Title

Evolving spike-protein *N*-glycosylation in SARS-CoV-2 variants

Authors

Sabyasachi Baboo*[,1], Jolene K. Diedrich[1], Jonathan L. Torres[2], Jeffrey Copps[2], Bhavya Singh[1,$], Patrick T. Garrett[1], Andrew B. Ward[2], James C. Paulson[1,4,], John R. Yates III*[,1]

Affiliations

1 Department of Molecular Medicine, The Scripps Research Institute, La Jolla, California 92037, United States
2 Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, California 92037, United States
4 Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California 92037, United States

$ Present address: Department of Computer Science and Engineering, University of California San Diego, La Jolla 92093, United States

Corresponding Authors

* Sabyasachi Baboo (sbaboo@scripps.edu), John R. Yates III (jyates@scripps.edu)

Abstract

Since >3 years, SARS-CoV-2 has plunged humans into a colossal pandemic. Henceforth, multiple waves of infection have swept through the human population, led by variants that were able to partially evade acquired immunity. The co-evolution of SARS-CoV-2 variants with human immunity provides an excellent opportunity to study the interaction between viral pathogens and their human hosts. The heavily *N*-glycosylated spike-protein of SARS-CoV-2 plays a pivotal role in initiating infection and is the target for host immune-response, both of which are impacted by host-installed *N*-glycans. Using highly-sensitive DeGlyPHER approach, we compared the *N*-glycan landscape on spikes of the SARS-CoV-2 Wuhan-Hu-1 strain to seven WHO-defined variants of concern/interest, using recombinantly expressed, soluble spike-protein trimers, sharing same stabilizing-mutations. We found that *N*-glycan processing is conserved at most sites. However, in multiple variants, processing of *N*-glycans from high mannose- to complex-type is reduced at sites N165, N343 and N616, implicated in spike-protein function.

Keywords

N-glycan, heterogeneity, SARS-CoV-2, COVID-19, spike-protein, variants, proteomics, mass spectrometry, vaccine, virus

Introduction

The COVID-19 pandemic caused by SARS-CoV-2[1,2] has highlighted the threat that zoonotic viruses pose to human health and socio-economic sustainability[3,4]. The current pandemic is one of the most severe humans have experienced, with a global death toll of 6.9 million and infection of more than 9.6% of global human population (https://covid19.who.int/, https://www.census.gov/popclock/). The original SARS-2 strain is believed to have begun as a moderately transmissive virus, infecting its first human host(s) in China in late 2019[5-7]. Since then, SARS-CoV-2 has infected the human population in surges, mutating

frequently and evolving into increasingly more transmissive variants[4,8-13] **(Figure 1A)**. Wuhan-Hu-1, the original SARS-CoV-2 strain, was the dominant strain throughout 2020, and has since produced 5 variants of concern (VoC), as defined by World Health Organization (WHO): Alpha, Beta, Gamma, Delta, and Omicron[14]. The Alpha variant was first detected in the United Kingdom, was highly transmissible, and quickly became the most dominant variant outside of South America and Africa from December 2020 until July 2021[15,16]. Almost simultaneously, Beta was detected in South Africa and became dominant in Africa[17], and Gamma was detected in Brazil and became dominant in South America[18]. The highly virulent and transmissible Delta[19] was first detected in India in mid-2021 and became the dominant variant until it was overtaken in January 2022 by the currently dominant Omicron[20]. GISAID (https://gisaid.org/) followed the spread of 2 WHO-recognized variants of interest (VoI), Mu[21,22] (first detected in Colombia) and Lambda[23] (first detected in Peru), which emerged in 2021 and partially overlapped with Gamma in South America.

Thanks to decades of scientific research devoted to the improvement of vaccine technology and advances in an understanding of the immunology of viral infection (especially that of Coronaviruses, Human Immunodeficiency Virus (HIV) and Influenza)[24-33] successful vaccines to counter SARS-CoV-2 infection were developed at an unprecedented pace[34-41]. These vaccines were designed against heavily glycosylated spike-protein (S-protein) trimer, commonly known as "spike", anchored to the lipid-bilayer of the SARS-CoV-2 viral envelope[34,42]. This non-covalently associated trimeric form of spike-protein mediates virus entry by binding to its host cell-surface receptor, human ACE2 (Angiotensin converting enzyme 2)[43,44]. After fusing with a host cell, SARS-2 replicates its genetic material, and the host's translation apparatus is hijacked to synthesize viral proteins[45,46]. The spike-protein is glycosylated with high mannose to complex type *N*-linked glycans at sequons encoded by Asn-X-Thr/Ser or NX(S|T), where X ≠ P[47]. *N*-glycans are added co-translationally by *en bloc* transfer from a lipid-containing intermediate to yield $Glc_3Man_9GlcNAc_2$-Asn (Glc = Glucose, Man = Mannose, GlcNAc = *N*-acetylglucosamine). The monomeric spike proteins fold and assemble into a trimeric spike in the endoplasmic reticulum (ER) lumen[48-50]. Glycan addition at each NX(S/T) site occurs co-translationally through the *en bloc* transfer of a large high mannose oligosaccharide ($Glc_3Man_9GlcNAc_2$). Trimming of outer Glc and Man residues is initiated in the ER, and further processing to complex type glycans occurs as the spike protein moves through the ER and Golgi complex (GC) enroute to the cell surface[48-50]. The degree of conversion of high mannose to complex type *N*-glycans at each glycosylation site is influenced by the dynamics of the underlying protein structure and the proximity of other *N*-glycans that can restrict access to the processing enzymes[51]. While most *N*-glycosylation sites on SARS-CoV-2 spike-protein, glycans are processed to complex type, there is minimal processing at a few sites (*e.g.*, N234), resulting in high mannose type glycans in the mature spike protein.

Interest in site specific glycosylation of the spike-protein stems from the functional significance of these glycans in infection and the host immune response. The opening up of the spike-protein receptor binding domain (RBD) for initiating interaction with ACE2 on host cells, is shown to be regulated by glycans on N122, N165, N234 and N343[52-54]. In addition, *N*-glycans of viral spike glycoproteins shield the underlying protein structure facilitating escape from host immune responses[50,55,56]. Thus, for rational design of vaccines capable of offering protection against evolving SARS-CoV-2 variants, it is important to assess the *N*-glycan landscape on the SARS-CoV-2 spike[57-61]. Multiple critical mutations that have appeared on spike-protein of SARS-CoV-2 variants, are known to facilitate escape from humoral immunity[62,63], and some of these mutations might be responsible for a change in the *N*-glycan landscape of the spike, partially explaining the immune escape strategies of SARS-CoV-2 variants[9,42].

Here, we present a systematic evaluation of the *N*-glycan landscape of the spike-protein from the original SARS-CoV-2 strain and 7 of its WHO-defined major variants (Alpha, Beta, Gamma, Delta, Omicron, Mu, and Lambda). Although a number of studies, using mass spectrometry (MS)-based glycoproteomic approaches, have compared glycosylation of SARS-COV-2 variant spike-proteins[64-69], with the glycosylation of the original Wuhan-Hu-1 spike-protein[51,57,60,70-76], they have examined a smaller subset of variants and there is substantial differences with regards to the degree of glycan processing and glycan occupancy at each glycosylation site **(Figure S1)**. Such variation can be expected for sites that have insufficient glycopeptide sampling to establish an accurate representation of the "*N*-glycan demographic".

Here we have compared multiple SARS-CoV-2 variant spike-proteins, to make critical inferences from site-specific *N*-glycosylation changes with high statistical confidence – a property unique to our DeGlyPHER approach, owing to hugely enhanced peptide sampling due to Proteinase K digestion along with Endo H/PNGase F sequential deglycosylation[49,77-79]. We assess the impact on glycosylation of mutations needed to produce high-yielding and stable "soluble spike-protein trimers" for vaccines. We compare our results with the *N*-glycan landscape of the recombinant spike manufactured by Novavax (NVX-CoV2373), used as a vaccine against SARS-CoV-2[59], which has been recently authorized by United States Food and Drug Administration (US FDA) for emergency use.

Results

Analysis of *N*-glycan landscape on spike-protein

The spike-protein of Wuhan-Hu-1 is 1,273 amino acids long and *N*-glycosylated (we exclusively analyze *N*-glycans in this study) at 22 potential *N*-glycosylation sites (NGS)[57] **(Figure 1B)**. The spike trimer is comprised of 3 protomers and has an apical broad head consisting of the receptor binding domain (RBD) surrounded by the N-terminal domain (NTD) at its base, and tapers down to a long stalk[34,53,80] **(Figure 1C)**. The spike head of a single spike-protein protomer is >1,100 amino acid long, populated by 8 NGS in NTD, 2 NGS in RBD and another 9 NGS. The 10 NGS in NTD and RBD are some of the most extensively studied glycosylation sites[52,53,81,82]. The spike stalk consists of <150 amino acids per protomer, populated by only 3 NGS.

We used recombinantly expressed SARS-CoV-2 spike (M1-Q1208) in cultured human cells. These proteins lacked the transmembrane (TM) and C-terminal (CT) domains but included C-terminal tags to aid in trimerization and affinity purification. The Furin cleavage site (FCS, S1/S2) was abrogated (682RRAR substituted with 682GSAS) to stabilize the spike **(Figure 1B, Supplementary Tables 1,2)**[83]. The structural integrity and purity of the spikes in their prefusion state was validated with size-exclusion chromatography (SEC) and negative-stain electron microscopy (NS-EM) **(Figure S2)**. We employed the DeGlyPHER[77,79] approach that uses sequential deglycosylation followed by bottom-up MS-based proteomics to resolve the site-specific *N*-glycan heterogeneity on spikes from 7 SARS-CoV-2 variants and the Novavax full-length spike-protein (NVX-CoV2373; expressed in Sf9 cells *i.e.*, insect cells) and compared them with the Wuhan-Hu-1. DeGlyPHER allowed us to broadly resolve the abundance of high-mannose (including hybrid) and complex glycans at each NGS and to estimate glycan unoccupancy[79].

Original SARS-CoV-2 strain/Wuhan-Hu-1

Wuhan-Hu-1 is believed to have jumped from an undetermined animal reservoir(s) to humans at the end of 2019[5,6], and its spike is one of the of the most studied protein structures in recent times[80]. Most of the 22 NGS on spike-protein are predominantly complex type; only N122, N234, N603, N709, N717, N801 and N1074 show a majority of high mannose-type glycans **(Figure 2A)**. The results provide a robust dataset for detailed comparison among 7 spike proteins of SARS-CoV-2 VoC/Is defined by WHO. The site-specific glycan distribution observed for the Wuhan-Hu-1 spike protein is largely consistent with several reports from the Crispin group using a traditional glycoproteomic approach[57,64,76]. More significant differences have been reported by Shajahan *et. al.*[65,72], likely the result of a workflow that provides insufficient sampling at each site to quantitate relative proportions of the major glycan species **(Figure S1)**. We analyzed 3 different versions of the spike-protein that were mutated to stabilize the immunogen in its prefusion state **(Figure 1B)**. The first version, 2P, contains 2 residues (K986 and V987) that were substituted to proline and this mutant is used for most of the currently authorized SARS-CoV-2 vaccines[29,59]. The second version, 6P (or HexaPro), has 4 additional proline substitutions at positions F817, A892, A899, and A942 which help with stability and increase protein yield[83]. The third version, 6P-Mut7, adds an additional interprotomer disulfide bridge at positions V705C/T883C, thus further stabilizing the immunogen while keeping the RBD flexible[84]. When we compared the glycan distribution pattern between 2P, 6P, and 6P-Mut7, the 6P/6P-Mut7 mutants exhibited a significantly higher proportion of high mannose glycans at N603, N801 and N1074 relative to the 2P mutant, but there were no statistically significant differences between 6P and 6P-Mut7 mutants, **(Figures 2B,C,D)**. We see that different stabilizing mutations can differentially impact the glycosylation pattern on spikes, which could complicate comparison of variant spikes with differing stabilizing mutations. With this in mind, we only made

glycosylation comparisons between spikes that were produced using the same stabilizing mutations (6P-Mut7 when comparing all variants in this study with Wuhan-Hu-1; 6P only when comparing Delta and Omicron with Wuhan-Hu-1; 2P when comparing Wuhan-Hu-1 and NVX-CoV2373).

