The role of *N*-glycosylation in spike antigenicity for the SARS-CoV-2 Gamma variant

Cassandra L. Pegg[1*], Naphak Modhiran[1,2 *], Rhys Parry[1], Benjamin Liang[1], Alberto A. Amarilla[1], Alexander A. Khromykh[1,4], Lucy Burr[3], Paul R. Young [1,2,4] Keith Chappell[1,2,4], Benjamin L. Schulz[1,4 #], Daniel Watterson[1,2,4 #]

[1] School of Chemistry and Molecular Bioscience, University of Queensland, Brisbane, Australia
[2] Australian Institute for Bioengineering and Nanotechnology, University of Queensland, Brisbane, Australia
[3] Department of Respiratory Medicine, Mater Health Services, South Brisbane, QLD, Australia.
[4] Australian Infectious Disease Research Centre, Brisbane, Australia

[*]These authors contributed equally
#Corresponding authors: d.watterson@uq.edu.au and b.schulz@uq.edu.au

Abstract

The emergence of SARS-CoV-2 variants alters the efficacy of existing immunity towards the viral spike protein, whether acquired from infection or vaccination. Mutations that impact *N*-glycosylation of spike may be particularly important in influencing antigenicity, but their consequences are difficult to predict. Here, we compare the glycosylation profiles and antigenicity of recombinant viral spike of ancestral Wu-1 and the Gamma strain, which has two additional *N*-glycosylation sites due to amino acid substitutions in the N-terminal domain (NTD). We found that a mutation at residue 20 from threonine to asparagine within the NTD caused the loss of NTD-specific antibody binding. Glycan site-occupancy analyses revealed that the mutation resulted in *N*-glycosylation switching to the new sequon at N20 from the native N17 site. Site-specific glycosylation profiles demonstrated distinct glycoform differences between Wu-1, Gamma, and selected NTD variant spike proteins, but these did not affect antibody binding. Finally, we evaluated the specificity of spike proteins against convalescent COVID-19 sera and found reduced cross-reactivity against some mutants, but not Gamma spike compared to Wuhan spike. Our results illustrate the impact of viral divergence on spike glycosylation and SARS-CoV-2 antibody binding profiles.

Introduction

In December of 2019, a new coronavirus was detected in Wuhan, China which has subsequently been named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). SARS-CoV-2, the causative agent of COVID-19, has accounted for more than 651 million infections and more than

1

six million deaths worldwide (World Health Organization Coronavirus 2019 (COVID-19) Dashboard; https://covid19.who.int/, Accessed 03 January 2023). Virus genomic sequences are being generated and shared at an unprecedented rate, with more than one million SARS-CoV-2 sequences available via the Global Initiative on Sharing All Influenza DATA(GISAID), permitting near real-time surveillance of the unfolding pandemics.

The SARS-CoV-2 spike (S) glycoprotein is composed of two subunits: S1, which contains the receptor-binding domain (RBD) responsible for interaction with receptors on host cells, and S2, which mediates membrane fusion and viral entry. The S1 subunit contains two highly immunogenic domains, the N-terminal domain (NTD) and the RBD, which are the major targets of neutralising antibodies. While the RBD binds to the host cell receptor angiotensin-converting enzyme 2 (ACE2), the NTD is proposed to interact with auxiliary receptors including DC-SIGN/L-SIGN[1,2]. Serological analysis of plasma or serum from SARS-CoV-2 infected individuals has revealed only ~6-20% of circulating antibodies target the NTD compared to ~65-80% that target the RBD, with the remaining ~4-20% targeting the S2 subunit[3]. Nonetheless, NTD-targeting mAbs can neutralise SARS-CoV2-2 infection *in vitro* and *in vivo*, suggesting they could be useful for COVID-19 prophylaxis or treatment[3-8]. The NTD-directed monoclonal antibodies identified to-date recognise a glycan-free epitope named the NTD-supersite (residues 14-20, 140-158 and 245-264[3,9]). NTD-directed antibodies are a major selective pressure against the virus, and promote the emergence of NTD escape mutations of SARS-CoV-2 variants[3].

In late 2020, the SARS-CoV-2 variant of concern (VOC), designated P.1 or Gamma, was first detected in Manaus, Amazonas state, Brazil. This lineage was first detected in four travellers returning to Japan from Amazonas state on 2 January 2021[10] and was soon recognised as an emergent lineage in Manaus[11]. Gamma evolved from a local B.1.1.28 clade and replaced the parental lineage in <2 months. The Gamma spike protein harbours multiple substitutions: L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, H655Y, T1027I and V1176F (**Fig. 1A**). Most of these mutations localise to the NTD and RBD, which are the major targets of neutralising antibodies in convalescent and vaccinated individuals, raising concerns about the efficacy of available vaccines and therapeutic monoclonal antibodies towards this lineage. Some of these mutations had occurred in other VOCs: E484K was shared with B.1.351 (Beta), and N501Y was shared with B.1.1.7 (Alpha) and B.1.351 (Beta). Both these mutations reduce the neutralisation potency of some monoclonal antibodies[12,13].

A key post-translational modification that can also modulate the biophysical properties of proteins is glycosylation, whereby carbohydrate moieties (glycans) are attached to nascent proteins during synthesis in the endoplasmic reticulum (ER)[14]. Asparagine- (N-) glycans are attached at acceptor asparagine residues within an amino acid consensus motif or sequon (Asn-Xaa-Thr/Ser; Xaa≠Pro). This peptide motif binds with high affinity to the active site of the oligosaccharyltransferase, the enzyme that catalyses N-glycosylation in the ER[15]. Glycans play critical roles in virus-host interactions, including stabilising the conformation of viral and host proteins; serving as viral attachment factors, co-receptors, or receptors; promoting structural conformations that facilitate or enhance receptor binding; shielding or presenting underlying viral epitopes; and acting as antigens to direct immune respones[16-18]. Indeed, the spike protein is heavily glycosylated, with 22 occupied N-glycosylation sequons present per protomer[19] and glycosylation has been shown to influence SARS-CoV-2 infection[20]. Host glycans act as ligands, facilitating virus binding and uptake[21,22] and inhibiting cellular glycosylation blocks SARS-CoV-2 entry[23]. Furthermore, molecular dynamics (MD) studies indicate that spike N-glycosylation sites N165 and N234[24], and in addition N343[25], stabilise the protein in an open conformation allowing more favourable interactions with the ACE receptor, and that glycan-protein and glycan-glycan interactions form between ACE and spike[26].

The glycosylation profile of spike has been extensively investigated[27,28]. Minimal differences in glycosylation have been observed between vaccine constructs of trimeric ancestral spike held in a pre-fusion conformation and native (virion derived) spike[29]. Conversely, glycosylation differs substantially between native spike and a monomeric form of the S1 subunit from a non-stabilised protein construct[30]. These studies indicate that the glycosylation profile of spike is relatively stable when proteins are produced in similar cell lines, and overall spike architecture is conserved. However, specific amino acid substitutions in spike can modify site-specific glycosylation. A single substitution D614G (present in the Alpha variant and subsequent VOCs) and mutations specific to the Alpha variant, alter glycosylation profiles at selected sites compared to the ancestral Wuhan-Hu-1 strain[31,32].

In addition to sequence variations that may promote conformational changes, mutations can also introduce, remove, or change N-glycosylation sequons (NxT/S; x≠Pro) which may alter the structure and function of spike. For instance, a mutation ($NST_{372}>A$) in spike of human SARS-CoV-2 results in the loss of an N-glycosylation sequon when compared to other related coronaviruses[33]. MD

3

simulations[25] and binding assays[33] reveal that the loss of this site increases the binding affinity of the spike RBD to host ACE2 by stabilising the open conformation of the RBD. In VOCs, a mutation at position 19 in the Delta variant abolishes the *N*-glycosylation sequon containing site N17 (NLT$_{19}$>R) but has little impact on the overall glycosylation profile of spike compared to the ancestral strain[34]. The Gamma variant mutations T20N and R190S introduce two new *N*-glycosylation sequons, N$_{20}$RT and NLS$_{190}$, respectively. In addition, the Gamma substitution L18F changes the middle amino acid within the native *N*-glycosylation sequon containing site N17 (NF$_{18}$T). The introduced site N188 in Gamma is occupied with oligomannose type glycans[34,35], and occupancy of the second introduced site, N20, has been confirmed through cryo-EM mapping[13,35] and was found to contain predominantly complex type glycans[34,36]. Intriguingly, molecular modelling of the Gamma spike NTD predicts additional N20 and N188 glycans could sterically block NTD-specific antibodies targeting the supersite[35]. To our knowledge, the impact of these Gamma NTD mutations on spike-antibody binding has not been described.

With new and emerging variants and the threat of reduced neutralising antibody capacity, understanding the effects of different mutations on the antibody and glycosylation profiles of the NTD is critical to fully understand SARS-CoV-2 immunity during pandemics. Here, we produced nine trimeric prefusion-stabilised spike proteins that included the ancestral spike (Wu-1), the Gamma variant with all 12 substitutions and Wu-1 spike with single, double or triple mutagenesis of L18F, T20N or R190S. We analysed the mAb binding profiles of each spike protein and measured site-specific *N*-glycosylation occupancy and glycoform abundance at the acquired N20 and N188 *N*-glycosylation sites, and at the native *N*-glycosylation site N17. In addition, we measured NTD-specific IgG titers in sera from COVID-19 convalescent individuals.

