# Biochemistry

# Site-Specific Steric Control of SARS-CoV-2 Spike Glycosylation

Joel D. Allen, Himanshi Chawla, Firdaus Samsudin, Lorena Zuzic, Aishwary Tukaram Shivgan, Yasunori Watanabe, Wan-ting He, Sean Callaghan, Ge Song, Peter Yong, Philip J. M. Brouwer, Yutong Song, Yongfei Cai, Helen M. E. Duyvesteyn, Tomas Malinauskas, Joeri Kint, Paco Pino, Maria J. Wurm, Martin Frank, Bing Chen, David I. Stuart, Rogier W. Sanders, Raiees Andrabi, Dennis R. Burton, Sai Li, Peter J. Bond,* and Max Crispin*

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

**ABSTRACT:** A central tenet in the design of vaccines is the display of native-like antigens in the elicitation of protective immunity. The abundance of N-linked glycans across the SARS-CoV-2 spike protein is a potential source of heterogeneity among the many different vaccine candidates under investigation. Here, we investigate the glycosylation of recombinant SARS-CoV-2 spike proteins from five different laboratories and compare them against S protein from infectious virus, cultured in Vero cells. We find patterns that are conserved across all samples, and this can be associated with site-specific stalling of glycan maturation that acts as a highly sensitive reporter of protein structure. Molecular dynamics simulations of a fully glycosylated spike support a model of steric restrictions that shape enzymatic processing of the glycans. These results suggest that recombinant spike-based SARS-CoV-2 immunogen glycosylation reproducibly recapitulates signatures of viral glycosylation.
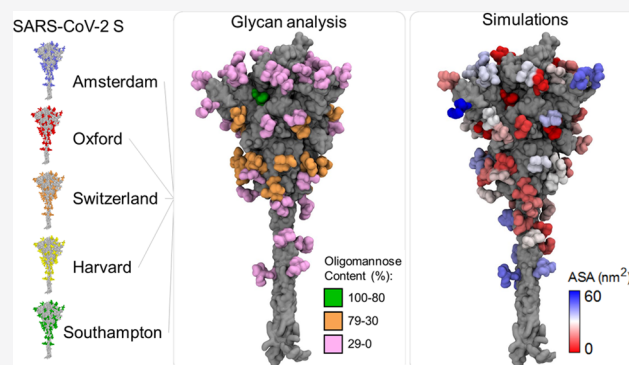


The Coronavirus Disease 2019 (COVID-19) pandemic has prompted the development of an unprecedented array of vaccine candidates against the causative pathogen, severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2). All approaches aim to deliver molecular features of the virus to induce immunity. The viral spike glycoprotein, also termed S protein, has emerged as the principal focus of vaccine design efforts as antibodies against this target can offer robust immunity.[1−6] Encouragingly, neutralization can readily occur despite the extensive array of N-linked glycans distributed across the viral spike consistent with numerous vulnerabilities in this so-called glycan shield.[7] Despite these observations, glycosylation has emerged as an important parameter in vaccine development for SARS-CoV-1 and SARS-CoV-2.[8,9] The glycosylation processing state can influence immunogen trafficking in the lymphatic system,[10] influence the presentation of both native and unwanted cryptic epitopes,[11] and reveal the extent to which immunogens recapitulate native viral architecture. Evidence that glycosylation can somewhat influence the interaction between SARS-CoV-2 and its target receptor, angiotensin converting enzyme 2 (ACE2), is also emerging.[12−15] Each protomer of the trimeric SARS-CoV-2 S protein contains at least 22 N-linked glycosylation sequons that direct the attachment of host glycans to specific Asn residues. This extensive glycosylation is important in lectin-

mediated protein folding and direct stabilization of the protein fold.[16] In addition, certain glycans incompletely mature during biogenesis and can lead to the presentation of immature glycans terminating with mannose residues that can act as ligands for innate immune recognition.[17−19]

Despite the focus on the S protein in vaccine development efforts, there has been considerable divergence in the mechanisms of delivery. In one approach, a nucleic acid encoding the spike is delivered through mRNA or with a viral vector.[4−6,20−23] The resulting S protein is assembled and glycosylated by the host tissue. In a contrasting approach, the S protein can be recombinantly manufactured as either a recombinant protein using mammalian or insect cell lines or using inactivated virus-based approaches that allow detailed characterization of the immunogen prior to delivery.[24−28] Immunogen glycosylation can be influenced by factors specific to the manufacturing conditions such as cell-type or cell culture conditions;[29,30] however, construct design and protein

architecture can also have a substantial impact. For example, underprocessed oligomannose sites can occur at sites sterically hidden from the host mannosidase enzymes by the tertiary or quaternary architecture, including obfuscation by the neighboring protein and glycan structure.[18,31] Immunogens displaying native-like architecture recapitulate these sites of oligomannose glycosylation. Conversely, immunogen design can adversely impact the presentation of native-like glycosylation. Importantly, despite the differences in the biosynthesis of S protein in virions and from mammalian expression systems, they seem to generate broadly similar glycosylation.[32] However, the success of a broad range of different vaccine platforms exhibiting different S protein glycosylation indicates that native-like glycosylation is not a prerequisite for a successful vaccine. Despite this observation, understanding S protein glycosylation will help benchmark material employed in different serological and vaccine studies and help define the impact of this extensive features of the protein surface.

The flexible and heterogeneous nature of N-linked glycosylation necessitates auxiliary methodologies in addition to cryo-electron microscopy or X-ray crystallography to characterize this key part of the S protein structure. Site-specific glycan analysis employing liquid chromatography-mass spectrometry is a widely used approach for obtaining this information.[33−37] As research into the structure and function of the SARS-CoV-2 S protein has progressed, more details about the glycan shield of S protein have become apparent. Analyses of recombinant trimeric S protein revealed divergent N-linked glycosylation from host glycoproteins in the presence of underprocessed oligomannose-type glycans at several sites.[7,15,38] Comparative analyses with monomeric and trimeric S proteins have revealed site-specific differences in glycosylation with regard to both oligomannose-type glycans and the presentation of sialic acid.[38] Analysis of S protein from insect cells demonstrated that oligomannose-type glycans were conserved on trimeric S protein, notably at N234.[39] In addition, molecular dynamics (MD) studies have proposed that the N234 site plays a role in stabilizing the receptor binding domain (RBD) in an exposed "up" conformation.[31] The presence of larger underprocessed oligomannose-type glycans, such as $Man_9GlcNAc_2$, on both mammalian and insect-derived S protein provides an indication that the structure of the S protein is driving the presentation of these glycans. Subsequent studies have investigated the presentation of N-linked glycans on S protein produced for vaccination, notably the Novavax full length S protein and S protein isolated following administration of the ChAdOx-nCoV-19 vector.[39,40] The observed glycan signatures were broadly in agreement with previous analyses. However, these studies involved the truncation of glycan structures using glycosidase treatment, which is useful for categorizing glycans into high-mannose or complex-type glycans and determining the potential N-glycosylation site (PNGS) occupancy on small amounts of material but does not allow for the identification of changes in terminal glycan processing such as sialylation. The glycan processing of the two N-linked glycan sites located on the RBD has also been investigated for the monomeric RBD. These sites present high levels of complex-type glycans,[41,42] and as a majority of antibodies raised against SARS-CoV-2 S protein target the RBD, it is important to fully characterize the structure of the RBD, including the presentation of glycans.

N-Linked glycans are highly dynamic and can substantially vary in chemical composition within a single batch of protein.

While glycosylation is heterogeneous, it is important for therapeutic and vaccine design to understand whether similar glycoforms arise across different protein expression platforms from different sources to ensure the antigenic surface of the S protein remains consistent when it is used as an immunogen or in serological assays.[43−46] This is particularly important as glycan processing can be impacted by adverse protein conformations.[47] Here, we describe an integrated approach that includes liquid chromatography-mass spectrometry (LC-MS) and MD simulations to understand the glycosylation features across recombinant S proteins from different sources. We then explore the extent to which these recombinant proteins reproduce the glycosylation of virally derived SARS-CoV-2 S protein produced from cultured Vero cells that was obtained from a previous study.[32] We have employed an identical analytical approach to determine the glycan composition at each site, illustrating conservation across recombinant protein and virion-derived material. Furthermore, we have analyzed the glycosylation of monomeric RBD recombinant protein, as it has previously been explored as a subunit vaccine and a candidate for serological testing.[48,49] We then compared the site-specific glycan analysis of the two sites located in the RBD of SARS-CoV-2, N331 and N343, between the monomer and trimer and revealed a broad consensus in glycan processing, with some modest change in the processing of complex-type glycans. We also performed a comparative analysis on the MERS-CoV RBD. This contrasts with glycan processing of the SARS-CoV-2 RBD monomer when compared with S protein, with trimeric MERS-CoV RBD glycan sites presenting restricted glycan processing, likely due to conformational masking of these sites, either by proximal glycans or by nearby protein clashes, on trimeric MERS-CoV S protein.[50,51] We have further combined the site-specific glycan data of SARS-CoV-2 S protein with MD simulations. These simulations reveal distinct degrees of accessibility between different glycan sites across the protein that broadly correlate to their processing states. Taken together, our results reveal the conserved structural N-glycan sites in S protein when compared with native viral spike, which drives similarity in glycosylation among S proteins from disparate sources and manufacturing methods. Understanding S protein glycosylation will aid in the analysis of the vaccine and serological work of the global COVID-19 response.

## ■ MATERIALS AND METHODS

**SARS-CoV-2 S Protein Purification (Harvard).** To express a stabilized ectodomain of spike protein, a synthetic gene encoding residues 1−1208 of SARS-CoV-2 Spike with the furin cleavage site (residues 682−685) replaced by a "GGSG" sequence, proline substitutions at residues 986 and 987, and a foldon trimerization motif followed by a C-terminal six-His tag was created and cloned into mammalian expression vector pCMV-IRES-puro (Codex BioSolutions, Inc., Gaithersburg, MD). The expression construct was transiently transfected in HEK 293T cells using polyethylenimine (Polysciences, Inc., Warrington, PA). Protein was purified from cell supernatants using Ni-NTA resin (Qiagen), and the eluted fractions containing S protein were pooled, concentrated, and further purified by gel filtration chromatography on a Superose 6 column (GE Healthcare).

**SARS-CoV-2 S Protein Purification (Oxford).** The extracellular domain of SARS-CoV-2 spike protein was cloned into the pHLsec vector[52] and encompassed residues M1−

Q1208 (UniProtKB P0DTC2) with mutations R682G/R683S/R685S (furin recognition sequence) and K986P/V987P, followed by the Fibritin trimerization region, an HRV3C protease cleavage site, an eight-His tag, and a Twin-Strep tag at the C-terminus.[28,53] The spike ectodomain was expressed by transient transfection of HEK293T (ATCC, CRL-3216) cells for 6 days at 30 °C. The conditioned medium was dialyzed against 2× phosphate-buffered saline buffer (pH 7.4) (Sigma-Aldrich). The spike ectodomain was purified using Strep-Tactin Superflow resin (IBA Lifesciences) followed by size exclusion chromatography using a Superose 6 Increase column (GE Healthcare) equilibrated in 200 mM NaCl, 2 mM Tris-HCl (pH 8.0), and 0.02% $NaN_3$ at 21 °C.

**SARS-CoV-2 S Protein Purification (Amsterdam).** The SARS-CoV-2 S protein construct was designed and expressed as described previously.[1] Briefly, a SARS-CoV-2 S gene encoding residues 1−1138 (WuhanHu-1; GenBank entry MN908947.3) was ordered (Genscript) and cloned into a pPPI4 plasmid containing a T4 trimerization domain followed by a hexahistidine tag by *PstI*-*Bam*HI digestion and ligation. Modifications consist of substituting amino acid positions 986 and 987 with prolines and removing the furin cleavage site by substituting amino acids 682−685 with glycines. S protein was expressed in HEK 293F cells. Approximately 1.0 million cells/mL, maintained in Freestyle medium (Gibco), were transfected by adding a mix of PEImax (1 $\mu$g/$\mu$L) and SARS-CoV-2 S plasmid (312.5 $\mu$g/L) in a 3:1 ratio in OptiMEM. After 6 days, supernatants were centrifuged for 30 min at 4000 rpm and filtered using a Steritop filter (Merck Millipore, 0.22 $\mu$m). S protein was purified from the supernatant by affinity purification using Ni-NTA agarose beads, and eluates were concentrated and buffer exchanged with PBS using Vivaspin filters with a 100 kDa molecular weight cutoff (GE Healthcare). The S protein concentration was determined by the Nanodrop method using the proteins' peptidic molecular weight and extinction coefficient as determined by the online ExPASy software (ProtParam).

**SARS-CoV-2 Spike Trimer Protein Production and Purification from CHO Cells.** For CHO-based production of a trimeric spike protein, the Spike_ΔCter_ΔFurin_2P_-T4_His construct was used.[54] Briefly, on the basis of the CHO vector pXLG-6 (ExcellGene SA), containing a puromycin resistance marker and optimized expression elements, the variant spike sequence was inserted, containing a scambled furin cleavage site sequence, a two-proline sequence for stabilizing the protein in the prefusion form, and a 3′ T4 trimerization DNA, followed by a hexahistidine tag sequence. From chemical transfection (CHO4Tx, ExcellGene) to clonal selection and isolation of protein-producing cell lines to subsequent expansion of cells and production, all steps were performed under a suspension culture in a chemically defined medium with strict adherence to regulatory recommendations.