SARS-CoV-2 variants of concern/interest

For the first year of the COVID-19 pandemic the only mutation acquired by the SARS-CoV-2 spike-protein that could be positively selected to establish stable VoC/I was D614G[9,85]. The N501Y mutation in the spike RBD established 3 new VoCs, Alpha (501Y.V1 or B.1.1.7 or 20I), Beta (501Y.V2 or B.1.351 or 20H) and Gamma (501Y.V3 or P.1 or 20J)[86]. N501Y, which is reported to enhance binding affinity of spike RBD to human ACE2, is shared by Alpha, Beta and Gamma[87]. The other spike RBD mutations include E484K on both Beta and Gamma, and K417N on Beta and K417T on Gamma. The spike NTD mutations on Alpha (del69-70 and del144), Beta (L18F, D80A, D215G and del242-244) and Gamma (L18F, T20N, P26S, D138Y and R190S) are also known to affect neutralization[88]. None of the mutations in the Alpha or Beta resulted in gain or loss of an NGS, but there were notable changes in glycan processing at several NGS. In the Alpha spike-protein, there were significantly more complex glycans at N122, and more high mannose glycans at N616 **(Figure 3A)**. In the Beta spike, there were more high mannose glycans at N165 and N616 **(Figure 3A)**. The biggest changes in glycosylation occur in the Gamma spike-protein. Two of mutations resulted in a gain of NGS, T20N creating N20, and R190S creating N188. In the Gamma spike-protein **(Figure 3A)**, the new NGS at N20 contained predominantly complex type glycans, and the adjacent N17 NGS which was largely unoccupied. Moreover, the newly created NGS at N188 is close to N17 and N20 and had largely high mannose glycans. In contrast, glycan composition at N603 and N801 shifted significantly to complex type glycans.

The Delta VoC (B.1.617.2 or 21A) emerged in 2021, spreading globally and replacing the predominance of 501Y lineage[19]. D614G is retained in the Delta spike-protein as well. In the Delta spike-protein, T19R mutation in NTD caused the loss of N17 NGS. There was a significant shift to high mannose glycans at N165 and N616 in both 6P and 6P-Mut7, and additionally at N343 in only 6P version. **(Figures 3A;4A,C)**.

The Omicron VoC (B.1.1.529 or 21M) replaced the Delta and spread globally in 2022, and Omicron subvariants still dominate[20]. The Omicron spike-protein hosts at least 33 mutations, none of which caused gain or loss of any NGS **(Figure 1B)** (though the newer Omicron subvariants are missing N17[89]). D614G is retained in the Omicron spike-protein. Despite this large number of mutations, when we compared the Omicron variant spike to spike from Wuhan-Hu-1 **(Figures 3A;4B,C)**, there was no significant shift in site-specific glycan processing. However, we note a shift from high mannose to complex glycans at N122 and *vice-versa* at N165. Additionally, we observed higher unoccupancy at N149 and N657.

The spread of Mu (B.1.1.621 or 21H) and Lambda (C.37 or 21G) VoIs was limited to South America[21,22]. The Mu spike-protein hosts 10 mutations while Lambda spike-protein hosts 7 mutations, the D614G mutation is retained in both **(Figure 1B)**, and include some of the known mutations (at L452, E484, N501 and P681) that help SARS-CoV-2 escape neutralization[21,22]. In the Mu spike-protein, there was a significant shift to high mannose glycans at N165, N343 and N616, while glycan unoccupancy increased significantly at N149 **(Figure 3A)**. In the Lambda spike-protein, T76I mutation in NTD caused the loss of N74 NGS and there was a significant shift to high mannose glycans only at N1098, though a similar trend was observed at N165 and N343 also **(Figure 3A)**.

S2-protein

The post-fusion S2-protein structure that extends from the CT at its base to fusion peptide (FP) (or beyond if S2' cleavage has not occurred) at the top[75,90], hosts 3 stalk NGS (N1158, N1173 and N1194) in addition to the 5 NGS (N709, N717, N801, N1074, N1098 and N1134) that are structurally conserved in SARS-CoV[91]. The 5 NGS from N1098 to N1194 are spaced at almost regular intervals of 40Å[75,90], suggesting that glycans at these NGS and at N1074 (assuming S2' cleavage) are involved in "caging" the degrees of freedom of the stalk when the viral and host membranes are undergoing fusion[92]. We observed a statistically significant increase in high mannose glycans at N1098 in Lambda **(Figure 3A)**. Glycans anchored to N1098 on the apical rim of the post-fusion S2-protein[75,90] could be important determinants of SARS-CoV-2's infectivity, as observed in bronchial cells expressing ACE2 when N1098 is

occupied with high-mannose *N*-glycans[93]. In this post-fusion S2-protein structure, N1074, N709 and N717 sit just below the apical rim as it extends into the Connecting domain (CD) that hosts N1134[75,90].

The stalk of the spike is highly conserved across Beta-coronaviruses[94], and thus it provides a promising target for immune intervention[95]. The 3 NGS (N1158, N1173 and N1194) on the stalk of the SARS-CoV-2 spike are fully occupied by complex glycans in all the variants **(Figure 3A)**. Complex glycans on these 3 NGS are known to massively shield the stalk at the flexible hinges that are prominent epitopes for immune recognition by antibodies[53,96]. Their conservation across evolving SARS-CoV-2 variants suggests that there is no apparent selection pressure from host immunity to modify the glycan shield around the stalk.

NTD

NTD, which forms a whorl around the RBD, has 8 NGS (N17, N61, N74, N122, N149, N165, N234 and N282), the most of any of the structural domains in SARS-CoV-2 spike. Five of these NGS (N17, N61, N74, N234 and N282) have highly conserved glycan distributions across variants **(Figure 3A)**. Glycans at N17 are difficult to resolve on the spike structure using cryo-EM since it is in the N-terminal unstructured region[96], except when stabilized by antibodies[80,81]. N17 is fully occupied by complex glycans **(Figure 3A)**, and this trend is conserved across all variants except Gamma, where it is mostly unoccupied, and in Delta and newer subvariants of Omicron, where it is absent[89]. In Gamma, the newly created N20 is known to be glycosylated and exterior to N17 on the N1 loop[81,97-99], thus explaining why we observe it to be fully occupied by complex glycans, functionally replacing the adjacent N17, possibly impacting NTD antigenicity[68]. N17, along with N74, N122 and N149 border a multi-epitope "neutralization supersite" on NTD[81,100]. Absence of N17 in Delta is known to reduce neutralization by antibodies[97,100,101] and absence of N74 (on N2 loop) in Lambda is known to increase infectivity (the spike-protein we used for Lambda did not include the new N254 NGS in N5 loop which is associated with reduced neutralization by antibodies[102-104]). N74 is fully occupied by complex glycans across all variants except Lambda, where it is absent **(Figure 3A)**. N122 is mostly occupied by high mannose glycans across variants, except in Alpha where N122 is significantly more occupied by complex glycans **(Figure 3A)**. N122 and N165 are proposed to affect an "RBD-up" conformation[54]. When the spike is in an "RBD-up" conformation[34], glycans on N122 are predicted to significantly clash with the highly neutralizing S309 monoclonal antibody binding to its epitope in the "open" RBD. However, its binding is not expected to be affected[105,106]; indeed, no significant loss in neutralization of Alpha by S309 has been reported[107,108], even though we see a shift in the type of glycans on N122 compared to Wuhan-Hu-1 **(Figure 3A)**. We found N149 (on surface-accessible N3 loop) to be mostly occupied by complex glycans **(Figure 3A)**. When these glycans include $GlcNAc_2Man_3$, N149 is reported to interact with host cell-surface lectin receptors, enhancing ACE2/spike-protein interaction in lung tissue and consequently driving cell-fusion and syncytia formation that is known to facilitate infection[109,110]. We observed that N149 is almost completely occupied by glycans in variants where the N3 loop is not mutated (Beta and Lambda), while the glycan unoccupancy at N149 is most significant on Mu **(Figure 3A)**. The glycosylation pattern on N165 shifts from almost completely complex type in Wuhan-Hu-1 virus to high mannose glycans in all variants except Alpha and Gamma **(Figure 3A)**. The inferred steric occlusion of N165 in most variants (the shift being highest in Delta) is highly relevant given the role of glycans at N165 and N234 in modulating conformational transitions of RBD and its stabilization in an "up" conformation and the RBM (receptor binding motif) is shielded in a "down" conformation[53,111]. The near-complete occupancy at N234 with high mannose glycans and at N282 with complex glycans is conserved across all variants we studied **(Figure 3A)**, with N282 strategically positioned in close proximity to the FP[82] and also implicated in formation of dimeric spike-trimers[59]. In the Gamma variant only, N188 is at the C-terminal edge of the N4 loop, and we find it to be mostly occupied with high mannose glycans **(Figure 3A)**, which are reported to increase the glycan shielding effect[64]. Emerging mutations in recent Omicron subvariants, such as H245N mutation in BA.2.86 lineage, causes gain of N245 as NGS, and occurs along with loss of N17 (https://gisaid.org/hcov19-mutation-dashboard/), both in close proximity on NTD, possibly suggesting a functional role of glycans at this newly-emerged NGS.

RBD

RBD, which physically interacts with ACE2 and initiates infection, hosts only 2 NGS (N331 and N343). Even in the heavily mutated RBD of Omicron (15 mutations), N331 and N343 are conserved and no new NGS arise. However, some emerging mutations in recent Omicron subvariants affect known *N*-glycosylation sites and/or create new ones (https://gisaid.org/hcov19-mutation-dashboard/) on spike-protein[112]. K529N mutation in an Omicron XBB.1.5 subvariant spike-protein RBD creates a new glycosylation site at the RBD-hinge, an otherwise conserved region[113], close to N331, and its glycosylation can potentially regulate RBD "up/down" dynamics and indirectly reduce access to RBD epitopes[114,115]. N331 is populated almost completely with complex glycans across all variants we studied **(Figure 3A)**. N331 and N343 glycans are reported to shield the RBM in a "down" conformation[53]. We observed that N343 is occupied almost completely with complex glycans in the Wuhan-Hu-1 strain, as well as in some of the variants, but a shift towards high mannose glycans was seen in Lambda, Beta, and this shift was statistically significant in Mu and Delta (for 6P versions) variants **(Figures 3A;4A,C)**. These changes are highly relevant because glycans at N343 are implicated in pushing the RBD "up" into an "open" conformation[52], and is targeted by neutralizing antibodies[106,108].

S1-protein CTD

The NTD-associated subdomain-2 of S1-protein *i.e.*, CTD-1 and CTD-2 **(Figure 1B)** hosts 3 NGS (N603, N616 and N657). These 3 NGS, along with N61 and N801 in the S2-protein, are proximal to the proteolysis sites S1/S2 and S2'[116]. The glycans anchored to these NGS may affect access to Furin-like proteases through glycan holes around S1/S2, S2' and FP[117]. Though we abrogated the S1/S2 cleavage site in the spike-proteins used in this study, any fluctuation in glycosylation state at these NGS could provide evidence for an interaction of the S1/S2 region with glycans on N61/N603[118], and glycans on all 5 NGS that affect spike incorporation on virions and their capability to infect cells[116]. We observed a significantly higher proportion of high-mannose glycans on N603 and N801 in the 6P and 6P-Mut7 versus the 2P versions of Wuhan-Hu-1, as well as in in all variants we studied except Gamma, where complex glycans are significantly higher at N603/N801 **(Figure 3A)**. We find N61 and N657 to be almost completely occupied with complex glycans across all variants we studied **(Figure 3A)**.

N616 is adjacent to D614G, which is conserved in all SARS-CoV-2 variants that have emerged since Wuhan-Hu-1. The D614G mutation is known to have significantly increased SARS-CoV-2 infectivity by reducing S1-shedding, thus ensuring that the spike is intact and available for receptor binding[119-121]. The stabilization of spike in D614G variants is attributed to the loss of an interaction between G614 in one protomer and a region immediately downstream of FP (amino acids 823-862) in a neighboring protomer, and the effect this plays on an adjacent loop (loop 630) that is proposed to play a critical role in modulating/regulating spike structure for viral fusion with host[59,90,122,123]. Compared to Wuhan-Hu-1 where N616 is mostly occupied with complex glycans, Alpha, Beta, Delta, and Mu show significant shifts towards high-mannose glycans **(Figure 3A)**. The absence of this shift maybe explained by the FP-proximal T859N mutation in Lambda and N856K mutation in Omicron, which could possibly counter the loss of inter-protomer interaction caused due to D614G[124-127]. The change in glycosylation state at N616 in variants can contribute to or be affected by the structural dynamics of D614G[70]. This relationship has been postulated[85,128], but not systematically evaluated.

Discussion

The COVID-19 pandemic is driven by evolving variants of SARS-CoV-2, which are characterized by critical mutations that cause changes in the structure of the spike, hence affecting infectivity[129]. These structural modifications cause and are also affected by *N*-glycosylation changes. Using the 6P-Mut7 form of spike-protein from 7 WHO-defined SARS-CoV-2 variants, we present here a comprehensive analysis of changes in the site-specific *N*-glycan processing relative to Wuhan-Hu-1 using our highly-sensitive DeGlyPHER approach[77,79] **(Figures 3A,B,C)**. We benchmarked the *N*-glycosylation pattern on Wuhan-Hu-1 6P-Mut7 spike-protein to its 2P and 6P versions that have been previously studied[57,59,60,65,72]. We note that the *N*-glycosylation pattern remains similar for most sites across these 3 mutant versions. However, there are a few site-specific differences (*e.g.*, at 603, 1074, 1098) **(Figure 2)**, so we have taken

care to compare SARS-CoV-2 variant spike-proteins using the same stabilizing mutations to ensure that any additional differences are due to variant mutations.

