ARTICLE

Check for updates

Vulnerabilities in coronavirus glycan shields despite extensive glycosylation

Yasunori Watanabe[1,2,3], Zachary T. Berndsen[4], Jayna Raghwani[5], Gemma E. Seabright[1,2], Joel D. Allen [1], Oliver G. Pybus[6], Jason S. McLellan [7], Ian A. Wilson [4,8], Thomas A. Bowden [3], Andrew B. Ward [4] & Max Crispin [1]✉

Severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS) coronaviruses (CoVs) are zoonotic pathogens with high fatality rates and pandemic potential. Vaccine development focuses on the principal target of the neutralizing humoral immune response, the spike (S) glycoprotein. Coronavirus S proteins are extensively glycosylated, encoding around 66–87 N-linked glycosylation sites per trimeric spike. Here, we reveal a specific area of high glycan density on MERS S that results in the formation of oligomannose-type glycan clusters, which were absent on SARS and HKU1 CoVs. We provide a comparison of the global glycan density of coronavirus spikes with other viral proteins including HIV-1 envelope, Lassa virus glycoprotein complex, and influenza hemagglutinin, where glycosylation plays a known role in shielding immunogenic epitopes. Overall, our data reveal how organisation of glycosylation across class I viral fusion proteins influence not only individual glycan compositions but also the immunological pressure across the protein surface.

[1] School of Biological Sciences, University of Southampton, Southampton SO17 1BJ, UK. [2] Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK. [3] Division of Structural Biology, University of Oxford, Wellcome Centre for Human Genetics, Oxford OX3 7BN, UK. [4] Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. [5] Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7LF, UK. [6] Department of Zoology, University of Oxford, Oxford OX1 3PS, UK. [7] Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA. [8] Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA. ✉email: max.crispin@soton.ac.uk

Coronaviruses (CoVs) are enveloped pathogens responsible for multiple respiratory disorders of varying severity in humans[1]. Certain CoVs represent a significant threat to global human health, as illustrated by outbreaks of severe acute respiratory syndrome coronavirus (SARS-CoV) in 2003[2], Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012[3], and most recently of SARS-CoV-2[4]. Given their mortality rates, the current lack of targeted treatments and licensed vaccines, and their capacity to transmit between humans and across species barriers[5,6], there is an urgent need for effective countermeasures to combat these pathogens. Ongoing vaccine development efforts focus on the spike (S) proteins that protrude from the viral envelope and constitute the main target of neutralizing antibodies[7,8].

These trimeric S proteins mediate host-cell entry with the S1 and S2 subunits responsible for binding to the host-cell receptor and facilitating membrane fusion, respectively[9–11]. MERS S binds to dipeptidyl-peptidase 4 (DPP4)[12], whereas SARS S[13] and SARS-CoV-2[14,15] utilize angiotensin-converting enzyme 2 (ACE2) as a host cellular receptor. CoV S proteins are the largest class I viral fusion proteins known[9], and are extensively glycosylated, with SARS and MERS S glycoproteins both encoding 69 N-linked glycan sequons per trimeric spike with SARS-CoV-2 containing 66 sites. These modifications often mask immunogenic protein epitopes from the host humoral immune system by occluding them with host-derived glycans[16–18]. This phenomenon of immune evasion by molecular mimicry and glycan shielding has been well characterised across other viral glycoproteins, such as HIV-1 envelope protein (Env)[19–21], influenza hemagglutinin (HA)[22,23] and Lassa virus glycoprotein complex (LASV GPC)[24–26].

Previous analyses of viral glycan shields have revealed the presence of underprocessed oligomannose-type glycans that seemingly arise due to steric constraints that prevent access of glycan processing enzymes to substrate glycans[24,27,28], especially when the viral glycoprotein has evolved to mask immunogenic epitopes with a particularly dense array of host-derived glycans[26,29–34]. Restricted access to these glycan sites or interference with surrounding protein surface or neighbouring glycan residues can render glycan processing enzymes ineffective in specific regions[27,28,35]. Glycan processing on soluble glycoproteins has also been shown to be a strong reporter of native-like protein architecture and thus immunogen integrity[36–38]; and glycan processing on a successful immunogen candidate should therefore mimic, as closely as possible, the structural features observed on the native virus[39,40].

Here, we provide global and site-specific analyses of N-linked glycosylation on soluble SARS, MERS and HKU1 CoV S glycoproteins and reveal extensive heterogeneity, ranging from oligomannose-type glycans to highly-processed complex-type glycosylation. The structural mapping of glycans of trimeric S proteins revealed that some of these glycans contribute to the formation of a cluster of oligomannose-type glycans at specific regions of high glycan density on MERS-CoV S. Molecular evolution analysis of SARS and MERS S genes also reveals a higher incidence of amino-acid diversity on the exposed surfaces of the S proteins that are not occluded by N-linked glycans. In addition, we compare the structures of the respective glycan coats of SARS and HIV-1 envelope proteins using cryo-electron microscopy (cryo-EM) and computational modelling, which delineate a sparse glycan shield exhibited on SARS S compared with other viral glycoproteins. We therefore undertook a comparative analysis of viral glycan shields from characterized class I fusion proteins to highlight how glycosylation density influences oligomannose-type glycan abundance, and the relationship between effective glycan shields and viral evasion ability.

Together, these data underscore the importance of glycosylation in viral immune evasion.

Results and discussion

Glycan processing of trimeric SARS and MERS spike proteins. To generate a soluble mimic of the viral S proteins, we used the 2P-stabilised native-like SARS and MERS S protein antigens, the design and structures of which have been described previously by Pallesen et al.[41]. SARS, MERS and HKU1 S genes encode many N-linked glycan sequons; 23, 23 and 29, respectively (Fig. 1a). We initially sought to quantitatively assess the composition of the carbohydrate structures displayed on the S glycoproteins. N-linked glycans were enzymatically released, fluorescently labelled, and subjected to hydrophilic interaction chromatography-ultra-performance liquid chromatography (HILIC-UPLC). Treatment with endoglycosidase H (Endo H) revealed a population (SARS 32.2%; MERS 33.8%, HKU1 25.0%) of underprocessed oligomannose-type glycans (Fig. 1b). This observation of both complex and oligomannose-type glycans reveals that the majority of N-linked glycans can be processed, although there is limited processing at specific sites across the S proteins. It is also interesting to note that the distribution of oligomannose-type glycans was broad, with $Man_5GlcNAc_2$ to $Man_9GlcNAc_2$ glycans all present, without one particular dominant peak, as is the case for some viral glycoproteins, such as HIV-1 Env[36]. The proportion of oligomannose-type glycans on recombinant coronavirus S proteins is consistent with previous studies performed on virally derived MERS and SARS coronavirus S proteins[17,42]. Coronaviruses have been previously been reported to form virions by budding into the lumen of endoplasmic reticulum-Golgi intermediate compartments (ERGIC)[43,44]. Observations of hybrid- and complex-type glycans on virally derived material[17,42] would, however, suggest that it is likely that coronavirus virions travel through the Golgi apparatus after virion formation in the ERGIC en route to the cell surface, thus supporting recombinant immunogens as models of viral glycoproteins.

To ascertain the precise structures of N-linked glycans, glycan pools of each coronavirus S protein were analysed by negative-ion ion-mobility-electrospray ionisation mass spectrometry (IM-ESI MS) (Supplementary Fig. 1). Consistent with the UPLC data, IM-ESI MS confirmed an array of complex-type glycans ranging from mono- to tetra-antennary, but also oligomannose- and hybrid-type glycans. The glycan compositions characterised in the spectra were largely invariant among the coronaviruses with no major structural differences observed.