Production was performed in orbitally shaken containers (TubeSpin bioreactor 600, TPP, Trasadingen, Switzerland) or in 5 L Erlenmeyer-type shaker flasks in a Kuhner Shaker ISF4-X incubator (Adolf Kühner AG, Birsfelden, Switzerland), set to 37 °C, 5% $CO_2$, and humidity control and shaking at 150 rpm with a displacement radius of 50 mm. A fed-batch process was implemented until the supernatants were harvested on day 10. The viability of the cell culture, reaching ∼25 × 10^6 cells/mL, remained in the range of 90−100% until harvest. Production culture fluids were subjected to purification by affinity chromatography after depth filtration to remove cells. Loading

onto, washing of, and elution from a Ni-Sepharose column (Cytiva) were optimized, following the suggestions of the resin producers. The eluted product stream was loaded on a size exclusion column (SEC, Superdex 200 pg, Cytiva) for further purification, following a concentration step through tangential flow filtration.

**Plasmid Construction and Expression of Receptor Binding Domain Proteins.** To express the receptor binding domains (RBDs) of SARS-CoV-2 (residues 320−527), SARS-CoV-1 (residues 307−513), and MERS-CoV (residues 368−587), we amplified the DNA fragments by polymerase chain reaction using the SARS-CoV-2, SARS-CoV-1, and MERS-CoV S protein ectodomain plasmids as templates. The DNA fragments were cloned into the phCMV3 (Genlantis) vector in frame with the tissue plasminogen activator (TPA) leader sequence. To facilitate the purification and biotinylation of the protein, we fused the protein with a C-terminal six-His tag and AviTag spaced by GS linkers.

The RBD constructs were expressed by transient transfection of FreeStyle293F cells (Thermo Fisher). For a 1 L transfection, we added 350 $\mu$g of the RBD-encoding plasmids to 16 mL of Transfectagro (Corning) and 1.5 mL of 40K PEI (1 mg/mL) in 16 mL of Transfectagro in separate 50 mL conical tubes. The media containing the plasmids were filtered with a 0.22 $\mu$m Steriflip Disposable Vacuum Filter Unit (Millipore Sigma) and gently mixed with PEI. The plasmid/PEI mixture was incubated for 30 min and gently added to 1 L of cells at a concentration of 1 million cells/mL. The culture supernatants were harvested after 4 days and loaded onto a HisPur Ni-NTA resin column (Thermo Fisher). Following a wash with 25 mM imidazole, the protein from the column was eluted with 250 mM imidazole. The eluate was buffer exchanged with PBS and concentrated using 10K Amicon tubes. The RBD proteins were further purified via size exclusion chromatography using a Superdex 200 Increase 10/300 GL column (GE Healthcare).

**Sample Preparation and Analysis by LC-MS.** Mass spectrometry data for the virion-derived glycosylation were obtained using the protocols described by Yao et al.[32] and reanalyzed. For recombinant S protein and the RBD, 30 $\mu$g aliquots of SARS-CoV-2 S proteins were denatured for 1 h in 50 mM Tris-HCl (pH 8.0) containing 6 M urea and 5 mM dithiothreitol (DTT). Next, the S proteins were reduced and alkylated by adding 20 mM IAA and incubated for 1 h in the dark, followed by incubation with DTT to eliminate any residual IAA. The alkylated S proteins were buffer-exchanged into 50 mM Tris-HCl (pH 8.0) using Vivaspin columns (3 kDa) and digested separately overnight using trypsin, chymotrypsin (mass spectrometry grade, Promega), or α-lytic protease (Sigma-Aldrich) at a ratio of 1:30 (w/w). The peptides were dried and extracted using a C18 Zip-tip (Merck Millipore). The peptides were dried again, resuspended in 0.1% formic acid, and analyzed by nanoLC-ESI MS with an Easy-nLC 1200 (Thermo Fisher Scientific) system coupled to an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific) using higher-energy collision-induced dissociation (HCD) fragmentation. Peptides were separated using an EasySpray PepMap RSLC C18 column (75 $\mu$m × 75 cm). A trapping column [PepMap 100 C18 3 $\mu$m (particle size), 75 $\mu$m × 2 cm] was used in line with the liquid chromatograph prior to separation with the analytical column. The LC conditions were as follows: 275 min linear gradient consisting of 0% to 32% acetonitrile in 0.1% formic acid over 240 min

followed by 35 min of 80% acetonitrile in 0.1% formic acid. The flow rate was set to 300 nL/min. The spray voltage was set to 2.7 kV, and the temperature of the heated capillary was set to 40 °C. The ion transfer tube temperature was set to 275 °C. The scan range was $m/z$ 400−1600. The HCD collision energy was set to 50%. Precursor detection and fragment detection were performed using Orbitrap at a resolution MS1 of 100000 and a resolution MS2 of 30000. The AGC target for an MS1 of $4 \times 10^5$ and an MS2 of $5 \times 10^4$ and injection times of 50 ms for MS1 and 54 ms of MS2.

**Data Processing of LC-MS Data.** Glycopeptide fragmentation data were extracted from the raw file Byos (version 3.5, Protein Metrics Inc.). The following parameters were used for data searches in Byonic. The precursor mass tolerance was set at 4 and 10 ppm for fragments. Peptide modifications included in the search include Cys carbamidomethyl, Met oxidation, Glu → pyroGlu, Gln → pyroGln, and N deamidation. For each protease digest, a separate search node was used with digestion parameters that are appropriate for each protease (trypsin RK, chymotrypsin YFW, and ALP TASV) using a semispecific search with two missed cleavages. A 1% false discovery rate (FDR) was applied. All three digests were combined into a single file for downstream analysis. All charge states for a single glycopeptide were summed. The glycopeptide fragmentation data were evaluated manually for each glycopeptide; the peptide was scored as true positive when the correct b and y fragment ions were observed along with oxonium ions corresponding to the glycan identified. The protein metrics 309 mammalian N-glycan library was modified to include sulfated glycans and phosphorylated mannose species, although no phosphorylated mannose glycans were detected on any of the samples analyzed. The relative amounts (determined by comparing the XIC of each glycopeptide, summing charge states) of each glycan at each site, as well as the unoccupied proportion, were determined by comparing the extracted chromatographic areas for different glycotypes with that of an identical peptide sequence. Glycans were categorized according to the composition detected. HexNAc(2), Hex(9−3) was classified as M9 to M3. Any of these compositions that were detected with a fucose are classified as FM. HexNAc(3)-Hex(5−6)NeuAc(0−1) was classified as hybrid with HexNAc(3)Hex(5−6)Fuc(1)NeuAc(0−1) classified as Fhybrid. Complex-type glycans were classified according to the number of processed antenna and fucosylation state. Complex glycans are categorized as HexNAc(3)(X), HexNAc(3)(F)(X), HexNAc(4)(X), HexNAc(4)(F)(X), HexNAc(5)(X), HexNAc(5)(F)(X), HexNAc(6+)(X), and HexNAc(6+)(F)-(X). Core glycans are any glycan smaller than HexNAc(2)-Hex(3). Glycans were further classified according to their processing state. A glycan was classified as fucosylated or sialylated if it contained at least one fucose or sialic acid residue. Agalactosylated glycans consisted of compositions in the categories mentioned above for complex-type glycans that possessed only three hexose residues for complex-type glycans, or if the composition was a hybrid-type glycan, then this was increased to five hexoses. If the complex-type glycan did not contain sialic acid but more than three hexoses for complex-type glycans or five hexoses for hybrid-type glycans, then it was counted as galactosylated. Sulfated glycans were included in the categories described above, so a glycan can be in both galactosylated and sulfated categories, for example.

**Integrative Modeling and Molecular Dynamics Simulation.** The model of S protein was built using Modeler
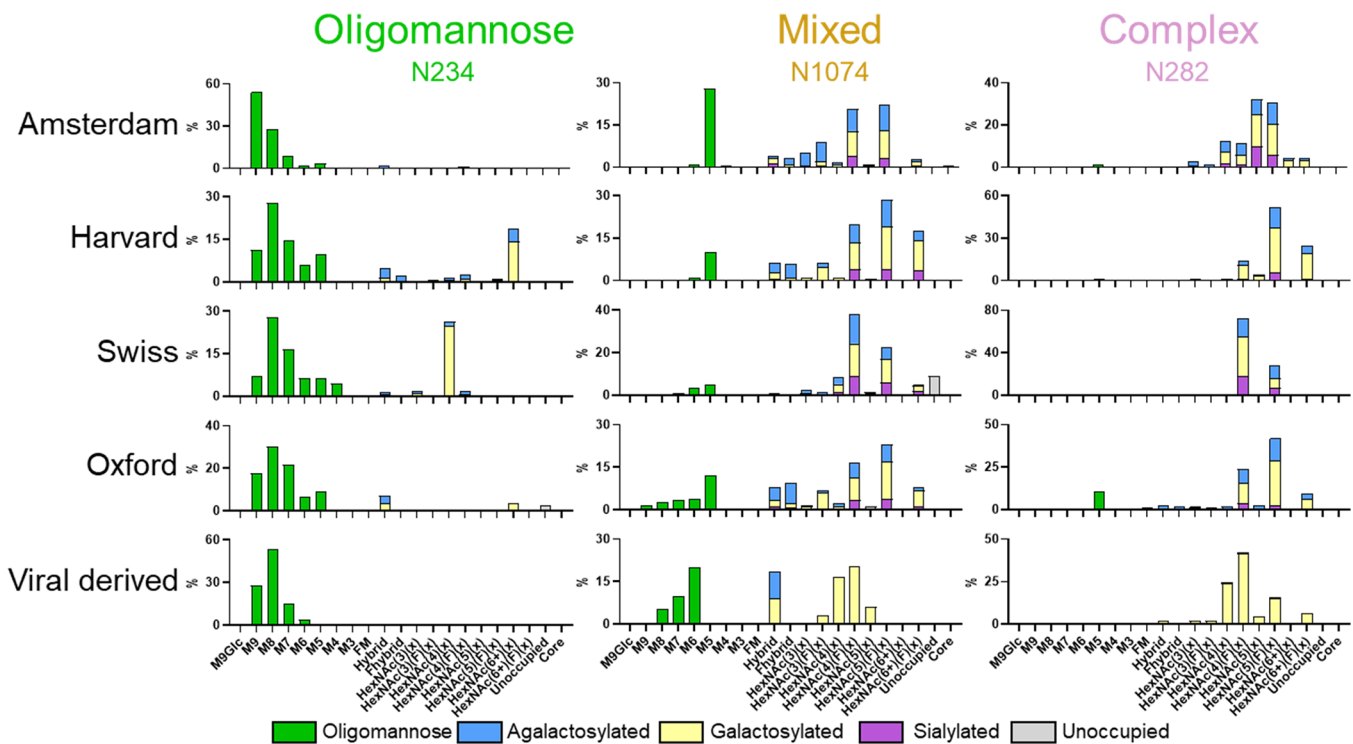
9.21[55] with three structural templates: (i) the cryo-EM structure of SARS-CoV-2 S ECD in the open state [Protein Data Bank (PDB) entry 6VSB],[28] (ii) the NMR structure of the SAR-CoV S HR2 domain (PDB entry 2FXP),[56] and (iii) the NMR structure of the HIV-1 gp-41 TM domain (PDB entry 5JYN).[57] Missing loops in the N-terminal domain and the C-terminal domain of the ECD were modeled using the cryo-EM structure of the S ECD in the closed state, which was resolved at a higher resolution (PDB entry 6XR8).[58] A total of 10 models were built, and three top models (based on the lowest discrete optimized protein energy score[59]) were subjected to further stereochemical assessment using Ramachandran analysis.[60] The model with the smallest number of outliers was then selected. Man-9 glycans were added to all 22 glycosylation sites using the CHARMM-GUI Glycan Reader and Modeler web server.[61] Three cysteine residues on the C-terminal domain (C1236, C1240, and C1243) were palmitoylated as they have been shown to be important in cell fusion.[62] The glycosylated S protein model was then embedded into a model ER/Golgi intermediate complex (ERGIC) membrane built using CHARMM-GUI Membrane Builder,[63] which contained 47% phosphatidylcholine (PC), 20% phosphatidylethanolamine (PE), 11% phosphatidylinositol phosphate (PIP), 7% phosphatidylserine (PS), and 15% cholesterol.[64−66]

The system was parametrized using the CHARMM36 force field.[67] TIP3P water molecules were used to solvate the system, and 0.15 M NaCl salt was added to neutralize it. The system was then subjected to stepwise energy minimization and equilibration with decreasing positional and dihedral restraints, following the standard CHARMM-GUI protocol.[68] A 200 ns production simulation was conducted with the temperature maintained at 310 K using the Nosé−Hoover thermostat.[69,70] The pressure was maintained at 1 atm using a semi-isotropic coupling to the Parrinello−Rahman barostat.[71] Electrostatics were calculated using the smooth particle mesh Ewald method[72] with a real-space cutoff of 1.2 nm, while the van der Waals interactions were truncated at 1.2 nm with a force switch smoothing function applied between 1.0 and 1.2 nm. The LINCS algorithm was utilized to constrain all covalent bonds involving hydrogen atoms,[73] and a 2 fs integration time step was employed. All simulations were performed using GROMACS 2018[74] and visualized in VMD.[75] Accessible surface area (ASA) and glycan contact analyses were performed using the built-in GROMACS tools *gmx sasa* and *gmx select*.