Results and Discussion

### Sensitivity of Gamma spike and its mutations to NTD-specific mAbs

The emergence of the Brazilian VOC (Gamma lineage, P.1) impacted the epidemiological profile of COVID-19 cases due to its higher transmissibility and immune evasion ability[11,37-40]. It emerged after a period of rapid genetic diversification[38] and accumulated 17 non-synonymous defining mutations, ten of which are in the S gene. Mutations in the RBD at K417T, E484K and N501Y are involved in immune escape[9], while the NTD also contained five mutations, two of which, T20N and R109S, introduce new *N*-glycosylation sequons (**Fig.1A**). Interestingly, sequence analysis revealed

4

the early detection of mutations at T20N and R190S prior to the description of the gamma variant (**Fig.1B**). Although several studies documented Gamma's increased transmissibility and immune evasion, there is limited data about the glycosylation profile of Gamma and its impact on antibody binding, particularly within the NTD. Such data may help us better understand immune responses to emerging variants and help mitigate the severe impact of the ongoing pandemic.
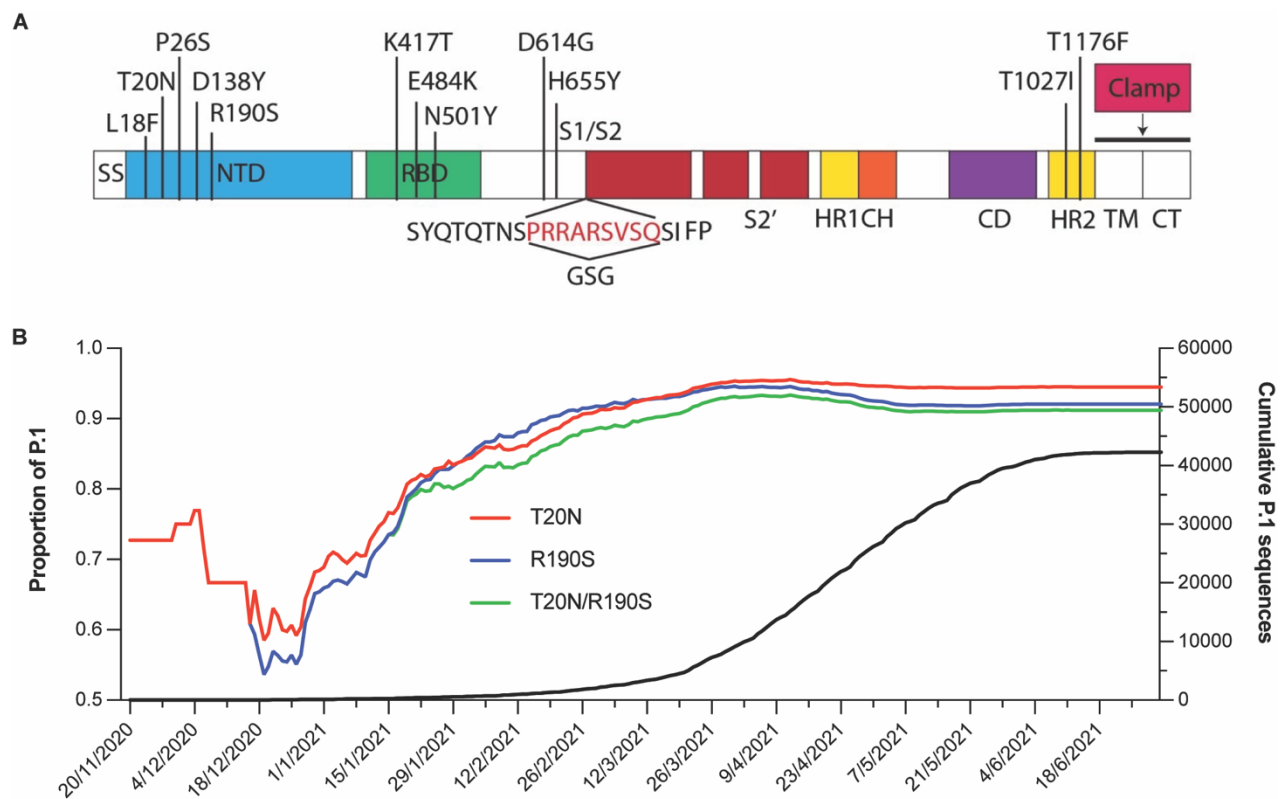
Fig. 1. Mutations of the spike protein of the SARS-CoV-2 Gamma variant and evolution of T20N and R190S mutations. (A) Schematic of the key domains of spike and locations of the 12 mutations present in the Gamma variant. SS, signal sequence; HR, heptad repeat; CH, central helix; CD, connector domain; TM, transmembrane domain; CT, cytoplasmic domain. The spike construct used in this study comprised residues 1204 of SARS-CoV-2 S with 'GSG' substituted at the furin cleavage site and a molecular clamp substituted for the TM and CT domains. **(B)** Early evolution and fixation of SARS-CoV-2 spike glycoforms in Gamma (P.1 and descendent lineages). The left y-axis shows the proportion of sequences with complete collection dates on GISAID harbouring T20N (red), R190S (blue) or both T20N/R190S (green). The right y-axis shows the cumulative number of P.1 sequences deposited on GISAID between 20 Nov 2020 to 1 Jul 2021.

To examine potential changes in antibody binding of Gamma spike, we generated recombinant trimeric SARS-CoV-2 spike proteins of the ancestral Wu-1 strain and Gamma with a modified furin cleavage site from Chinese Hamster Ovary (CHO) cells, as described previously[41] and assayed their binding to defined neutralising mAbs originally isolated from SARS-CoV-2 patients (**Fig. 2**). We found that structurally defined epitope-specific monoclonal antibodies (mAbs) including anti-RBD (B38, CB6, CR3022, 2M10B11 and S309) and anti-NTD (4A8, COVA2-17 and COVA1-22) bound differently to Wu-1 and Gamma spike proteins, consistent with previous studies[42-44]. All class I (B38 and CB6) and class IV (CR3022 and 2M10B11) RBD-specific mAbs exhibited lower binding affinity to Gamma spike than to Wu-1 spike, whereas the class III (S309) mAb showed no difference in binding. NTD-specific mAbs also had lower binding affinity to Gamma spike than to Wu-1 spike: 4A8, recognises the NTD supersite; and COVA1-22, where the binding mode has not yet been determined. Interestingly, the COVA2-17, NTD-specific neutralising antibody, bound strongly to Wu-1 spike, but its binding was completely abolished for Gamma spike. In addition to changes in antibody binding, Gamma spike showed slightly lower affinity to monomeric ACE2 (ACE2 FcM). Influenza hemagglutinin specific C05[45] and anti-clamp[46] antibodies served as negative and positive controls for these binding assays, and showed no differences in binding. Together, these data showed substantial changes to antibody binding across the RBD and NTD of Gamma spike.

6
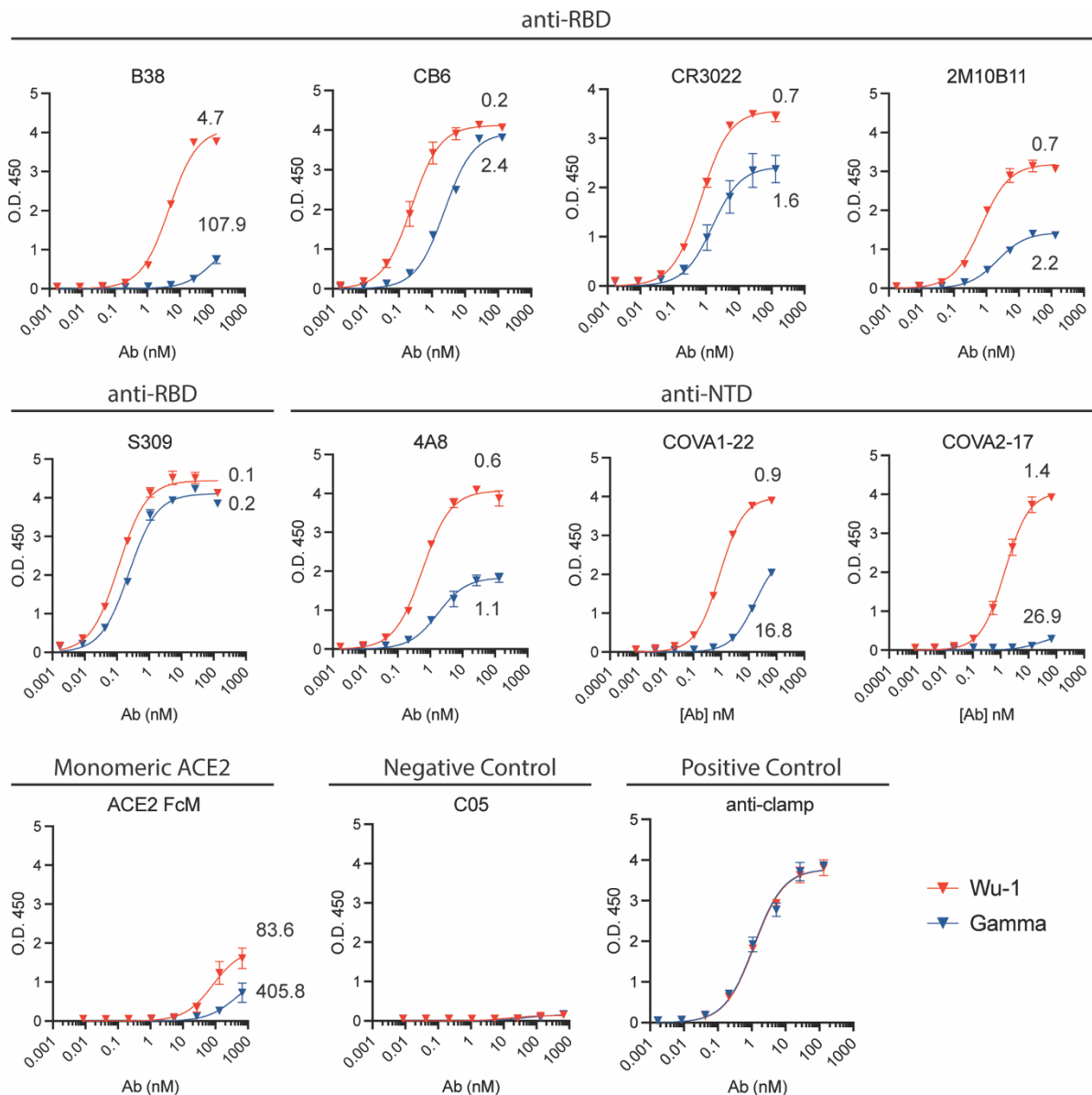
Fig. 2. Wu-1 and Gamma SARS-CoV-2 spike proteins exhibit different sensitivities to RBD and NTD specific mAbs. (A) Indirect ELISA show binding curves of Wu-1 (Red) and Gamma (Blue) spike to RBD- and NTD-specific antibodies as indicated. Kd shown for each curve, nM.

