Beyond Shielding: The Roles of Glycans in SARS-CoV-2 Spike Protein

Lorenzo Casalino+[1], Zied Gaieb+[1], Jory A. Goldsmith [2], Christy K. Hjorth [2], Abigail C. Dommer[1], Aoife M. Harbison[3], Carl A. Fogarty[3], Emilia P. Barros[1], Bryn C. Taylor[1,4], Jason S. McLellan,[2] Elisa Fadda[3], Rommie E. Amaro*[1,4]

[1] Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093, USA
[2] Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA
[3] Department of Chemistry and Hamilton Institute, Maynooth University, Dublin, Ireland
[4] Biomedical Sciences Graduate Program, University of California San Diego, La Jolla, CA, 92093, USA

+ Shared first authorship
* Corresponding author

Abstract

The ongoing COVID-19 pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has resulted in more than 15,000,000 infections and 600,000 deaths worldwide to date. Antibody development efforts mainly revolve around the extensively glycosylated SARS-CoV-2 spike (S) protein, which mediates the host cell entry by binding to the angiotensin-converting enzyme 2 (ACE2). Similar to many other viruses, the SARS-CoV-2 spike utilizes a glycan shield to thwart the host immune response. Here, we built a full-length model of glycosylated SARS-CoV-2 S protein, both in the open and closed states, augmenting the available structural and biological data. Multiple microsecond-long, all-atom molecular dynamics simulations were used to provide an atomistic perspective on the roles of glycans, and the protein structure and dynamics. We reveal an essential structural role of N-glycans at sites N165 and N234 in modulating the conformational dynamics of the spike's receptor binding domain (RBD), which is responsible for ACE2 recognition. This finding is corroborated by biolayer interferometry experiments, which show that deletion of these glycans through N165A and N234A mutations significantly reduces binding to ACE2 as a result of the RBD conformational shift towards the "down" state. Additionally, end-to-end accessibility analyses outline a complete overview of the vulnerabilities of the glycan shield of SARS-CoV-2 S protein, which may be exploited by therapeutic efforts targeting this molecular machine. Overall, this work presents hitherto unseen functional and structural insights into the SARS-CoV-2 S protein and its glycan coat, providing a strategy to control the conformational plasticity of the RBD that could be harnessed for vaccine development.

Introduction

COVID-19 is an infectious respiratory disease that started in Wuhan, China, near the end of 2019, and has now spread worldwide as a global pandemic.[1] This is not the first time that a coronavirus (CoV) has posed a threat to human health. SARS-CoV-2 (the virus that causes COVID-19) is in the same family of viruses, *Coronaviridae*, as severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS) related coronaviruses, which have resulted in previous epidemics.[2–4] Owing to the lack of immunity, COVID-19 has already caused a catastrophic loss of human life worldwide[5] as well as significant economic damage.[6]

Coronaviruses, including SARS-CoV-2, are lipid-enveloped positive-sense RNA viruses. Together with the host-derived membrane, a set of structural proteins provide an organizational scaffold that wraps and contains the positive-sense viral RNA. Among them, the most critical is the spike, or S, protein, which is conserved to varying degrees across the *Coronaviridae* family and plays a key role in the virus' initial attachment and fusion with the host cell. The S protein is a class I fusion protein, synthesized as a single 1273 amino acid polypeptide chain, which associates as a trimer. Each monomer is made of two subunits, S1 and S2, and can be divided into three main topological domains, namely the head, stalk, and cytoplasmic tail (CT) (**Figure 1A**). One particularly interesting feature of the SARS-CoV-2 S protein is its adoption of a novel furin cleavage site between S1 and S2 (S1/S2), likely cleaved by the TMPRSS2 protease,[7] and believed to prime the spike for infection.[8,9] A second proteolytic cleavage at site S2′ releases the fusion peptide (FP), which penetrates the host cell membrane, gearing it up for fusion.[10] A number of recently published structural studies has provided an atomic or a near-atomic understanding of the head portion of SARS-CoV-2 spike, which comprises multiple domains (**Figures 1A** and **1B**).[11,12] The S1 subunit contains an N-terminal domain (NTD) and the receptor binding domain (RBD), where the receptor binding motif (RBM) is responsible for the interaction with the angiotensin-converting enzyme 2 (ACE2) receptor to gain entry into the host.[13] The S2 subunit has been aptly described as "a metastable spring-loaded fusion machine" because it plays a key role in integrating the viral and host cell membranes.[14] It contains the FP, the central helix (CH), and the connecting domain (CD). Additional domains within the S2 subunit that are not resolved in the pre-fusion state via cryo-EM or X-ray experiments include the heptad repeat 2 (HR2) and the transmembrane (TM) domains forming the stalk, and the CT (**Figures 1A** and **1B**).