## ■ RESULTS

**Expression and Purification of Recombinant S Protein from Multiple Sources.** To define the variability in trimeric recombinant S protein glycosylation and compare recombinant and virally derived S protein, we obtained preparations of recombinant S protein from a range of laboratories. These include the Amsterdam Medical Centre (Amsterdam), Harvard Medical School (Harvard), Switzerland, The Wellcome Centre for Human Genetics (Oxford), and Biological Sciences of The University of Texas at Austin (Southampton/Texas).[1,28,53,54,76,77] While there are minor differences between the constructs used to produce the S protein, the overall design is similar. This involves a truncation of the S protein prior to the C-terminus, replacement of the furin cleavage site between S1 and S2 with a "GSAS" or otherwise mutated linker, a C-terminal T4 fibritin trimerization motif, and stabilization of the prefusion S protein conforma-

**Figure 1.** Site-specific glycosylation of recombinant and virally derived S protein from multiple laboratories. (A) Site-specific glycan analysis of recombinant S protein expressed and purified at different locations. The bar charts represent the relative proportions of glycoforms present at each site, including the proportion of PNGs that were not modified by an N-linked glycan. The proportions of oligomannose- and hybrid-type glycans are colored green. Processed complex-type glycans are colored pink, and the proportions of unoccupied sites are colored gray. The institution that provided the S protein for analysis is listed above each chart. The Texas/Southampton data are reproduced from Chawla et al. (unpublished data). The average compositions of recombinant S protein were calculated using all samples. Bars represent the mean ± the standard error of the mean of all recombinant samples analyzed. (B) Virally derived site-specific analysis was performed using data acquired by Yao et al. and categorized in the same manner as described above.[32] Data for sites N657 and N1158 could not be obtained and are not represented. (C) Full length model displaying the site-specific oligomannose glycosylation of virally derived S protein. A description of how this model was generated can be found in Materials and Methods. Both protein and glycans are shown in surface representation; the former is colored gray, and the latter colored on the basis of the oligomannose content as shown in Table S1 (green for 80−100%, orange for 30−79%, and pink for 0−29%).

tions using proline substitutions.[28,50,78] The recombinant proteins analyzed contain 22 N-linked glycan sites with the exception of the "Amsterdam" preparation that was truncated prior to the last three potential N-linked glycosylation sites. An alignment of the recombinant proteins was performed, and glycan sites are numbered according to the numbering used in previous reports of the site-specific glycosylation of the 2P-stabilized S protein (Figure S1).[7] All recombinant proteins were expressed in human embryonic kidney 293 (HEK293) cells with the Amsterdam, Harvard, and Southampton/Texas S protein preparations produced in HEK293F cells and the Oxford S protein preparation produced in HEK293T cells. The "Swiss" S protein was expressed in Chinese hamster ovary (CHO) cells. Following expression and purification, an identical approach was utilized to prepare the recombinant samples for analysis by LC-MS, involving three separate protease digests (trypsin, chymotrypsin, and α-lytic protease). After LC-MS analysis, the site-specific glycosylation of these recombinant samples was compared to that of previously published virally derived S protein that we searched using the same analytical parameters that were used for the recombinant proteins.[32]

**Conservation of Underprocessed Glycans on Trimeric Recombinant and Virally Derived S Protein.** Underprocessed oligomannose-type and hybrid-type glycans are common on viral glycoproteins and commonly arise due to either glycan- or protein-mediated steric clashes with ER and Golgi resident mannosidase enzymes, terminating the glycan processing pathway. As the presence of these glycans is linked to the quaternary structure of the protein, changes in the abundance of these glycoforms can indicate changes in the fine structure of the glycoprotein. To investigate the abundance of these glycans, we simplified the heterogeneous glycan compositions detected by LC-MS into three categories: (1) oligomannose-type and hybrid-type glycans, (2) complex-type glycans, and (3) the proportion of PNGs lacking an N-linked glycan. Overall, the recombinant samples recapitulated the glycan processing observed on the virally derived S protein (Figure 1A,B). Sites on the virally derived material with an abundance of underprocessed glycans of >30% include N61, N122, N234, N603, N709, N717, N801, and N1074 (Figure 1B). With a few exceptions, all recombinant samples analyzed also present at least 30% oligomannose-type glycans at these sites (Table S1).

**Figure 2.** Comparison of detailed compositions between sites with differential processing states. Glycan compositions at N234, N1074, and N282. For all samples analyzed, glycans were categorized and colored according to the detected compositions. Oligomannose-type glycans (M9-M4) are colored green. Hybrid-type glycans, those containing three HexNAcs and at least five hexoses, were colored as for complex-type glycans because one arm can be processed in a similar manner. Complex-type glycans were categorized according to the number of HexNAc residues detected and the presence or absence of fucose. Core glycans represent any detected composition smaller than $HexNAc_2Hex_3$. For hybrid and complex-type glycans, bars are colored to represent the terminal processing present. Blue represents agalactosylated, yellow galactosylated (containing at least one galactose), and purple sialylated (containing at least one sialic acid). The proportion of unoccupied PNGSs is colored gray.

To compare the variability of oligomannose- and hybrid-type glycosylation across all recombinant proteins, the compositions were averaged and displayed with the standard error of the mean (SEM) of S protein preparations from different laboratories (Figure 1A,B). This analysis revealed a broad consensus of glycan processing regarding high-mannose glycans, with localized variations occurring. One example of remarkable homogeneity is the N234 site, which presents high levels of underprocessed glycans in all samples analyzed. This processing was conserved on the virally derived S protein. The recombinant material possessed low levels of oligomannose-type glycans at several sites across the protein not present on virally derived material. For example, material analyzed from Oxford had at least 5% oligomannose-type glycans at every site except N343 (Figure 1A). This global moderate increase in the level of oligomannose-type glycans could potentially arise from two sources; the first is that the recombinant preparations utilize stabilizing mutations that might generate a more compact structure that limits the ability of glycan processing enzymes to act, and this is supported by the observation that virally derived S protein exhibits metastable conformations in the NTD, likely increasing accessibility to particular glycan sites.[32] The second could be that the higher yield of protein obtained during recombinant protein expression places limitations on the capacity of the cell to process the large number of glycan sites present on the S protein.

Glycan occupancy is another important parameter to consider for immunogen design. The underoccupancy of recombinant glycoprotein PNGSs compared to that of their viral counterparts has been shown to result in the presentation

of non-neutralizing distracting epitopes and is important to monitor.[79,80] As is the case with virally derived HIV-1 Env, virally derived SARS-CoV-2 S protein displayed high levels of glycan occupancy at all sites analyzed (Figure 1B). Upon comparison of these data to the consensus site-specific glycosylation data, a majority of PNGSs on recombinant S protein are highly occupied. The exceptions to this are at the C-terminus, where several samples analyzed display reduced glycan occupancy denoted by a larger population of unoccupied sites (Figure 1A). These sites displaying larger proportions of unoccupied glycans at the C-terminus are likely due to the truncation of the recombinant proteins required for solubilization; it has previously been reported that the proximity of a glycan sequon to the C-terminus can influence its occupancy.[81]

Modeling the site-specific glycosylation of virally derived material enables the glycosylation of S protein to be contextualized spatially in three dimensions. Using the cryo-EM structure of the S protein ECD in the open state (one RBD in the "up conformation" and two RBDs in the "down conformation") (PDB entry 6VSB),[28] as well as the NMR structures of the SARS-CoV HR2 domain (PDB entry 2FXP)[56] and the HIV-1 gp-41 TM domain (PDB entry 5JYN),[57] a complete model of the S protein was generated (details in Materials and Methods). This model includes glycosylation sites that are often not resolved as they are present on variable loops or along the flexible stem region of the S protein. This model demonstrated the spread of different glycan processing states across the S protein (Figure 1C). As with other published models of spike glycosylation, the N234

site is extensively buried within the protein surface proximal to the RBD and rationalizes the conserved presentation of oligomannose-type glycans on recombinant and virally derived material as the glycan processing enzymes are unable to completely process these sites. Glycan sites located on variable loops and on the exposed stem of the S protein are much more processed as these sites are more exposed to glycan processing enzymes.

**Divergent Glycan Processing at Sites Presenting Complex-Type Glycans.** While glycan underprocessing can provide information pertaining to the quaternary structure of the glycoprotein, a majority of glycans across the S protein can be processed beyond these glycoforms. Glycan sites not under the same structural pressures, such as at restricted sites like N234, will be processed in a manner analogous to those presented on the majority of host glycoproteins. By comparing the site-specific glycan compositions at example sites, we sought to understand the variability in glycan elaboration across multiple preparations of S protein. For this analysis, we compared the recombinant protein glycosylation data presented in this study (Amsterdam, Harvard, Switzerland, and Oxford) with data for virally derived S protein previously reported but analyzed them using the same search parameters that were used for the recombinant proteins. To aid comparison, we selected one site that presented high levels of underprocessed oligomannose-type glycans (N234), one that presented a mixture of processing states (N1074), and a site that was populated by complex-type glycans (N282) (Figure 2). Comparing the site-specific glycosylation of N234 demonstrated the conservation of oligomannose-type glycan processing. On four of the five samples presented in Figure 2, the most abundant glycan detected was $Man_8GlcNAc_2$, and on the Amsterdam S protein, $Man_9GlcNAc_2$ predominated, which is marginally less processed. The Harvard and Swiss samples displayed elevated levels of glycan processing at N234, but these glycans were only ~20% abundant and were still the least processed sites on these samples (Figure 2 and Table S1).

The N1074 site presented a diverse mixture of oligomannose-, hybrid-, and complex-type glycans across all samples analyzed. The abundance of oligomannose-type glycans was more variable at this site with the total oligomannose-type glycan content varying from 11% to 35% (Figure 2). The predominant composition at N1074 varied from sample to sample. For the Amsterdam recombinant preparation and the virally derived S protein, an oligomannose-type glycan was the most abundant glycan at N1074; however, for the other samples, a fully processed glycan was more abundant. For all samples analyzed, N1074 presented diverse glycan processing states. The elaboration of complex-type glycans with different monosaccharides can influence the function of the glycoprotein to which they are attached; for example, sialylation can extend the half-life of a glycoprotein in the body.[82] As they progress through the Golgi apparatus, glycans can be elaborated by the addition of fucose, galactose, and sialic acid. The abundance of these monosaccharides is influenced by the cell from which they are produced and by the culture conditions or medium compositions under which recombinant proteins are manufactured. Typically, glycoproteins produced from HEK293F and CHO cells present complex-type glycans with high levels of fucosylation and galactosylation but lower levels of sialylation.[83,84] The processing of N1074 and N282 demonstrates this with the majority of complex-type glycans bearing at least one fucose on all recombinant S preparations.

Likewise, a majority of glycans are galactosylated, although populations of glycans that lack any elaboration beyond N-acetylglucosamine branching are present. Hybrid-type glycans, where one arm of the glycan is processed and elaborated as for complex-type glycans and one remains underprocessed, are present in lower abundances across N1074 in recombinant preparations. These hybrid-type glycans are also generally of low abundance on mammalian glycoproteins.[85]

While the complex- and hybrid-type glycan compositions are variable, there are some visible trends when comparing glycan processing between recombinant and virally derived S protein. The starkest difference is the lack of sialic acid residues across not only N1074 and N282 but also all PNGSs of virally derived material (Figure 2 and Table S1). Upon averaging all of the recombinantly produced S proteins and comparing them to virally derived S proteins, we find a 21 percentage point (p.p.) decrease in sialylation (Table S2). The fucosylation of virally derived material is also lower across both N1074 and N282, and this trend is again mirrored across all sites, with virally derived S protein possessing a 16 p.p. decrease in fucosylation (Table S2). The final difference is in glycan branching. The number of processed glycan antennae can be inferred from the detected number of N-acetylhexosamine (HexNAc) residues determined by LC-MS. The more HexNAc present on a glycan composition, the more branched that glycan tends to be; for example, HexNAc(6) can correspond to a tetra-antennary glycan, whereas HexNAc(4) is biantennary. As glycan processing is heterogeneous, this change is more subtle. However, the complex-type glycans of N1074 and N282 are less branched on viral S protein than on the recombinant S proteins, with an ~30 p.p. decrease in the abundance of HexNAc(5) and HexNAc(6) corresponding to tri- and tetra-antennary glycans (Figure 2). These changes are also apparent on other processed sites such as N165 and N1158 (Table S2).