Antibody binding of trimeric spike from the Gamma variant with selected mutations

Although multiple antigenic sites, including one at the surface of the NTD, are present on both Wu-1 and Gamma, a single supersite of vulnerability is targeted by neutralising Abs elicited upon infection and vaccination[3,47]. This antigenic supersite (designated site i) comprises the NTD N-terminus (residues 14 to 20), a β-hairpin (residues 140 to 158), and a loop (residues 245 to 264). In the context of this antigenic supersite, the Gamma spike substitutions L18F and T20N are of particular interest, as they alter the native N-glycosylation sequon at N17 ($NL_{18}T > NF_{18}T$) and

introduce a new *N*-glycosylation sequon at N20 ($T_{20}RT > N_{20}RT$). Given the proximity of the additional Gamma mutation R190S to the antigenic supersite and predictions that glycans at the resulting introduced *N*-glycosylation sequon (NLS_{190}) could alter antibody binding[35], this substitution is also of high interest. To dissect the consequences of these substitutions, we performed all combinations of single, double or triple mutagenesis at L18F, T20N, and R190S on Wu-1 spike, resulting in seven proteins that were expressed and purified in addition to Wu-1 and Gamma. We then tested binding affinities for selected mAbs to these variant spike proteins (**Fig. 3**). As expected, these substitutions in the NTD did not affect binding of the anti-RBD mAb CB6. Despite anti-NTD mAb COVA1-22 having reduced binding to Gamma, its binding was not reduced to the variant spike proteins with targeted substitutions. In contrast, while anti-NTD COVA2-17 mAb bound Wu-1 spike with high affinity (Kd ~0.77 nM), the spike variants T20N (Kd ~8.35 nM), L18F/T20N (Kd >10 nM), and L18F/T20N/R190S (Kd >10 nM) showed substantially reduced or completely abrogated binding for this mAb. COVA2-17 binding affinity did not change with the L18F single mutation (Kd ~0.5 nM). This data suggests that introduction of a new *N*-glycosylation sequon with the T20N substitution, particularly in combination with L18F, eliminated binding of the anti-NTD COVA2-17 mAb.
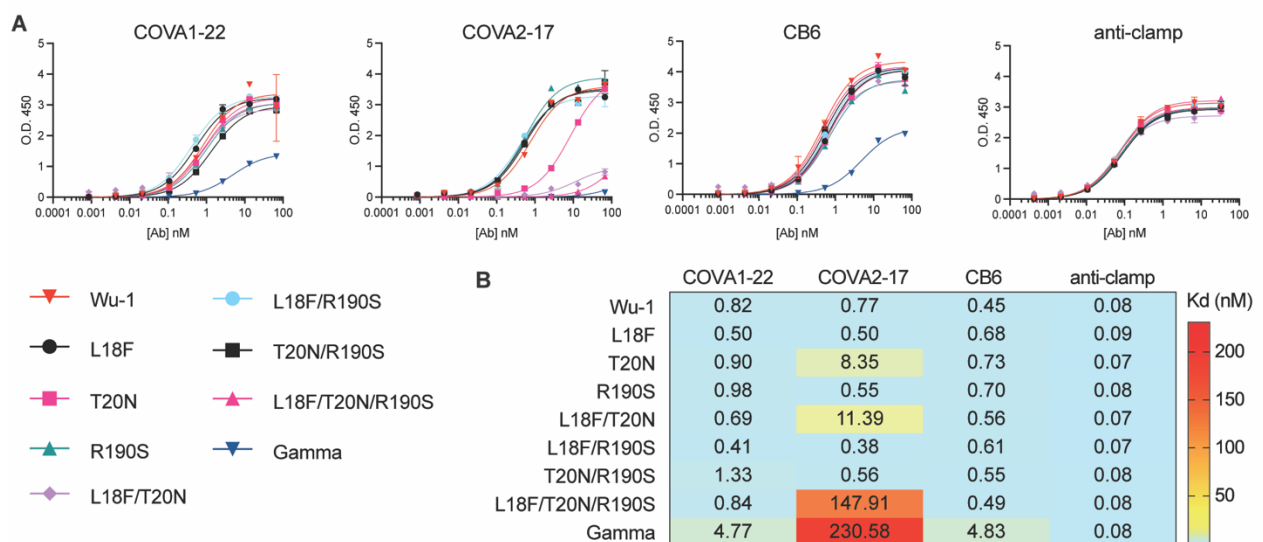
Fig. 3. Binding characterisation of SARS-CoV-2 spike proteins of Wu-1, Gamma and NTD mutants. (A). Indirect ELISA show binding curves of RBD- and NTD-specific antibodies as indicated. **(B)** Dissociation constant of mAbs against Wu-1 or Gamma, mutated SARS-CoV-2 spike proteins.

	COVA1-22	COVA2-17	CB6	anti-clamp
Wu-1	0.82	0.77	0.45	0.08
L18F	0.50	0.50	0.68	0.09
T20N	0.90	8.35	0.73	0.07
R190S	0.98	0.55	0.70	0.08
L18F/T20N	0.69	11.39	0.56	0.07
L18F/R190S	0.41	0.38	0.61	0.07
T20N/R190S	1.33	0.56	0.55	0.08
L18F/T20N/R190S	0.84	147.91	0.49	0.08
Gamma	4.77	230.58	4.83	0.08

N-glycosylation occupancy of introduced sites N20 and N188 and native N17 in the NTD of Gamma

The introduction of the Gamma mutations T20N and R190S creates two additional *N*-glycosylation sequons (N_{20}RT and NLS$_{190}$) in the spike protein, while L18F changes the middle residue of the native N17 sequon proximal to N20 ($N_{17}F_{18}TN_{20}$RT). To confirm if the decreased antibody binding we observed (**Fig. 3**) was associated with changes in glycosylation at this region, we performed biochemical confirmation of glycosylation site-occupancy in spike from Wu-1, Gamma, and variant spike proteins with selected Gamma substitutions introduced by site-directed mutagenesis. We measured site-specific *N*-glycosylation occupancy and glycoform abundance at the acquired N20 and N188 *N*-glycosylation sites, and at the native *N*-glycosylation site N17. To increase the accuracy with which we could measure site-specific *N*-glycosylation occupancy, we aliquoted pools of peptides and *N*-glycopeptides into two equal portions, one for endoglycosidase treatments with peptide-*N*-Glycosidase F (PNGase F) in ^{18}O water to obtain de-glycosylated peptides, and one left untreated for intact glycopeptide analyses. PNGase F removes *N*-glycans converting the previously glycosylated Asn to Asp which results in a mass shift of +0.984 compared to unmodified Asn. In the presence of ^{18}O water, the mass difference is increased to +2.984 Da which increases the confidence of site assignment, particularly when multiple *N*-glycosylation sites are potentially present on the same peptide. Using LC-MS/MS, we then measured the abundance of the resulting peptides and de-glycosylated peptides (Supplementary Data 1) and validated *N*-glycosylation occupancy through the identification of glycopeptides (Supplementary Data 2). We observed high occupancy at N17 in the Wu-1 spike. High occupancy was also observed at N17 in Wu-1 spike proteins with one or both of the introduced L18F and R190S mutations (**Fig. 4A**). However, in Gamma spike and all variant proteins with the T20N substitution, the introduction of the sequon at N20 resulted in a switch in site usage, with the new sequon being fully occupied and minimal occupancy at the original N17 site (**Fig. 4A and B**). Similarly, occupancy was high at all variants with the introduced *N*-glycosylation sequon at N188 (**Fig. 4C**). Together, this data showed that loss of binding of the COVA2-17 mAb was associated with *N*-glycosylation site switching from N17 to N20.