Another key structural feature of the S protein that eludes detailed experimental structural characterization is its extensive glycosylation, shown in **Figure 1C**. Protein glycosylation plays a crucial role in viral pathogenesis,[15–17] as demonstrated by the characteristically thick N-glycan coating of the viral fusion proteins.[18–21] In the HIV-1 envelope spike (Env), for example, the protein-accessible surface area is almost entirely covered in N-glycans.[20,22] These are so densely packed that they account for more than half of the protein's molecular weight.[23] The biological roles of the N-glycans expressed on the surface of viral envelope glycoproteins are very diverse[16] and are all inextricably linked to their nature. Viral entry through membrane fusion is initiated by envelope glycoproteins through molecular recognition events involving cell surface receptors, which are often mediated by specific N-glycan epitopes.[16,24–26] Moreover, a highly dense coating of non-immunogenic or weakly immunogenic complex carbohydrates on otherwise dangerously exposed viral proteins constitutes a perfect camouflage (or shield) to evade the immune system.[16,18,19,27,28] To this end, the HIV-1 Env glycan shield, which is largely structured by oligomannose (Man5-9) N-glycans,[18,20,27,29] has been shown to be quite effective in allowing the virus to thwart the immune system;[15,16] it has also been found to be responsible for the virus'

interactions with DC-SIGN C-type lectins.[30] Contrary to HIV-1 Env, the betacoronaviruses SARS and MERS S proteins are not shielded as effectively.[15] Furthermore, both SARS-CoV and SARS-CoV-2 spikes present a rather different glycosylation pattern from that of HIV-1 Env, with a large presence of complex N-glycans relative to oligomannose type.[11,15,31] More specifically, the SARS-CoV-2 spike has 22 predicted N-glycosylation sites per protomer,[11,12] of which at least 17 have been found to be occupied,[11,31] plus at least two predicted O-glycosylation sites (see **Figures 1C-1E**).[31]