These changes contrast with similar analyses comparing virally derived Env and soluble recombinant variants. For HIV-1 Env, an increase in glycan branching and sialylation is observed across sites presenting complex-type glycans.[86] These changes were also present on recombinant full length Env.[87] Several factors could be influencing these changes. The first is that the membrane tether afforded to the viral Env brings the glycans into the proximity of the glycan processing enzymes, and the second is that the producer cell had greater expression levels of the glycosyltransferase enzymes involved in glycan processing. The SARS-CoV-2 S protein in a viral context is likely under similar constraints; however, the glycan processing is distinct. One factor that could be important is the early budding of the SARS-CoV-2 virion into the ER/Golgi intermediate complex (ERGIC) that may distance the spike protein from glycosyltransferases present in the trans-Golgi compared to Env, which remains attached to the membrane, proximal to glycosyltransferases.[88] The choice of the producer cell and culture condition of the virus, in this case Vero cells, which are derived from the kidney cells of *Cercopithecus aethiops* (African green monkey), may also influence the glycan processing as the expression levels of glycosyltransferase enzymes may account for the weakened attachment of sialic acid and fucose observed on viral S protein compared to recombinant S protein. Previous studies have highlighted the low abundance of sialic acid moieties on Vero cells;[89] however, these changes could also occur due to the culture conditions used for viral production. As these changes in glycosylation likely are not impacting the immunogenicity of the viral spike
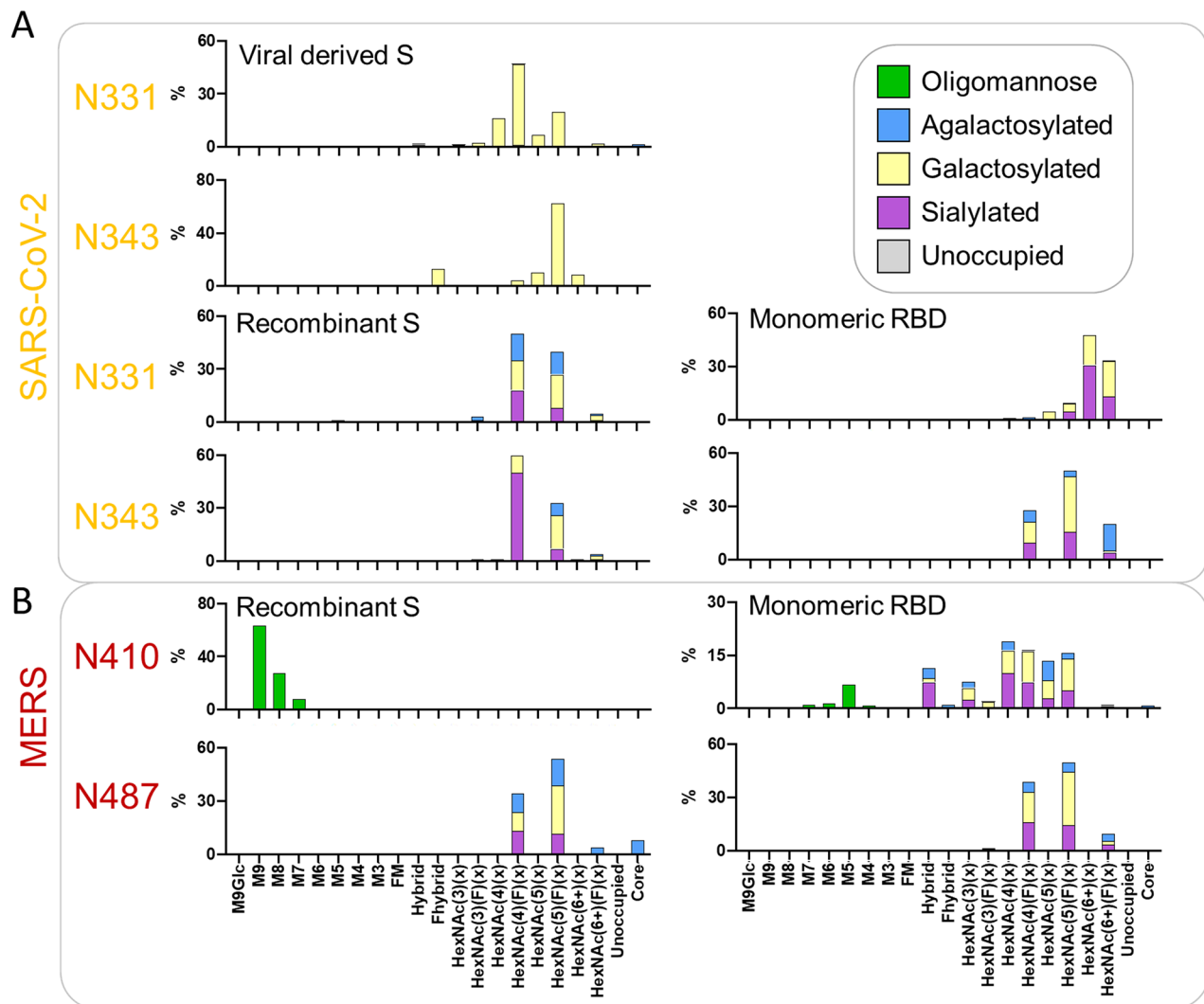
**Figure 3.** Comparative analysis of the glycosylation of the two PNGSs on the RBD for virally derived S protein, recombinant S protein, and the monomeric RBD. (A) Detailed site-specific glycan compositions of the two sites located in the RBD of SARS-CoV-2. Recombinant S protein data are reproduced from Chawla et al. (unpublished data), and data for the viral S protein RBD sites were obtained from ref 32. Site-specific glycan data are presented as outlined in Figure 2. (B) Site-specific compositions for N-glycan sites located in the RBD of MERS-CoV when expressed as part of a soluble recombinant S protein compared to the RBD only. Data for the MERS-CoV S protein were obtained from ref 51. Site-specific glycan data are presented as outlined in Figure 2.

mimetics, as evidenced by the high efficacy of several vaccines, these observations remain important when considering how the virus may be interacting with the immune system via lectin interactions and may be informative when considering antigenic tests and purifications using glycan binding reagents. Despite the observed differences, we note that the site-specific glycosylation of the virally derived material outlined here is consistent with previously reported glycan analysis from virally derived S protein produced in Calu-3 lung epithelial cells.[90]

**The Expression of Monomeric RBD Constructs Impacts Glycan Processing.** Our analysis demonstrates that S protein glycosylation is influenced by quaternary protein architecture, and other factors. We next sought to compare the glycosylation of the soluble recombinant monomeric RBD with that of recombinant and virally derived trimeric S protein. We expressed and purified the recombinant RBD and compared the site-specific glycosylation of the two glycan sites located in the RBD, N331 and N343, with those observed on recombinant S protein and virally derived S protein reported

previously.[7,32] Overall, the N331 and N343 glycans across all expression formats were highly processed, with little to no oligomannose-type glycans detected (Figure 3A,B and Table S3). For the RBD sites presented on virally derived S protein, the complex-type glycans observed were similar to those observed at N282 with most of the PNGSs occupied by bi- and triantennary glycans [HexNAc(4) and HexNAc(5)] (Figure 3A). A majority of these glycans were fucosylated; however, a large proportion of the complex-type glycans on virally derived S lacked fucosylation (24% N331 and 20% N343). As with other sites on the virally derived S protein, minimal levels of sialylation were observed and a majority of glycans possessed at least one galactose residue.

In comparison, the two RBD sites of recombinant soluble S protein are highly fucosylated, with close to 100% of the glycans at N331 and N343 containing at least one fucose (Figure 3B). As with the virally derived S protein, a majority of the glycans are bi- and triantennary complex-type glycans. The recombinant S protein RBD sites are also more sialylated than
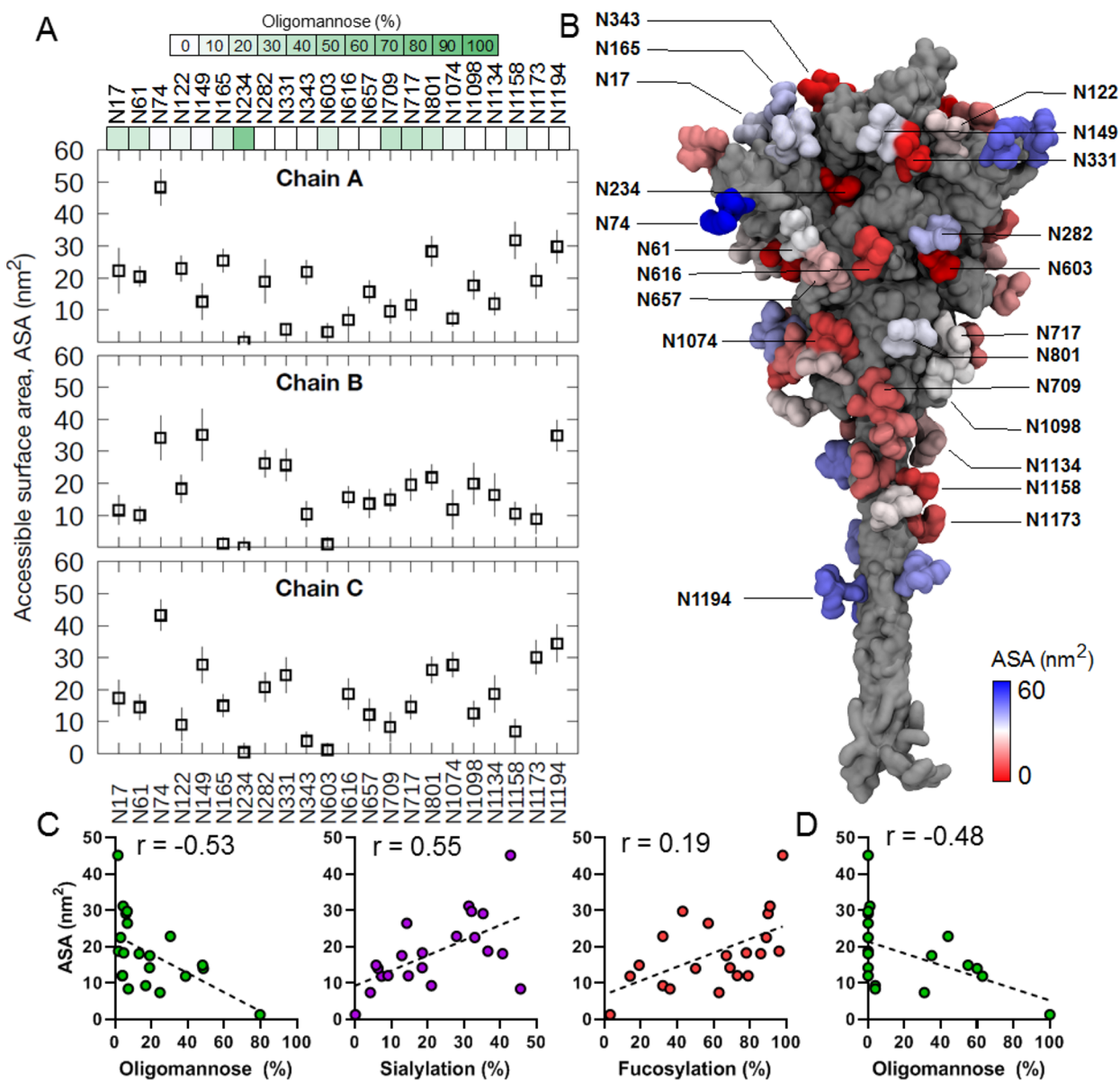
**Figure 4.** Accessible surface area (ASA) of oligomannose-type glycans from MD simulations. (A) The ASA values were calculated for each oligomannose-type (M9) glycan for all three S protein chains. In this model, chain A is modeled in the RBD "up" conformation. The last 50 ns of the simulation was used for calculation, and error bars indicate the standard deviation along the trajectory. The probe size used was 1.5 nm. The green color bars represent the oligomannose content of respective sites, calculated as the average oligomannose-type glycan content of the recombinant S proteins analyzed in this work (Amsterdam, Harvard, Swiss, and Oxford S proteins). (B) Structure of the S protein (gray) with glycans colored on the basis of their ASA values shown in surface representation. (C) Graphs comparing glycan processing to calculated ASA values, averaged for chains A−C, for the recombinant protein average for the oligomannose-type glycan content, the proportion of glycans containing at least one sialic acid, and the proportion of glycans containing at least one fucose. The reported $r$ values represent the Spearman's rank correlation coefficient. (D) Oligomannose-type glycan content vs ASA for virally derived SARS-CoV-2 S protein presented as in panel C.