The *N*-linked sites N17 and N20 are adjacent to each other in Gamma ($N_{17}FTN_{20}$RT), and glycosylation sites can be skipped when sequons are closely spaced[48]. The competitive higher occupancy at N20 may be due to the type of amino acid in the middle of the *N*-linked sequon (NxS/T;

x≠P) as sequons with Leu/L or Phe/F, as at position X of the N17 sequon in the variant proteins, are less efficiently glycosylated[49].
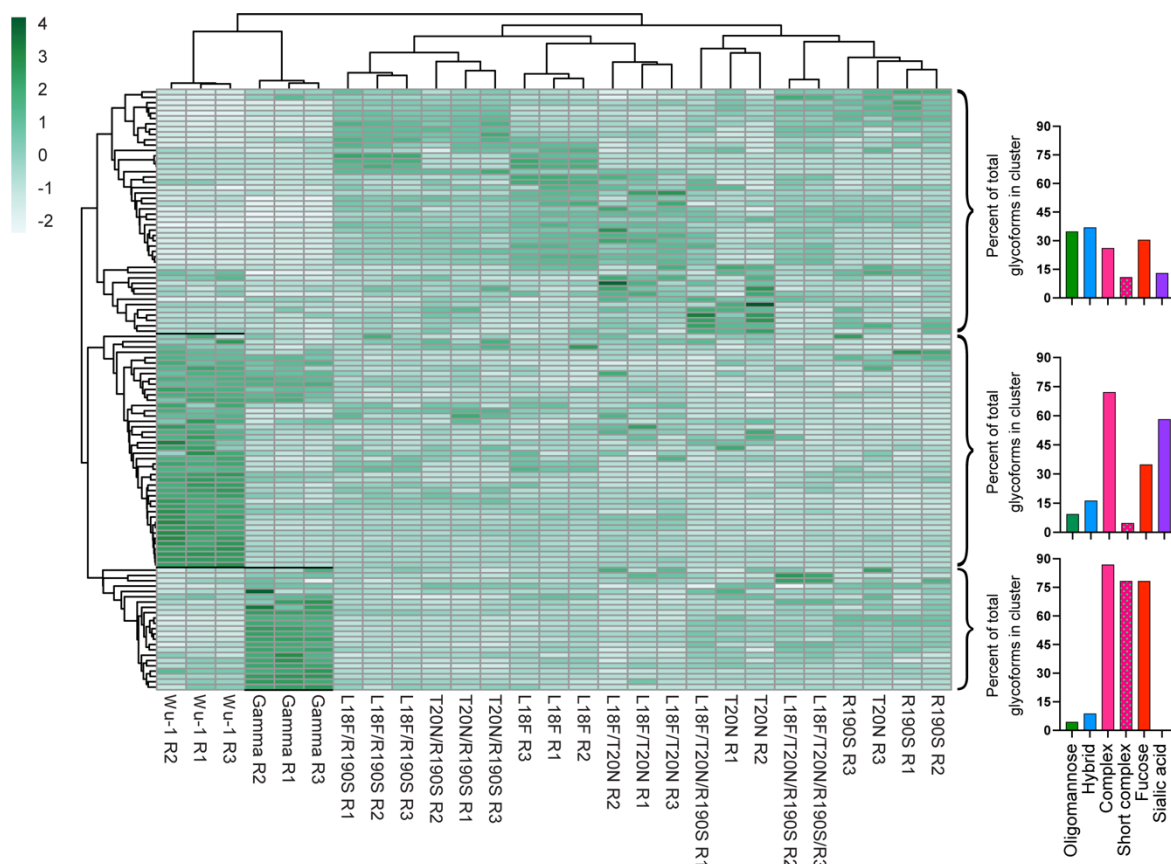
Glycoform abundance

The precise glycan structures present at specific *N*-glycosylation sites can be critical for various aspects of spike's function. We therefore used intact glycopeptide analyses to determine the glycosylation profiles at the native N17, the newly introduced N20 and N188, and six other native *N*-linked sites reliably measured after trypsin digestion (N122, N165, N234, N282, N801, N1098) (Supplementary Data 3 and Supplementary Figure 1). Principal component analysis (PCA) revealed clear separation of glycoforms between Wu-1, Gamma, and the variant spike proteins (**Fig. 4D**). The site-specific glycosylation profile of both N17 and N20 showed predominately complex *N*-glycans with a high degree of fucosylation. These profiles are similar to those reported at N17 after production of trimeric spike in mammalian cell lines[19,26,29,30,34,50,51], and are consistent with the surface-exposed location of N17 and N20 on the NTD, which would allow processing of glycans to more mature forms[34,35]. Glycan heterogeneity was greater at N17, with more sialylation (NeuAc, *N*-acetylneuraminic acid). The glycans at N20 were more homogenous and contained two predominant forms: $HexNAc_4Hex_3Fuc_1$ and $HexNAc_5Hex_3Fuc_1$ (HexNAc, *N*-acetylhexosamine; Hex, hexose; Fuc, fucose) (**Fig. 4E**). Compared to Wu-1, there was a substantial decrease in sialylation in Gamma and all seven variants, with the greatest difference apparent in the variants with the T20N mutation (**Fig. 4F**). Site-specific glycosylation analyses of N188 revealed predominantly oligomannose *N*-glycans for all proteins in which it was present (**Fig. 4G**), in agreement with previous reports[34,35], and consistent with its inaccessible location within a cleft in the NTD[34].
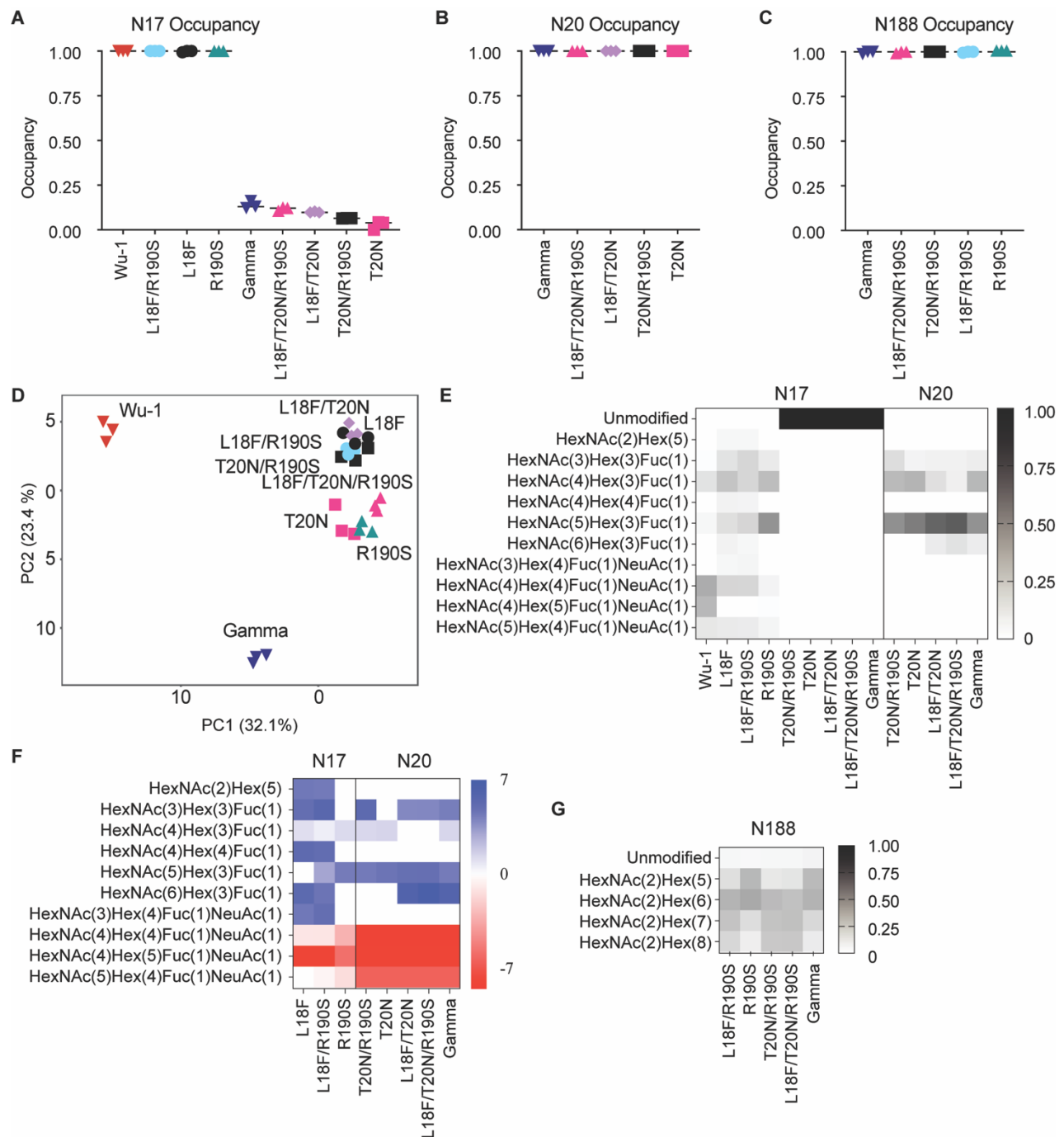
Fig. 4. Site-specific *N*-linked glycosylation occupancy and structural heterogeneity of SARS-CoV-2 spike proteins of Wu-1, Gamma and NTD mutants. Site-specific *N*-glycosylation occupancy at sites **(A)** N17, **(B)** N20 and **(C)** N188. **(D)** Principal component analysis (PCA) of glycoform abundances excluding N188. Site-specific glycoform analysis at sites **(E)** N17 and N20 with the **(F)** $\log_2$ fold change relative to Wu-1 (blue increased abundance, red decreased abundance, $P<0.05$). For N17 in (E) the unmodified intensity for T20N, L18F/T20N, L18F/T20N/R190S and Gamma were inferred from the occupancy analysis. **(G)** Site-specific glycoform analysis at site N188. HexNAc, *N*-acetylhexosamine; Hex, hexose; Fuc, fucose; NeuAc, *N*-acetylneuraminic acid.