Using stabilized soluble spike-protein trimers hosting known variant-specific mutations, we observed that many NGS in NTD, RBD and stalk (N17/N20, N61, N74, N282, N331, N1158, N1173, N1194) exhibited near-complete occupancy with complex glycans, while N234 is occupied with high mannose glycans **(Figure 3A)**. Glycosylation changes were prominent **(Figure 3A)** at many NGS shown/predicted to be directly involved in: [1] RBD dynamics (N122, N149, N165 and N343)[53,54,81,109-111], [2] proteolysis/spike-stabilization by virtue of their proximity to S1/S2 and/or S2' (N603, N616 and N801)[70,85,116,118], and [3] stabilization of post-fusion functions of S2-protein (N1098)[75,90,93]. Newby *et. al.*[64] have made similar observations using 6P-stabilized spikes from 4 VoCs (Beta, Gamma, Delta and Omicron) using traditional MS-based glycoproteomic approaches. However, the results of Newby *et*. *al*. differ at a number of sites from those reported by Shajahan *et. al.*[65] for 5 VoCs (Alpha, Beta, Gamma, Delta and Omicron) and Wang *et. al.*[66] for 4 VoCs (Alpha, Beta, Gamma and Delta) using similar MS-based traditional glycoproteomic approaches. Moreover, in these previous studies, the results for the glycosylation pattern of the benchmark Wuhan-Hu-1 spike protein, against which the glycosylation pattern of variant spikes are being compared, also differ at multiple sites[57,64-66,70,72,76] **(Figure S1)**. For example, using the same 6P-stabilized version of the Wuhan-Hu-1 spike one study found complete glycan occupancy at site N1173[76] while another found no glycan occupancy[64]. Such variation is most likely explained by insufficient glycopeptide sampling per site in some studies, which we believe points to a strength of DeGlyPHER for more confident site-specific comparative analyses from extensive sampling at each site.

While our results on changes in glycosylation of variants largely agree with those of Newby *et*. *al*.[64], there are several notable differences. While we concluded that several variants showed significantly increased high mannose at N165 relative to Wuhan-Hu-1, Newby *et. al.* [64] concluded there was little change as a result of greater high mannose in their analysis of Wuhan-Hu-1 and vice-versa for N61 where we do not observe any significant change in glycosylation **(Figures 3A;S1)**. Also, glycan unoccupancy reported by Newby *et*. *al*.[64] was substantial at N331 (in Beta), N657 (in Wuhan-Hu-1, Beta, Gamma, Delta and Omicron), N1074 (Gamma), N1158 (Gamma), N1173 (in Wuhan-Hu-1, Beta, Gamma, Delta and Omicron; almost never occupied by glycans) and 1194 (in Beta, Gamma, Delta and Omicron), while we find all these sites to be almost completely occupied with glycans across all variants analyzed here **(Figures 3A;S1)**. Sparse peptide sampling at N709, N717 and N1134 **(Figure S3A)** precludes us from reading too much into the observed glycosylation pattern at these sites **(Figure S3A)**, though in all variants we studied, glycans decorating N1134 are mostly complex type **(Figure S3B)** and glycans at N709 and N717 on Wuhan-Hu-1 spike (2P/6P stabilized mutants, where we recorded sufficient peptide sampling) are mostly high mannose type **(Figures 3A;S1)**, as also reported by Newby *et. al.*[64]. The sparse sampling that we recorded may be attributable to structural constraints that limit protease access since we avoid harsh denaturants in our sample preparation. Alternatively, peptides mapping to these regions might have reduced ionization in the mass spectrometer[77,79]. Speculating from structural analysis of post-fusion spike[90,130], glycans on N709/N717 may possibly affect proteolysis/spike-stabilization, and glycans on N1134 may play a role in post-fusion S2-protein stability.

Using a hierarchical clustering analysis for similarity in glycosylation pattern among shared and "well-sampled" NGS in 6P-Mut7, we see that despite the huge number mutations in Omicron, it is glycosylated more like Wuhan-Hu-1 than other variants, with Gamma and Delta being the most different **(Figure 3D)**. This hierarchical cluster differs from the inferred evolutionary pattern of the SARS-CoV-2 variants determined by nucleic acid mutation analysis (**Figure 3E**, modified from https://covariants.org/), suggesting that selective pressure on glycan changes is more dependent on higher-order structure than primary sequence of a protein, as evident from Wuhan-Hu-1 and Omicron spike-protein trimers being structurally more similar to each other than Alpha, Beta and Delta[127].

We observed many prominent shifts from complex glycans on Wuhan-Hu-1 to high mannose glycans on "variant" spikes, at multiple NGS **(Figure 3B)**. This trend hints at reduced *N*-glycan processivity at these NGS, hence implying an increase in their steric occlusion. It also suggests that variant spikes may be better at using mannose-binding membrane proteins as receptors for host-cell entry/fusion[131], thus

explaining SARS-CoV-2 tropism in the lower human respiratory tract, where infection is most prominent but the primary host receptor (ACE2) is minimally expressed[97,132]. This spike-glycan-based attachment of infectious virus to host cells can increase virus residence time on host surface, thus providing a better chance of host-virus fusion and consequent infection. Unlike the spike on an infectious virus, vaccine action is independent of cellular attachment or viral fusion, so the variant-based protein-structure changes are more likely to explain immunogenicity than the changes they effect in glycan landscapes. This notion was supported by studies that reported broader immunogenicity of partially-deglycosylated immunogens[133-135]. We benchmarked the Novavax SARS-CoV-2 spike and found its glycan landscape to be very similar to the 2P version of the Wuhan-Hu-1 spike (except significantly higher complex glycans at N801 and lower glycan occupancy at N657) **(Figure 5A,B)**, despite the absence of the TM and CT domains in the 2P version and both trimers being expressed in different cell types (2P version in human cells and Novavax vaccine in insect cells[59]). We note that this site-specific *N*-glycosylation pattern observed on Novavax SARS-CoV-2 spike-protein recapitulates very well the results from our previous attempt[59] at analyzing this glycoprotein (the only substantial differences being at N17, N343, N657 and N717, which can be explained by scant sampling or high variance in our previous attempt), that was performed with an earlier version of DeGlyPHER[49,78].

We observed the most conspicuous and consistent shifts from complex to high mannose glycans at N165, N343 and N616 **(Figure 3B)**. Newby *et. al.* see similar increase in high-mannose (mostly M5) glycans at N165 and N616, and hybrid glycans at N616[64], and though not substantial, but Shajahan *et. al.*[65] and Wang *et. al.*[66] observe similar trends. Owing to structural changes in variant spike-proteins, Golgi α-mannosidase II (~130 kDa), that initiates processing of all complex glycans[136], may have limited access to process glycans at these NGS by virtue of its size, when compared to other mannosidases[137,138] and glycosyl transferases[139,140] (~40-70 kDa) that form M4-M6 high-mannose and hybrid glycans[50]. N165 and N343 are implicated in RBD opening[52-54], and N616 is a single amino acid away from D614G, which is the most consistent mutation in SARS-CoV-2 variants[9] and implicated in critical structural changes that determine S1-S2 integrity after Furin cleavage[122,141]. So, there is very high probability that N616 glycans can regulate the integrity of S1-S2 coupling in the spike-protein trimer[85]. Presence of *O*-glycans has been reported at N618, that may also contribute to this effect[142,143]. P681R mutation may enhance Furin cleavage by adding a basic residue to 682RRAR in Delta and D614G stabilizes the Furin-cleaved S1-protein on S2-protein to facilitate fusogenicity, thus explaining Delta's high disease severity[122,125,144,145]. The significant decrease in glycan processing at N616, N165 and N343 that we observed in Delta **(Figures 3A;4A,C)** may further explain fusogenicity and syncytia formation *via* mannose-binding receptors on respiratory tract cells that are the primary tissue affected by Delta but which sparsely express ACE2[93,97,110,131,146,147]. Further interrogation of this possibility may bring a critical advance in understanding Delta's severe pathogenicity in human respiratory tissue.

SARS-CoV-2 spike is one of the most perused glycoprotein structures and a huge amount of structure-based immunology has been published on this since the COVID-19 pandemic[80,85,148-152]. Though these numerous studies often acknowledge the indispensable role of glycans and propose molecular mechanisms involving glycans that could possibly explain their observations and inferences, but there have been few studies evaluating site-specific glycosylation on spike-proteins from SARS-CoV-2 variants[64-66,153]. These MS-based traditional glycoproteomic approaches have provided insights into various glycoforms present at each site on spike-protein from SARS-CoV-2 variants, but their estimation of glycan occupancy and proportion of *N*-glycan types lacks precision, as evident from variation in estimation between studies and their inability to plot any metric of error to represent confidence in estimating proportion of glycan types they observe[57,60,64-66,70,76,135]; presumably due to inherent challenges of measuring sufficient glycopeptides mapping to each site, thus failing to recapitulate the accurate site-specific representative distribution of glycans with high degree of statistical confidence. This also limits their scope to make head-to-head or pairwise site-specific statistical comparison between variants. DeGlyPHER[77,79] overcomes these limitations with excellent sampling per glycosylation site, providing the statistical power to make precise estimations and confident comparisons, and in rare cases where sampling dwindles, we abstain from making inferences. Thus, we used DeGlyPHER to compare the various VoC/Is of SARS-CoV-2 for their spike-protein *N*-glycosylations. DeGlyPHER's advantage is evident from the large number of statistically confident glycosylation state differences we identified, and

their critical importance is validated by many structural and immunological studies that are cited throughout this report.

Acknowledgements

We thank Claire Delahunty and Sandhya Bangaru for critically reading the manuscript, Charles Bowman for computational assistance, Salvador Martínez-Bartolomé for assistance with GlycoMSQuant, and Alessio D'Addabbo and Max Crispin (University of Southampton, UK) for pointing us to the use of percentage-point scale. We thank Nita Patel, Alyse D. Portnoff, and Gale E. Smith from Novavax, Inc. for supplying NVX-CoV2373 for this research. This work was supported by the grants P41GM103533 (NIH), UM1AI100663 (NIH, CHAVI-ID), UM1AI144462 (NIH, CHAVD), R01/R56AI113867 (NIH/NIAID) and Bill and Melinda Gates Foundation (INV-004923).

Contributions

S.B. conceptualized the project, performed glycan analysis, and wrote the manuscript with the contribution of all authors. J.K.D. performed LC-MS/MS. J.L.T and J.C. produced the spike-proteins. J.L.T performed the NS-EM. B.S. created the molecular structure graphics. P.T.G. performed the hierarchical clustering analysis. A.B.W., J.C.P, and J.R.Y supervised the project.

Competing Interests

The authors declare no competing financial interest.

Figure Title/Legends

Figure 1: Mutational landscape of spike-protein from SARS-CoV-2 variants. (A) Global waves of SARS-CoV-2 variants over last 3 years and their temporal proportions (modified from nextstrain.org). (B) Critical domains along the amino acid (aa) length of spike-protein *viz.*, N-terminal domain (NTD), Receptor-binding domain (RBD), C-terminal domains (CTD-1 and 2) of S1-protein, Fusion peptide (FP), Heptad repeats (HR-1 and 2), Central helix (CH), Connecting domain (CD), Transmembrane (TM) domain and Cytoplasmic tail (CT), with associated *N*-glycosylation sites (NGS, green), incorporated stabilizing mutations (SM, blue) and most frequent variant-specific mutations (deleted NGS in red). Furin cleavage site (FCS or S1/S2) and the subsequent cleavage site (S2') for other cellular proteases is shown. (C) Structure of the full-length spike-protein trimer with representative *N*-glycans in yellow.

Figure 2: Comparing glycosylation changes on Wuhan-Hu-1 spike-protein owing to stabilizing mutations. (A) *N*-glycan heterogeneity determined on 6P version with dominant proportion color-coded from high mannose (green) to complex (magenta). A percentage-point scale signifying increase (towards red) or decrease (towards blue) in high mannose *N*-glycans, when performing site-specific comparison between 6P and 2P (B), showing significantly reduced *N*-glycan processing at N603, N801 and N1074; and 6P-Mut7 and 6P (C), showing no significant site-specific changes (Poor sampling of peptides mapping N709, N717 and N1134 in 6P-Mut7, numbered in red). (D) Bar-graphs demonstrating site-specific differences in glycosylation, if any, between 3 differently-stabilized mutants. *N*-glycosylation states are color-coded. Error bars represent mean–SEM. Any significant difference (BH-corrected p-value <0.05) in proportion of a certain *N*-glycosylation state at an NGS is represented by color-coded ↕.