Clustering of underprocessed glycans on MERS S. We subsequently performed glycopeptide analysis to ascertain the compositions of glycans at all of the potential N-linked glycosylation sites (PNGs). MERS, SARS and HKU1 recombinant S proteins were reduced, alkylated and digested with an assortment of proteases to yield glycopeptides, which were subjected to in-line liquid chromatography-mass spectrometry (LC-MS). This revealed differential levels of oligomannose, hybrid, and complex-type glycan populations (Fig. 2a, b). Using structures of the trimeric MERS and SARS S proteins (PDB ID: 5X59 and 5X58, respectively), we generated models of fully glycosylated coronavirus spikes using experimentally determined glycan compositions (Fig. 3a, b). This revealed that oligomannose-type glycans on MERS S co-localize to specific clusters on the head of the S protein, consisting of glycans at Asn155, Asn166, and Asn236 (Fig. 3a). We hypothesized that the fully oligomannose-type glycan population in this cluster arises due to the hindered accessibility of glycan processing enzymes to access the substrate glycan[28]. As such, we performed mutagenesis to knock out
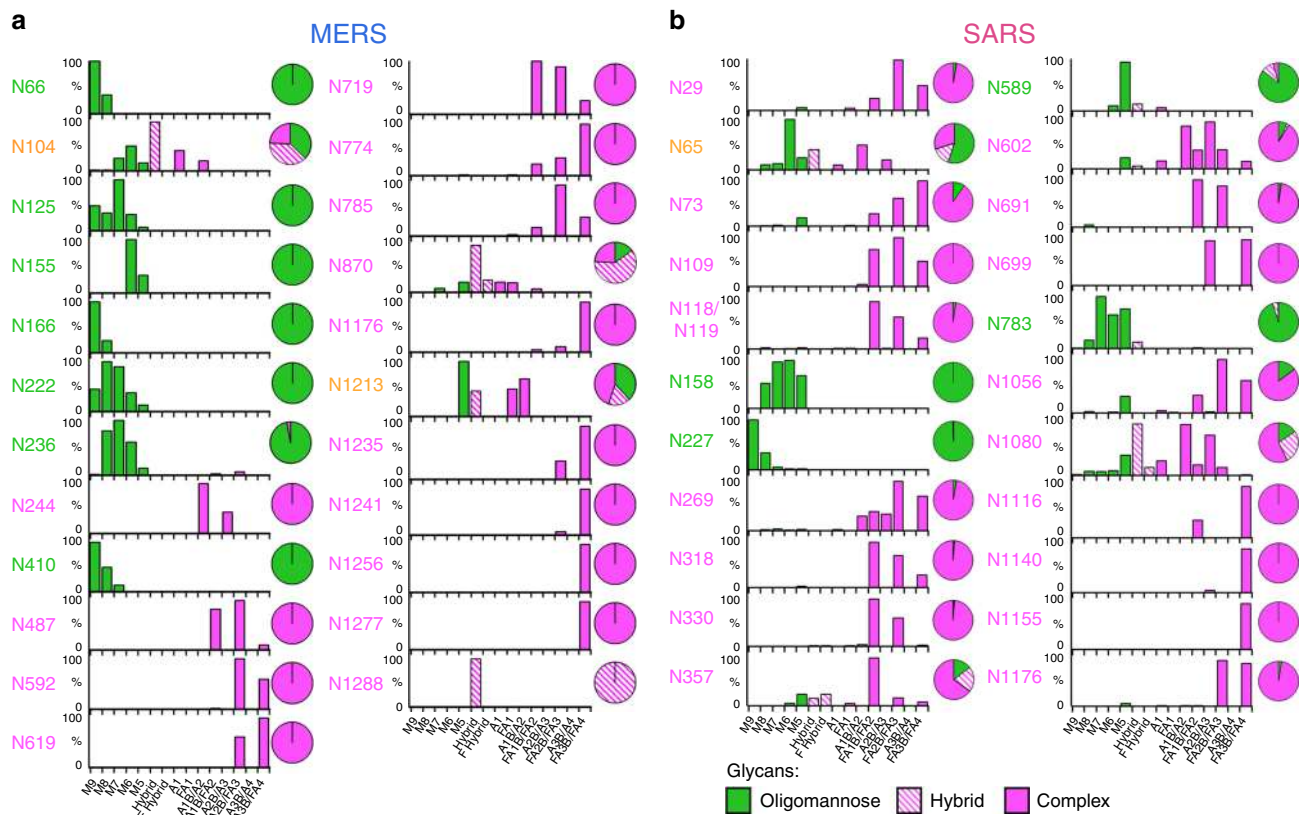
Fig. 1 Compositional analysis of SARS, MERS and HKU1 glycans. a Schematic representation of SARS, MERS and HKU1 coronavirus S glycoproteins, showing the positions of N-linked glycosylation amino-acid sequons (NXS/T, where X ≠ P) shown as branches. The domains of the S glycoproteins are illustrated: N-terminal domain (NTD), receptor-binding domain (RBD), sub-domain 1/2 (SD1/2), fusion peptide (FP), heptad repeat 1/2 (HR1/2), central helix (CH), and transmembrane domain (TM). **b** HILIC-UPLC chromatograms of fluorescently labelled N-linked glycans from SARS, MERS and HKU1 S. Oligomannose-type glycans (M5 to M9; $Man_5GlcNAc_2$–$Man_9GlcNAc_2$) (green) and complex-type glycans (magenta) were identified by Endo H digestion, with quantification of major glycan types summarised as a pie chart. Oligomannose-type glycans are schematically annotated with mannose residues as green circles and GlcNAc residues as blue squares.

glycosylation sites with N155A, N166A, and N236A mutations. Site-specific analysis of these glycan-KO mutants revealed enhanced trimming of mannose residues, i.e. increased processing, when glycan clustering was reduced (SI Fig. 4). The presence of clustered oligomannose-type glycans is reminiscent of that found on other viral glycoproteins, including HIV-1 Env and LASV GPC[24,31,34,36,45,46].

Interestingly, SARS and HKU1 (SI Fig. 2) S proteins did not exhibit specific mannose clusters that contribute to the overall mannose abundance, but only isolated glycans were underprocessed. We speculate that the oligomannose-type glycans here arise from protein-directed inhibition of glycan processing, as opposed to the glycan-influenced processing observed on MERS. Importantly, oligomannose-type glycans has also been implicated in innate immune recognition of coronaviruses by lectins[47,48] that recognise these underprocessed glycans as pathogen-associated molecular patterns.

Given that the receptor-binding domain is the main target of neutralising antibodies[8], it is surprising that the DPP4 receptor-binding site of MERS S was not occluded by glycans (Fig. 3a), as observed for other receptor-binding sites of class I viral fusion proteins, including SARS S (Fig. 3b), HIV-1 Env[49], LASV GPC[24] and influenza HA[50]. We suggest that this is likely due to the intrinsic functionality of the receptor-binding domain of MERS S, that would be sterically hindered by the presence of N-linked glycans, whereas other viruses are able to accommodate the post-translational modifications, without greatly perturbing functionality.

Sequence diversification of CoV spikes. We hypothesized that solvent-accessible, amino-acid residues on S proteins would be undergoing higher rates of mutations compared with buried residues and regions that are occluded by glycans, which are unable to be targeted by host immune responses. To that end, we performed an evaluation of amino-acid diversification on a residue-specific level, using publicly available gene sequences of SARS and MERS S, which was calculated as the number of observed pairwise differences divided by the total number of pairwise comparisons. Firstly, we found that amino-acid diversity

was elevated at known epitopes targeted by neutralizing antibodies, such as the N-terminal domain and the receptor-binding domains, and reduced in the regions in the S2 domain, such as the fusion peptide, heptad repeat one, and the central helix domains, which are likely subject to greater functional constraints (Fig. 4a).