To gain insights into the global glycosylation profiles of the various spike proteins, we performed a clustered heatmap analysis of their site-specific *N*-glycosylation profiles (**Fig. 5**). This analysis revealed that Gamma clustered distinctly from Wu-1, while the other variant spike glycoproteins clustered together. This clustering was primarily driven by abundant sialylation in Wu-1, in contrast to abundant fucosylation and short complex glycans (HexNAc$_{(3-6)}$Hex$_{(3)}$) in Gamma. Both Wu-1 and Gamma had similar afucosylation and oligomannose glycans, while these were more abundant in the site-directed mutagenesis variants. To confirm these differences were due to the protein sequence, rather than being expression artefacts, we performed independent expression, purification, and glycosylation analysis, which revealed these glycosylation profiles were robust and reproducible, and were driven by spike protein sequences (Supplementary Figure 1-3 and Supplementary Data 4). Together, our site-specific glycosylation analysis showed that in addition to causing *N*-glycosylation site switching from N17 to N20, the protein sequence changes in Gamma spike caused changes to the *N*-glycan structures at multiple sites across spike, with an overall decrease in sialylation compared to Wu-1.
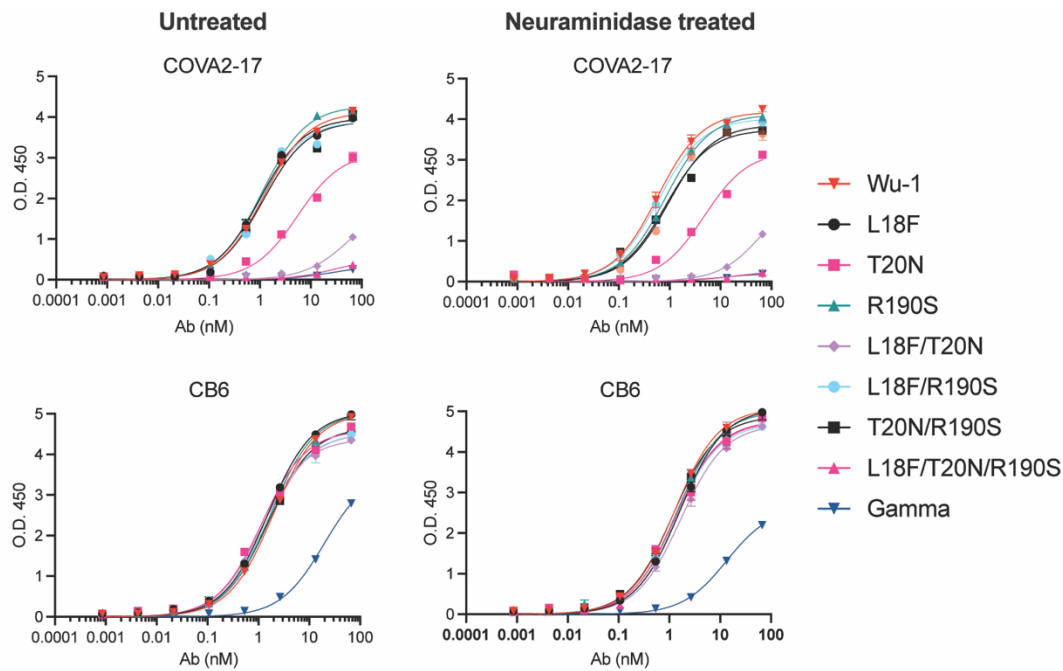
Fig. 5. SARS-CoV-2 spike proteins of Wu-1, Gamma and NTD mutants have diverse glycosylation profiles. Clustered heatmap of the relative abundance of all identified glycoforms in all replicates, excluding N188. Both rows and columns were clustered using correlation distance and average linkage. The percentage of types of glycans observed in each selected cluster are represented in bar graphs (Refer to the methods for the allocation of glycan-types). The full clustered heatmap with labelled rows (glycopeptide glycoforms) is shown in Supplementary Figure 3A.

Previous comparisons of the glycosylation of Gamma and Wu-1 spike produced in HEK cells observed less sialylation on Gamma, although less extreme than we observe here in spike produced in CHO cells[34,36]. It is therefore possible that the considerable differences in glycosylation of Gamma and Wu-1 we report here may be cell line specific. Nonetheless, the mechanisms underlying the reduced sialylation of *N*-glycans at diverse sites across Gamma spike compared with Wu-1 spike remain unclear.

Antibody binding of sialidase treated spike proteins

Our site-specific glycosylation analysis showed two key changes that correlated with reduced binding of select mAbs to Gamma compared to Wu-1 spike: glycosylation site switching from N17 to N20; and lower sialylation at *N*-glycosylation sites across Gamma spike. We therefore tested if this reduced binding was due to *N*-glycosylation site switching or changes in global sialylation. Variant spike proteins were treated with sialidase or left untreated, with removal of sialic acid confirmed by mass spectrometry (Supplementary Data 5). The proteins were then tested for anti-RBD (CB6) and anti-NTD (COVA2-17) binding. No change in mAb binding was observed in any neuraminidase-treated spike compared to untreated spike (**Fig. 6**). These results indicated that the significant changes in sialylation associated with the Gamma mutations were not responsible for the decreased binding of CB6 and COVA2-17 to Gamma, but that these changes in mAb binding were driven by protein sequence changes in the case of CB6, and *N*-glycosylation site switching in the case of COVA2-17.

13

Fig. 6. The binding affinities of SARS-CoV-2 spike proteins of Wu-1, Gamma and NTD mutants are not dependant on sialic acid content. Variant spike proteins were treated with neuraminidase for 1 h or left untreated. The proteins were then coated on ELISA plates and binding affinities determined for COVA2-17 or CB6 mAb.

These results are consistent with previous reports showing that global changes in spike glycosylation have minimal influence on its antibody binding profile. For instance, the binding profile of spike to sera from patients previously infected with COVID-19 was similar in spike with near-native glycosylation or which had been engineered to contain only oligomannose glycans[50]. Furthermore, glycosidase digestion of spike glycans to leave only the core GlcNAc elicited effective immune responses in *in vitro* and *in vivo* and protected mice from challenge with the Gamma variant[52]. Glycosylation of spike also likely affects its interactions with innate immune regulators and receptors. For example, spike lacking sialylation has decreased binding avidity for ACE2[52], while RBDs with complex glycans bind the receptor with greater affinity than RBDs with high mannose glycans[53]. Lectins, including the sialic acid-binding immunoglobulin-like lectin 1 (SIGLEC1) have also been demonstrated to function as attachment factors for SARS-CoV-2[54]. The global changes we observe in Gamma spike glycosylation may therefore have diverse unexpected biophysical or biological consequences.

14

ELISA of convalescence sera (3mths after exposure) against NTD domains

The mutations in Gamma spike reduced binding to mAbs with binding sites across the NTD and RBD (**Fig. 2** and **3**). To extend these observations, we characterised the total IgG antibody titers from infected individuals (3 months after infection in 2020, most probably with the Wu-1 strain) against Wu-1, Gamma, and the seven variant spikes. COVID-19 convalescence sera had high affinity to Wu-1, and approximately two to five times lower affinity to L18F, T20N and L18F/T20N/R190S, but not to Gamma (**Fig. 7**). It was unexpected that we observed a reduction in IgG binding titre to spike variants with few targeted site-directed mutations but not in Gamma spike, which contains those same mutations as well as additional sequence changes, including P26S and D138Y. It is possible that the single, double, or triple point mutant variants induce subtle global changes to spike structure or dynamics that perturb domain folding and IgG binding in isolation, but that these changes are rescued by the additional mutations in Gamma Spike.
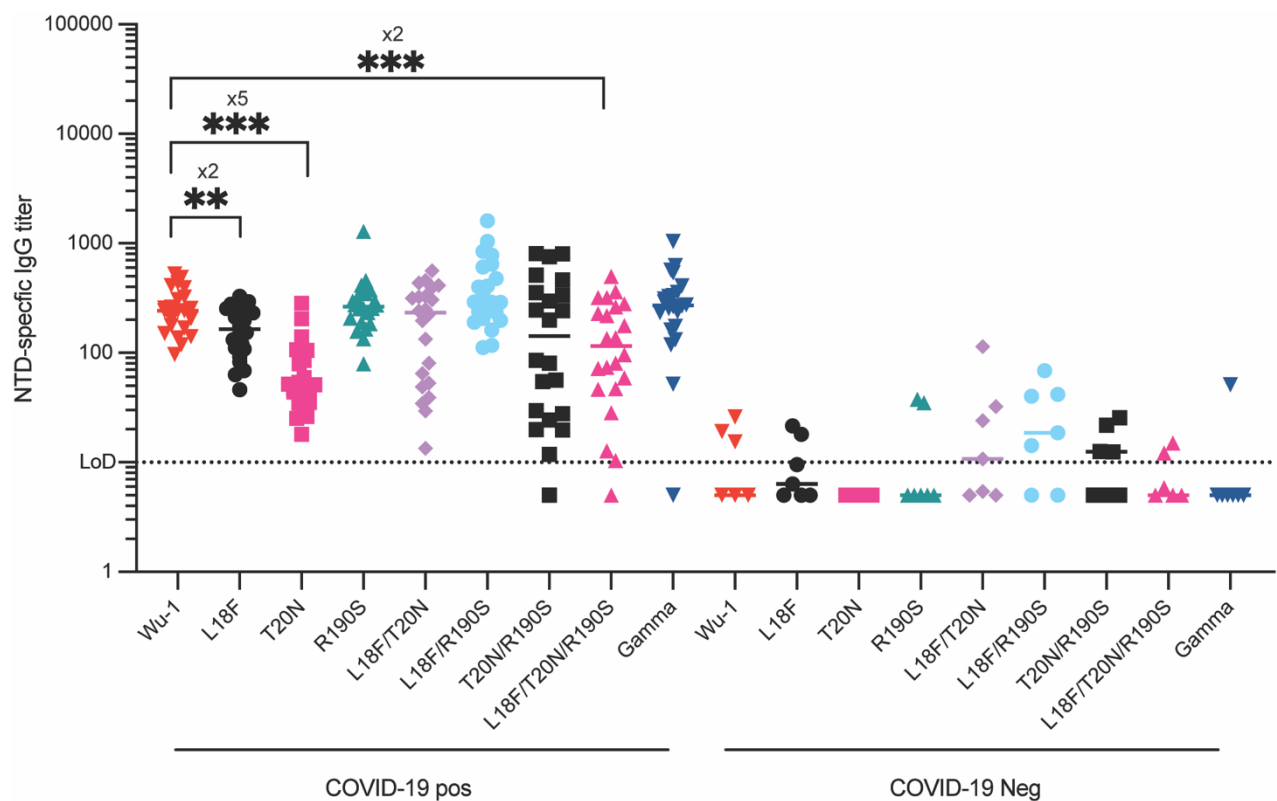
Fig. 7. NTD-specific IgG titer of 30 COVID-19 patients (COVID-19 pos) and close contact (COVID-19 neg). Blood samples were obtained from 30 COVID-19 patients and close contact at 3 months after infection IgG titers against the recombinant NTD domains (*** p-value <0.01, *** p-value<0.001*).