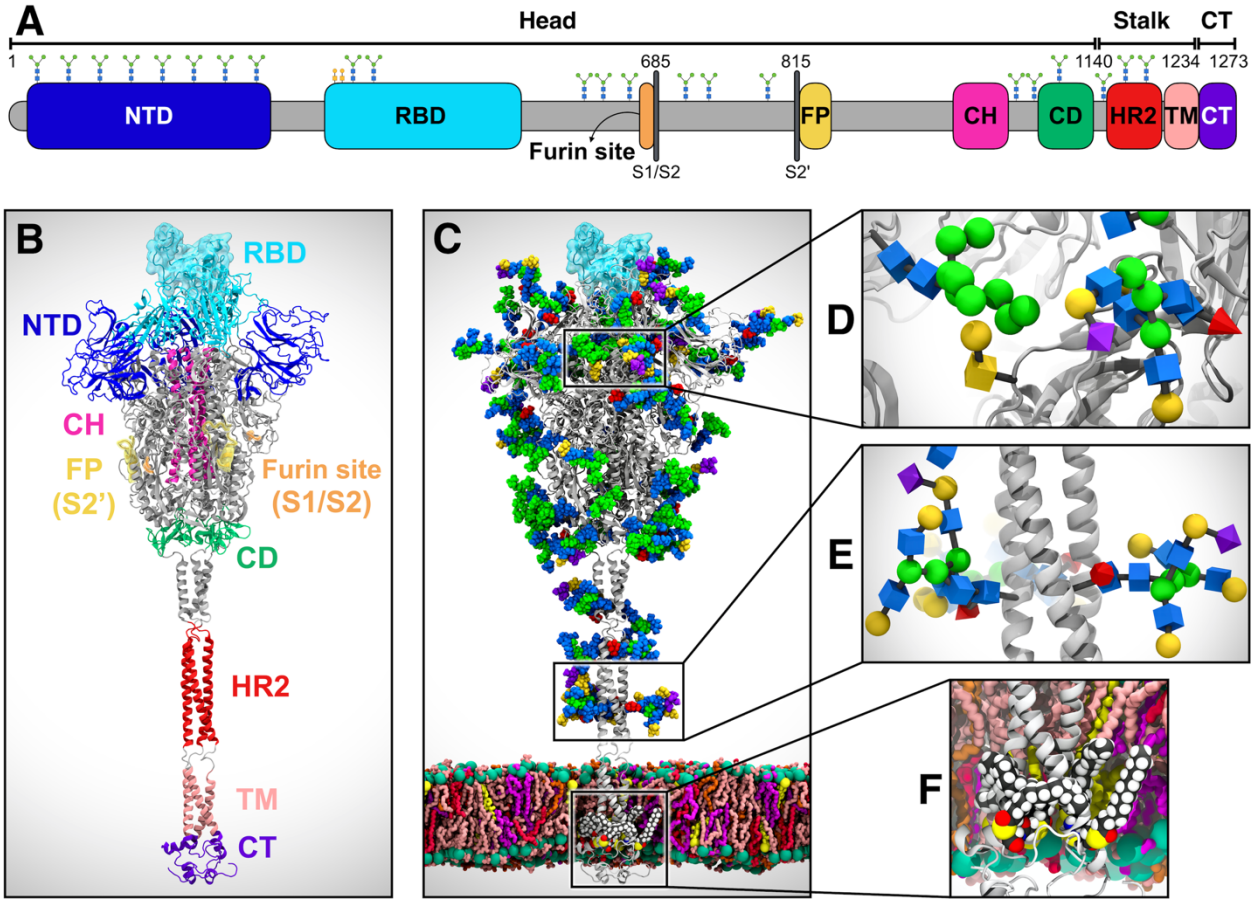


Figure 1. System overview. (**A**) Sequence of the full-length spike (S) protein contains the N-terminal domain (NTD, 16–291), receptor binding domain (RBD, 330–530), furin cleavage site (S1/S2), fusion peptide (FP, 817–834), central helix (CH, 987–1034), connecting domain (CD, 1080–1135), heptad repeat 2 (HR2, 1163–1210) domain, transmembrane domain (TM, 1214–1234), and cytoplasmic tail (CT, 1235–1273). Representative icons for N-glycans (blue and green) and O-glycan (yellow) are also depicted according to their position in the sequence. (**B**) Assembly of the head, stalk, and CT domains into the full-length model of the S protein in the open state. (**C**) Glycosylated and palmitoylated full-length model of the S protein in the open state embedded in a lipid bilayer mimicking the composition of the endoplasmic reticulum-Golgi intermediate compartment. Color code used for lipid tails: POPC (pink), POPE (purple), POPI (orange), POPS (red), cholesterol (yellow). P atoms of the lipid heads are shown with green spheres. Cholesterol's O3 atoms are shown with yellow spheres. (**D**) Magnified view of S protein head glycosylation, where glycans are depicted using the Symbol Nomenclature for Glycans (SNFG). (**E**) Magnified view of S protein stalk glycosylation. (**F**) Magnified view of S protein S-palmitoylation within CT.

In this work, we present multiple microsecond-long, all-atom, explicitly solvated molecular dynamics (MD) simulations of the full-length SARS-CoV-2 S glycoprotein embedded in the viral membrane, with a complete glycosylation profile consistent with glycomic data.[11,31] The

simulations discussed here augment and extend the available structural and biological data to provide an atomically detailed perspective on the full-length SARS-CoV-2 S protein's glycan coat, structure, and dynamics. Beyond shielding, our work reveals an essential structural role of N-glycans linked to N165 and N234 in modulating the conformational transitions of the RBD. Deletion of these glycans in our simulations elicits a destabilizing effect on the RBD "up" conformation. These results are corroborated via biolayer interferometry experiments showing a reduction of ACE2-binding, which underlies an increase of the RBD "down" population. Moreover, our simulations highlight how glycans camouflage SARS-CoV-2 S protein for thwarting the host immune response. A detailed analysis of the S protein's glycan shield discloses a lack of vulnerabilities in the stalk, particularly for large molecules, whereas the head region appears to be a more viable target. Importantly, an in-depth overview of the RBM accessibility indicates a remarkably different extent of glycan shield between the "up" and "down" RBD conformations. Overall, this work provides an atomic-level perspective on the SARS-CoV-2 S protein, highlighting the importance of glycans not only as shielding devices for immune evasion, but also as essential structural elements for virus infectivity. These insights lay the foundations for a possible strategy to modulate the RBD conformational plasticity and virus infectivity, which could be harnessed in the development of therapeutics aimed at fighting the pandemic threat.