Analysis of the relative ratio of non-synonymous to synonymous nucleotide substitutions (i.e. dN/dS ratios) revealed that exposed residues exhibited significantly higher dN/dS values (Fig. 4b). Buried residues on SARS had mean dN/dS ratios of 0.31 compared with 2.82 for exposed resides. Likewise, the buried residues on MERS had a calculated dN/dS ratio of 0.10 compared with exposed residues with a value of 0.45. Furthermore, when per-site amino-acid diversities were mapped onto the fully glycosylated structural model of the respective CoV S proteins (Fig. 4c), hotspots of mutations were highlighted on the protein surface throughout the trimer revealing extensive vulnerabilities permeating through the glycan shield of SARS and MERS CoVs. It is interesting to note the lack of amino-acid diversity on the receptor-binding domains of MERS S proteins that protrude away from the glycans. We would suggest that this may result from the intrinsic receptor-binding functionality of these domains.

Although dN/dS estimates are comparable within each viral outbreak, they are not directly comparable between viral families as they can only be considered in the environment in which they are measured (i.e. multiple differences in transmission ecology and host-virus interactions disallow meaningful comparisons). For example, differences in the epidemic behaviour and host immune environment of MERS and SARS outbreaks likely contribute to the observed genetic diversity and thus dN/dS. MERS was characterized by repeated spillover events from camels into humans, where it circulated transiently. In contrast, the SARS outbreak corresponded to a single zoonotic event followed by extensive human-to-human transmission. Consequently, inferring the degree of selection acting upon MERS and SARS from dN/dS analysis is extremely difficult. Importantly, while similar analyses of SARS-CoV-2 is desirable, due to the low genetic variation among the current SARS-CoV-2 sequences (as of 17 March 2020), which likely include deleterious mutations

Fig. 2 Quantitative site-specific N-linked glycan analysis of SARS and MERS coronavirus spike glycoproteins. Purified (**a**) MERS and **b** SARS S proteins were digested. Quantitative site-specific N-linked glycan analysis of **a** MERS and **b** SARS S glycoproteins. Purified S proteins were digested with trypsin, chymotrypsin, alpha-lytic protease, Glu-C, and trypsin plus chymotrypsin, then analysed by LC-ESI MS. Glycan compositions are based on the glycan library generated from negative-ion mass spectrometry of released N-glycans. The bar graphs represent the relative quantities of each glycan group with oligomannose-type glycan series (M9 to M5; Man$_9$GlcNAc$_2$ to Man$_5$GlcNAc$_2$) (green), afucosylated and fucosylated hybrid glycans (Hybrid & F Hybrid) (dashed pink), and complex glycans grouped according to the number of antennae and presence of core fucosylation (A1 to FA4) (pink). Left to right; least processed to most processed. The pie charts summarise the quantification of these glycans. Additional compositional information regarding the distribution of fucosylation and sialylation can be found in SI Fig. 3.

that will be removed by selection over time, the resulting bioinformatic analyses would be unreliable.

Visualising the HIV-1 and SARS glycan shields by cryo-EM. HIV-1 Env is a prototypic viral class I fusion protein that exhibits extensive surface glycosylation, resulting in an effective glycan shield to aid evasion from the host adaptive immune response[21,31]. In order to visualize the structure of the respective glycan "shields" of HIV-1 and SARS coronavirus we used single-particle cryo-electron microscopy (cryo-EM). The results for HIV-1 Env were reproduced directly from Berndsen et al.[51] while the previously published SARS 2P dataset[52] was reprocessed for this study. Although cryo-EM datasets of fully glycosylated MERS S[41] and chimpanzee simian immunodeficiency virus (SIVcpz)[53] are also available, only the HIV and SARS data were of sufficient quality (Fig. 5). We recently showed[51] that dynamics in surface exposed glycans HIV-1 Env leads to an extensive network of interactions that drive higher-order structuring in the glycan shield. This structure defines diffuse boundaries between buried and exposed surface protein surface, which can serve to define potential sites of vulnerability. Cryo-EM captures the ensemble-average structure of biomolecules and therefore glycan dynamics results in blurred density at the resolutions necessary for building atomic structure. However, we showed how a simple combination of low-pass filtering and auto-thresholding, as well as 3D varia-bility analysis, can reveal the previously hidden structure of the

SARS glycan shield and compare it with the HIV-1 Env glycan shield[51] (Fig. 5). We observe the nearly all-encompassing glycan density on HIV-1 Env and evidence for extensive glycan–glycan interactions, especially in the oligomannose patch regions, whereas the glycans on SARS S appear more isolated and lack the wide-ranging glycan networks that are the hallmark of an effec-tive glycan shield[54,55]. The 3D variability maps are more sensitive to low intensity signal and reveal additional glycan–glycan interactions in both maps, however the S1 receptor-binding domains in the SARS dataset were shown to exist in both up and down conformations[52], leading to poor resolution and significant 2D-variability which is convolved with the variability coming from glycans and limits the interpretability of glycan shielding effects in this region of the map.

Disparate shielding efficacies of viral glycosylation. Viral envelope proteins are glycosylated to varying degrees, but depending on their overall mass, surface area, and volume, the overall density of glycan shielding may differ significantly. For example, both LASV GPC and coronavirus S proteins consist of 25% glycan by molecular weight. However, given the significantly larger protein surface area and volume of coronavirus S proteins, coverage of the glycan "shield" over the proteinaceous surface is considerably sparser in comparison to the smaller LASV GPC, which occludes a far greater proportion of the protein surface with fewer glycans. To demonstrate that the presence of
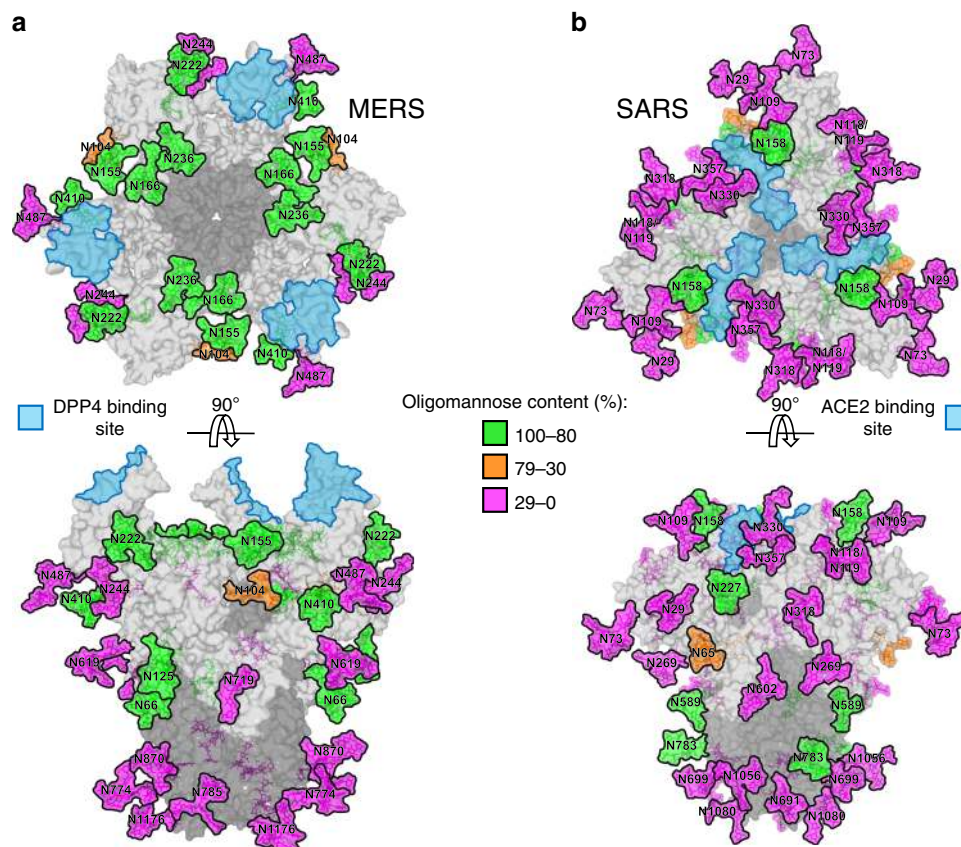
Fig. 3 Structure-based mapping of N-linked glycans on MERS and SARS S proteins. The modelling of the experimentally observed glycosylation is illustrated on the prefusion structure of trimeric **a** MERS S (PDB ID 5X59)[11] and **b** SARS S (PDB ID 5X58)[11] glycoproteins. Structural-based mapping of N-linked glycans on **a** MERS and **b** SARS S proteins. The modelling of the experimentally observed glycosylation is illustrated on the prefusion structure of trimeric MERS S (PDB ID 5X59)[11] and SARS S (PDB ID 5X58)[11] glycoproteins. The glycans are colored according to oligomannose content, as defined by the key. DPP4 receptor-binding sites and ACE2 receptor-binding sites for MERS and SARS, respectively, are indicated in light blue. The S1 and S2 subunits are colored light grey and dark grey, respectively.

glycosylation plays a major role in the immune response to these different glycoproteins, we studied the glycome of several biomedically important coronaviruses and compared their glycan compositions in a structural context.