N331 and N343 on virally derived S protein, with 28% of N331 glycans containing at least one sialic acid and 60% of those on N343 (Table S3). Interestingly, when the RBD is expressed as a monomer, there are additional subtle changes compared to virally derived and recombinant S protein. The most prominent change is in the glycan branching; whereas recombinant and virally derived trimeric S protein RBD glycan sites possess small quantities of tetra-antennary glycans, on the monomeric RBD approximately one-third of the glycans at N331 and one-fifth of the glycans at N343 consist of these larger branched structures (Figure 3). When compared to recombinant S protein, the monomeric RBD sites also possess

lower levels of biantennary glycans. Despite these changes, both recombinant trimeric S protein and monomeric RBD sites have high levels of fucosylation. These results suggest that complex-type glycans are under differential control hierarchies in which certain forms of glycan processing could be influenced by the structural presentation of glycan sites. The attachment of fucose, galactose, and sialic acid for these RBD sites appears to be controlled by more global phenomena, such as the producer cell, as the attachment of these monosaccharides is similar when comparing the monomeric RBD and trimeric recombinant S protein that were produced in identical cell lines. The branching of the RBD sites on

monomeric RBD is greater than that of trimeric S protein, both viral and recombinant, and suggests that the quaternary structure of the glycoprotein may play a small role in the elaboration of complex-type glycans of the RBD. These results are similar to those of previous analyses that have compared trimeric S protein with monomeric S1.[38]

To further explore differences in the glycosylation of RBD PNGSs, we performed a similar comparative analysis on MERS-CoV. The site-specific glycan analysis of recombinant MERS-CoV S protein has been reported previously.[51] For the comparative analysis of trimeric recombinant protein and the RBD, the MS files obtained in the previous study were searched using an identical version of the analysis software using the same glycan libraries as for the analysis displayed in Figures 1 and 2. This analysis revealed differences at the glycan sites present on the MERS RBD when expressed monomerically compared to recombinant trimeric soluble S protein. One of the sites, N487, is similar between the two platforms, presenting glycoforms typical of sites populated by complex-type glycans (Figure 3B). In contrast, N410 is occupied by exclusively oligomannose-type glycans when present on recombinant trimeric S protein. When the monomeric MERS RBD is expressed, the level of processing of the N410 site is markedly increased, with complex-type glycans dominating the glycan profile, although a subpopulation of oligomannose-type glycans remain (Figure 3B). Modeling the N410 glycan onto published structures of MERS S protein reveals that the N410 glycan is protected from processing by mannosidase enzymes such as ER α-mannosidase I when buried within the trimer, but when presented on the monomeric RBD, it is readily accessible to glycan processing enzymes (Figure S2). These observations further highlight how the quaternary structure of a glycoprotein is a key determinant of the glycan processing state and demonstrate how glycan sites can provide information about protein folding and quaternary structure.

**Molecular Dynamics Simulations Reveal the Relationship between Accessibility and Glycan Processing.** Using site-specific glycan analysis, it is possible to infer the structural restraints placed on the glycan processing of particular PNGSs by the quaternary structure of the protein. MD simulations can help to understand how protein flexibility can influence glycan processing. The recombinant proteins analyzed in this study all use proline substitutions to stabilize the prefusion S protein conformation. Structural analyses of S protein containing these proline substitutions have shown that one of the three subunits frequently displays its RBD in an up conformation.[28] By performing simulations using models containing one RBD up, one can investigate how a differential RBD presentation can impact site-specific glycosylation.

To this end, we performed 200 ns MD simulations of fully glycosylated trimeric S protein embedded within an ERGIC membrane model (details in Materials and Methods). The $Man_9GlcNAc_2$ glycan (Man-9), which represents the primary substrate for glycan processing enzymes, was added to each PNGS to understand the effect of protein and glycan dynamics on glycan processing. Glycan accessibility to enzymes was then elucidated by calculating the accessible surface area (ASA). To ensure the correct probe size was used for ASA calculation, we first looked at the structure of mannosidases and glycosyltransferases bound to its substrate or substrate analogue.[91,92] By measuring the distance between the substrate binding site and the outer surface of the enzymes, we found a probe with a 1.25−1.5 nm radius would best approximate the size of the

enzymes (Figure S3). We measured ASA values for chain A of S protein using probes with 1.25 and 1.5 nm radii and found they both gave very similar results. Plotting accessible points of a 1.5 nm probe around individual glycans also indicated that a 1.5 nm radius would be required for the Man-9 glycan to be accessed by mannosidases and glycosyltransferases. We therefore used a 1.5 nm probe to measure the ASA of each Man-9 glycan on the S protein model during the MD simulation.

Comparing the ASA of each glycan across each protomer reveals a diverse range of accessibility (Figure 4A,B). Certain sites such as N234 were determined to possess low ASA values (Figure S4A), indicating that this site is extremely buried, correlating with observations that N234 is the least processed on all recombinant and virally derived S protein. Conversely, the most exposed site across all three chains is N74 (Figure S4B). This glycan is highly processed when analyzed by LC-MS. Interestingly, previous analyses have shown that N74 possesses sulfated glycans.[15] We have included sulfated compositions in our glycan library to search the MS data and observe that N74 contains multiple sulfated glycan compositions on all samples produced in HEK293 cells (Table S1). The higher accessibility of this site may explain why sulfated glycans are observed at higher abundances at this site, but not others. Some glycans show a distinctly bimodal accessibility pattern, wherein high and low ASA values were observed along the trajectory (Figure S5). For example, two of the N122 glycans were buried within a crevice between the N-terminal domain of one chain and the RBD of an adjacent chain during a portion of the simulation. Similarly, one of the N603 glycans was inserted into a large interprotomeric groove near the S1/S2 cleavage site. This observation correlates with LC-MS data from both recombinant and virally derived S protein, showing that these two sites are populated by approximately half oligomannose/hybrid and half complex glycans, potentially due to the ability of these glycans to be either buried or exposed on the protein surface. The glycan on N165 also showed bimodal ASA values. However, this is due to the RBD up configuration, which allows N165 glycan to insert into the gap between the RBD and the N-terminal domain, which was not accessible in the RBD down configuration. N165, along with N234 glycans, has been shown to modulate the RBD's conformational dynamics by maintaining the up configuration necessary for ACE2 recognition.[31]

To understand the relationship between oligomannose-type glycan presentation and the accessible surface area of each glycan site, we used the average oligomannose-type glycan content of each site as determined in Figure 1, for both the recombinant and virally derived material, and compared it to the average ASA values for chains A−C, as LC-MS cannot distinguish which glycan came from which chain (Figure 4C,D). For both recombinant and virally derived S protein, there was a negative correlation between ASA and oligomannose content (Spearman's correlation coefficients of −0.53 for recombinant and −0.48 for virally derived). This demonstrates the link between glycan processing and quaternary architecture, as sites that are less accessible to the glycan processing enzymes present larger populations of immaturely processed oligomannose-type glycans. Conversely, the processing of complex-type glycans such as the addition of sialic acid correlated with the accessbile surface area at a site-specific level ($r = 0.55$) and also fucosylation, albeit to a lesser extent ($r = 0.19$). While there is an observed correlation

between oligomannose-type glycans and ASA, the nanosecond time scale captured during MD simulations is shorter than the overall time taken for glycan processing in the ER and Golgi apparatus to occur, and therefore, the S protein will be able to sample more conformations not resolved during simulation.

Finally, to determine the potential effect of glycan clustering on enzyme accessibility and subsequent glycan processing, we calculated the number of contacts made by each glycan at a given site with its surrounding glycans during the simulation (Figure S6). Interestingly, in all three chains, the sites on the S2 subunit show a higher number of glycan—glycan contacts, suggesting that they are most likely to form local clusters. An overlay of consecutive simulation snapshots of glycans also clearly shows that the S2 region is more densely packed with glycans compared to S1 and HR2 domains. This agrees with the LC-MS data showing that the sites on the S2 subunit, specifically, N709, N717, and N801, have a high degree of unprocessed oligomannose-type glycans. The high density of glycan—glycan contacts within the mannose glycan patch has previously been implicated in reduced processivity by glycan processing enzymes in HIV-1 gp120,[93] and a similar mechanism might explain the increased oligomannose content at S2 sites of SARS-CoV-2. A recent study showed that glycan binding Fab-dimerized antibodies recognize underprocessed oligomannose-type glycans on SARS-CoV-2 via a common epitope around N709 on the S2 subunit,[94] further suggesting the formation of a local glycan cluster. Collectively, our MD simulation sampling shows that glycan accessibility to host enzymes, which is affected by protein quaternary structure and local glycan density, is an important determinant of glycan processing.

## ◼ PERSPECTIVES

The global impact of COVID-19 has resulted in laboratories across the globe producing recombinant spike protein for vaccine design, antigenic testing, and structural character-ization. While the site of glycan attachment is encoded by the viral genome, the processing of the attached glycans can also be influenced by a wide range of exogenous phenomena, including recombinant host systems and processes for production in cell culture. Using immediately available materials, we have compared the glycosylation of Vero-produced virus preparations for spike proteins with those of HEK-293 and CHO-derived recombinant spike proteins. We are aware that these preparations are likely to be different from spike proteins as produced in a diversity of cells in infected patients. However, here we show that when a glycan site is located in regions of the SARS-CoV-2 S protein that are not readily accessible the glycan site will possess high levels of underprocessed oligomannose-type glycans, a phenomenon that will likely be of a general nature. This conclusion is derived from the observation that the underprocessing of glycans is found in protein preparations from a range of institutions across the globe and they are also present on Vero cell-derived S protein derived from infectious SARS-CoV-2. While the presence of oligomannose-type glycans is rare on the majority of host glycoproteins, their abundance on SARS-CoV-2 is lower than that of other viral glycoproteins and means that the glycan shield density of this virus is low compared to those of HIV-1 Env and Lassa GPC. This is likely to mean that the immunodominant protein epitopes remain exposed. The processing equivalence between recombinant and virally derived S protein indicates that recombinant S protein

glycosylation for vaccinations will likely mimic those occurring in human infections and will remain antigenically comparable. By analyzing complex-type glycan processing across multiple samples, we demonstrate that this glycosylation is driven more by other parameters, including the producer cell and culturing conditions. One observed exception to this was that the expression of monomeric RBD as opposed to trimeric S protein does increase the level of glycan branching on the two glycan sites located on the SARS-CoV-2 RBD. This suggests that the glycan processing of complex-type glycans, in addition to oligomannose-type glycans, may also be under structural constraints, albeit to a much lesser extent. Overall, these results demonstrate that while N-linked glycosylation is highly diverse at certain regions of the S protein, there is a broad consensus of glycan processing with regard to oligomannose-type glycans between virus and immunogen S protein. This is something that cannot be taken for granted, as upon comparison of recombinant and virally derived HIV-1 Env, the reduced glycan occupancy of immunogens can induce an immune response that is incapable of neutralizing the virus. The reproducibility of S protein glycosylation from many different sources is of significant benefit for immunogen design, serology testing, and drug discovery and will mean that the glycans of SARS-CoV-2 are unlikely to provide a barrier to combating the COVID-19 pandemic.

## ◼ ASSOCIATED CONTENT

### ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.1c00279.

> Average recombinant and virally derived S protein site-specific glycan compositions averaged (Table S1), percentage point change in glycosylation on virally derived S protein compared to recombinant S protein (Table S2), site-specific glycan compositions of the two RBD sites, N331 and N343 (Table S3), BLAST alignment of recombinant S proteins (Figure S1), model demonstrating the differential accessibility of the N410 glycan for glycan processing enzymes between trimeric MERS S protein and the monomeric MERS RBD (Figure S2), validation of probe size for ASA measurement (Figure S3), examples of glycans either buried or exposed in all three chains (Figure S4), examples of glycans with bimodal accessibility properties (Figure S5), and glycan—glycan contacts from MD simulations (Figure S6) (PDF)

### Accession Codes

Spike glycoprotein, UniProtKB P0DTC2 (SPIKE_SARS2); MERS CoV Spike glycoprotein, UniProtKB R9UQ53 (R9UQ53_MERS); raw file accession, MassIVE MSV000087308.