15

Conclusion

The emergence of new SARS-CoV-2 variants and evolution of existing variants will result in ongoing diversification of spike protein sequences in circulating virus. Many of these sequence changes may impact a key feature of spike – its glycan shield. The associated impacts of changes in glycosylation on spike function and antibody-mediated immunity are yet to be fully understood. Our results here emphasise that biochemical validation is required to confirm how protein sequence changes that introduce or alter *N*-glycosylation sequons actually affect glycosylation occupancy and associated antibody binding and neutralisation. We found that the structures of *N*-glycans across spike were altered in Gamma compared to Wu-1, highlighting that sequence changes in proteins can globally alter their glycosylation profile. We found no evidence that glycan structure affects antibody binding to spike, but it is likely to impact other spike functions including receptor binding, protein stability, and diverse protein-protein interactions. Our results highlight the importance of continued monitoring of the glycosylation profiles of spike from emerging variants and consideration of the structural, functional, and immunological functions of glycosylation in infection, and in vaccine design and production.

Methods

Analysis of P.1 SARS-CoV-2 Spike T20N and R190S variation over time

The metadata of SARS-CoV-2 sequences from the P.1 lineage a with complete collection dates (n=75,879) were downloaded from the GISAID EpiCoV™ portal. Accession IDs with combinations of individual T20N and R190S mutations and combined T20N/R190S mutations were programmatically identified using base Unix grep and sort commands. Cumulative proportions of sequences were calculated using Excel and visualised using GraphPad Prism (v9.0.0).

Recombinant protein production

For spike protein, soluble, trimeric spikes (residue 1-1204 amino acid) of SARS-CoV 2 /human/China/Wuhan-Hu-1/2019 (GenBank: MN908947), Gamma variant (P.1) (L18F, T20N, P26S, D138Y, R190S, K417T, E484K, N501Y, D614G, D655Y, T1027I and V1176F) were generated using an additional trimerisation domain[41,55]. Mutations at residue 18, (L18F), 20, (T20N) or 190 (R190S) were introduced as single/double/triple mutations by mutagenesis. Spike mutations were added through PCR into the codon-optimised Wuhan reference stain and were cloned into

16

pNBF plasmid via standard infusion cloning procedures. Spike proteins contain substitutions at the furin cleavage site (residues 682–685)[41]. For monoclonal antibodies, heavy and light chains of S309[56], CB6[57], B38[58], CR3022[58], 2M-10B11[5], COVA1-22, COVA2-17[59] and C05[45] were cloned into a human IgG1 expression vector as described previously[60]. For N-terminal domain, sequence encoding residue 1-305 amino acid of SARS-CoV 2 was cloned in-frame with mannose-binding protein for purification purpose. In brief, proteins were produced in ExpiCHO cells transfected with Expifectamine^{TM} reagent (Thermo Fisher Scientific) as per manufacturer's protocol[41]. Clarified supernatants were then filtered using 0.22 µm and were affinity purified, concentrated and buffer exchanged into PBS pH7.4.

Enzyme-linked immunosorbent assay (ELISA)

To test antibodies, SARS-CoV-2 spike variant proteins in PBS pH7.4 were immobilised on Maxisorb ELISA (Nunc) plates at a concentration of 2 µg/mL overnight. Serial 5-fold dilutions of Fc-fusion nanobodies or antibodies in 1X KPL in blocking buffer were incubated with the immobilised antigen, followed by incubation with HRP-coupled anti-human IgG (MilleniumScience) before adding the chromogenic substrate TMB (ThermoFisher). Reactions were stopped with 2 M H_2SO_4 and absorption measured at 450 nm.

For sialic acid removal, 15-20 µg of each sample was treated with Neuraminidase A (α2-3,6,8,9 specific, New England Bioloabs, catalogue number P0722S) using 4 U/µg of protein in 100 µL of PBS for 1 h at 37 °C. Confirmation of sialic acid removal was confirmed by mass spectrometry using 2.5-5 µg of protein, which was proteolytically digested as described in *Proteomic and glycoproteomic sample preparation*.

Proteomic and glycoproteomic sample preparation

Spike protein samples (10 µg of protein) were prepared in triplicate (n=27). Proteins were denatured, reduced, alkylated, quenched and methanol/acetone precipitated overnight as previously described[61]. Protein pellets were resuspended in 50 µL of 50 mM NH_4HCO_3 and digested for 16 h with trypsin (Sigma-Aldrich, Product code T6567) with an enzyme to protein ratio of 1:20. The proteolytic enzyme was inactivated by heating for 5 min at 95 °C before incubation with 1 mM phenylmethylsulfonyl fluoride at room temperature for 10 min. The proteolytic digests were aliquoted into two equal volumes (containing 5 µg of protein each, n=54) and one aliquot was dried.

17

At the same time, 15,000 U of glycerol free PNGase F (New England Bioloabs, catalogue number P0705S) was also dried. From here, all steps were preformed quickly to prevent exposure to air. The freshly dried PNGase F was resuspended in 50 mM NH_4HCO_3 prepared in 18O water (Sigma-Aldrich, Product code 329878) and 500 U of PNGase F was added to the dried proteolytic digests. The tubes were flushed with nitrogen and the samples were incubated at 37 °C for 4 h with gentle shaking. The peptides were desalted and concentrated with a C18 ZipTip (10 μL pipette tip with a 0.6 μL resin bed; Millipore, Part No: ZTC18S960) according to the manufacture's recommendations. Samples were dried and stored at -20 °C then reconstituted in 0.1% formic acid directly before mass spectrometry analysis. Protein samples (5 μg of protein) from a second batch of spike were prepared in triplicate (n=27) on a different day in the same manner described above except that the PNGase F steps were omitted.

Mass spectrometry for site occupancy analysis

The samples (+/- PNGase F) were randomised and run on an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) as three full experimental technical replicates (n=54) with ~100 ng of peptides injected for each chromatographic run using a Dionex UltiMate 3000 uHPLC system (Thermo Fisher Scientific, Bremen, Germany). Solvent A was 1% CH_3CN in 0.1% (v/v) aqueous formic acid and solvent B was 80% (v/v) CH_3CN containing 0.1% (v/v) formic acid. Samples were loaded onto a C18 Acclaim™ PepMap™ trap column (100 Å, 5 μm x 0.3 mm x 5 mm, Thermo Fisher Scientific) and washed for 3 min at 30 μL/min before peptides were eluted onto a C18 Acclaim™ PepMap™ column (100 Å, 5 μm x 0.75 mm x 150 mm, Thermo Fisher Scientific) at a flow rate of 0.3 μL/min. Peptides and glycopeptides were separated with a gradient of 3-8% solvent B in 5 min to 50% solvent B over 40 min. Survey scans of peptide precursors from *m/z* of 300 to 1800 were acquired in the Orbitrap at a resolution of 60K (full width at half-maximum, FWHM) at 400 *m/z* using an automatic gain control target of 1,000,000 and maximum injection time of 200 ms. The ten most intense precursors with an intensity over 1,000 and charge states above two were selected for fragmentation by beam-type collision-induced dissociation (CID) using a normalised collision energy of 35% with a precursor isolation window of 2 Da. Fragment ions were acquired in the Orbitrap at a resolution of 30K using an automatic gain control target of 100,000 and maximum injection time of 200 ms.

18

Mass spectrometry for glycoform analyses and confirmation of sialic acid removal

For glycoform analyses of each batch, samples were randomised and run as technical replicates with ~300 ng of peptides injected. Batch 1 (n=27) and 2 (n=27) were run on different days. For confirmation of sialic acid removal, each of the samples (+\- Neuraminidase A) were analysed (n=18). LC-ESI-MS/MS was performed using a Prominence nanoLC system (Shimadzu) coupled to a TripleTof 5600 instrument (SCIEX) using a Nanospray III interface using a data-dependent acquisition method as described[62] with minor alterations. Peptides and glycopeptides were separated with solvent A (1% CH_3CN in 0.1% (v/v) aqueous formic acid) and solvent B (80% (v/v) CH_3CN containing 0.1% (v/v) formic acid) with a gradient of 2-60% solvent B in 45 min. Full MS scans were obtained with a range of m/z 350-1800 with accumulation times of 0.5 s. High sensitivity mode was used where the top 20 most intense precursors with charge states of 2-5 and intensities greater than 100 were selected for fragmentation with a collision energy (CE) of 40 V and a 15 V spread. An accumulation time of 0.05 s was used with a scan range of m/z 40-1800 and precursors were excluded for 5 s after two selections.