We then investigated the glycan shield densities of seven viral class I fusion proteins using a global structural approach which was calculated by dividing the number of amino-acids that interact with glycans by the number of solvent-accessible amino-acid residues of each respective glycoprotein and plotted this against oligomannose abundance. A strong correlation was observed (Fig. 6) and viruses historically classified as "evasion strong"[56] had significantly elevated glycan shield densities and oligomannose abundance, which underscores the importance of glycan shielding in immune evasion.

Whether the restricted glycan shielding observed on coronaviruses is linked to the zoonosis of the pathogens is unknown. However, it is tempting to speculate, for example, that MERS has not evolved a dense shield since it would not offer as much of a protective advantage against camel nanobodies (also known as single-domain antibodies) which could more easily penetrate it. Investigation of the host immune response to viruses in their natural reservoirs may offer a route to understanding why coronavirus glycosylation does not reach the density of other viruses such as HIV-1. In addition, it may be that functional constraints, such as maintaining flexibility of the receptor-binding domains, limit the accretion of glycans on coronavirus spikes, which would render it incapable of performing its primary

functions, including receptor-binding and membrane fusion. This phenomenon has been observed on other viral glycoproteins, including influenza HAs, where there is a limit to the accumulation of glycosylation sites that can be incorporated in vivo[57,58], compared with in vitro[59], with H3N2 and H1N1 HAs replacing existing PNGs rather than continually adding them upon the glycoprotein[22,58]. The importance of glycosylation in modulation of viral infectivity and immune responses have also been investigated during influenza vaccine research[22,60] and should be considered in coronavirus vaccine research.

More topically, it is interesting to note the conservation of N-linked glycosylation sites on S proteins from the SARS-CoV-2 and SARS (SI Fig. 6). SARS-CoV-2 possesses a total of 22 N-linked glycan sites compared with 23 on SARS, with 18 of these sites being in common. As such, it is likely that these glycans on this novel coronavirus would shield similar immunogenic epitopes that are observed on SARS S. As expected, most of the differences between the two viruses are observed on the S1 subunit, due to its amenability to substitutions while still remaining functionally competent. Furthermore, likely targets for the majority of antibodies targeting the spike are located on S1, resulting in greater levels of immune pressure upon this subunit. This notion is further reflected in terms of glycosylation, with all of the glycan sites conserved on the S2 subunit between SARS and SARS-CoV-2, whereas the S1 subunit exhibits glycan site additions and deletions (SI Fig. 7). Bioinformatic analysis of current SARS-CoV-2 spike genes ($n = 566$ as of 17 March 2020)
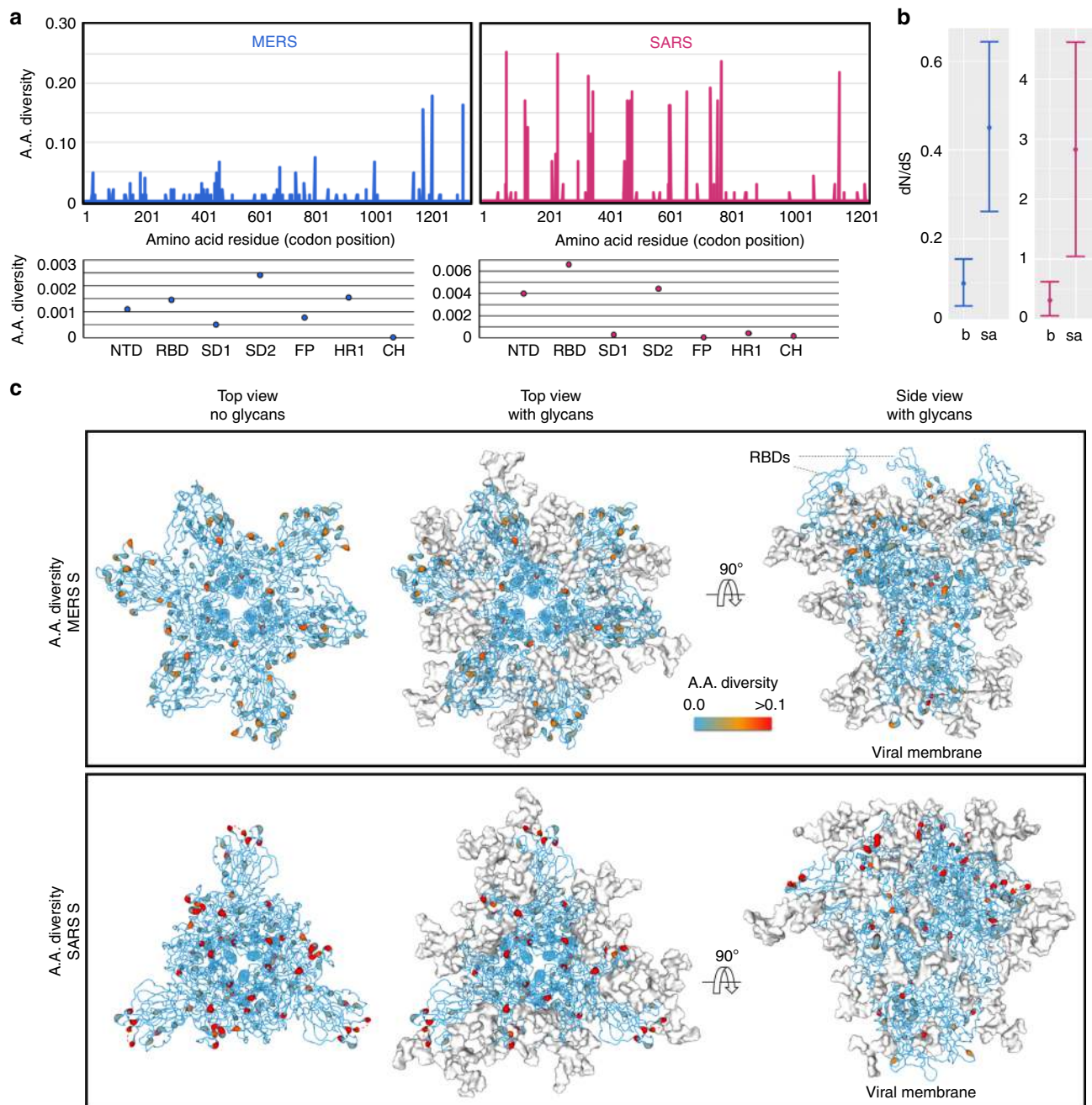
Fig. 4 Amino-acid sequence diversification across SARS and MERS spikes. a Amino-acid diversity in SARS and MERS S gene sequences. Averaged values for each domain are also shown below. **b** Comparison of dN/dS values between buried and exposed residues across SARS and MERS S ($n = 70$ and 100 for SARS and MERS, respectively). The error bars correspond to the 95% highest posterior density intervals while the circles indicate mean dN/dS values. **c** Mapping of the per residue amino-acid diversity shown in A onto the structures of SARS and MERS S (PDB ID 5X58 and 5X59, respectively)[11]. S proteins are presented as backbone traces with residues colored according to amino-acid diversity. Residues with elevated diversity are colored in red, and N-linked glycans are presented as white surfaces.

from nextstrain[61] (https://nextstrain.org/ncov) revealed low sequence diversity and no changes in glycosylation sites (SI Fig. 8).