## ◼ AUTHOR INFORMATION

### Corresponding Authors

**Max Crispin** — *School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, U.K.;* ⓞ orcid.org/0000-0002-1072-2694; Email: max.crispin@soton.ac.uk

**Peter J. Bond** — *Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), Singapore 138671; Department of Biological Sciences, National University of Singapore, Singapore 117543;* ⓞ orcid.org/0000-0003-2900-098X; Email: peterjb@bii.a-star.edu.sg

## Authors

**Joel D. Allen** − *School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, U.K.;* orcid.org/0000-0003-2547-968X

**Himanshi Chawla** − *School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, U.K.;* orcid.org/0000-0001-9828-6593

**Firdaus Samsudin** − *Bioinformatics Institute, Agency for Science, Technology and Research (A\*STAR), Singapore 138671*

**Lorena Zuzic** − *Bioinformatics Institute, Agency for Science, Technology and Research (A\*STAR), Singapore 138671; Department of Chemistry, Faculty of Science and Engineering, Manchester Institute of Biotechnology, The University of Manchester, Manchester M1 7DN, U.K.;* orcid.org/0000-0002-7834-612X

**Aishwary Tukaram Shivgan** − *Bioinformatics Institute, Agency for Science, Technology and Research (A\*STAR), Singapore 138671; Department of Biological Sciences, National University of Singapore, Singapore 117543;* orcid.org/0000-0002-2032-8738

**Yasunori Watanabe** − *School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, U.K.;* orcid.org/0000-0002-7952-1268

**Wan-ting He** − *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California 92037, United States; IAVI Neutralizing Antibody Center and Consortium for HIV/AIDS Vaccine Development (CHAVD), The Scripps Research Institute, La Jolla, California 92037, United States*

**Sean Callaghan** − *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California 92037, United States; IAVI Neutralizing Antibody Center and Consortium for HIV/AIDS Vaccine Development (CHAVD), The Scripps Research Institute, La Jolla, California 92037, United States*

**Ge Song** − *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California 92037, United States; IAVI Neutralizing Antibody Center and Consortium for HIV/AIDS Vaccine Development (CHAVD), The Scripps Research Institute, La Jolla, California 92037, United States*

**Peter Yong** − *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California 92037, United States; IAVI Neutralizing Antibody Center and Consortium for HIV/AIDS Vaccine Development (CHAVD), The Scripps Research Institute, La Jolla, California 92037, United States;* orcid.org/0000-0002-4899-3270

**Philip J. M. Brouwer** − *Department of Medical Microbiology, Amsterdam UMC, University of Amsterdam, Amsterdam Infection & Immunity Institute, 1007 MB Amsterdam, The Netherlands*

**Yutong Song** − *Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China; Beijing Advanced Innovation Center for Structural Biology and Frontier Research Center for Biological Structure, Beijing 100084, China;* orcid.org/0000-0002-2206-9549

**Yongfei Cai** − *Division of Molecular Medicine, Boston Children's Hospital, Boston, Massachusetts 02115, United States*

**Helen M. E. Duyvesteyn** − *Division of Structural Biology, University of Oxford, The Wellcome Centre for Human Genetics, Oxford OX3 7BN, U.K.*

**Tomas Malinauskas** − *Division of Structural Biology, University of Oxford, The Wellcome Centre for Human Genetics, Oxford OX3 7BN, U.K.;* orcid.org/0000-0002-4847-5529

**Joeri Kint** − *ExcellGene SA, CH1870 Monthey, Switzerland*

**Paco Pino** − *ExcellGene SA, CH1870 Monthey, Switzerland*

**Maria J. Wurm** − *ExcellGene SA, CH1870 Monthey, Switzerland*

**Martin Frank** − *Biognos AB, 41705 Göteborg, Sweden;* orcid.org/0000-0002-1006-6746

**Bing Chen** − *Division of Molecular Medicine, Boston Children's Hospital, Boston, Massachusetts 02115, United States; Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, United States*

**David I. Stuart** − *Division of Structural Biology, University of Oxford, The Wellcome Centre for Human Genetics, Oxford OX3 7BN, U.K.; Diamond Light Source Ltd., Harwell Science & Innovation Campus, Didcot OX11 0DE, U.K.;* orcid.org/0000-0002-3426-4210

**Rogier W. Sanders** − *Department of Medical Microbiology, Amsterdam UMC, University of Amsterdam, Amsterdam Infection & Immunity Institute, 1007 MB Amsterdam, The Netherlands; Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, New York 10065, United States*

**Raiees Andrabi** − *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California 92037, United States; IAVI Neutralizing Antibody Center and Consortium for HIV/AIDS Vaccine Development (CHAVD), The Scripps Research Institute, La Jolla, California 92037, United States*

**Dennis R. Burton** − *Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California 92037, United States; IAVI Neutralizing Antibody Center and Consortium for HIV/AIDS Vaccine Development (CHAVD), The Scripps Research Institute, La Jolla, California 92037, United States; Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard University, Cambridge, Massachusetts 02139, United States*

**Sai Li** − *Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China; Beijing Advanced Innovation Center for Structural Biology and Frontier Research Center for Biological Structure, Beijing 100084, China;* orcid.org/0000-0002-9353-0355

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.biochem.1c00279

## Author Contributions

J.D.A., H.C., and F.S. contributed equally to this work. J.D.A., Y.W., and H.C. performed site-specific analysis. J.D.A., F.S., H.C., P.J.B., and M.C. wrote and edited the manuscript. F.S., L.Z., A.T.S., M.F., and P.J.B. performed MD simulations. W.H., S.C., G.S., P.Y., R.A., and D.R.B. provided monomeric RBD proteins. P.J.M.B., Y.C., H.M.E.D., T.M., J.K., P.P., M.J.W., B.C., D.I.S., and R.W.S. provided recombinant S protein. Y.S. and S.L. produced and collected LC-MS data for viral derived S protein. All authors edited and reviewed the manuscript prior to submission.

## Notes

The authors declare the following competing financial interest(s): ExcellGene sells purified trimeric spike protein

preparations from CHO cells to commercial companies for internal research and for use in diagnostic applications.

## ■ REFERENCES

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

(2) Pinto, D., Park, Y. J., Beltramello, M., Walls, A. C., Tortorici, M. A., Bianchi, S., Jaconi, S., Culap, K., Zatta, F., De Marco, A., Peter, A., Guarino, B., Spreafico, R., Cameroni, E., Case, J. B., Chen, R. E., Havenar-Daughton, C., Snell, G., Telenti, A., Virgin, H. W., Lanzavecchia, A., Diamond, M. S., Fink, K., Veesler, D., and Corti, D. (2020) Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* 583, 290–295.

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

(4) Polack, F. P., Thomas, S. J., Kitchin, N., Absalon, J., Gurtman, A., Lockhart, S., Perez, J. L., Pérez Marc, G., Moreira, E. D., Zerbini, C., Bailey, R., Swanson, K. A., Roychoudhury, S., Koury, K., Li, P., Kalina, W. V., Cooper, D., Frenck, R. W., Hammitt, L. L., Türeci, Ö., Nell, H., Schaefer, A., Ünal, S., Tresnan, D. B., Mather, S., Dormitzer, P. R., Şahin, U., Jansen, K. U., and Gruber, W. C. (2020) Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N. Engl. J. Med.* 383, 2603–2615.

(5) Baden, L. R., El Sahly, H. M., Essink, B., Kotloff, K., Frey, S., Novak, R., Diemert, D., Spector, S. A., Rouphael, N., Creech, C. B., McGettigan, J., Khetan, S., Segall, N., Solis, J., Brosz, A., Fierro, C., Schwartz, H., Neuzil, K., Corey, L., Gilbert, P., Janes, H., Follmann, D., Marovich, M., Mascola, J., Polakowski, L., Ledgerwood, J., Graham, B. S., Bennett, H., Pajon, R., Knightly, C., Leav, B., Deng, W., Zhou, H., Han, S., Ivarsson, M., Miller, J., and Zaks, T. (2021) Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N. Engl. J. Med.* 384, 403–416.

(6) Ramasamy, M. N., Minassian, A. M., Ewer, K. J., Flaxman, A. L., Folegatti, P. M., Owens, D. R., Voysey, M., Aley, P. K., Angus, B., Babbage, G., Belij-Rammerstorfer, S., Berry, L., Bibi, S., Bittaye, M., Cathie, K., Chappell, H., Charlton, S., Cicconi, P., Clutterbuck, E. A., Colin-Jones, R., Dold, C., Emary, K. R. W., Fedosyuk, S., Fuskova, M., Gbesemete, D., Green, C., Hallis, B., Hou, M. M., Jenkin, D., Joe, C. C. D., Kelly, E. J., Kerridge, S., Lawrie, A. M., Lelliott, A., Lwin, M. N., Makinson, R., Marchevsky, N. G., Mujadidi, Y., Munro, A. P. S., Pacurar, M., Plested, E., Rand, J., Rawlinson, T., Rhead, S., Robinson, H., Ritchie, A. J., Ross-Russell, A. L., Saich, S., Singh, N., Smith, C. C., Snape, M. D., Song, R., Tarrant, R., Themistocleous, Y., Thomas, K. M., Villafana, T. L., Warren, S. C., Watson, M. E. E., Douglas, A. D., Hill, A. V. S., Lambe, T., Gilbert, S. C., Faust, S. N., Pollard, A. J., Aboagye, J., Adams, K., Ali, A., Allen, E. R., Allen, L., Allison, J. L., Andritsou, F., Anslow, R., Arbe-Barnes, E. H., Baker, M., Baker, N., Baker, P., Baleanu, I., Barker, D., Barnes, E., Barrett, J. R., Barrett, K., Bates, L., Batten, A., Beadon, K., Beckley, R., Bellamy, D., Berg, A., Bermejo, L., Berrie, E., Beveridge, A., Bewley, K., Bijker, E. M., Birch, G., Blackwell, L., Bletchly, H., Blundell, C. L., Blundell, S. R., Bolam, E., Boland, E., Bormans, D., Borthwick, N., Boukas, K., Bower, T., Bowring, F., Boyd, A., Brenner, T., Brown, P., Brown-O'Sullivan, C., Bruce, S., Brunt, E., Burbage, J., Burgoyne, J., Buttigieg, K. R., Byard, N., Cabera Puig, I., Camara, S., Cao, M., Cappuccini, F., Carr, M., Carroll, M. W., Cashen, P., Cavey, A., Chadwick, J., Challis, R., Chapman, D., Charles, D., Chelysheva, I., Cho, J. S., Cifuentes, L., Clark, E., Collins, S., Conlon, C. P., Coombes, N. S., Cooper, R., Cooper, C., Crocker, W. E. M., Crosbie, S., Cullen, D., Cunningham, C., Cuthbertson, F., Datoo, B. E., Dando, L., Datoo, M. S., Datta, C., Davies, H., Davies, S., Davis, E. J., Davis, J., Dearlove, D., Demissie, T., Di Marco, S., Di Maso, C., DiTirro, D., Docksey, C., Dong, T., Donnellan, F. R., Douglas, N., Downing, C., Drake, J., Drake-Brockman, R., Drury, R. E., Dunachie, S. J., Edwards, C. J., Edwards, N. J., El Muhanna, O., Elias, S. C., Elliott, R. S., Elmore, M. J., English, M. R., Felle, S., Feng, S., Ferreira Da Silva, C., Field, S., Fisher, R., Fixmer, C., Ford, K. J., Fowler, J., Francis, E., Frater, J., Furze, J., Galian-Rubio, P., Galloway, C., Garlant, H., Gavrila, M., Gibbons, F., Gibbons, K., Gilbride, C., Gill, H., Godwin, K., Gordon-Quayle, K., Gorini, G., Goulston, L., Grabau, C., Gracie, L., Graham, N., Greenwood, N., Griffiths, O., Gupta, G., Hamilton, E., Hanumunthadu, B., Harris, S. A., Harris, T., Harrison, D., Hart, T. C., Hartnell, B., Haskell, L., Hawkins, S., Henry, J. A., Hermosin Herrera, M., Hill, D., Hill, J., Hodges, G., Hodgson, S. H. C., Horton, K. L., Howe, E., Howell, N., Howes, J., Huang, B., Humphreys, J., Humphries, H. E., Iveson, P., Jackson, F., Jackson, S., Jauregui, S., Jeffers, H., Jones, B., Jones, C. E., Jones, E., Jones, K., Joshi, A., Kailath, R., Keen, J., Kelly, D. M., Kelly, S., Kelly, D., Kerr, D., Khan, L., Khozoee, B., Killen, A., Kinch, J., King, L. D. W., King, T. B., Kingham, L., Klenerman, P., Knight, J. C., Knott, D., Koleva, S., Lang, G., Larkworthy, C. W., Larwood, J. P. J., Law, R., Lee, A., Lee, K. Y. N., Lees, E. A., Leung, S., Li, Y., Lias, A. M., Linder, A., Lipworth, S., Liu, S., Liu, X., Lloyd, S., Loew, L., Lopez Ramon, R., Madhavan, M., Mainwaring, D. O., Mallett, G., Mansatta, K., Marinou, S., Marius, P., Marlow, E., Marriott, P., Marshall, J. L., Martin, J., Masters, S., McEwan, J., McGlashan, J. L., McInroy, L., McRobert, N., Megson, C., Mentzer, A. J., Mirtorabi, N., Mitton, C., Moore, M., Moran, M., Morey, E., Morgans, R., Morris, S. J., Morrison, H. M., Morshead, G., Morter, R., Moya, N. A., Mukhopadhyay, E., Muller, J., Munro, C., Murphy, S., Mweu, P., Noé, A., Nugent, F. L., O'Brien, K., O'Connor,

D., Oguti, B., Olchawski, V., Oliveira, C., O'Reilly, P. J., Osborne, P., Owen, L., Owino, N., Papageorgiou, P., Parracho, H., Parsons, K., Patel, B., Patrick-Smith, M., Peng, Y., Penn, E. J., Peralta-Alvarez, M. P., Perring, J., Petropoulos, C., Phillips, D. J., Pipini, D., Pollard, S., Poulton, I., Pratt, D., Presland, L., Proud, P. C., Provstgaard-Morys, S., Pueschel, S., Pulido, D., Rabara, R., Radia, K., Rajapaska, D., Ramos Lopez, F., Ratcliffe, H., Rayhan, S., Rees, B., Reyes Pabon, E., Roberts, H., Robertson, I., Roche, S., Rollier, C. S., Romani, R., Rose, Z., Rudiansyah, I., Sabheha, S., Salvador, S., Sanders, H., Sanders, K., Satti, I., Sayce, C., Schmid, A. B., Schofield, E., Screaton, G., Sedik, C., Seddiqi, S., Segireddy, R. R., Selby, B., Shaik, I., Sharpe, H. R., Shaw, R., Shea, A., Silk, S., Silva-Reyes, L., Skelly, D. T., Smith, D. J., Smith, D. C., Smith, N., Spencer, A. J., Spoors, L., Stafford, E., Stamford, I., Stockdale, L., Stockley, D., Stockwell, L. V., Stokes, M., Strickland, L. H., Stuart, A., Sulaiman, S., Summerton, E., Swash, Z., Szigeti, A., Tahiri-Alaoui, A., Tanner, R., Taylor, I., Taylor, K., Taylor, U., te Water Naude, R., Themistocleous, A., Thomas, M., Thomas, T. M., Thompson, A., Thompson, K., Thornton-Jones, V., Tinh, L., Tomic, A., Tonks, S., Towner, J., Tran, N., Tree, J. A., Truby, A., Turner, C., Turner, R., Ulaszewska, M., Varughese, R., Verbart, D., Verheul, M. K., Vichos, I., Walker, L., Wand, M. E., Watkins, B., Welch, J., West, A. J., White, C., White, R., Williams, P., Woodyer, M., Worth, A. T., Wright, D., Wrin, T., Yao, X. L., Zbarcea, D. A., and Zizi, D. (2020) Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial. *Lancet 396*, 1979−1993.