Data analysis *N*-glycosylation occupancy at N17, N20 and N188

The Sequest HT node in Proteome Discoverer (v. 2.0.0.802 Thermo Fisher Scientific) was used to search RAW files from the PNGase F treated samples. The protein FASTA files contained the relevant SARS-CoV-2 spike-clamp protein sequence without the signal peptide combined with a custom contaminants protein database that included porcine trypsin and PNGase F and the UniProt proteome for Chinese hamster (*Cricetulus griseus*, UP000001075, downloaded 20 March 2018). Cleavage specificity was set to specific, allowing one missed cleavage. Mass tolerances of 10 ppm and 0.02 Da were applied to precursor and fragment ions, respectively. Cys-S-beta-propionamide was set as a static modification and dynamic modifications were set to deamidation of Asn (Asn>Asp, +0.984, and +2.984), pyroglutamic acid formation from N-terminal Gln and mono-oxidised Met. A maximum of four dynamic modifications were allowed per peptide. Confident peptide-to-spectrum matches (PSMs) were assigned using the "Fixed PSM Validator" node and a maximum Delta Cn of 0.05 was applied. Precursor peak areas were calculated using the Precursor Ions Area Detector node. The results from the Sequest HT search were investigated using a modified version of an in-house Python script [63]. Each Asn residue in an *N*-glycosylation consensus site from identified PSMs (with AUC values greater than 1x10^6) was assigned U (unmodified), D

(deamidated), ^{18}O-D (deamidated). Occupancy was defined as the proportion of the ion intensity peak area of a peptide with U or D (unmodified) or ^{18}O-D (occupied) identified in all charge states to the sum of the intensities of all peptide ions. All PSMs were manually validated.

Data analysis glycoforms

Searches of the proteolytic digests without PNGase F were used to investigate site-specific glycosylation. Glycopeptide identification was performed using Byonic (v2.13.2, Protein Metrics). Cleavage specificity was fully specific with two missed cleavages allowed. Mass tolerances of 50 ppm and 75 ppm were applied to precursor and fragment ions, respectively. For the samples acquired on an Elite Orbitrap instrument to validate occupancy data, mass tolerances of 10 ppm and 20 ppm were applied to precursor and fragment ions, respectively. Cys-S-beta-propionamide was set as a static modification and dynamic modifications were set to deamidation of Asn (common 2), pyroglutamic acid formation from N-terminal Gln (common 1) and mono-oxidised Met (common 1). The *N*-linked database (rare 1) contained 243 *N*-glycans (Supplementary Table **1**), the *O*-linked database (rare 1) was "6 most common" (Supplementary Table **2**) with the addition of three NeuGc glycans - HexNAc(1)Hex(1)NeuAc(1)NeuGc(1), HexNAc(1)Hex(1)NeuGc(1) and HexNAc(1)Hex(1)NeuGc(2). A maximum of three common modifications and one rare modification were allowed per peptide. Before the glycopeptide searches were conducted, a focused protein database was created in Byonic using the protein FASTA files described in the Sequest HT searches, and the parameters described above but without the dynamic glycan modifications. Glycopeptide PSMs were manually validated and occupancy was calculated using the Molecule Interface of Skyline (v20.2.0.343). A unique list of glycoforms was produced using the Byonic results. The unique list of glycopeptides contained the precursor name, precursor (*m/z*), charge and retention time were inserted into the Transition List in Skyline. A retention time window of 4 min was applied and quantification was performed at MS1 level with a mass tolerance of 0.05 (m/z). Instrument settings were specific to an ABSCIEX 5600 instrument. The results from the Skyline were converted to a readable format for GlypNirO[64]. Heatmaps and bar graphs were produced using PRISM v9.1.0 (GraphPad Software, La Jolla California USA). Principle component analysis (PCA) and clustered heatmaps were produced with ClustVis[65] where glycoform abundances were normalised to the summed abundance of all detected forms with the same site. Glycans were categorised as follows: Oligomannose, contains HexNAc(2); Hybrid, contains HexNAc(3); Complex, contains more than 3 HexNAc; Complex short, contains Hex(3) and more than 3 HexNAc; Fucose, contains at least one Fuc; Sialic acid, contains at least one NeuAc.

20

Human convalescent sera

Subjects were identified from samples stored in the David Serisiser Research Biobank (DSRB) (HREC/14/QPAH/275) at Mater Misericordiae Ltd. Previously biobanked sera from a total of 30 subjects who had been infected with SARS-CoV-2, and 6 non-infected close contacts were used in this project. Subjects were at least 3 months following a PCR defined SARS-CoV-2 infection and were not unwell at the time of sample collection. Samples were predominantly collected during mid to late 2020 and likely represent infection with the original "Wuhan" strain. In addition to sample collection, demographic data and clinical details pertaining to the original infection were also noted. All data and biological samples held within the DSRB are de-identified prior to sharing between collaborating sites. Access to samples for this project was approved by the DSRB biobank advisory committee under project number HREC/MML/68320.

Human ethics statement

All human sera collections were approved by the relevant ethics committees either by Mater Misericordiae Ltd Human Research Ethics Committee (Reference: HREC/MML/68320) or the University of Queensland (Reference: 2021/HE000139)

References

1   Thepaut, M. *et al.* DC/L-SIGN recognition of spike glycoprotein promotes SARS-CoV-2 trans-infection and can be inhibited by a glycomimetic antagonist. *PLoS Pathog* **17**, e1009576, doi:10.1371/journal.ppat.1009576 (2021).

2   Chiodo, F. *et al.* Novel ACE2-independent carbohydrate-binding of SARS-CoV-2 spike protein to host lectins and lung microbiota. *bioRxiv*, 2020.2005.2013.092478, doi:10.1101/2020.05.13.092478 (2020).

3   McCallum, M. *et al.* N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* **184**, 2332-2347 e2316, doi:10.1016/j.cell.2021.03.028 (2021).

4   Cerutti, G. *et al.* Potent SARS-CoV-2 neutralizing antibodies directed against spike N-terminal domain target a single supersite. *Cell Host Microbe* **29**, 819-833.e817, doi:10.1016/j.chom.2021.03.005 (2021).

5   Chi, X. *et al.* A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* **369**, 650-655, doi:10.1126/science.abc6952 (2020).

6   Liu, L. *et al.* Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450-456, doi:10.1038/s41586-020-2571-7 (2020).

7   Sun, D. *et al.* Potent neutralizing nanobodies resist convergent circulating variants of SARS-CoV-2 by targeting diverse and conserved epitopes. *Nat Commun* **12**, 4676, doi:10.1038/s41467-021-24963-3 (2021).

8   Voss, W. N. *et al.* Prevalent, protective, and convergent IgG recognition of SARS-CoV-2 non-RBD spike epitopes. *Science* **372**, 1108-1112, doi:10.1126/science.abg5268 (2021).

9   Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* **19**, 409-424, doi:10.1038/s41579-021-00573-0 (2021).

10   Fujino, T. *et al.* Novel SARS-CoV-2 variant in travelers from Brazil to Japan. *Emerg Infect Dis* **27**, 1243-1245, doi:10.3201/eid2704.210138 (2021).

11   Faria, N. R. *et al.* Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815-821, doi:10.1126/science.abh2644 (2021).

12   Wang, P. *et al.* Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130-135, doi:10.1038/s41586-021-03398-2 (2021).

13   Zhang, J. *et al.* Membrane fusion and immune evasion by the spike protein of SARS-CoV-2 Delta variant. *Science* **374**, 1353-1360, doi:10.1126/science.abl9463 (2021).

14   Lee, H. S., Qi, Y. & Im, W. Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. *Sci Rep* **5**, 8926, doi:10.1038/srep08926 (2015).

15   Shrimal, S. & Gilmore, R. Oligosaccharyltransferase structures provide novel insight into the mechanism of asparagine-linked glycosylation in prokaryotic and eukaryotic cells. *Glycobiology* **29**, 288-297, doi:10.1093/glycob/cwy093 (2019).

16   Watanabe, Y., Bowden, T. A., Wilson, I. A. & Crispin, M. Exploitation of glycosylation in enveloped virus pathobiology. *Biochim Biophys Acta Gen Subj* **1863**, 1480-1497, doi:10.1016/j.bbagen.2019.05.012 (2019).

17   Bagdonaite, I. & Wandall, H. H. Global aspects of viral glycosylation. *Glycobiology* **28**, 443-467, doi:10.1093/glycob/cwy021 (2018).

18   Dugan, A. E., Peiffer, A. L. & Kiessling, L. L. Advances in glycoscience to understand viral infection and colonization. *Nat Methods* **19**, 384-387, doi:10.1038/s41592-022-01451-0 (2022).

19   Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S. & Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330-333, doi:10.1126/science.abb9983 (2020).

20   Reis, C. A., Tauber, R. & Blanchard, V. Glycosylation is a key in SARS-CoV-2 infection. *J Mol Med (Berl)* **99**, 1023-1031, doi:10.1007/s00109-021-02092-0 (2021).

21   Nguyen, L. *et al.* Sialic acid-containing glycolipids mediate binding and viral entry of SARS-CoV-2. *Nat Chem Biol* **18**, 81-90, doi:10.1038/s41589-021-00924-1 (2022).

22   Clausen, T. M. *et al.* SARS-CoV-2 infection depends on cellular heparan sulfate and ACE2. *Cell* **183**, 1043-1057 e1015, doi:10.1016/j.cell.2020.09.033 (2020).

23   Yang, Q. *et al.* Inhibition of SARS-CoV-2 viral entry upon blocking N- and O-glycan elaboration. *Elife* **9**, e61552, doi:10.7554/eLife.61552 (2020).

24   Casalino, L. *et al.* Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent Sci* **6**, 1722-1734, doi:10.1021/acscentsci.0c01056 (2020).