Although it is difficult to directly compare viruses in terms of immunogenic responses, on the one hand, SARS and MERS coronaviruses readily elicit neutralizing antibodies following infection or immunization[62–65]. Indeed, many potential MERS-CoV vaccine candidates are able to elicit high titres of serum IgG upon immunization but fail to produce sufficient mucosal immunity[65]. In contrast, the high mutation rate[66] and the evolving glycan shield of HIV-1[20,39], which firmly exemplifies it

as "evasion strong" virus, hinders the development of broadly neutralizing antibodies[67].Viruses classified as "evasion strong"[26,56] may then differ due to varied efficacies of protein surface shielding by glycans.

Overall, this study adds further evidence suggesting that extensive N-linked glycan modifications of SARS and MERS CoV S proteins do not constitute an effective shield, in comparison to glycan shields of certain other viruses, which is reflected by the overall structure, density and oligomannose abundance across the corresponding trimeric glycoproteins. We also demonstrate that amino-acid diversification indeed occurs at antibody accessible
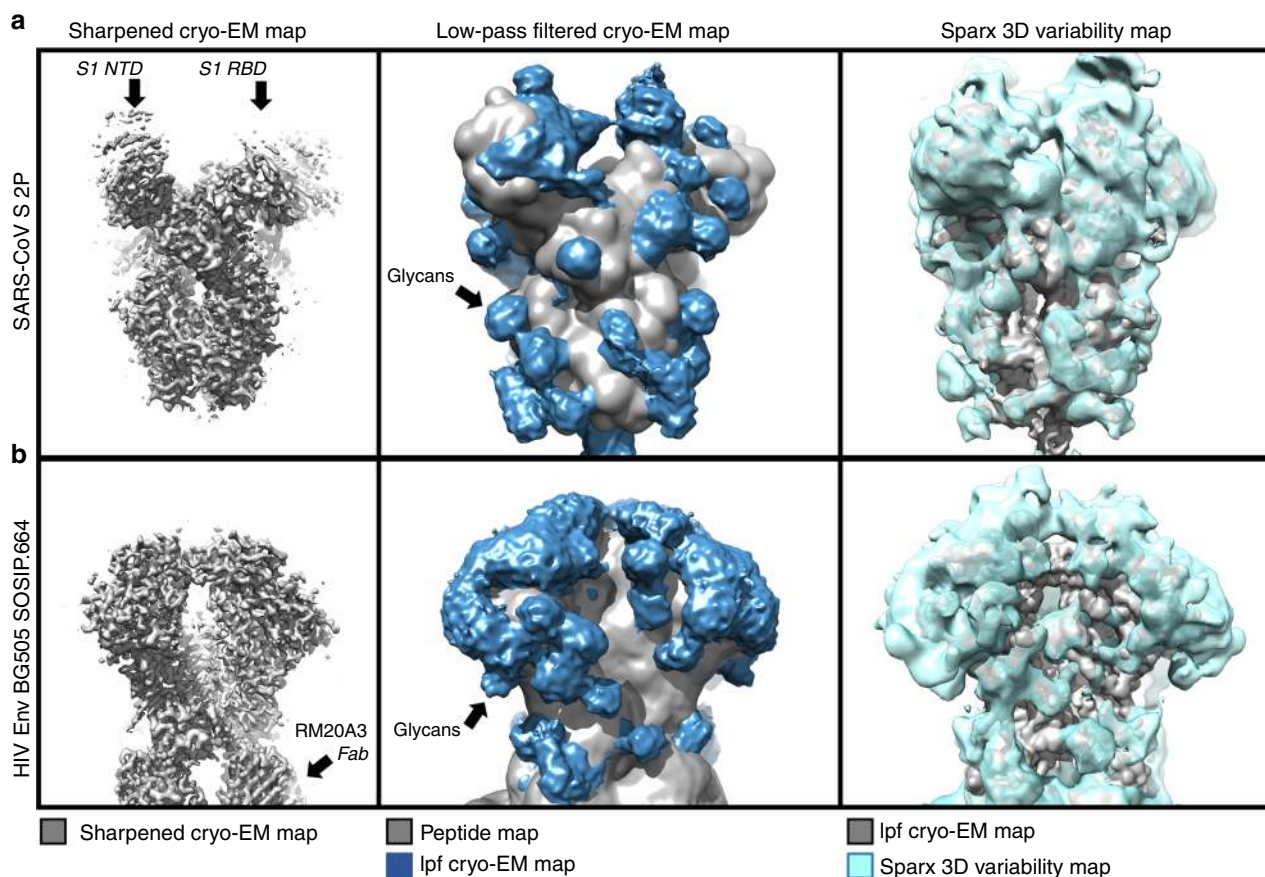
Fig. 5 Comparative cryo-EM analysis of SARS S and HIV-1 Env glycan shields. a Left panel: Sharpened 3.2-Å-resolution C3-symmetric cryo-EM map of SARS S 2P ectodomain[52] visualized at a high contour level with disordered S1 receptor-binding and N-terminal domains extending out from the central core. Middle panel: Low-pass filtered (lpf) cryo-EM map of the glycoprotein visualised at a low contour level along with a simulated peptide-only map overlaid. Right panel: SPARX 3D variability map[51]. **b** Same as in (**a**) but for HIV-1 Env BG505 SOSIP.664 construct in complex with three copies of RM20A3 base-specific Fabs[51].

regions on the trimer, which confirms that glycans play a role in occluding specific regions if vulnerability on the glycoprotein. Furthermore, comparisons between glycan shields from a number of viruses highlight the importance of a glycan shield in immune evasion and reveal structural principles that govern glycosylation status.

Methods

Expression and purification of coronavirus spike glycoproteins. Human embryonic kidney 293 Freestyle (HEK293F) cells were transfected with mammalian-codon-optimised genes encoding 2P-stabilised SARS MERS and HKU1 S proteins containing a C-terminal T4 fibritin trimerization domain, an HRV3C cleavage site, an 8xHis-tag and a Twin-Strep-tag[41]. H3N2 Victoria 2011 hemagglutinin was also expressed in the HEK293F cells. The 200 ml cultures were harvested 6 days after transfection, filtered and purified by nickel-affinity chromatography and size exclusion chromatography using a Superdex™ 16/600 75 pg column (GE Healthcare).

Release and labelling of N-linked glycans. Excised coronavirus S gel bands were washed alternately with acetonitrile and water before drying in a vacuum centrifuge. The bands were rehydrated with 100 μL of water and incubated with PNGase F at 37 °C overnight. Aliquots of released N-linked glycans were also fluorescently labelled with procainamide, by adding 100 μL of labelling mixture (110 mg/mL procainamide and 60 mg/mL sodium cyanoborohydride in 70% DMSO and 30% glacial acetic acid) and incubating for 4 h at 65 °C. Procainamide labelled glycans were purified using Spe-ed Amide 2 columns (Applied Separations).

Glycan analysis by HILIC-UPLC. Labelled glycans were analysed using a 2.1 mm × 150 mm Acquity BEH Glycan column (Waters) on an Acquity H-Class UPLC instrument (Waters), with fluorescence measurements occurring at $\lambda_{ex} = 310$ nm

and $\lambda_{em} = 370$ nm. The following gradient was used: time (t) = 0: 22% A, 78% B (flow rate = 0.5 mL/min); t = 38.5: 44.1% A, 55.9% B (0.5 mL/min); t = 39.5: 100% A, 0% B (0.25 mL/min); t = 44.5: 100% A, 0% B (0.25 mL/min); t = 46.5: 22% A, 78% B (0.5 mL/min), where solvent A was 50 mM ammonium formate (pH 4.4) and B was acetonitrile. Quantification of oligomannose-type glycans was achieved by digestion of fluorescently labelled glycans with Endo H, and clean-up using a PVDF protein-binding membrane (Millipore). Empower 3 software (Waters) was used for data processing.