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

(8) Chen, W. H., Du, L., Chag, S. M., Ma, C., Tricoche, N., Tao, X., Seid, C. A., Hudspeth, E. M., Lustigman, S., Tseng, C. T. K., Bottazzi, M. E., Hotez, P. J., Zhan, B., and Jiang, S. (2014) Yeast-expressed recombinant protein of the receptor-binding domain in SARS-CoV spike protein with deglycosylated forms as a SARS vaccine candidate. *Hum. Vaccines Immunother. 10*, 648−658.

(9) Kumar, S., Maurya, V. K., Prasad, A. K., Bhatt, M. L. B., and Saxena, S. K. (2020) Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). *VirusDisease 31*, 13−21.

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

(11) Carbone, F. R., and Gleeson, P. A. (1997) Carbohydrates and antigen recognition by T cells. *Glycobiology 7*, 725−730.

(12) Chu, H., Hu, B., Huang, X., Chai, Y., Zhou, D., Wang, Y., Shuai, H., Yang, D., Hou, Y., Zhang, X., Yuen, T. T.-T., Cai, J.-P., Zhang, A. J., Zhou, J., Yuan, S., To, K. K.-W., Chan, I. H.-Y., Sit, K.-Y., Foo, D. C.-C., Wong, I. Y.-H., Ng, A. T.-L., Cheung, T. T., Law, S. Y.-K., Au, W.-K., Brindley, M. A., Chen, Z., Kok, K.-H., Chan, J. F.-W., and Yuen, K.-Y. (2021) Host and viral determinants for efficient SARS-CoV-2 infection of the human lung. *Nat. Commun. 12*, 134.

(13) Yang, Q., Hughes, T. A., Kelkar, A., Yu, X., Cheng, K., Park, S. J., Huang, W. C., Lovell, J. F., and Neelamegham, S. (2020) Inhibition of SARS-CoV-2 viral entry upon blocking N-and O-glycan elaboration. *eLife 9*, e61552.

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

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

(16) Xu, C., and Ng, D. T. W. (2015) Glycosylation-directed quality control of protein folding. *Nat. Rev. Mol. Cell Biol. 16*, 742.

(17) Pritchard, L. K., Harvey, D. J., Bonomelli, C., Crispin, M., and Doores, K. J. (2015) Cell- and Protein-Directed Glycosylation of Native Cleaved HIV-1 Envelope. *J. Virol. 89*, 8932−8944.

(18) Behrens, A. J., and Crispin, M. (2017) Structural principles controlling HIV envelope glycosylation. *Curr. Opin. Struct. Biol. 44*, 125.

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

(20) McKay, P. F., Hu, K., Blakney, A. K., Samnuan, K., Brown, J. C., Penn, R., Zhou, J., Bouton, C. R., Rogers, P., Polra, K., Lin, P. J. C., Barbosa, C., Tam, Y. K., Barclay, W. S., and Shattock, R. J. (2020) Self-amplifying RNA SARS-CoV-2 lipid nanoparticle vaccine candidate induces high neutralizing antibody titers in mice. *Nat. Commun. 11*, 3523.

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

(22) Jackson, L. A., Anderson, E. J., Rouphael, N. G., Roberts, P. C., Makhene, M., Coler, R. N., McCullough, M. P., Chappell, J. D., Denison, M. R., Stevens, L. J., Pruijssers, A. J., McDermott, A., Flach, B., Doria-Rose, N. A., Corbett, K. S., Morabito, K. M., O'Dell, S., Schmidt, S. D., Swanson, P. A., Padilla, M., Mascola, J. R., Neuzil, K. M., Bennett, H., Sun, W., Peters, E., Makowski, M., Albert, J., Cross, K., Buchanan, W., Pikaart-Tautges, R., Ledgerwood, J. E., Graham, B. S., and Beigel, J. H. (2020) An mRNA Vaccine against SARS-CoV-2 — Preliminary Report. *N. Engl. J. Med. 383*, 1920−1931.

(23) Corbett, K. S., Edwards, D. K., Leist, S. R., Abiona, O. M., Boyoglu-Barnum, S., Gillespie, R. A., Himansu, S., Schäfer, A., Ziwawo, C. T., DiPiazza, A. T., Dinnon, K. H., Elbashir, S. M., Shaw, C. A., Woods, A., Fritch, E. J., Martinez, D. R., Bock, K. W., Minai, M., Nagata, B. M., Hutchinson, G. B., Wu, K., Henry, C., Bahl, K., Garcia-Dominguez, D., Ma, L. Z., Renzi, I., Kong, W. P., Schmidt, S. D., Wang, L., Zhang, Y., Phung, E., Chang, L. A., Loomis, R. J., Altaras, N. E., Narayanan, E., Metkar, M., Presnyak, V., Liu, C., Louder, M. K., Shi, W., Leung, K., Yang, E. S., West, A., Gully, K. L., Stevens, L. J., Wang, N., Wrapp, D., Doria-Rose, N. A., Stewart-Jones, G., Bennett, H., Alvarado, G. S., Nason, M. C., Ruckwardt, T. J., McLellan, J. S., Denison, M. R., Chappell, J. D., Moore, I. N., Morabito, K. M., Mascola, J. R., Baric, R. S., Carfi, A., and Graham, B. S. (2020) SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature 586*, 567−571.

(24) Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020) Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell 181*, 281−292.e6.

(25) Sternberg, A., and Naujokat, C. (2020) Structural features of coronavirus SARS-CoV-2 spike protein: Targets for vaccination. *Life Sci. 257*, 118056.

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

(27) Gao, Q., Bao, L., Mao, H., Wang, L., Xu, K., Yang, M., Li, Y., Zhu, L., Wang, N., Lv, Z., Gao, H., Ge, X., Kan, B., Hu, Y., Liu, J., Cai, F., Jiang, D., Yin, Y., Qin, C., Li, J., Gong, X., Lou, X., Shi, W., Wu, D., Zhang, H., Zhu, L., Deng, W., Li, Y., Lu, J., Li, C., Wang, X., Yin, W.,

Zhang, Y., and Qin, C. (2020) Development of an inactivated vaccine candidate for SARS-CoV-2. *Science 369*, 77−81.

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

(29) Butler, M., and Spearman, M. (2014) The choice of mammalian cell host and possibilities for glycosylation engineering. *Curr. Opin. Biotechnol. 30*, 107.

(30) Butler, M. (2006) Optimisation of the cellular metabolism of glycosylation for recombinant proteins produced by mammalian cell systems. *Cytotechnology 50*, 57.

(31) Casalino, L., Gaieb, Z., Goldsmith, J. A., Hjorth, C. K., Dommer, A. C., Harbison, A. M., Fogarty, C. A., Barros, E. P., Taylor, B. C., Mclellan, J. S., Fadda, E., and Amaro, R. E. (2020) Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent. Sci. 6*, 1722−1734.

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

(33) Behrens, A. J., Struwe, W. B., and Crispin, M. (2017) Glycosylation profiling to evaluate glycoprotein immunogens against HIV-1. *Expert Rev. Proteomics 14*, 881.

(34) Watanabe, Y., Bowden, T. A., Wilson, I. A., and Crispin, M. (2019) Exploitation of glycosylation in enveloped virus pathobiology. *Biochim. Biophys. Acta, Gen. Subj. 1863*, 1480.

(35) Behrens, A. J., Vasiljevic, S., Pritchard, L. K., Harvey, D. J., Andev, R. S., Krumm, S. A., Struwe, W. B., Cupo, A., Kumar, A., Zitzmann, N., Seabright, G. E., Kramer, H. B., Spencer, D. I. R, Royle, L., Lee, J. H., Klasse, P. J., Burton, D. R., Wilson, I. A., Ward, A. B., Sanders, R. W., Moore, J. P., Doores, K. J., and Crispin, M. (2016) Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein. *Cell Rep. 14*, 2695−2706.

(36) Cipollo, J. F., and Parsons, L. M. (2020) Glycomics and glycoproteomics of viruses: Mass spectrometry applications and insights toward structure-function relationships. *Mass Spectrom. Rev. 39*, 371−409.

(37) Bagdonaite, I., and Wandall, H. H. (2018) Global aspects of viral glycosylation. *Glycobiology*, Oxford University Press.

(38) Wang, D., Baudys, J., Bundy, J. L., Solano, M., Keppel, T., and Barr, J. R. (2020) Comprehensive Analysis of the Glycan Complement of SARS-CoV-2 Spike Proteins Using Signature Ions-Triggered Electron-Transfer/Higher-Energy Collisional Dissociation (EThcD) Mass Spectrometry. *Anal. Chem. 92*, 14730−14739.

(39) Bangaru, S., Ozorowski, G., Turner, H. L., Antanasijevic, A., Huang, D., Wang, X., Torres, J. L., Diedrich, J. K., Tian, J. H., Portnoff, A. D., Patel, N., Massare, M. J., Yates, J. R., Nemazee, D., Paulson, J. C., Glenn, G., Smith, G., and Ward, A. B. (2020) Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. *Science 370*, 1089−1094.

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

(41) Gstöttner, C., Zhang, T., Resemann, A., Ruben, S., Pengelley, S., Suckau, D., Welsink, T., Wuhrer, M., and Domínguez-Vega, E. (2021) Structural and Functional Characterization of SARS-CoV-2 RBD Domains Produced in Mammalian Cells. *Anal. Chem. 93*, 6839−6847.

(42) Antonopoulos, A., Broome, S., Sharov, V., Ziegenfuss, C., Easton, R. L., Panico, M., Dell, A., Morris, H. R., and Haslam, S. M. (2021) Site-specific characterization of SARS-CoV-2 spike glycoprotein receptor-binding domain. *Glycobiology 31*, 181−187.

(43) Morley, G. L., Taylor, S., Jossi, S., Perez-Toledo, M., Faustini, S. E., Marcial-Juarez, E., Shields, A. M., Goodall, M., Allen, J. D.,

Watanabe, Y., Newby, M. L., Crispin, M., Drayson, M. T., Cunningham, A. F., Richter, A. G., and O'Shea, M. K. (2020) Sensitive detection of SARS-CoV-2-Specific antibodies in dried blood spot samples. *Emerging Infect. Dis. 26*, 2970−2973.

(44) Faustini, S. E., Jossi, S. E., Perez-Toledo, M., Shields, A. M., Allen, J. D., Watanabe, Y., Newby, M. L., Cook, A., Willcox, C. R., Salim, M., Goodall, M., Heaney, J. L., Marcial-Juarez, E., Morley, G. L., Torlinska, B., Wraith, D. C., Veenith, T. V., Harding, S., Jolles, S., Ponsford, M. J., Plant, T., Huissoon, A., O'Shea, M. K., Willcox, B. E., Drayson, M. T., Crispin, M., Cunningham, A. F., and Richter, A. G. (2021) Development of a high sensitivity ELISA detecting IgG, A & M antibodies to the SARS-CoV-2 spike glycoprotein in serum and saliva. *Immunology*, DOI: 10.1111/imm.13349.

(45) Cook, A. M., Faustini, S. E., Williams, L. J., Cunningham, A. F., Drayson, M. T., Shields, A. M., Kay, D., Taylor, L., Plant, T., Huissoon, A., Wallis, G., Beck, S., Jossi, S. E., Perez-Toledo, M., Newby, M. L., Allen, J. D., Crispin, M., Harding, S., and Richter, A. G. (2021) Validation of a combined ELISA to detect IgG, IgA and IgM antibody responses to SARS-CoV-2 in mild and moderate non-hospitalised patients. *J. Immunol. Methods 494*, 113046.

(46) Shields, A., Faustini, S. E., Perez-Toledo, M., Jossi, S., Aldera, E., Allen, J. D., Al-Taei, S., Backhouse, C., Bosworth, A., Dunbar, L. A., Ebanks, D., Emmanuel, B., Garvey, M., Gray, J., Kidd, I. M., McGinnell, G., McLoughlin, D. E., Morley, G., O'Neill, J., Papakonstantinou, D., Pickles, O., Poxon, C., Richter, M., Walker, E. M., Wanigasooriya, K., Watanabe, Y., Whalley, C., Zielinska, A. E., Crispin, M., Wraith, D. C., Beggs, A. D., Cunningham, A. F., Drayson, M. T., and Richter, A. G. (2020) SARS-CoV-2 seroprevalence and asymptomatic viral carriage in healthcare workers: A cross-sectional study. *Thorax 75*, 1089−1094.