Figure 3: Site-specific comparison of SARS-CoV-2 6P-Mut7 spike-protein trimer from VoC/I with Wuhan-Hu-1. (A) Bar-graphs demonstrating site-specific differences in glycosylation, if any. N17 (glycosylation pattern transferred to N20 in Gamma), N61, N74, N149, N282, N331, N657, N1158, N1173 and N1194 are consistently occupied with mostly complex *N*-glycans across all variants (N17 and N74 absent in Delta and Lambda, respectively). N188 (only in Gamma), N234 and N801 are consistently occupied with mostly high mannose *N*-glycans across all variants. Across all variants, N1074 and N1098 are occupied by a more "even" distribution of high mannose and complex *N*-glycans. Variant specific

changes in glycosylation pattern are most apparent at N122, N165, N343, N603 and N616. Across variants, NGS are mostly occupied with *N*-glycans. *N*-glycosylation states are color-coded. Error bars represent mean–SEM. Any significant difference (BH-corrected p-value <0.05) in proportion of a certain *N*-glycosylation state at an NGS, compared to the corresponding state in Wuhan-Hu-1, is represented by color-coded ↕. (B) A percentage-point scale signifying increase (towards red) or decrease (towards blue) in high mannose *N*-glycans, averaging across all variants, at each NGS, demonstrating highest shifts at N165, N343 and N616. (C) Variance in glycosylation state across all variants and Wuhan-Hu-1, with highest variance seen at N165 and N343. (D) Hierarchical clustering of SARS-CoV-2 variants and Wuhan-Hu-1 by similarity in glycosylation state shows Omicron is most similar to Wuhan-Hu-1, and Gamma and Delta the most dissimilar. (E) An evolutionary dendrogram of SARS-CoV-2 variants based on GISAID data considering nucleic acid mutations (covariants.org). None of the above analyses consider N709, N717 and N1134 owing to inconsistent sampling.

Figure 4: Site-specific comparison of SARS-CoV-2 6P spike-protein trimer from Wuhan-Hu-1, Delta and Omicron. A percentage-point scale signifying increase (towards red) or decrease (towards blue) in high mannose *N*-glycans, when performing site-specific comparison between Delta and Wuhan-Hu-1 (A), showing significantly reduced *N*-glycan processing at N165, N343 and N616; and Omicron and Wuhan-Hu-1 (B), showing prominently reduced *N*-glycan processing at N165, but not statistically significant. (C) Bar-graphs demonstrating site-specific differences in glycosylation, if any, between Wuhan-Hu-1 and Delta, and Delta and Omicron. *N*-glycosylation states are color-coded. Error bars represent mean–SEM. Any significant difference (BH-corrected p-value <0.05) in proportion of a certain *N*-glycosylation state at an NGS is represented by color-coded ↕. The significant shift from complex *N*-glycans in Wuhan-Hu-1 to high mannose *N*-glycans in Delta at N165, N343 and N616, reverts from Delta to Omicron.

Figure 5: Comparing glycosylation changes between Wuhan-Hu-1 2P spike-protein and NVX-CoV2373. A percentage-point scale signifying increase (towards red) or decrease (towards blue) in high mannose *N*-glycans, shows not much difference between the 2 spike-protein trimers (A), even though many of the smaller differences are significant (B) owing much higher sampling from NVX-CoV2373. *N*-glycosylation states are color-coded. Error bars represent mean–SEM. Any significant difference (BH-corrected p-value <0.05) in proportion of a certain *N*-glycosylation state at an NGS is represented by color-coded ↕.

Methods

Glycoprotein expression and Purification

Spike-proteins (amino acid sequences in **Supplementary Information**) were expressed and purified as previously described[84]. Briefly, 500 µg of each plasmid encoding spike-proteins used in this study were incubated with 1.5 mg of polyethylenimine (Polysciences, Inc.) for 20 min and the mixture was deposited in 1L of HEK293F cells. Protein was expressed for 6 days, cells were harvested, and the protein was purified with a StrepTactin XT 4FLOW column (IBA Biosciences). The eluted protein was then size exclusion purified over a Superose 6 HiLoad 16/600 pg 120 mL or Superose 6 Increase 10/300 GL columns (Cytiva).

Electron microscopy

To assess trimer homogeneity, 3 µL of spike-protein at 0.03 mg/ml in 1X TBS was deposited on a glow-discharged carbon coated copper mesh grid (Electron Microscopy Sciences) and blotted off. 3 µL of 2% uranyl formate (Electron Microscopy Sciences) was deposited on the grid for 90 s then blotted off. Data collection was done on a FEI Tecnai Spirit (120 keV) and a FEI Eagle 4k x 4k CCD camera (Thermo Fisher) and automated with the Leginon software[154]. Raw micrographs were stored in the Appion[155] database, and 2D classes were generated with RELION[156].

Mass spectrometry analysis

DeGlyPHER[77,79] was used to ascertain site-specific glycan occupancy and processivity on glycoproteins.

Proteinase K treatment and deglycosylation

Purified SARS-CoV-2 spike glycoproteins were exchanged to water using Microcon Ultracel PL-10 centrifugal filter (Millipore Sigma). Glycoproteins were reduced with 5 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl; Thermo Scientific) and alkylated with 10 mM 2-Chloroacetamide (Sigma-Aldrich) in 100 mM ammonium acetate (Sigma Aldrich) for 20 min at room temperature (RT, 24□C). Initial protein-level deglycosylation was performed using 250 U of Endo H (New England Biolabs) for 5 µg trimer, for 1 h at 37□C. Glycoproteins were digested with 1:25 Proteinase K (PK; Sigma-Aldrich) for 30 min at 37□C (For NVX-CoV2373, 9 different conditions were used *viz.*, 16 h of trypsin (Promega) digestion, and variations in PK treatment – 1:25 PK for 30 min, 1 h, 2 h, 4 h; 1:100, 1:500, 1:1000, 1:2000 PK for 4 h). PK was denatured by incubating at 90□C for 15 min, then cooled to RT. Peptides were deglycosylated again with 250 U Endo H for 1 h at 37□C, then frozen at –80□C and lyophilized. 100 U PNGase F (New England Biolabs) was lyophilized, resuspended in 20 µl 100 mM ammonium bicarbonate (Sigma-Aldrich) prepared in $H_2{}^{18}O$ (Sigma-Aldrich), and added to the lyophilized peptides. Reactions were then incubated for 1 h at 37□C, and subsequently analyzed by LC-MS/MS.

LC-MS/MS

Samples were analyzed on an Q Exactive HF-X mass spectrometer (Thermo Scientific). Samples were injected directly onto a 25 cm, 100 µm ID column packed with BEH 1.7 µm C18 resin (Waters). Samples were separated at a flow rate of 300 nL/min on an EASY-nLC 1200 UHPLC (Thermo Scientific). Buffers A and B were 0.1% formic acid in 5% and 80% acetonitrile, respectively. The following gradient was used: 1–25% B over 160 min, an increase to 40% B over 40 min, an increase to 90% B over another 10 min and 30 min at 90% B for a total run time of 240 min. Column was re-equilibrated with Buffer A prior to the injection of sample. Peptides were eluted from the tip of the column and nanosprayed directly into the mass spectrometer by application of 2.8 kV at the back of the column. The mass spectrometer was operated in a data dependent mode. Full MS1 scans were collected in the Orbitrap at 120,000 resolution. The ten most abundant ions per scan were selected for HCD MS/MS at 25 NCE. Dynamic exclusion was enabled with exclusion duration of 10 s and singly charged ions were excluded.

Data Processing

Protein and peptide identifications were made with Integrated Proteomics Pipeline (IP2). Tandem mass spectra were extracted from raw files using RawConverter[157] and searched with ProLuCID[158] against a database comprising UniProt reviewed (Swiss-Prot) proteome for Homo sapiens (UP000005640), UniProt amino acid sequences for Endo H (P04067), PNGase F (Q9XBM8), and Proteinase K (P06873), amino acid sequences for the examined proteins, and a list of general protein contaminants. The search space included no cleavage-specificity. Carbamidomethylation (+57.02146 C) was considered a static modification. Deamidation in presence of $H_2{}^{18}O$ (+2.988261 N), GlcNAc (+203.079373 N), oxidation (+15.994915 M) and N-terminal pyroglutamate formation (–17.026549 Q) were considered differential modifications. Data was searched with 50 ppm precursor ion tolerance and 50 ppm fragment ion tolerance. Identified proteins were filtered using DTASelect2[159], utilizing a target-decoy database search strategy to limit the false discovery rate to 1%, at the spectrum level[160]. A minimum of 1 peptide per protein and no tryptic end per peptide were required and precursor delta mass cut-off was fixed at 15 ppm. Statistical models for peptide mass modification (modstat) were applied. Census2[161] label-free analysis was performed based on the precursor peak area, with a 15 ppm precursor mass tolerance and 0.1 min retention time tolerance. “Match between runs” was used to find missing peptides between runs. Data analysis using GlycoMSQuant[77] was implemented to automate the analysis. GlycoMSQuant summed precursor peak areas across replicates, discarded peptides without NGS, discarded misidentified peptides when *N*-glycan remnant-mass modifications were localized to non-NGS asparagines and corrected/fixed *N*-glycan mislocalization where appropriate. The results were aligned to NGS in spike of Wuhan-Hu-1 (UniProt #P0DTC2). The data can be accessed from MassIVE (MSV000091921)

A glycosylated model[53] based on the “closed state” of SARS-CoV-2 spike-protein trimer (PDB: 6VXX)[42] and *N*-glycosylation profile determined by Watanabe *et. al.*[57] was uploaded on ChimeraX v.1.5

(https://www.cgl.ucsf.edu/chimerax/). Glycans were separated into separate PDB files using The PyMOL Molecular Graphics System v.2.5.2 (https://pymol.org/). Coordinates for membrane structure were hidden and glycans were color-coded using Python v.3.9.7 (https://www.python.org/) in multiple formats: [1] monochrome to highlight all glycans over protein sub-structure, [2] a 5-color gradient to classify the dominant proportion of unoccupied glycosylation sites, high-mannose/hybrid glycans and complex glycans, and [3] a 21-color gradient to represent percentage-point shift in proportion of high-mannose/hybrid glycans at a particular glycosylation site[162], when comparing 2 immunogens. The code used to create visualizations can be accessed from GitHub (https://github.com/BhavyaSingh-Rolly/SARS-Visualization)

Proportional distribution of glycans at NGS common to all variants (*i.e.*, except N17, N20, N74 and N188, and the inconsistently sampled N709, N717 and N1134), as calculated by GlycoMSQuant[77], was condensed into a 1D feature vector, and a pairwise correlation analysis was conducted between all variants in their 6P-Mut7 versions. Subsequently, we used Scipy's hierarchical linkage function to cluster based on similarity and then used the hierarchical dendrogram function to visualize the results[163].

Supplementary Information Titles/Legends

Figure S1: Previous quantitation of site-specific *N*-glycosylation on SARS-CoV-2 spike-protein trimer vary. Shown are the results from site-specific *N*-glycosylation analysis by standard MS-based glycoproteomic method on Wuhan-Hu-1 as reported by Watanabe *et. al.*, 2020 [S1] (Report 1); Chawla *et. al.*, 2022 [S2] (Report 2); Newby *et. al.*, 2023 [S3] (Report 3); Shajahan *et. al.*, 2020 [S4] (Report 4); Shajahan *et. al.*, 2023 [S5] (Report 5); and Wang *et. al.*, 2021 [S6] (Report 6), compared to results from DeGlyPHER (top 2 rows, 2P followed by 6P). A visual estimate from these reports is plotted here and compared at each *N*-glycosylation site. Variation in glycosylation pattern observed in different reports are apparent at many sites. Color-coding of “Report #”, groups the reports based on research laboratories authoring them. Cited studies used either 2P stabilized mutant of SARS-CoV-2 spike-protein [S7, S8] or 6P stabilized mutant [S9], as indicated, except report 4 that used independently expressed S1 and S2 subunits of S-protein. Error bars are absent here because all the cited reports here except this study, have not estimated error in their calculations, presumably owing to low sampling. *Report 4 does not claim quantitation; thus, estimated values are from the types of glycoforms reported. Only this study and Report 1 have validated that S-proteins examined are well-folded trimers (using negative-stain electron microscopy). n.d.: glycosylation not determined.

Figure S2: Evaluating SARS-CoV-2 spike-protein trimer purity and integrity. (A) Size exclusion chromatography of SARS-CoV-2 spike-protein trimers on Superose 6 columns. Trimer elution peaks at 68- and 13-mL post-injection with Superose 6 Hiload 16/600 pg and Superose 6 Increase 10/300 GL columns, respectively. (B) Representative negative-stain TEM micrographs and 2D classes of SARS-CoV-2 spike-protein trimers show highly homogenous particles in various orientations.

Figure S3: Peptide sampling at each NGS across all spike-protein trimers analyzed. (A) Number of unique peptides mapping to each NGS shows that N709, N717 and N1134 are inconsistently sampled, but all other NGS are sufficiently sampled. (B) Bar graph demonstrating *N*-glycan heterogeneity determined at N1134. The *N*-glycans present at N1134 are mostly complex as seen in previous studies, though the sampling at this NGS is inconsistent in our study. *N*-glycosylation states are color-coded. Error bars represent mean–SEM.