Mass spectrometry of glycans. Prior to ion-mobility electrospray ionisation MS and tandem MS analysis, PNGase F released N-linked glycans were purified on a Nafion® 117 membrane (Sigma-Aldrich) and a trace amount of ammonium phosphate was added to promote phosphate adduct formation. Glycans were analyzed by direct infusion using a Synapt G2Si instrument (Waters) with the following settings: capillary voltage, 0.8–1.0 kV; sample cone, 150 V; extraction cone, 150 V; cone gas, 40 l/h; source temperature, 80 °C; trap collision voltage, 4–160 V; transfer collision voltage, 4 V; trap DC bias, 60 V; IMS wave velocity, 450 m/s; IMS wave height, 40 V; trap gas flow, 2 ml/min; IMS gas flow, 80 ml/min. Data were acquired and processed with MassLynx v4.1 and Driftscope version 2.8 software (Waters).

Mass spectrometry of glycopeptides. Aliquots of 30–50 μg of coronavirus spikes were denatured, reduced and alkylated as described previously[36]. Proteins were proteolytically digested with trypsin (Promega), chymotrypsin (Promega), alpha-lytic protease (Sigma-Aldrich) and Glu-C (Promega). Reaction mixtures were dried and peptides/glycopeptides were extracted using C18 Zip-tip (MerckMilipore) following the manufacturer's protocol. Samples were resuspended in 0.1% formic acid prior to analysis by liquid chromatography-mass spectrometry using an Easy-nLC 1200 system coupled to an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific). Glycopeptides were separated using an EasySpray PepMap RSLC C18 column (75 μm × 75 cm) with a 240-min linear solvent gradient of 0–32% acetonitrile in 0.1% formic acid, followed by 35 min of 80% acetonitrile in 0.1% formic acid. Other settings include an LC flow rate of 200 nL/min, spray voltage of
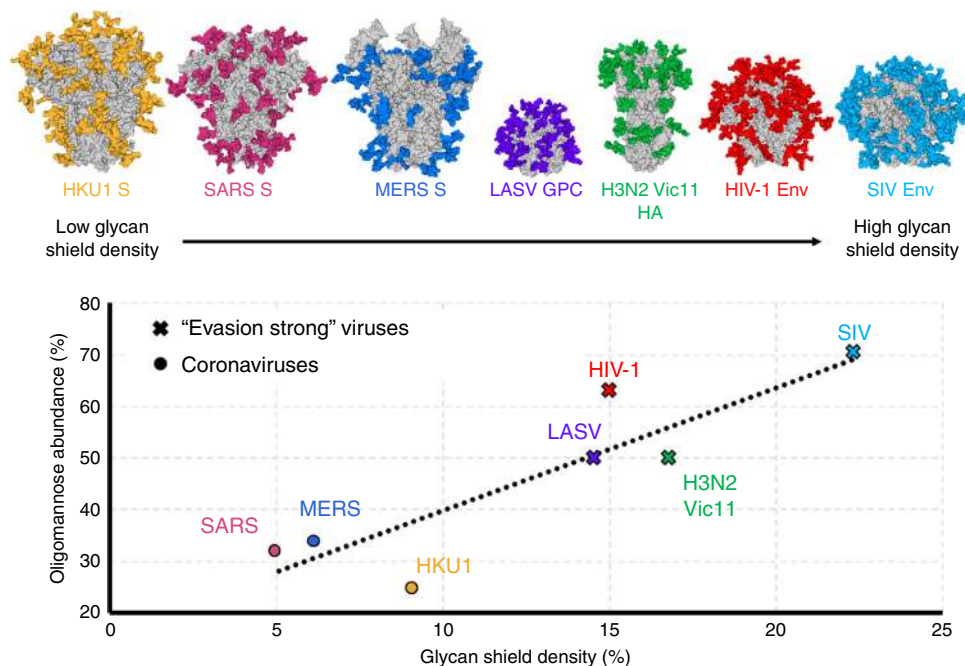
Fig. 6 Comparison of the glycan shields of viral class I fusion proteins. Glycan shield densities were calculated using Proteins, Interfaces, Structures and Assemblies (PISA)[83] analyses of fully glycosylated models of SARS S, MERS S, HKU1 S, LASV GPC, HIV-1 Env (BG505), Influenza H3N2 hemagglutinin (Victoria 2011), SIV Env (PDB ID 5X58, 5X59, 5I08, 5VK2, 4ZMJ, 4O5N, 6OHY, respectively)[9,11,53,84–86]. Oligomannose abundances of viral glycoproteins were ascertained by HILIC-UPLC analysis of PNGase F released N-linked glycans that were fluorescently labelled with procainamide[24,45,53] (SI Fig. 5). The number of amino-acid residues interacting with N-linked glycans was divided by the number of solvent-accessible amino-acid residues of the glycoprotein as a measure for global glycan shield density. All viral glycoproteins analysed were expressed as trimers in HEK293F cells apart from LASV GPC, which was derived from virus-like particles from Madin-Darby canine kidney II cells.

2.8 kV, capillary temperature of 275 °C, and an HCD collision energy of 50%. Precursor and fragmentation detection were performed using an Orbitrap at the following resolution: MS1 = 100,000 and MS2 = 30,000. The automatic gain control (AGC) targets were MS1 = 4e[5] and MS2 = 5e[4], and injection times were MS1 = 50 and MS2 = 54. The following cleavage sites were used for the respective proteases; trypsin=R/K, chymotrypsin=F/Y/W, alpha lytic protease=T/A/S/V, Glu C=E/D. Number of missed cleavages were set at 3. The following modifications were also included: Carbamidomethyl (+57.021464, target=C, fine control=fixed), Oxidation (+15.994915, target=M, fine control=variable rare 1), Glu to pyro-Glu (−18.010565, target=peptide N-term E, fine control=variable rare 1), and Gln to pyro-Glu (−17.026549, target peptide N-term Q, fine control=variable rare 1). Glycopeptide fragmentation data were extracted form raw files using Byonic[TM] (Version 3.5.0) and Byologic[TM] (Version 3.5-15; Protein Metrics Inc.). Glycopeptide fragmentation data were manually evaluated with true-positive assignments given when correct b- and y-fragments and oxonium ions corresponding to the peptide and glycan, respectively, were observed. The precursor mass tolerance was set at 4 ppm for precursor ions and 10 ppm for fragment ions. MS data were searched using a glycan library (SI Fig. 9) with the identical peptide sequence. A 1% false discovery rate (FDR) was applied. The extracted ion chromatographic areas for each true-positive glycopeptide, with the same amino-acid sequence, were compared to determine the relative quantitation of glycoforms at each specific N-linked glycan site.

Model construction. Structural models of N-linked glycan presentation on SARS, MERS and HKU1 S were created using electron microscopy structures (PDB ID 5X58, 5X59, and 5I08, respectively)[9,11], along with complex-, hybrid-, and oligomannose-type N-linked glycans (PDB ID 4BYH, 4B7I, and 2WAH). The most dominant glycoform presented at each site was modelled on to the N-linked carbohydrate attachment sites in Coot[68].

Molecular evolution analysis. Publicly available sequences encoding full-length GPC spike gene for SARS-CoV (3765 bp) were downloaded from GenBank and manually aligned. For MERS-CoV, we leveraged the whole genome alignment collated by Dudas et al.[69]. Specifically, the alignment corresponding to the spike gene was extracted (4059 bp), excluding sequences isolated from humans. Final alignments for SARS- and MERS-CoV corresponded to 70 and 100 sequences, respectively.