(47) Behrens, A.-J., Harvey, D. J., Milne, E., Cupo, A., Kumar, A., Zitzmann, N., Struwe, W. B., Moore, J. P., and Crispin, M. (2017) Molecular Architecture of the Cleavage-Dependent Mannose Patch on a Soluble HIV-1 Envelope Glycoprotein Trimer. *J. Virol. 91*, No. e01894-16.

(48) Tai, W., He, L., Zhang, X., Pu, J., Voronin, D., Jiang, S., Zhou, Y., and Du, L. (2020) Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell. Mol. Immunol. 17*, 613−620.

(49) Stadlbauer, D., Amanat, F., Chromikova, V., Jiang, K., Strohmeier, S., Arunkumar, G. A., Tan, J., Bhavsar, D., Capuano, C., Kirkpatrick, E., Meade, P., Brito, R. N., Teo, C., McMahon, M., Simon, V., and Krammer, F. (2020) SARS-CoV-2 Seroconversion in Humans: A Detailed Protocol for a Serological Assay, Antigen Production, and Test Setup. *Curr. Protoc. Microbiol. 57*, e100.

(50) Pallesen, J., Wang, N., Corbett, K. S., Wrapp, D., Kirchdoerfer, R. N., Turner, H. L., Cottrell, C. A., Becker, M. M., Wang, L., Shi, W., Kong, W. P., Andres, E. L., Kettenbach, A. N., Denison, M. R., Chappell, J. D., Graham, B. S., Ward, A. B., and McLellan, J. S. (2017) Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci. U. S. A. 114*, E7348−E7357.

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

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

(53) Huo, J., Zhao, Y., Ren, J., Zhou, D., Duyvesteyn, H. M. E., Ginn, H. M., Carrique, L., Malinauskas, T., Ruza, R. R., Shah, P. N. M., Tan, T. K., Rijal, P., Coombes, N., Bewley, K. R., Tree, J. A., Radecke, J., Paterson, N. G., Supasa, P., Mongkolsapaya, J., Screaton, G. R., Carroll, M., Townsend, A., Fry, E. E., Owens, R. J., and Stuart, D. I. (2020) Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell Host Microbe 28*, 445−454.e6.

(54) Pino, P., Kint, J., Kiseljak, D., Agnolon, V., Corradin, G., Kajava, A. V., Rovero, P., Dijkman, R., den Hartog, G., McLellan, J. S., Byrne,

P. O., Wurm, M. J., and Wurm, F. M. (2020) Trimeric SARS-CoV-2 spike proteins produced from CHO cells in bioreactors are high-quality antigens. *Processes 8*, 1539.

(55) Šali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234*, 779−815.

(56) Hakansson-McReynolds, S., Jiang, S., Rong, L., and Caffrey, M. (2006) Solution structure of the severe acute respiratory syndrome-coronavirus heptad repeat 2 domain in the prefusion state. *J. Biol. Chem. 281*, 11965−11971.

(57) Dev, J., Park, D., Fu, Q., Chen, J., Ha, H. J., Ghantous, F., Herrmann, T., Chang, W., Liu, Z., Frey, G., Seaman, M. S., Chen, B., and Chou, J. J. (2016) Structural basis for membrane anchoring of HIV-1 envelope spike. *Science 353*, 172−175.

(58) Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S. M., Walsh, R. M., Rawson, S., Rits-Volloch, S., and Chen, B. (2020) Distinct conformational states of SARS-CoV-2 spike protein. *Science 369*, 1586.

(59) Eramian, D., Shen, M., Devos, D., Melo, F., Sali, A., and Marti-Renom, M. A. (2006) A composite score for predicting errors in protein structure models. *Protein Sci. 15*, 1653−1666.

(60) Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations. *J. Mol. Biol. 7*, 95.

(61) Park, S. J., Lee, J., Qi, Y., Kern, N. R., Lee, H. S., Jo, S., Joung, I., Joo, K., Lee, J., and Im, W. (2019) CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology 29*, 320−331.

(62) Petit, C. M., Chouljenko, V. N., Iyer, A., Colgrove, R., Farzan, M., Knipe, D. M., and Kousoulas, K. G. (2007) Palmitoylation of the cysteine-rich endodomain of the SARS-coronavirus spike glycoprotein is important for spike-mediated cell fusion. *Virology 360*, 264−274.

(63) Lee, J., Patel, D. S., Ståhle, J., Park, S. J., Kern, N. R., Kim, S., Lee, J., Cheng, X., Valvano, M. A., Holst, O., Knirel, Y. A., Qi, Y., Jo, S., Klauda, J. B., Widmalm, G., and Im, W. (2019) CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *J. Chem. Theory Comput. 15*, 775−786.

(64) Casares, D., Escribá, P. V., and Rosselló, C. A. (2019) Membrane lipid composition: Effect on membrane and organelle structure, function and compartmentalization and therapeutic avenues. *Int. J. Mol. Sci. 20*, 2167.

(65) Van Meer, G., Voelker, D. R., and Feigenson, G. W. (2008) Membrane lipids: Where they are and how they behave. *Nat. Rev. Mol. Cell Biol. 9*, 112.

(66) Van Meer, G. (1998) Lipids of the Golgi membrane. *Trends Cell Biol. 8*, 29−33.

(67) Huang, J., and Mackerell, A. D. (2013) CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem. 34*, 2135−2145.

(68) Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., MacKerell, A. D., Klauda, J. B., and Im, W. (2016) CHARMM-GUI Input Generator for NAMD, GRO-MACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput. 12*, 405−413.

(69) Eslami, H., and Müller-Plathe, F. (2007) Molecular dynamics simulation in the grand canonical ensemble. *J. Comput. Chem. 28*, 1763−1773.

(70) Hoover, W. G. (1985) Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A: At., Mol., Opt. Phys. 31*, 1695−1697.

(71) Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys. 52*, 7182−7190.

(72) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys. 103*, 8577−8593.

(73) Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A Linear Constraint Solver for molecular simulations. *J. Comput. Chem. 18*, 1463−1472.

(74) Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX 1−2*, 19−25.

(75) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 33−38.

(76) Johari, Y. B., Jaffé, S. R. P., Scarrott, J. M., Johnson, A. O., Mozzanino, T., Pohle, T. H., Maisuria, S., Bhayat-Cammack, A., Lambiase, G., Brown, A. J., Tee, K. L., Jackson, P. J., Wong, T. S., Dickman, M. J., Sargur, R. B., and James, D. C. (2021) Production of trimeric SARS-CoV-2 spike protein by CHO cells for serological COVID-19 testing. *Biotechnol. Bioeng. 118*, 1013−1021.

(77) Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S. M., Walsh, R. M., Rawson, S., Rits-Volloch, S., and Chen, B. (2020) Distinct conformational states of SARS-CoV-2 spike protein. *Science 369*, 1586−1592.

(78) Kirchdoerfer, R. N., Cottrell, C. A., Wang, N., Pallesen, J., Yassine, H. M., Turner, H. L., Corbett, K. S., Graham, B. S., McLellan, J. S., and Ward, A. B. (2016) Pre-fusion structure of a human coronavirus spike protein. *Nature 531*, 118−121.

(79) Derking, R., Allen, J. D., Cottrell, C. A., Sliepen, K., Seabright, G. E., Lee, W.-H., Aldon, Y., Rantalainen, K., Antanasijevic, A., Copps, J., Yasmeen, A., Cupo, A., Cruz Portillo, V. M., Poniman, M., Bol, N., van der Woude, P., de Taeye, S. W., van den Kerkhof, T. L. G. M., Klasse, P. J., Ozorowski, G., van Gils, M. J., Moore, J. P., Ward, A. B., Crispin, M., and Sanders, R. W. (2021) Enhancing glycan occupancy of soluble HIV-1 envelope trimers to mimic the native viral spike. *Cell Rep. 35*, 108933.

(80) Cottrell, C. A., van Schooten, J., Bowman, C. A., Yuan, M., Oyen, D., Shin, M., Morpurgo, R., van der Woude, P., van Breemen, M., Torres, J. L., Patel, R., Gross, J., Sewall, L. M., Copps, J., Ozorowski, G., Nogal, B., Sok, D., Rakasz, E. G., Labranche, C., Vigdorovich, V., Christley, S., Carnathan, D. G., Sather, D. N., Montefiori, D., Silvestri, G., Burton, D. R., Moore, J. P., Wilson, I. A., Sanders, R. W., Ward, A. B., and van Gils, M. J. (2020) Mapping the immunogenic landscape of near-native HIV-1 envelope trimers in non-human primates. *PLoS Pathog. 16*, No. e1008753.

(81) Bañó-Polo, M., Baldin, F., Tamborero, S., Marti-Renom, M. A., and Mingarro, I. (2011) N-Glycosylation efficiency is determined by the distance to the C-terminus and the amino acid preceding an Asn-Ser-Thr sequon. *Protein Sci. 20*, 179−186.

(82) Higel, F., Sandl, T., Kao, C. Y., Pechinger, N., Sörgel, F., Friess, W., Wolschin, F., and Seidl, A. (2019) N-glycans of complex glycosylated biopharmaceuticals and their impact on protein clearance. *Eur. J. Pharm. Biopharm. 139*, 123−131.

(83) Goh, J. B., and Ng, S. K. (2018) Impact of host cell line choice on glycan profile. *Crit. Rev. Biotechnol. 38*, 851−867.

(84) Dumont, J., Euwart, D., Mei, B., Estes, S., and Kshirsagar, R. (2016) Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives. *Crit. Rev. Biotechnol. 36*, 1110−1122.

(85) Dotz, V., and Wuhrer, M. (2019) N-glycome signatures in human plasma: associations with physiology and major diseases. *FEBS Lett. 593*, 2966−2976.

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

(87) Torrents de la Peña, A., Rantalainen, K., Cottrell, C. A., Allen, J. D., van Gils, M. J., Torres, J. L., Crispin, M., Sanders, R. W., and Ward, A. B. (2019) Similarities and differences between native HIV-1 envelope glycoprotein trimers and stabilized soluble trimer mimetics. *PLoS Pathog. 15*, No. e1007920.

(88) Boson, B., Legros, V., Zhou, B., Siret, E., Mathieu, C., Cosset, F.-L., Lavillette, D., and Denolly, S. (2021) The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles. *J. Biol. Chem. 296*, 100111.

(89) Li, N., Qi, Y., Zhang, F. Y., Yu, X. H., Wu, Y. G., Chen, Y., Jiang, C. L., and Kong, W. (2011) Overexpression of $\alpha$-2,6-sialyltransferase stimulates propagation of human influenza viruses in Vero cells. *Acta Virol. 55*, 147−153.

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

(91) Xiang, Y., Karaveg, K., and Moremen, K. W. (2016) Substrate recognition and catalysis by GH47 $\alpha$-mannosidases involved in Asn-linked glycan maturation in the mammalian secretory pathway. *Proc. Natl. Acad. Sci. U. S. A. 113*, E7890−E7899.

(92) Gordon, R. D., Sivarajah, P., Satkunarajah, M., Ma, D., Tarling, C. A., Vizitiu, D., Withers, S. G., and Rini, J. M. (2006) X-ray Crystal Structures of Rabbit N-acetylglucosaminyltransferase I (GnT I) in Complex with Donor Substrate Analogues. *J. Mol. Biol. 360*, 67−79.

(93) Pritchard, L. K., Spencer, D. I. R., Royle, L., Bonomelli, C., Seabright, G. E., Behrens, A.-J., Kulp, D. W., Menis, S., Krumm, S. A., Dunlop, D. C., Crispin, D. J., Bowden, T. A., Scanlan, C. N., Ward, A. B., Schief, W. R., Doores, K. J., and Crispin, M. (2015) Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. *Nat. Commun. 6*, 7479.

(94) Williams, W. B., Meyerhoff, R. R., Edwards, R.J., Li, H., Manne, K., Nicely, N. I., Henderson, R., Zhou, Y., Janowska, K., Mansouri, K., Gobeil, S., Evangelous, T., Hora, B., Berry, M., Abuahmad, A. Y., Sprenz, J., Deyton, M., Stalls, V., Kopp, M., Hsu, A. L., Borgnia, M. J., Stewart-Jones, G. B.E., Lee, M. S., Bronkema, N., Moody, M. A., Wiehe, K., Bradley, T., Alam, S. M., Parks, R. J., Foulger, A., Oguin, T., Sempowski, G. D., Bonsignori, M., LaBranche, C. C., Montefiori, D. C., Seaman, M., Santra, S., Perfect, J., Francica, J. R., Lynn, G. M., Aussedat, B., Walkowicz, W. E., Laga, R., Kelsoe, G., Saunders, K. O., Fera, D., Kwong, P. D., Seder, R. A., Bartesaghi, A., Shaw, G. M., Acharya, P., and Haynes, B. F. (2021) Fab-dimerized glycan-reactive antibodies are a structural category of natural antibodies. *Cell 184*, 2955−2972.e25.