Supplementary Table-1: SARS-CoV-2 variant spike proteins.

Supplementary Table-2: SARS-CoV-2 spike-protein stabilizing mutations.

SARS-CoV-2 spike-protein sequences.

References

1 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020). https://doi.org:10.1038/s41586-020-2012-7
2 Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269 (2020). https://doi.org:10.1038/s41586-020-2008-3
3 Cutler, D. M. & Summers, L. H. The COVID-19 Pandemic and the $16 Trillion Virus. *Jama* **324**, 1495-1496 (2020). https://doi.org:10.1001/jama.2020.19759
4 Koelle, K., Martin, M. A., Antia, R., Lopman, B. & Dean, N. E. The changing epidemiology of SARS-CoV-2. *Science* **375**, 1116-1121 (2022). https://doi.org:10.1126/science.abm4915
5 Worobey, M. *et al.* The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. *Science* **377**, 951-959 (2022). https://doi.org:10.1126/science.abp8715
6 Pekar, J. E. *et al.* The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2. *Science* **377**, 960-966 (2022). https://doi.org:10.1126/science.abp8337
7 Gao, G. *et al.* (Research Square, 2022).
8 Tao, K. *et al.* The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat Rev Genet* **22**, 757-773 (2021). https://doi.org:10.1038/s41576-021-00408-x
9 Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* **19**, 409-424 (2021). https://doi.org:10.1038/s41579-021-00573-0
10 Tegally, H. *et al.* Emergence of SARS-CoV-2 Omicron lineages BA.4 and BA.5 in South Africa. *Nat Med* **28**, 1785-1790 (2022). https://doi.org:10.1038/s41591-022-01911-2
11 Carabelli, A. M. *et al.* SARS-CoV-2 variant biology: immune escape, transmission and fitness. *Nat Rev Microbiol* **21**, 162-177 (2023). https://doi.org:10.1038/s41579-022-00841-7
12 Suzuki, R. *et al.* Attenuated fusogenicity and pathogenicity of SARS-CoV-2 Omicron variant. *Nature* **603**, 700-705 (2022). https://doi.org:10.1038/s41586-022-04462-1
13 Manathunga, S. S., Abeyagunawardena, I. A. & Dharmaratne, S. D. A comparison of transmissibility of SARS-CoV-2 variants of concern. *Virol J* **20**, 59 (2023). https://doi.org:10.1186/s12985-023-02018-x
14 Konings, F. *et al.* SARS-CoV-2 Variants of Interest and Concern naming scheme conducive for global discourse. *Nat Microbiol* **6**, 821-823 (2021). https://doi.org:10.1038/s41564-021-00932-w
15 Tang, J. W., Tambyah, P. A. & Hui, D. S. Emergence of a new SARS-CoV-2 variant in the UK. *J Infect* **82**, e27-e28 (2021). https://doi.org:10.1016/j.jinf.2020.12.024
16 Washington, N. L. *et al.* Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell* **184**, 2587-2594.e2587 (2021). https://doi.org:10.1016/j.cell.2021.03.052
17 Tegally, H. *et al.* Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438-443 (2021). https://doi.org:10.1038/s41586-021-03402-9
18 Faria, N. R. *et al.* Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815-821 (2021). https://doi.org:10.1126/science.abh2644
19 Mlcochova, P. *et al.* SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature* **599**, 114-119 (2021). https://doi.org:10.1038/s41586-021-03944-y
20 Viana, R. *et al.* Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* **603**, 679-686 (2022). https://doi.org:10.1038/s41586-022-04411-y
21 Laiton-Donato, K. *et al.* Characterization of the emerging B.1.621 variant of interest of SARS-CoV-2. *Infect Genet Evol* **95**, 105038 (2021). https://doi.org:10.1016/j.meegid.2021.105038
22 Halfmann, P. J. *et al.* Characterization of the SARS-CoV-2 B.1.621 (Mu) variant. *Sci Transl Med* **14**, eabm4908 (2022). https://doi.org:10.1126/scitranslmed.abm4908
23 Romero, P. E. *et al.* The Emergence of Sars-CoV-2 Variant Lambda (C.37) in South America. *Microbiol Spectr* **9**, e0078921 (2021). https://doi.org:10.1128/Spectrum.00789-21
24 Haynes, B. F. *et al.* Strategies for HIV-1 vaccines that induce broadly neutralizing antibodies. *Nat Rev Immunol*, 1-17 (2022). https://doi.org:10.1038/s41577-022-00753-w
25 Burton, D. R. & Hangartner, L. Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design. *Annu Rev Immunol* **34**, 635-659 (2016). https://doi.org:10.1146/annurev-immunol-041015-055515
26 Leggat, D. J. *et al.* Vaccination induces HIV broadly neutralizing antibody precursors in humans. *Science* **378**, eadd6502 (2022). https://doi.org:10.1126/science.add6502
27 Wilson, I. A., Skehel, J. J. & Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. *Nature* **289**, 366-373 (1981). https://doi.org:10.1038/289366a0
28 Kirchdoerfer, R. N. *et al.* Pre-fusion structure of a human coronavirus spike protein. *Nature* **531**, 118-121 (2016). https://doi.org:10.1038/nature17200

29 Pallesen, J. *et al.* Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc Natl Acad Sci U S A* **114**, E7348-e7357 (2017). https://doi.org:10.1073/pnas.1707304114
30 Lyumkis, D. *et al.* Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1484-1490 (2013). https://doi.org:10.1126/science.1245627
31 Julien, J. P. *et al.* Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1477-1483 (2013). https://doi.org:10.1126/science.1245625
32 Kanekiyo, M. & Graham, B. S. Next-Generation Influenza Vaccines. *Cold Spring Harb Perspect Med* **11** (2021). https://doi.org:10.1101/cshperspect.a038448
33 Karikó, K., Buckstein, M., Ni, H. & Weissman, D. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* **23**, 165-175 (2005). https://doi.org:10.1016/j.immuni.2005.06.008
34 Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020). https://doi.org:10.1126/science.abb2507
35 Mulligan, M. J. *et al.* Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. *Nature* **586**, 589-593 (2020). https://doi.org:10.1038/s41586-020-2639-4
36 Jackson, L. A. *et al.* An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. *N Engl J Med* **383**, 1920-1931 (2020). https://doi.org:10.1056/NEJMoa2022483
37 Sadoff, J. *et al.* Interim Results of a Phase 1-2a Trial of Ad26.COV2.S Covid-19 Vaccine. *N Engl J Med* **384**, 1824-1835 (2021). https://doi.org:10.1056/NEJMoa2034201
38 Voysey, M. *et al.* Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* **397**, 99-111 (2021). https://doi.org:10.1016/s0140-6736(20)32661-1
39 Ella, R. *et al.* Efficacy, safety, and lot-to-lot immunogenicity of an inactivated SARS-CoV-2 vaccine (BBV152): interim results of a randomised, double-blind, controlled, phase 3 trial. *Lancet* **398**, 2173-2184 (2021). https://doi.org:10.1016/s0140-6736(21)02000-6
40 Logunov, D. Y. *et al.* Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. *Lancet* **397**, 671-681 (2021). https://doi.org:10.1016/s0140-6736(21)00234-8
41 Heinz, F. X. & Stiasny, K. Distinguishing features of current COVID-19 vaccines: knowns and unknowns of antigen presentation and modes of action. *NPJ Vaccines* **6**, 104 (2021). https://doi.org:10.1038/s41541-021-00369-6
42 Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e286 (2020). https://doi.org:10.1016/j.cell.2020.02.058
43 Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215-220 (2020). https://doi.org:10.1038/s41586-020-2180-5
44 Wolf, K. A., Kwan, J. C. & Kamil, J. P. Structural Dynamics and Molecular Evolution of the SARS-CoV-2 Spike Protein. *mBio* **13**, e0203021 (2022). https://doi.org:10.1128/mbio.02030-21
45 V'Kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol* **19**, 155-170 (2021). https://doi.org:10.1038/s41579-020-00468-6
46 Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat Rev Mol Cell Biol* **23**, 3-20 (2022). https://doi.org:10.1038/s41580-021-00418-x
47 Marshall, R. D. Glycoproteins. *Annu Rev Biochem* **41**, 673-702 (1972). https://doi.org:10.1146/annurev.bi.41.070172.003325
48 Schjoldager, K. T., Narimatsu, Y., Joshi, H. J. & Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat Rev Mol Cell Biol* **21**, 729-749 (2020). https://doi.org:10.1038/s41580-020-00294-x
49 Cao, L. *et al.* Global site-specific analysis of glycoprotein N-glycan processing. *Nat Protoc* **13**, 1196-1212 (2018). https://doi.org:10.1038/nprot.2018.024
50 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 (2019). https://doi.org:10.1016/j.bbagen.2019.05.012
51 Chawla, H., Fadda, E. & Crispin, M. Principles of SARS-CoV-2 glycosylation. *Curr Opin Struct Biol* **75**, 102402 (2022). https://doi.org:10.1016/j.sbi.2022.102402
52 Sztain, T. *et al.* A glycan gate controls opening of the SARS-CoV-2 spike protein. *Nat Chem* **13**, 963-968 (2021). https://doi.org:10.1038/s41557-021-00758-3

53 Casalino, L. *et al.* Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent Sci* **6**, 1722-1734 (2020). https://doi.org:10.1021/acscentsci.0c01056
54 Pang, Y. T., Acharya, A., Lynch, D. L., Pavlova, A. & Gumbart, J. C. SARS-CoV-2 spike opening dynamics and energetics reveal the individual roles of glycans and their collective impact. *Commun Biol* **5**, 1170 (2022). https://doi.org:10.1038/s42003-022-04138-6
55 Thompson, A. J., de Vries, R. P. & Paulson, J. C. Virus recognition of glycan receptors. *Curr Opin Virol* **34**, 117-129 (2019). https://doi.org:10.1016/j.coviro.2019.01.004
56 Watanabe, Y. *et al.* Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat Commun* **11**, 2688 (2020). https://doi.org:10.1038/s41467-020-16567-0
57 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 (2020). https://doi.org:10.1126/science.abb9983
58 Zhang, S. *et al.* Analysis of Glycosylation and Disulfide Bonding of Wild-Type SARS-CoV-2 Spike Glycoprotein. *J Virol* **96**, e0162621 (2022). https://doi.org:10.1128/jvi.01626-21
59 Bangaru, S. *et al.* Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. *Science* **370**, 1089-1094 (2020). https://doi.org:10.1126/science.abe1502
60 Allen, J. D. *et al.* Site-Specific Steric Control of SARS-CoV-2 Spike Glycosylation. *Biochemistry* **60**, 2153-2169 (2021). https://doi.org:10.1021/acs.biochem.1c00279
61 Zhao, P. *et al.* Virus-Receptor Interactions of Glycosylated SARS-CoV-2 Spike and Human ACE2 Receptor. *Cell Host Microbe* **28**, 586-601.e586 (2020). https://doi.org:10.1016/j.chom.2020.08.004
62 Gobeil, S. M. *et al.* Effect of natural mutations of SARS-CoV-2 on spike structure, conformation, and antigenicity. *Science* **373** (2021). https://doi.org:10.1126/science.abi6226
63 Yuan, M. *et al.* Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. *Science* **373**, 818-823 (2021). https://doi.org:10.1126/science.abh1139
64 Newby, M. L. *et al.* Variations within the glycan shield of SARS-CoV-2 impact viral spike dynamics. *J Mol Biol*, 167928 (2022). https://doi.org:10.1016/j.jmb.2022.167928
65 Shajahan, A., Pepi, L. E., Kumar, B., Murray, N. B. & Azadi, P. Site specific N- and O-glycosylation mapping of the spike proteins of SARS-CoV-2 variants of concern. *Sci Rep* **13**, 10053 (2023). https://doi.org:10.1038/s41598-023-33088-0
66 Wang, D., Baudys, J., Osman, S. H. & Barr, J. R. Analysis of the N-glycosylation profiles of the spike proteins from the Alpha, Beta, Gamma, and Delta variants of SARS-CoV-2. *Anal Bioanal Chem* **415**, 4779-4793 (2023). https://doi.org:10.1007/s00216-023-04771-y
67 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 (2022). https://doi.org:10.1093/glycob/cwab102
68 Pegg, C. L. *et al.* The role of N-glycosylation in spike antigenicity for the SARS-CoV-2 Gamma variant. *bioRxiv*, 2023.2004.2003.535004 (2023). https://doi.org:10.1101/2023.04.03.535004
69 Xie, Y. & Butler, M. Quantitative profiling of N-glycosylation of SARS-CoV-2 spike protein variants. *Glycobiology* **33**, 188-202 (2023). https://doi.org:10.1093/glycob/cwad007
70 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 (2021). https://doi.org:10.1038/s41598-021-02904-w
71 Sanda, M., Morrison, L. & Goldman, R. N- and O-Glycosylation of the SARS-CoV-2 Spike Protein. *Anal Chem* **93**, 2003-2009 (2021). https://doi.org:10.1021/acs.analchem.0c03173
72 Shajahan, A., Supekar, N. T., Gleinich, A. S. & Azadi, P. Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* **30**, 981-988 (2020). https://doi.org:10.1093/glycob/cwaa042
73 Shajahan, A., Pepi, L. E., Rouhani, D. S., Heiss, C. & Azadi, P. Glycosylation of SARS-CoV-2: structural and functional insights. *Anal Bioanal Chem* **413**, 7179-7193 (2021). https://doi.org:10.1007/s00216-021-03499-x
74 Gong, Y., Qin, S., Dai, L. & Tian, Z. The glycosylation in SARS-CoV-2 and its receptor ACE2. *Signal Transduct Target Ther* **6**, 396 (2021). https://doi.org:10.1038/s41392-021-00809-8
75 Yao, H. *et al.* Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **183**, 730-738.e713 (2020). https://doi.org:10.1016/j.cell.2020.09.018
76 Chawla, H. *et al.* Glycosylation and Serological Reactivity of an Expression-enhanced SARS-CoV-2 Viral Spike Mimetic. *J Mol Biol* **434**, 167332 (2022). https://doi.org:10.1016/j.jmb.2021.167332