For the dN/dS analysis, we first estimated Bayesian molecular clock phylogenies for SARS- and MERS-CoV independently using BEAST v 1.8.4[70]. For both viruses, we assumed an uncorrelated log-normal distributed molecular clock[71], Bayesian Skyline

coalescent prior[72] and a codon-structured substitution model[73]. Multiple independent MCMC runs of 10–20 million steps were executed to ensure that stationarity and convergence had been achieved. Empirical distributions of time-scaled phylogenies were obtained by combining (after the removal of burnin) the posterior tree distributions from the separate runs, which were subsequently used to estimate dN/dS ratios using the renaissance counting approach[74,75] implemented in BEAST v 1.8.4. We also estimated per-site amino-acid diversity, which was calculated as the average number of amino-acid difference between two sequences at an amino-acid position in all possible pairs in the sequence alignment.

Cryo-EM data analysis and visualization. Single-particle cryo-EM data analysis of BG505 SOSIP.664 in complex with RM20A3 Fab was reproduced directly from Berndsen et al.[51]. Data for the SARS-CoV S 2P ectodomain was previously published[52] and the final particle stack and alignment parameters from the published reconstruction were used for 3D variability analysis in the SPARX software package[76,77]. All metadata for these reconstructions along with raw data images and FSC resolution curves can be found in the original publications. In summary, both datasets were acquired on a FEI Titan Krios (Thermo Fisher) operating at 300 KeV equipped with a K2 Summit Direct Electron Detector (Gatan). Movie micrographs were aligned and dose weighted with MotionCor2[78] and CTF estimation was performed with Gctf[79]. Single-particle data processing was performed using CryoSparc v.2[80] and Relion v.3[81]. Maps were low-pass filtered using a Gaussian kernel and visualized in UCSF chimera[82]. A detailed description of the auto-thresholding method used to set the isosurface value for visualisation of low-pass filtered maps can be found in Berndsen et al.[51].

Clustering analysis of viral glycan shields. Solvent-accessible residues and interactions between N-linked glycans and amino-acid residues were calculated using Proteins, Interfaces, Structures and Assemblies (PISA) European Bioinformatics Institute (EBI)[83]. Glycan shield density was calculated by the number of amino-acid residues interacting with glycans divided by the total number of solvent-accessible amino-acid residues.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

In addition to the data reported in this paper and accompanying supplementary materials, the raw mass spectrometric data that support the findings of this study

presented have been deposited on the MassIVE server (https://massive.ucsd.edu) with the accession codes (MSV000084993 for glycopeptides analysis [https://doi.org/10.25345/C58T21]; and MSV000085152 for N-linked glycans [https://doi.org/10.25345/C54705]).

Code availability

All code is available upon request.

References

1. Cui, J., Li, F. & Shi, Z. L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192 (2019).
2. Peiris, J. S. M., Guan, Y. & Yuen, K. Y. Severe acute respiratory syndrome. *Nat. Med.* **10**, S88–S97 (2004).
3. Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E. & Fouchier, R. A. M. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* **367**, 1814–1820 (2012).
4. Huang, C. et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **6736**, 1–10 (2020).
5. Azhar, E. I. et al. Evidence for camel-to-human transmission of MERS coronavirus. *N. Engl. J. Med.* **370**, 2499–2505 (2014).
6. Li, W. et al. Bats are natural reservoirs of SARS-like coronaviruses. *Science* **310**, 676–679 (2005).
7. Zhou, Y., Jiang, S. & Du, L. Prospects for a MERS-CoV spike vaccine. *Expert Rev. Vaccines* **17**, 677–686 (2018).
8. Xu, J. et al. Antibodies and vaccines against Middle East respiratory syndrome coronavirus. *Emerg. Microb. Infect.* **8**, 841–856 (2019).
9. Kirchdoerfer, R. N. et al. Pre-fusion structure of a human coronavirus spike protein. *Nature* **531**, 118–121 (2016).
10. Tortorici, M. A. & Veesler, D. Structural insights into coronavirus entry. in. *Adv. Virus Res.* **105**, 93–116 (2019).
11. Yuan, Y. et al. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat. Commun.* **8**, 15092 (2017).
12. Raj, V. S. et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* **495**, 251–254 (2013).
13. Li, W. et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
14. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 1–4, https://doi.org/10.1038/s41586-020-2012-7 (2020).
15. Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* https://doi.org/10.1126/science.abb2507 (2020).
16. Walls, A. C. et al. Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat. Struct. Mol. Biol.* **23**, 899–905 (2016).
17. Walls, A. C. et al. Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. *Cell* **176**, 1026-1-39.e5 (2019).
18. Yang, T. J. et al. Cryo-EM analysis of a feline coronavirus spike protein reveals a unique structure and camouflaging glycans. *Proc. Natl Acad. Sci. USA* **117**, 1438–1446 (2020).
19. Stewart-Jones, G. B. E. et al. Trimeric HIV-1-Env structures define glycan shields from Clades A, B, and G. *Cell* **165**, 813–826 (2016).
20. Wei, X. et al. Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).
21. Crispin, M., Ward, A. B. & Wilson, I. A. Structure and immune recognition of the HIV glycan shield. *Annu. Rev. Biophys.* **47**, 499–523 (2018).
22. Wu, N. C. & Wilson, I. A. A perspective on the structural and functional constraints for immune evasion: insights from influenza virus. *J. Mol. Biol.* **429**, 2694–2709 (2017).
23. Tate, M. D. et al. Playing hide and seek: how glycosylation of the influenza virus hemagglutinin can modulate the immune response to infection. *Viruses* **6**, 1294–1316 (2014).
24. Watanabe, Y. et al. Structure of the Lassa virus glycan shield provides a model for immunological resistance. *Proc. Natl Acad. Sci. USA* **115**, 7320–7325 (2018).
25. Sommerstein, R. et al. Arenavirus glycan shield promotes neutralizing antibody evasion and protracted infection. *PLoS Pathog.* **11**, e1005276 (2015).
26. Watanabe, Y., Bowden, T. A., Wilson, I. A. & Crispin, M. Exploitation of glycosylation in enveloped virus pathobiology. *Biochim. Biophys. Acta* **1863**, 1480–1497 (2019).
27. Behrens, A.-J. & Crispin, M. Structural principles controlling HIV envelope glycosylation. *Curr. Opin. Struct. Biol.* **44**, 125–133 (2017).
28. Pritchard, L. K. et al. Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. *Nat. Commun.* **6**, 7479 (2015).
29. Zhang, M. et al. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology* **14**, 1229–1246 (2004).
30. Yu, W. H. et al. Exploiting glycan topography for computational design of Env glycoprotein antigenicity. *PLoS Comput. Biol.* **14**, e1006093 (2018).
31. Zhou, T. et al. Quantification of the impact of the HIV-1-glycan shield on antibody elicitation. *Cell Rep.* **19**, 719–732 (2017).
32. Go, E. P. et al. Glycosylation benchmark profile for HIV-1 envelope glycoprotein production based on eleven env trimers. *J. Virol.* **91**, e02428-16 (2017).
33. Go, E. P. et al. Comparative analysis of the glycosylation profiles of membrane-anchored HIV-1 envelope glycoprotein trimers and soluble gp140. *J. Virol.* **89**, 8245–8257 (2015).
34. Hargett, A. A. et al. Defining HIV-1 envelope N-glycan microdomains through site-specific heterogeneity profiles. *J. Virol.* **93**, e01177-18 (2018).
35. Crispin, M. D. et al. Monoglucosylated glycans in the secreted human complement component C3: implications for protein biosynthesis and structure. *FEBS Lett.* **566**, 270–274 (2004).
36. Struwe, W. B. et al. Site-specific glycosylation of virion-derived HIV-1 Env Is mimicked by a soluble trimeric immunogen. *Cell Rep.* **24**, 1958–1966.e5 (2018).
37. Pritchard, L. et al. Structural constraints determine the glycosylation of HIV-1 envelope trimers. *Cell Rep.* **11**, 1604–1613 (2015).
38. Cao, L. et al. Differential processing of HIV envelope glycans on the virus and soluble recombinant trimer. *Nat. Commun.* **9**, 3693 (2018).
39. Wagh, K. et al. Completeness of HIV-1 envelope glycan shield at transmission determines neutralization breadth. *Cell Rep.* **25**, 893–908.e7 (2018).
40. Seabright, G. E., Doores, K. J., Burton, D. R. & Crispin, M. Protein and glycan mimicry in HIV vaccine design. *J. Mol. Biol.* https://doi.org/10.1016/j.jmb.2019.04.016 (2019).
41. Pallesen, J. et al. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. *Proc. Natl Acad. Sci. USA* **114**, E7348–E7357 (2017).
42. Ritchie, G. et al. Identification of N-linked carbohydrates from severe acute respiratory syndrome (SARS) spike glycoprotein. *Virology* **399**, 257–269 (2010).
43. Stertz, S. et al. The intracellular sites of early replication and budding of SARS-coronavirus. *Virology* **361**, 304–315 (2007).
44. Ng, M. L., Tan, S. H., See, E. E., Ooi, E. E. & Ling, A. E. Proliferative growth of SARS coronavirus in Vero E6 cells. *J. Gen. Virol.* **84**, 3291–3303 (2003).
45. Behrens, A.-J. et al. Composition and antigenic effects of individual glycan sites of a trimeric HIV-1 envelope glycoprotein. *Cell Rep.* **14**, 2695–2706 (2016).
46. Cao, L. et al. Global site-specific N-glycosylation analysis of HIV envelope glycoprotein. *Nat. Commun.* **8**, 14954 (2017).
47. Marzi, A. et al. DC-SIGN and DC-SIGNR interact with the glycoprotein of Marburg virus and the S protein of Severe Acute Respiratory Syndrome coronavirus. *J. Virol.* **78**, 12090–12095 (2004).
48. Zhou, Y. et al. A single asparagine-linked glycosylation site of the severe acute respiratory syndrome Coronavirus spike glycoprotein facilitates inhibition by mannose-binding lectin through multiple mechanisms. *J. Virol.* **84**, 8753–8764 (2010).
49. Jardine, J. et al. Rational HIV immunogen design to target specific germline B cell receptors. *Science* **340**, 711–716 (2013).
50. Xu, R. et al. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science* **328**, 357–360 (2010).
51. Berndsen, Z. T. et al. Visualization of the HIV-1 Env glycan shield across scales. *bioRxiv*, https://doi.org/10.1101/839217 (2019).
52. Kirchdoerfer, R. N. et al. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. *Sci. Rep.* **8**, 15701 (2018).
53. Andrabi, R. et al. The chimpanzee SIV envelope trimer: structure and deployment as an HIV vaccine template. *Cell Rep.* **27**, 2426–2441.e6 (2019).
54. Lemmin, T., Soto, C., Stuckey, J. & Kwong, P. D. Microsecond dynamics and network analysis of the HIV-1 SOSIP Env trimer reveal collective behavior and conserved microdomains of the glycan shield. *Structure* **25**, 1631–1639.e2 (2017).
55. Chakraborty, S. et al. A network-based approach for quantifying the resilience and vulnerability of HIV-1 native glycan shield. *bioRxiv*, https://doi.org/10.1101/846071 (2019).
56. Burton, D. R. What are the most powerful immunogen design vaccine strategies? Reverse vaccinology 2.0 shows great promise. *Cold Spring Harb. Perspect. Biol.* **9**, a030262 (2017).
57. Das, S. R. et al. Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. *Proc. Natl Acad. Sci. USA* **108**, E1417–E1422 (2011).

