, T.S., et al., (2021). Vaccination with SARS-CoV-2 Spike Protein and AS03 Adjuvant Induces Rapid Anamnestic Antibodies in the Lung and Protects Against Virus Challenge in Nonhuman Primates. *BioRxiv*. https://doi.org/10.1101/2021.03.02.433390. 2021.03.02.433390.
74. Vogel, A.B., Kanevsky, I., Che, Y., Swanson, K.A., Muik, A., Vormehr, M., Kranz, L.M., Walzer, K.C., et al., (2021). BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature* **592**, 283–289. https://doi.org/10.1038/s41586-021-03275-y.
75. K.S. Corbett, B. Flynn, K.E. Foulds, J.R. Francica, S. Boyoglu-Barnum, A.P. Werner, B. Flach, S. O'Connell, et al., Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates, Https://Doi.Org/10.1056/NEJMoa2024671. 383 (2020) 1544–1555. https://doi.org/10.1056/NEJMOA2024671.
76. Kuo, T.-Y., Lin, M.-Y., Coffman, R.L., Campbell, J.D., Traquina, P., Lin, Y.-J., Liu, L.T.-C., Cheng, J., et al., (2020). Development of CpG-adjuvanted stable prefusion SARS-CoV-2 spike antigen as a subunit vaccine against COVID-19. *Sci. Rep.* **101** (10), 1–10. https://doi.org/10.1038/s41598-020-77077-z.
77. Amanat, F., Stadlbauer, D., Strohmeier, S., Nguyen, T.H. O.O., Chromikova, V., McMahon, M., Jiang, K., Arunkumar, G.A., et al., (2020). A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nature Med.* **26**, 1033–1036 (accessed July 7, 2021) https://www.nature.com/articles/s41591-020-0913-5.
78. Kontou, P.I., Braliou, G.G., Dimou, N.L., Nikolopoulos, G., Bagos, P.G., (2020). Antibody Tests in Detecting SARS-CoV-2 Infection: A Meta-Analysis. *Diagnostics (Basel, Switzerland)* **10**, 319. https://doi.org/10.3390/diagnostics10050319.
79. Ladner, J.T., Henson, S.N., Boyle, A.S., Engelbrektson, A. L., Fink, Z.W., Rahee, F., D'ambrozio, J., Schaecher, K.E., et al., (2021). Epitope-resolved profiling of the SARS-CoV-2 antibody response identifies cross-reactivity with endemic human coronaviruses. *Cell Reports Med.* **2**, https://doi.org/10.1016/j.xcrm.2020.100189 100189.
80. Šali, A., Blundell, T.L., (1993). Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **234**, 779–815. https://doi.org/10.1006/jmbi.1993.1626.
81. Trabuco, L.G., Villa, E., Mitra, K., Frank, J., Schulten, K., (2008). Flexible Fitting of Atomic Structures into Electron Microscopy Maps Using Molecular Dynamics. *Structure.* **16**, 673–683. https://doi.org/10.1016/j.str.2008.03.005.
82. Humphrey, W., Dalke, A., Schulten, K., (1996). VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.
83. Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., Schulten, K., (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802. https://doi.org/10.1002/jcc.20289.
84. Huang, J., MacKerell, A.D., (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145. https://doi.org/10.1080/10810730902873927.Testing.
85. Hakansson-McReynolds, S., Jiang, S., Rong, L., Caffrey, M., (2006). Solution Structure of the Severe Acute Respiratory Syndrome-Coronavirus Heptad Repeat 2 Domain in the Prefusion State. *J. Biol. Chem.* **281**, 11965–11971. https://doi.org/10.1074/jbc.M601174200.
86. Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S.M., Walsh, R.M., Rawson, S., Rits-Volloch, S., et al., (2020). Distinct conformational states of SARS-CoV-2 spike protein. *Science* **21**, 1586–1592. https://doi.org/10.1016/j.solener.2019.02.027.
87. Zuzic, L., Samsudin, F., Shivgan, A.T., Raghuvamsi, P. V., Marzinek, J.K., Boags, A., Pedebos, C., Tulsian, N. K., et al., (2021). Uncovering cryptic pockets in the SARS-CoV-2 spike glycoprotein. *BioRxiv.* 2021.05.05.442536.
88. Park, S.J., Lee, J., Qi, Y., Kern, N.R., Lee, H.S., Jo, S., Joung, I., Joo, K., et al., (2019). CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* **29**, 320–331. https://doi.org/10.1093/glycob/cwz003.
89. van Meer, G., Voelker, D.R., Feigenson, G.W., Meer, G.V., (2008). Membrane lipids: where they are and how they behave. *Nature Rev. Mol. Cell Biol.* **9**, 112–124. https://doi.org/10.1038/nrm2330.Membrane.

90. Lee, J., Patel, D.S., Ståhle, J., Park, S.-J., Kern, N.R., Kim, S., Lee, J., Cheng, X., et al., (2019). CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *J. Chem. Theory Comput.* **15**, 775–786. https://doi.org/10.1021/acs.jctc.8b01066.
91. Lee, J., Cheng, X., Swails, J.M., Yeom, M.S., Eastman, P. K., Lemkul, J.A., Wei, S., Buckner, J., et al., (2016). CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **12**, 405–413. https://doi.org/10.1021/acs.jctc.5b00935.
92. Hoover, W.G., (1985). Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697. https://doi.org/10.1103/PhysRevA.31.1695.
93. Nosé, S., (1984). A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268. https://doi.org/10.1080/00268978400101201.
94. Parrinello, M., Rahman, A., (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190. https://doi.org/10.1063/1.328693.
95. Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G., (1995). A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593. https://doi.org/10.1063/1.470117.
96. Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M., (1997). LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* **18**, 1463–1472.
97. Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., Lindahl, E., (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.