77 Baboo, S. *et al.* DeGlyPHER: An Ultrasensitive Method for the Analysis of Viral Spike N-Glycoforms. *Anal Chem* **93**, 13651-13657 (2021). https://doi.org:10.1021/acs.analchem.1c03059
78 Cao, L. *et al.* Global site-specific N-glycosylation analysis of HIV envelope glycoprotein. *Nat Commun* **8**, 14954 (2017). https://doi.org:10.1038/ncomms14954
79 Baboo, S. *et al.* DeGlyPHER: Highly sensitive site-specific analysis of N-linked glycans on proteins. *Methods Enzymol* **682**, 137-185 (2023). https://doi.org:10.1016/bs.mie.2022.09.004
80 Zhang, J., Xiao, T., Cai, Y. & Chen, B. Structure of SARS-CoV-2 spike protein. *Curr Opin Virol* **50**, 173-182 (2021). https://doi.org:10.1016/j.coviro.2021.08.010
81 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 (2021). https://doi.org:10.1016/j.chom.2021.03.005
82 Wang, Z. *et al.* Analysis of memory B cells identifies conserved neutralizing epitopes on the N-terminal domain of variant SARS-Cov-2 spike proteins. *Immunity* **55**, 998-1012.e1018 (2022). https://doi.org:10.1016/j.immuni.2022.04.003
83 Hsieh, C. L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501-1505 (2020). https://doi.org:10.1126/science.abd0826
84 Torres, J. L. *et al.* Structural insights of a highly potent pan-neutralizing SARS-CoV-2 human monoclonal antibody. *Proc Natl Acad Sci U S A* **119**, e2120976119 (2022). https://doi.org:10.1073/pnas.2120976119
85 Korber, B. *et al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812-827.e819 (2020). https://doi.org:10.1016/j.cell.2020.06.043
86 Martin, D. P. *et al.* The emergence and ongoing convergent evolution of the SARS-CoV-2 N501Y lineages. *Cell* **184**, 5189-5200.e5187 (2021). https://doi.org:10.1016/j.cell.2021.09.003
87 Liu, Y. *et al.* The N501Y spike substitution enhances SARS-CoV-2 infection and transmission. *Nature* **602**, 294-299 (2022). https://doi.org:10.1038/s41586-021-04245-0
88 Alkhatib, M. *et al.* SARS-CoV-2 Variants and Their Relevant Mutational Profiles: Update Summer 2021. *Microbiol Spectr* **9**, e0109621 (2021). https://doi.org:10.1128/Spectrum.01096-21
89 Iketani, S. *et al.* Antibody evasion properties of SARS-CoV-2 Omicron sublineages. *Nature* **604**, 553-556 (2022). https://doi.org:10.1038/s41586-022-04594-4
90 Cai, Y. *et al.* Distinct conformational states of SARS-CoV-2 spike protein. *Science* **369**, 1586-1592 (2020). https://doi.org:10.1126/science.abd4251
91 Fan, X., Cao, D., Kong, L. & Zhang, X. Cryo-EM analysis of the post-fusion structure of the SARS-CoV spike glycoprotein. *Nat Commun* **11**, 3618 (2020). https://doi.org:10.1038/s41467-020-17371-6
92 Dodero-Rojas, E., Onuchic, J. N. & Whitford, P. C. Sterically confined rearrangements of SARS-CoV-2 Spike protein control cell invasion. *Elife* **10** (2021). https://doi.org:10.7554/eLife.70362
93 Guo, L. *et al.* Epigenetic glycosylation of SARS-CoV-2 impact viral infection through DC&L-SIGN receptors. *iScience* **24**, 103426 (2021). https://doi.org:10.1016/j.isci.2021.103426
94 Allen, J. D. *et al.* The diversity of the glycan shield of sarbecoviruses related to SARS-CoV-2. *Cell Rep* **42**, 112307 (2023). https://doi.org:10.1016/j.celrep.2023.112307
95 Zhou, P. *et al.* A human antibody reveals a conserved site on beta-coronavirus spike proteins and confers protection against SARS-CoV-2 infection. *Sci Transl Med* **14**, eabi9215 (2022). https://doi.org:10.1126/scitranslmed.abi9215
96 Turoňová, B. *et al.* In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* **370**, 203-208 (2020). https://doi.org:10.1126/science.abd5223
97 Zhang, J. *et al.* Membrane fusion and immune evasion by the spike protein of SARS-CoV-2 Delta variant. *Science* **374**, 1353-1360 (2021). https://doi.org:10.1126/science.abl9463
98 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 (2020). https://doi.org:10.1126/science.abc6952
99 Yang, T.-J. *et al.* Structure-activity relationships of B.1.617 and other SARS-CoV-2 spike variants. *bioRxiv*, 2021.2009.2012.459978 (2021). https://doi.org:10.1101/2021.09.12.459978
100 McCallum, M. *et al.* N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* **184**, 2332-2347.e2316 (2021). https://doi.org:10.1016/j.cell.2021.03.028
101 McCallum, M. *et al.* Molecular basis of immune evasion by the Delta and Kappa SARS-CoV-2 variants. *Science* **374**, 1621-1626 (2021). https://doi.org:10.1126/science.abl8506
102 Kimura, I. *et al.* The SARS-CoV-2 Lambda variant exhibits enhanced infectivity and immune resistance. *Cell Rep* **38**, 110218 (2022). https://doi.org:10.1016/j.celrep.2021.110218

103 Pascarella, S. *et al.* Shortening Epitopes to Survive: The Case of SARS-CoV-2 Lambda Variant. *Biomolecules* **11** (2021). https://doi.org:10.3390/biom11101494
104 Liu, H. *et al.* The Receptor Binding Domain of SARS-CoV-2 Lambda Variant Has a Better Chance Than the Delta Variant in Evading BNT162b2 COVID-19 mRNA Vaccine-Induced Humoral Immunity. *Int J Mol Sci* **23** (2022). https://doi.org:10.3390/ijms231911325
105 Choi, Y. K. *et al.* Structure, Dynamics, Receptor Binding, and Antibody Binding of the Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein in a Viral Membrane. *J Chem Theory Comput* **17**, 2479-2487 (2021). https://doi.org:10.1021/acs.jctc.0c01144
106 Pinto, D. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290-295 (2020). https://doi.org:10.1038/s41586-020-2349-y
107 Chen, R. E. *et al.* In vivo monoclonal antibody efficacy against SARS-CoV-2 variant strains. *Nature* **596**, 103-108 (2021). https://doi.org:10.1038/s41586-021-03720-y
108 Fang, Y. *et al.* An antibody that neutralizes SARS-CoV-1 and SARS-CoV-2 by binding to a conserved spike epitope outside the receptor binding motif. *Sci Immunol*, eabp9962 (2022). https://doi.org:10.1126/sciimmunol.abp9962
109 Lempp, F. A. *et al.* Lectins enhance SARS-CoV-2 infection and influence neutralizing antibodies. *Nature* **598**, 342-347 (2021). https://doi.org:10.1038/s41586-021-03925-1
110 Soh, W. T. *et al.* The N-terminal domain of spike glycoprotein mediates SARS-CoV-2 infection by associating with L-SIGN and DC-SIGN. *bioRxiv*, 2020.2011.2005.369264 (2020). https://doi.org:10.1101/2020.11.05.369264
111 Henderson, R. *et al.* Glycans on the SARS-CoV-2 Spike Control the Receptor Binding Domain Conformation. *bioRxiv* (2020). https://doi.org:10.1101/2020.06.26.173765
112 Wang, Q. *et al.* Alarming antibody evasion properties of rising SARS-CoV-2 BQ and XBB subvariants. *Cell* **186**, 279-286.e278 (2023). https://doi.org:10.1016/j.cell.2022.12.018
113 Fallon, L. *et al.* Free Energy Landscapes from SARS-CoV-2 Spike Glycoprotein Simulations Suggest that RBD Opening Can Be Modulated via Interactions in an Allosteric Pocket. *J Am Chem Soc* **143**, 11349-11360 (2021). https://doi.org:10.1021/jacs.1c00556
114 Seow, J. *et al.* A neutralizing epitope on the SD1 domain of SARS-CoV-2 spike targeted following infection and vaccination. *Cell Rep* **40**, 111276 (2022). https://doi.org:10.1016/j.celrep.2022.111276
115 Liu, L. *et al.* Antibodies targeting a quaternary site on SARS-CoV-2 spike glycoprotein prevent viral receptor engagement by conformational locking. *Immunity* **56**, 2442-2455.e2448 (2023). https://doi.org:10.1016/j.immuni.2023.09.003
116 Yang, Q. *et al.* Role for N-glycans and calnexin-calreticulin chaperones in SARS-CoV-2 Spike maturation and viral infectivity. *Sci Adv* **8**, eabq8678 (2022). https://doi.org:10.1126/sciadv.abq8678
117 Xu, C. *et al.* Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *Sci Adv* **7** (2021). https://doi.org:10.1126/sciadv.abe5575
118 Lemmin, T., Kalbermatter, D., Harder, D., Plattet, P. & Fotiadis, D. Structures and dynamics of the novel S1/S2 protease cleavage site loop of the SARS-CoV-2 spike glycoprotein. *J Struct Biol X* **4**, 100038 (2020). https://doi.org:10.1016/j.yjsbx.2020.100038
119 Zhang, L. *et al.* SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat Commun* **11**, 6013 (2020). https://doi.org:10.1038/s41467-020-19808-4
120 Daniloski, Z. *et al.* The Spike D614G mutation increases SARS-CoV-2 infection of multiple human cell types. *Elife* **10** (2021). https://doi.org:10.7554/eLife.65365
121 Plante, J. A. *et al.* Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **592**, 116-121 (2021). https://doi.org:10.1038/s41586-020-2895-3
122 Zhang, J. *et al.* Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science* **372**, 525-530 (2021). https://doi.org:10.1126/science.abf2303
123 Benton, D. J. *et al.* Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* **588**, 327-330 (2020). https://doi.org:10.1038/s41586-020-2772-0
124 Benton, D. J. *et al.* The effect of the D614G substitution on the structure of the spike glycoprotein of SARS-CoV-2. *Proc Natl Acad Sci U S A* **118** (2021). https://doi.org:10.1073/pnas.2022586118
125 Juraszek, J. *et al.* Stabilizing the closed SARS-CoV-2 spike trimer. *Nat Commun* **12**, 244 (2021). https://doi.org:10.1038/s41467-020-20321-x
126 Yurkovetskiy, L. *et al.* Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **183**, 739-751.e738 (2020). https://doi.org:10.1016/j.cell.2020.09.032