58. Altman, M. O. et al. Human influenza a virus hemagglutinin glycan evolution follows a temporal pattern to a glycan limit. *mBio* **10**, e00204–e00219 (2019).

59. Bajic, G. et al. Influenza antigen engineering focuses immune responses to a subdominant but broadly protective viral epitope. *Cell Host Microbe* **25**, 827–835.e6 (2019).

60. Wu, C. Y. et al. Influenza A surface glycosylation and vaccine design. *Proc. Natl Acad. Sci. USA* **114**, 280–285 (2017).

61. Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).

62. Bisht, H. et al. Severe acute respiratory syndrome coronavirus spike protein expressed by attenuated vaccinia virus protectively immunizes mice. *Proc. Natl Acad. Sci. USA* **101**, 6641–6646 (2004).

63. Buchholz, U. J. et al. Contributions of the structural proteins of severe respiratory syndrome coronavirus to protective immunity. *Proc. Natl Acad. Sci. USA* **101**, 9804–9809 (2004).

64. Yang, Z. Y. et al. A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice. *Nature* **428**, 561–564 (2004).

65. Yong, C. Y., Ong, H. K., Yeap, S. K., Ho, K. L. & Tan, W. S. Recent advances in the vaccine development against Middle East Respiratory Syndrome-coronavirus. *Front. Microbiol.* **10**, 1781 (2019).

66. Cuevas, J. M., Geller, R., Garijo, R., López-Aldeguer, J. & Sanjuán, R. Extremely high mutation rate of HIV-1 in vivo. *PLoS Biol.* **13**, e1002251 (2015).

67. Subbaraman, H., Schanz, M. & Trkola, A. Broadly neutralizing antibodies: What is needed to move from a rare event in HIV-1 infection to vaccine efficacy? *Retrovirology* **15**, 52 (2018).

68. Emsley, P. & Crispin, M. Structural analysis of glycoproteins: building N-linked glycans with coot. *Acta Crystallogr. Sect. D. Struct. Biol.* **74**, 256–263 (2018).

69. Dudas, G., Carvalho, L. M., Rambaut, A. & Bedford, T. MERS-CoV spillover at the camel-human interface. *eLife* **7**, e37324 (2018).

70. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).

71. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).

72. Drummond, A. J. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**, 1185–1192 (2005).

73. Shapiro, B., Rambaut, A. & Drummond, A. J. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* **23**, 7–9 (2006).

74. Lemey, P., Minin, V. N., Bielejec, F., Pond, S. L. K. & Suchard, M. A. A counting renaissance: combining stochastic mapping and empirical Bayes to quickly detect amino acid sites under positive selection. *Bioinformatics* **28**, 3248–3256 (2012).

75. Minin, V. N. & Suchard, M. A. Counting labeled transitions in continuous-time Markov models of evolution. *J. Math. Biol.* **56**, 391–412 (2008).

76. Hohn, M. et al. SPARX, a new environment for cryo-EM image processing. *J. Struct. Biol.* **157**, 47–55 (2007).

77. Cheng, Y., Grigorieff, N., Penczek, P. A. & Walz, T. A primer to single-particle cryo-electron microscopy. *Cell* **161**, 438–449 (2015).

78. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

79. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).

80. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

81. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

82. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

83. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

84. Hastie, K. M. et al. Structural basis for antibody-mediated neutralization of Lassa virus. *Science* **356**, 923–928 (2017).

85. Lee, P. S. et al. Receptor mimicry by antibody F045-092 facilitates universal binding to the H3 subtype of influenza virus. *Nat. Commun.* **5**, 3614 (2014).

86. Kwon, Y. D. O. et al. Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat. Struct. Mol. Biol.* **22**, 522–531 (2015).

Acknowledgements

Competing interests

J.S.M. and A.B.W. are inventors on U.S. patent application no. 62/412,703 ("Prefusion Coronavirus Spike Proteins and Their Use"). J.S.M. is an inventor on U.S. patent application no. 62/972,886 ("2019-nCoV Vaccine").

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-16567-0.

Correspondence and requests for materials should be addressed to M.C.

Peer review information *Nature Communications* thanks Matthew B. Renfrow and other, anonymous, reviewers for their contributions to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.