25   Harbison, A. M. *et al.* 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**, 386-395, doi:10.1039/d1sc04832e (2022).

26   Zhao, P. *et al.* Virus-receptor interactions of glycosylated SARS-CoV-2 spike and human ACE2 receptor. *Cell Host Microbe* **28**, 586-601 e586, doi:10.1016/j.chom.2020.08.004 (2020).

27   Chawla, H., Fadda, E. & Crispin, M. Principles of SARS-CoV-2 glycosylation. *Curr Opin Struct Biol*, 102402, doi:10.1016/j.sbi.2022.102402 (2022).

28   Praissman, J. L. & Wells, L. Proteomics-based insights into the SARS-CoV-2-mediated COVID-19 pandemic: A review of the first year of research. *Mol Cell Proteomics* **20**, 100103, doi:10.1016/j.mcpro.2021.100103 (2021).

29   Allen, J. D. *et al.* Site-specific steric control of SARS-CoV-2 spike glycosylation. *Biochemistry* **60**, 2153-2169, doi:10.1021/acs.biochem.1c00279 (2021).

30   Brun, J. *et al.* Assessing antigen structural integrity through glycosylation analysis of the SARS-CoV-2 viral spike. *ACS Cent Sci* **7**, 586-593, doi:10.1021/acscentsci.1c00058 (2021).

31   Wang, D. *et al.* N-glycosylation profiles of the SARS-CoV-2 spike D614G mutant and its ancestral protein characterized by advanced mass spectrometry. *Sci Rep* **11**, 23561, doi:10.1038/s41598-021-02904-w (2021).

32   Kuo, C. W. *et al.* Distinct shifts in site-specific glycosylation pattern of SARS-CoV-2 spike proteins associated with arising mutations in the D614G and Alpha variants. *Glycobiology* **32**, 60-72, doi:10.1093/glycob/cwab102 (2022).

33   Kang, L. *et al.* A selective sweep in the Spike gene has driven SARS-CoV-2 human adaptation. *Cell* **184**, 4392-4400 e4394, doi:10.1016/j.cell.2021.07.007 (2021).

34   Newby, M. L. *et al.* Variations within the glycan shield of SARS-CoV-2 impact viral spike dynamics. *J Mol Biol*, 167928, doi:10.1016/j.jmb.2022.167928 (2022).

35   Yang, T.-J. *et al.* Structure-activity relationships of B.1.617 and other SARS-CoV-2 spike variants. *bioRxiv*, 2021.2009.2012.459978, doi:10.1101/2021.09.12.459978 (2021).

36   Shajahan, A., Pepi, L., Kumar, B., Murray, N. & Azadi, P. Site Specific N- and O-glycosylation mapping of the Spike Proteins of SARS-CoV-2 Variants of Concern. *Res Sq*, doi:10.21203/rs.3.rs-2188138/v1 (2022).

37   Souza, W. M. *et al.* Neutralisation of SARS-CoV-2 lineage P.1 by antibodies elicited through natural SARS-CoV-2 infection or vaccination with an inactivated SARS-CoV-2 vaccine: an immunological study. *The Lancet Microbe* **2**, e527-e535, doi:10.1016/S2666-5247(21)00129-4 (2021).

38   Naveca, F. G. *et al.* COVID-19 in Amazonas, Brazil, was driven by the persistence of endemic lineages and P.1 emergence. *Nat Med* **27**, 1230-1238, doi:10.1038/s41591-021-01378-7 (2021).

39   Imai, M. *et al.* Characterization of a new SARS-CoV-2 variant that emerged in Brazil. *Proc Natl Acad Sci U S A* **118**, e2106535118, doi:10.1073/pnas.2106535118 (2021).

40   Dejnirattisai, W. *et al.* Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell* **184**, 2939-2954 e2939, doi:10.1016/j.cell.2021.03.055 (2021).

41   Watterson, D. *et al.* Preclinical development of a molecular clamp-stabilised subunit vaccine for severe acute respiratory syndrome coronavirus 2. *Clin Transl Immunology* **10**, e1269, doi:10.1002/cti2.1269 (2021).

42   Tao, K. *et al.* The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat Rev Genet* **22**, 757-773, doi:10.1038/s41576-021-00408-x (2021).

43   Wang, P. *et al.* Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization. *Cell Host Microbe* **29**, 747-751 e744, doi:10.1016/j.chom.2021.04.007 (2021).

44   Caniels, T. G. *et al.* Emerging SARS-CoV-2 variants of concern evade humoral immune responses from infection and vaccination. *medRxiv*, 2021.2005.2026.21257441, doi:10.1101/2021.05.26.21257441 (2021).

45   Ekiert, D. C. *et al.* Cross-neutralization of influenza A viruses mediated by a single antibody loop. *Nature* **489**, 526-532, doi:10.1038/nature11414 (2012).

46   Young, A. *et al.* A platform technology for generating subunit vaccines against diverse viral pathogens. *Front Immunol* **13**, 963023, doi:10.3389/fimmu.2022.963023 (2022).

47   Lok, S. M. An NTD supersite of attack. *Cell Host Microbe* **29**, 744-746, doi:10.1016/j.chom.2021.04.010 (2021).

48   Shrimal, S. & Gilmore, R. Glycosylation of closely spaced acceptor sites in human glycoproteins. *J Cell Sci* **126**, 5513-5523, doi:10.1242/jcs.139584 (2013).

49   Malaby, H. L. & Kobertz, W. R. The middle X residue influences cotranslational N-glycosylation consensus site skipping. *Biochemistry* **53**, 4884-4893, doi:10.1021/bi500681p (2014).

50   Chawla, H. *et al.* Glycosylation and serological reactivity of an expression-enhanced SARS-CoV-2 viral spike mimetic. *J Mol Biol* **434**, 167332, doi:10.1016/j.jmb.2021.167332 (2022).

51   Yao, H. *et al.* Molecular architecture of the SARS-CoV-2 virus. *Cell* **183**, 730-738.e713, doi:10.1016/j.cell.2020.09.018 (2020).

52   Huang, H. Y. *et al.* Vaccination with SARS-CoV-2 spike protein lacking glycan shields elicits enhanced protective responses in animal models. *Sci Transl Med* **14**, eabm0899, doi:10.1126/scitranslmed.abm0899 (2022).

53   Bouwman, K. M. *et al.* Multimerization- and glycosylation-dependent receptor binding of SARS-CoV-2 spike proteins. *PLoS Pathog* **17**, e1009282, doi:10.1371/journal.ppat.1009282 (2021).

54   Lempp, F. A. *et al.* Lectins enhance SARS-CoV-2 infection and influence neutralizing antibodies. *Nature* **598**, 342-347, doi:10.1038/s41586-021-03925-1 (2021).

55   Chappell, K. J. *et al.* Safety and immunogenicity of an MF59-adjuvanted spike glycoprotein-clamp vaccine for SARS-CoV-2: a randomised, double-blind, placebo-controlled, phase 1 trial. *Lancet Infect Dis* **21**, 1383-1394, doi:10.1016/S1473-3099(21)00200-0 (2021).

56   Pinto, D. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290-295, doi:10.1038/s41586-020-2349-y (2020).

57   Shi, R. *et al.* A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120-124, doi:10.1038/s41586-020-2381-y (2020).

58   Yuan, M. *et al.* A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **368**, 630-633, doi:10.1126/science.abb7269 (2020).

59   Brouwer, P. J. M. *et al.* Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science* **369**, 643-650, doi:10.1126/science.abc5902 (2020).

60   Modhiran, N. *et al.* A broadly protective antibody that targets the flavivirus NS1 protein. *Science* **371**, 190-194, doi:10.1126/science.abb9425 (2021).

61   Pegg, C. L., Zacchi, L. F., Recinos, D. R., Howard, C. B. & Schulz, B. L. Identification of novel glycosylation events on human serum-derived factor IX. *Glycoconj J* **37**, 471-483, doi:10.1007/s10719-020-09922-2 (2020).

62   Zacchi, L. F. *et al.* Coagulation factor IX analysis in bioreactor cell culture supernatant predicts quality of the purified product. *Commun Biol* **4**, 390, doi:10.1038/s42003-021-01903-x (2021).

63   Pelingon, R. *et al.* Glycoproteomic measurement of site-specific polysialylation. *Anal Biochem* **596**, 113625, doi:10.1016/j.ab.2020.113625 (2020).

64   Phung, T. K., Pegg, C. L. & Schulz, B. L. GlypNirO: An automated workflow for quantitative N- and O-linked glycoproteomic data analysis. *Beilstein J Org Chem* **16**, 2127-2135, doi:10.3762/bjoc.16.180 (2020).

65   Metsalu, T. & Vilo, J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res* **43**, W566-570, doi:10.1093/nar/gkv468 (2015).

Acknowledgments

Author contributions

Competing interests

The authors declare no competing interests.

Data availability

All genome sequences and associated metadata used in this work are published in GISAID's EpiCoV database under GISAID Identifier: EPI_SET_230311as. To view the contributors of each individual sequence with details such as accession number, Virus name, Collection date, Originating Lab and Submitting Lab and the list of Authors, visit 10.55876/gis8.230311as

The *N*-occupancy and glycoproteomic datasets generated and analysed during the current study are available in the MassIVE repository, [ftp://MSV000091533@massive.ucsd.edu; Username for web access: MSV000091533_reviewer; Password: Gamma_Glyco]

Supplementary information

Supplementary Figures 1-3

Supplementary Data 1_N-linked Occupancy

Supplementary Data 2_Occupancy Validation

Supplementary Data 3_Batch 1 Glycoform Analysis

Supplementary Data 4_ Batch 2 Glycoform Analysis

Supplementary Data 5_Sialidase Glycoform Analysis