127 Calvaresi, V. *et al.* Structural dynamics in the evolution of SARS-CoV-2 spike glycoprotein. *Nat Commun* **14**, 1421 (2023). https://doi.org:10.1038/s41467-023-36745-0
128 Hu, J. *et al.* D614G mutation of SARS-CoV-2 spike protein enhances viral infectivity. *bioRxiv*, 2020.2006.2020.161323 (2020). https://doi.org:10.1101/2020.06.20.161323
129 Li, Q. *et al.* The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* **182**, 1284-1294.e1289 (2020). https://doi.org:10.1016/j.cell.2020.07.012
130 Tai, L. *et al.* Nanometer-resolution in situ structure of the SARS-CoV-2 postfusion spike protein. *Proc Natl Acad Sci U S A* **118** (2021). https://doi.org:10.1073/pnas.2112703118
131 Gao, C. *et al.* SARS-CoV-2 Spike Protein Interacts with Multiple Innate Immune Receptors. *bioRxiv* (2020). https://doi.org:10.1101/2020.07.29.227462
132 Hikmet, F. *et al.* The protein expression profile of ACE2 in human tissues. *Mol Syst Biol* **16**, e9610 (2020). https://doi.org:10.15252/msb.20209610
133 Wu, C. Y. *et al.* Glycosite-deleted mRNA of SARS-CoV-2 spike protein as a broad-spectrum vaccine. *Proc Natl Acad Sci U S A* **119** (2022). https://doi.org:10.1073/pnas.2119995119
134 Huang, H. C. *et al.* Targeting conserved N-glycosylation blocks SARS-CoV-2 variant infection in vitro. *EBioMedicine* **74**, 103712 (2021). https://doi.org:10.1016/j.ebiom.2021.103712
135 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 (2022). https://doi.org:10.1126/scitranslmed.abm0899
136 van den Elsen, J. M., Kuntz, D. A. & Rose, D. R. Structure of Golgi alpha-mannosidase II: a target for inhibition of growth and metastasis of cancer cells. *Embo j* **20**, 3008-3017 (2001). https://doi.org:10.1093/emboj/20.12.3008
137 Vallee, F., Karaveg, K., Herscovics, A., Moremen, K. W. & Howell, P. L. Structural basis for catalysis and inhibition of N-glycan processing class I alpha 1,2-mannosidases. *J Biol Chem* **275**, 41287-41298 (2000). https://doi.org:10.1074/jbc.M006927200
138 Tempel, W. *et al.* Structure of mouse Golgi alpha-mannosidase IA reveals the molecular basis for substrate specificity among class 1 (family 47 glycosylhydrolase) alpha1,2-mannosidases. *J Biol Chem* **279**, 29774-29786 (2004). https://doi.org:10.1074/jbc.M403065200
139 Harduin-Lepers, A. *et al.* The human sialyltransferase family. *Biochimie* **83**, 727-737 (2001). https://doi.org:10.1016/s0300-9084(01)01301-3
140 de Vries, T., Knegtel, R. M., Holmes, E. H. & Macher, B. A. Fucosyltransferases: structure/function studies. *Glycobiology* **11**, 119r-128r (2001). https://doi.org:10.1093/glycob/11.10.119r
141 Peacock, T. P. *et al.* The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. *Nat Microbiol* **6**, 899-909 (2021). https://doi.org:10.1038/s41564-021-00908-w
142 Bagdonaite, I. *et al.* Site-Specific O-Glycosylation Analysis of SARS-CoV-2 Spike Protein Produced in Insect and Human Cells. *Viruses* **13** (2021). https://doi.org:10.3390/v13040551
143 Tian, W. *et al.* O-glycosylation pattern of the SARS-CoV-2 spike protein reveals an "O-Follow-N" rule. *Cell Res* **31**, 1123-1125 (2021). https://doi.org:10.1038/s41422-021-00545-2
144 Saito, A. *et al.* Enhanced fusogenicity and pathogenicity of SARS-CoV-2 Delta P681R mutation. *Nature* **602**, 300-306 (2022). https://doi.org:10.1038/s41586-021-04266-9
145 Lubinski, B. *et al.* Spike Protein Cleavage-Activation in the Context of the SARS-CoV-2 P681R Mutation: an Analysis from Its First Appearance in Lineage A.23.1 Identified in Uganda. *Microbiol Spectr* **10**, e0151422 (2022). https://doi.org:10.1128/spectrum.01514-22
146 Ghorbani, M., Brooks, B. R. & Klauda, J. B. Exploring dynamics and network analysis of spike glycoprotein of SARS-COV-2. *Biophys J* **120**, 2902-2913 (2021). https://doi.org:10.1016/j.bpj.2021.02.047
147 Lu, Q. *et al.* SARS-CoV-2 exacerbates proinflammatory responses in myeloid cells through C-type lectin receptors and Tweety family member 2. *Immunity* **54**, 1304-1319.e1309 (2021). https://doi.org:10.1016/j.immuni.2021.05.006
148 Li, D., Sempowski, G. D., Saunders, K. O., Acharya, P. & Haynes, B. F. SARS-CoV-2 Neutralizing Antibodies for COVID-19 Prevention and Treatment. *Annu Rev Med* **73**, 1-16 (2022). https://doi.org:10.1146/annurev-med-042420-113838
149 Rapp, M., Shapiro, L. & Frank, J. Contributions of single-particle cryoelectron microscopy toward fighting COVID-19. *Trends Biochem Sci* **47**, 117-123 (2022). https://doi.org:10.1016/j.tibs.2021.10.005
150 Telenti, A., Hodcroft, E. B. & Robertson, D. L. The Evolution and Biology of SARS-CoV-2 Variants. *Cold Spring Harb Perspect Med* **12** (2022). https://doi.org:10.1101/cshperspect.a041390

151 Mykytyn, A. Z., Fouchier, R. A. & Haagmans, B. L. Antigenic evolution of SARS coronavirus 2. *Curr Opin Virol* **62**, 101349 (2023). https://doi.org:10.1016/j.coviro.2023.101349
152 Wrobel, A. G. Mechanism and evolution of human ACE2 binding by SARS-CoV-2 spike. *Curr Opin Struct Biol* **81**, 102619 (2023). https://doi.org:10.1016/j.sbi.2023.102619
153 Campos, D., Girgis, M. & Sanda, M. Site-specific glycosylation of SARS-CoV-2: Big challenges in mass spectrometry analysis. *Proteomics* **22**, e2100322 (2022). https://doi.org:10.1002/pmic.202100322
154 Suloway, C. *et al.* Automated molecular microscopy: the new Leginon system. *J Struct Biol* **151**, 41-60 (2005). https://doi.org:10.1016/j.jsb.2005.03.010
155 Lander, G. C. *et al.* Appion: an integrated, database-driven pipeline to facilitate EM image processing. *J Struct Biol* **166**, 95-102 (2009). https://doi.org:10.1016/j.jsb.2009.01.002
156 Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530 (2012). https://doi.org:10.1016/j.jsb.2012.09.006
157 He, L., Diedrich, J., Chu, Y. Y. & Yates, J. R., 3rd. Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter. *Anal Chem* **87**, 11361-11367 (2015). https://doi.org:10.1021/acs.analchem.5b02721
158 Xu, T. *et al.* ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *J Proteomics* **129**, 16-24 (2015). https://doi.org:10.1016/j.jprot.2015.07.001
159 Tabb, D. L., McDonald, W. H. & Yates, J. R., 3rd. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J Proteome Res* **1**, 21-26 (2002). https://doi.org:10.1021/pr015504q
160 Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J. & Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* **2**, 43-50 (2003). https://doi.org:10.1021/pr025556v
161 Park, S. K., Venable, J. D., Xu, T. & Yates, J. R., 3rd. A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat Methods* **5**, 319-322 (2008). https://doi.org:10.1038/nmeth.1195
162 Seabright, G. E. *et al.* Networks of HIV-1 Envelope Glycans Maintain Antibody Epitopes in the Face of Glycan Additions and Deletions. *Structure* **28**, 897-909.e896 (2020). https://doi.org:10.1016/j.str.2020.04.022
163 Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **17**, 261-272 (2020). https://doi.org:10.1038/s41592-019-0686-2

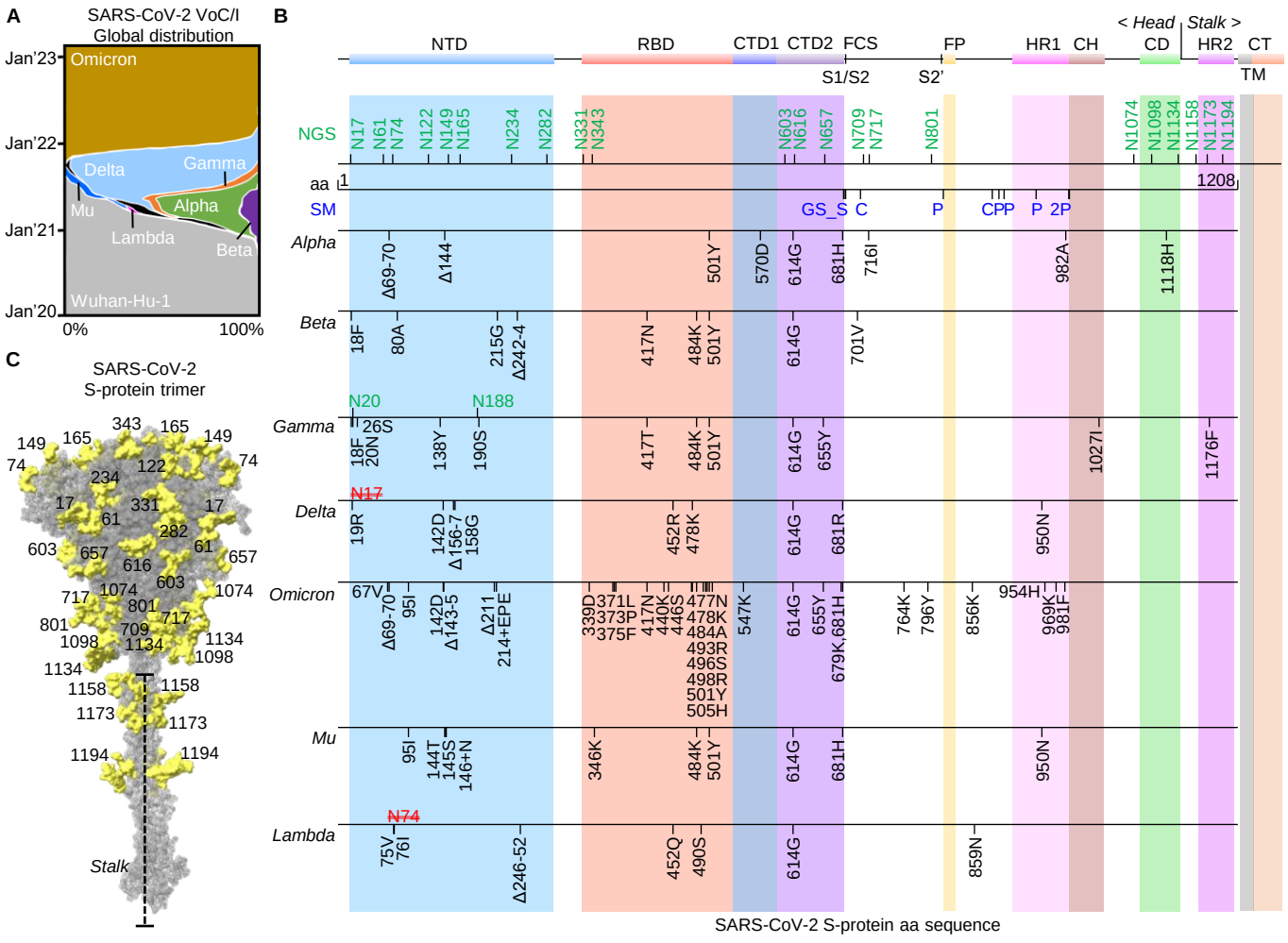
A
SARS-CoV-2 VoC/I
Global distribution
Jan'23
Jan'22
Jan'21
Jan'20
Omicron
Delta
Gamma
Alpha
Mu
Lambda
Beta
Wuhan-Hu-1
0%
100%
B
NTD
RBD
CTD1
CTD2
FCS
FP
HR1
CH
< Head
Stalk >
CD
HR2
CT
S1/S2
S2'
TM
NGS
N17 N61 N74 N122 N149 N165 N234 N282 N331 N343 N603 N616 N657 N709 N717 N801 N1074 N1098 N1134 N1158 N1173 N1194
aa
1
1208
SM
GS_S C P CPP P 2P
Alpha
Δ69-70 Δ144 501Y 570D 614G 681H 716I 982A 1118H
Beta
18F 80A 215G Δ242-4 417N 484K 501Y 614G 701V
Gamma
N20 N188
26S 18F 20N 138Y 190S 417T 484K 501Y 614G 655Y 1027I 1176F
Delta
N17
19R 142D Δ156-7 158G 452R 478K 614G 681R 950N
Omicron
67V Δ69-70 95I 142D Δ143-5 Δ211 214+EPE 339D 371L 373P 375F 417N 440K 446S 477N 478K 484A 493R 496S 498R 501Y 505H 547K 614G 655Y 679K 681H 764K 796Y 856K 954H 969K 981F
Mu
95I 144T 145S 146+N 346K 484K 501Y 614G 681H 950N
Lambda
N74
75V 76I Δ246-52 452Q 490S 614G 859N
SARS-CoV-2 S-protein aa sequence
C
SARS-CoV-2
S-protein trimer
Stalk