Results & Discussion

All-Atom MD Simulations of the Full-Length Model of the SARS-CoV-2 S Protein in Open and Closed States.

In this work, we built a complete, full-length model of the glycosylated SARS-CoV-2 S protein in both closed and open states, hereafter referred to as "Closed" and "Open," respectively. Closed is based on cryo-EM structure 6VXX, where all three RBDs are in a "down" conformation.[12] Open is based on cryo-EM structure 6VSB, where the RBD within chain A (RBD-A) is in an "up" conformation.[32] A detailed view of chains A, B, and C and the RBD "up/down" conformations as in Open/Closed is shown in **Figure S1** of the Supporting Information (SI). These models were built in three steps, as fully described in the Materials and Methods section in the SI: (i) the "head," comprising S1/S2 subunits until residue 1140 and based on the above-mentioned experimental coordinates;[12,32] (ii) the "stalk" (residues 1141–1234), comprising HR2 and TM domains, and constructed through homology modeling using Modeller;[33] (iii) the CT (residues 1235–1273) built using I-TASSER (**Figures 1A**).[34] The modeled constructs (see **Figure 1B** for Open) were fully glycosylated at N-/O-glycosylation sites[11,31] (**Figures 1C-E**), and further refined with cysteine palmitoylation within the CT (**Figure 1F**).[35,36] An asymmetric (i.e., not specular across monomers) site-specific glycoprofile totaling 70 glycans ($22 \times 3_{monomers}$ N-glycans and $2_{chainA} + 1_{chainB} + 1_{chainC}$ O-glycans) was derived according to available glycoanalytic data (**Tables S1-S3** for composition).[11,31] We remark that the saccharides originally solved in the cryo-EM structures have been generally retained and utilized as a basic scaffold, when possible, to build the full glycan structure using CHARMM-GUI (details in the Materials and Methods section in the SI).[37–39] The full-length structures were embedded into an equilibrated all-atom membrane bilayer with a composition mimicking the endoplasmic reticulum-Golgi intermediate compartment (ERGIC, **Table S4** for composition), where the virus buds. The membrane was generated through CHARMM-GUI.[40] Subsequently, explicit water molecules and ions were added, affording two final systems each tallying ~1.7 million atoms (**Figure 1B**). Using CHARMM36 all-atom additive

force fields[41,42] and NAMD 2.14,[43] multiple replicas of all-atom MD simulations of the Open (6x) and Closed (3x) systems were run on the NSF Frontera computing system at the Texas Advanced Computing Center (TACC), achieving benchmarks of ~60 ns/day on 256 nodes for cumulative extensive sampling of ~4.2 and ~1.7 µs, respectively (**Table S5**, **Movie S1**). Note that, since RBD-A within Open is captured in a metastable state,[32] Open was simulated for a longer time than Closed. Visual inspection of trajectories was performed with VMD[44] that, together with MDtraj,[45] was also used for analyses.

Topological domain-specific root-mean-square-deviation (RMSD) relative to the starting structures was calculated to examine the structural stability along the simulations (**Figures S2-S3**). Overall, the two systems showed structural convergence of the extracellular topological domains (head and stalk) within ~400 ns. No significant difference was observed in the RMSD values of the stalk between the systems (**Figure S3**). Besides displaying a replica-specific convergence of RMSD values, the triple-stranded coiled-coil structural motif of the stalk persists throughout the simulations (**Movie S1**). Notably, as a result of secondary structure prediction (see Material and Methods in the SI), the stalk model includes two loops that break the alpha-helix within each strand (**Figure 1B**). Along with linked N-glycans, these loops confer critical dynamic properties to the stalk, allowing it to bend to accommodate the head's wiggling motions while remaining anchored to the membrane (**Movie S1**). The RMSD values of the CT domain show extremely large deviations and a mixed range of dynamic behaviors across replicas (**Figure S3**). While the first section of CT remains anchored to the inner-leaflet of the lipid bilayer through the palmitoylated cysteines, the rest is highly flexible and solvent exposed.

The root-mean-square-fluctuation (RMSF) values of the glycans were measured to determine their average flexibility around their mean position across the trajectories of all replicas (**Figure S4**). These values indicate a diverse range of mobility for specific sites, which is mostly dependent on solvent exposure, branching, and sequence. Complex-type N-glycans located on the NTD exhibit larger RMSF values than oligomannose and other N-glycans populating the head, including the ones in the RBD at sites N331 and N343. Interestingly, the O-glycans linked in the immediate vicinity of the RBD, namely at T323 and S325, show the lowest RMSF values. This is mostly attributable to their small size; however, their position may be critical for modulating RBD conformational changes, although we note that both positions are not consistently occupied. Finally, the large tetrantennary N-glycans in the stalk region at N1173 and N1194, modeled consistently with experimental data,[11] show the most extensive fluctuations, suggesting their good shielding potential.