Figure 1

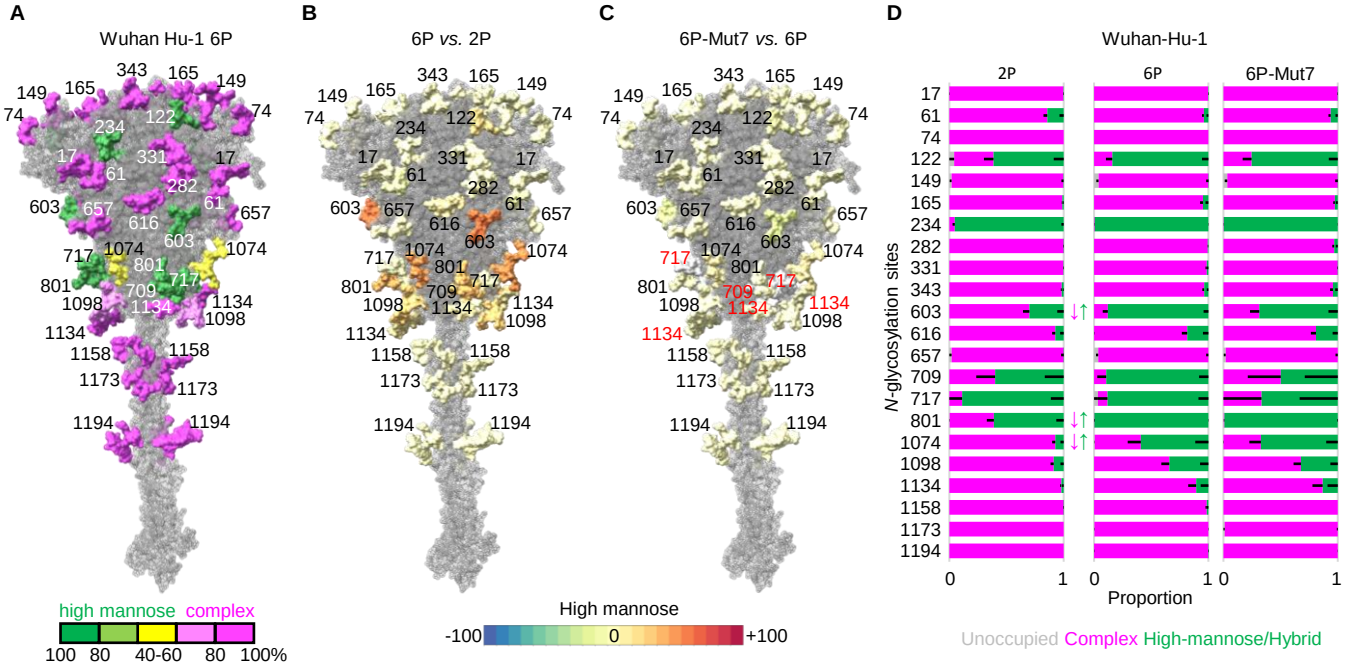
A
Wuhan Hu-1 6P
B
6P vs. 2P
C
6P-Mut7 vs. 6P
D
Wuhan-Hu-1
2P
6P
6P-Mut7
N-glycosylation sites
Proportion
high mannose
complex
100
80
40-60
80
100%
High mannose
-100
0
+100
Unoccupied
Complex
High-mannose/Hybrid

Figure 2

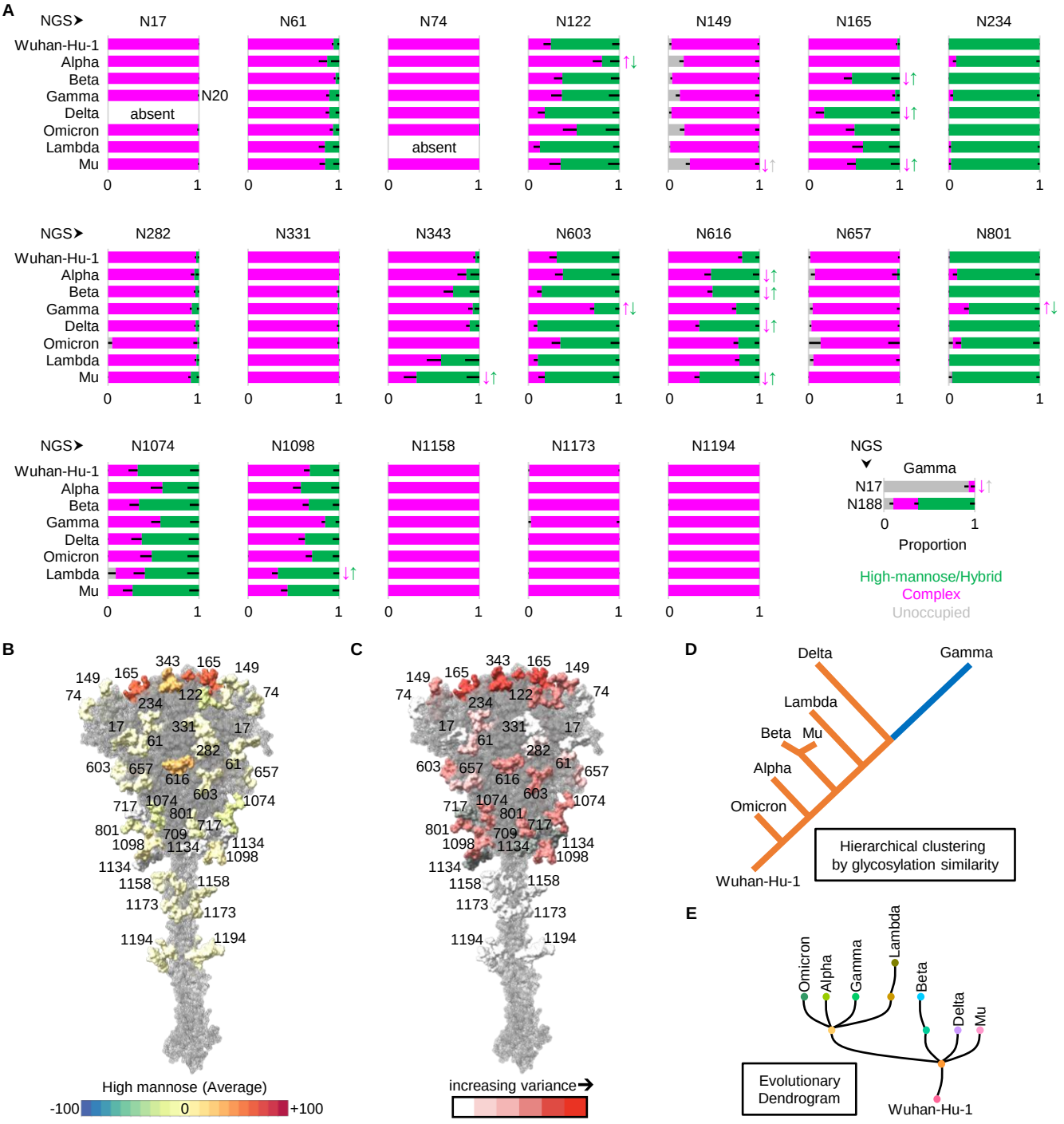
A
NGS
N17
N61
N74
N122
N149
N165
N234
Wuhan-Hu-1
Alpha
Beta
Gamma
Delta
Omicron
Lambda
Mu
N20
absent
absent
N282
N331
N343
N603
N616
N657
N801
N1074
N1098
N1158
N1173
N1194
0
1
Gamma
N17
N188
Proportion
High-mannose/Hybrid
Complex
Unoccupied
B
C
High mannose (Average)
-100
0
+100
increasing variance
D
Delta
Gamma
Lambda
Beta
Mu
Alpha
Omicron
Wuhan-Hu-1
Hierarchical clustering
by glycosylation similarity
E
Omicron
Alpha
Gamma
Lambda
Beta
Delta
Mu
Evolutionary
Dendrogram
Wuhan-Hu-1

Figure 3

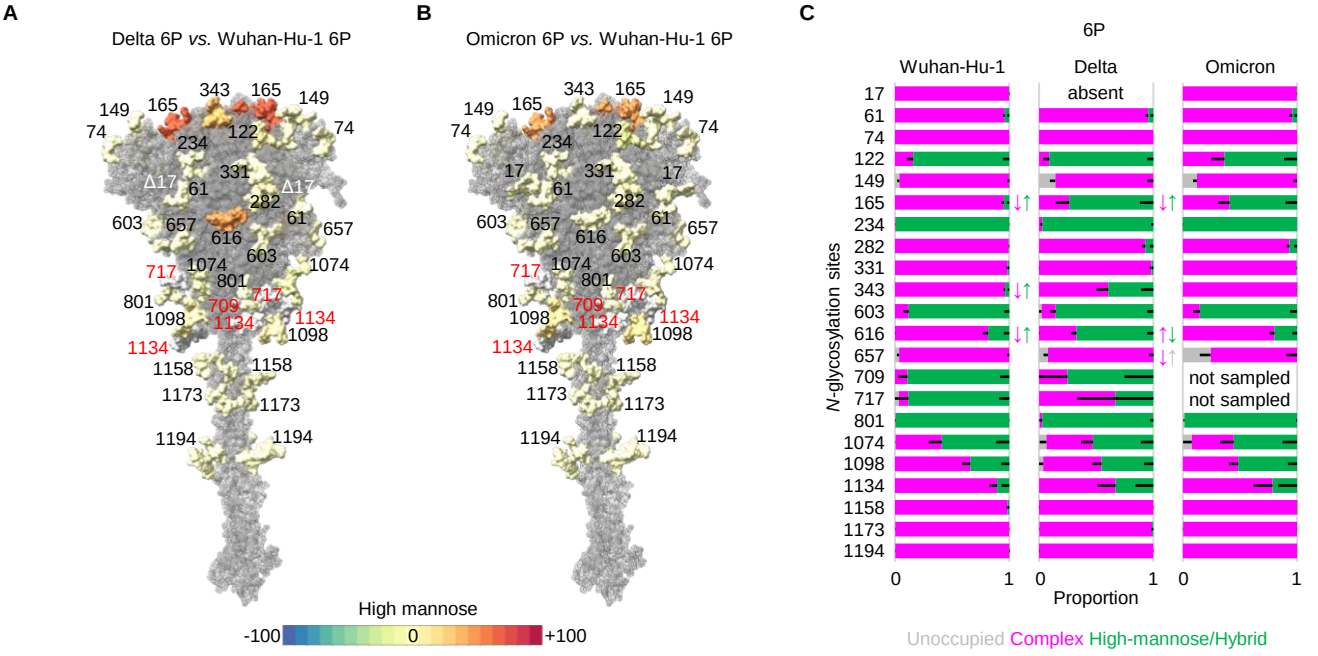
A
Delta 6P vs. Wuhan-Hu-1 6P
B
Omicron 6P vs. Wuhan-Hu-1 6P
High mannose
-100
0
+100
C
6P
Wuhan-Hu-1
Delta
Omicron
absent
not sampled
not sampled
N-glycosylation sites
Proportion
Unoccupied Complex High-mannose/Hybrid

Figure 4

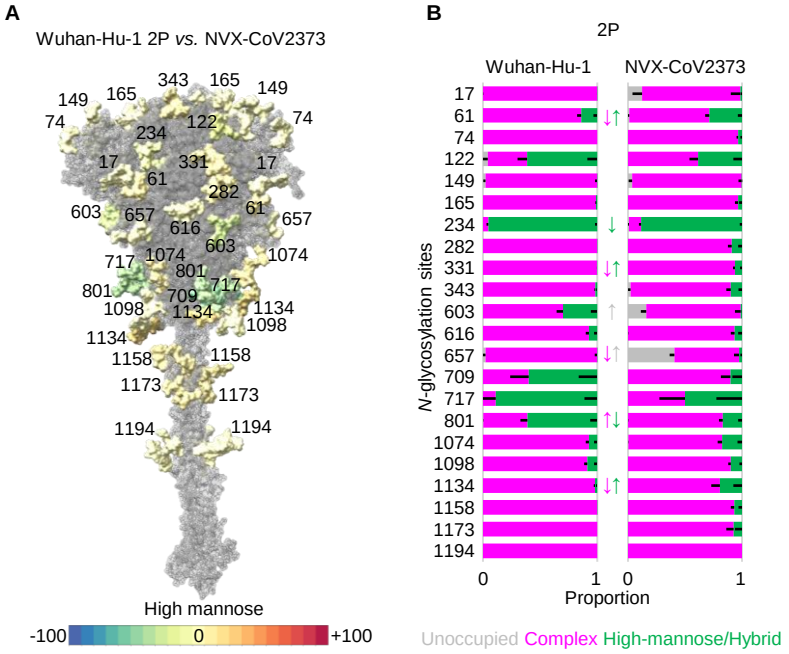
A
Wuhan-Hu-1 2P vs. NVX-CoV2373
343
165
149
165
149
74
74
122
234
17
331
17
61
282
61
603
657
657
616
603
717
1074
1074
801
801
717
709
1098
1134
1134
1098
1134
1158
1158
1173
1173
1194
1194
High mannose
-100
0
+100
B
2P
Wuhan-Hu-1
NVX-CoV2373
N-glycosylation sites
17
61
74
122
149
165
234
282
331
343
603
616
657
709
717
801
1074
1098
1134
1158
1173
1194
0
1
0
1
Proportion
Unoccupied
Complex
High-mannose/Hybrid

Figure 5