The stability of the membrane bilayer was also monitored by tracking the time evolution of a series of structural properties, including area per lipid, membrane thickness, order parameters, and phospholipid tilt angles. Equilibrium area per lipid and P–P membrane thickness plots are shown in **Figure S5**. These results are in agreement with known POPC and previously reported binary POPC + X trends, where X is any other lipid added to the mix.[46–52]

N-Glycans at N165 and N234 modulate the RBD Conformational Dynamics

SARS-CoV and SARS-CoV-2 S glycoproteins share 76% sequence identity,[12] and 18 out of 22 N-glycosylation sites found on SARS-CoV-2 are conserved from SARS-CoV. Based on the analysis of our simulations of the full-length model of SARS-CoV-2 S protein, we identified two N-glycans linked to N165 and N234 in the NTD, which, given their strategic position and structure, we hypothesized may play a role in the RBD conformational dynamics. Our simulations show that in Closed, the N-glycan at N234 within NTD-B, modeled as Man9 in agreement with glycoanalytic

data,[11] is pointing away from the core of the S protein and is directed toward the solvent (**Figure 2A**); in contrast, in Open, where the RBD of chain A is "up," the same Man9 linked to N234 of NTD-B is directed inward, inserting itself into the large volume left vacant by the opening of RBD-A (**Figures 2B** and **2C**),. Because the glycans were built by structural alignment on the linked GlcNAc (or chitobiose) residues, wherever available, the relative position of the existing GlcNAc or chitobiose may affect the orientation of the entire N-glycan. Indeed, the N-glycan core is quite rigid.[53] To further assess the possible impact of the force fields and/or of the starting cryo-EM structure on the simulations described here, we performed an additional set of simulations of the open SARS-CoV-2 S protein's head using AMBER ff14SB/GLYCAM06j-1 force fields[54,55] and an alternative initial cryo-EM structure (PDB ID: 6VYB),[56] which presents the GlcNAc linked to N234 in a slightly different orientation. These simulations, detailed in **Section 2** of SI, show that the Man9 glycan at N234 progressively inserts itself to reach the apical core of the trimer through interactions with the surface of the closed RBD domain located across from it (**Movie S2**), adopting a structural arrangement similar to the one described in **Figure 2B**. The highly conserved N-glycan at N165 also inserts itself between NTD-B and RBD-A, either making extensive interactions with RBD-A or occupying the volume left vacant following its opening (**Figure 2C**). This was also observed in the additional simulations described in **Section 2** of the SI.

In light of these observations, an additional third system, "Mutant," was generated from Open by introducing N165A and N234A mutations, which led to glycan deletions at the respective sites. Mutant was then simulated through all-atom MD for six replicas totaling ~4.2 μs (**Table S5**). To determine the structural consequences of removing these two glycans, we compared the conformational landscape of the RBD in Open, Mutant, and Closed systems using principal component analysis (PCA). Scatter plot projections of RBD-A dynamics onto the first two eigenvectors (PC1 vs. PC2), i.e. the two motions with the largest variance in the trajectories (65% and 15%, respectively), clearly indicate that the RBD-A in Mutant explores a larger conformational space than in Open, which in turn is considerably larger than in Closed (**Figure 2D**). Although the difference between Open and Closed is immediately explained by the RBD-A state ("up/down"), the PCA landscape pinpoints a more stable RBD "up" conformation in the presence of N165 and N234, thus suggesting a structural role for the respective N-linked glycans (**Movie S3**). Detailed analysis of the PCA space explored by RBD-A, including the comparison between Open and Mutant only and the per-replica contribution in all systems, is presented in **Figures S6** and **S7**. Analysis of RMSD along the trajectory further highlights the instability of RBD-A in Mutant (**Figure S2**). This behavior is confirmed even in the case of a single point mutation (N234A), as revealed by the additional independent simulations of the spike head described in **Section 2** of SI. In contrast to RBD-A, RBD-B and RBD-C are in the "down" conformation in all three systems, showing mostly stable overlapping PCA distributions (**Figure S7**).

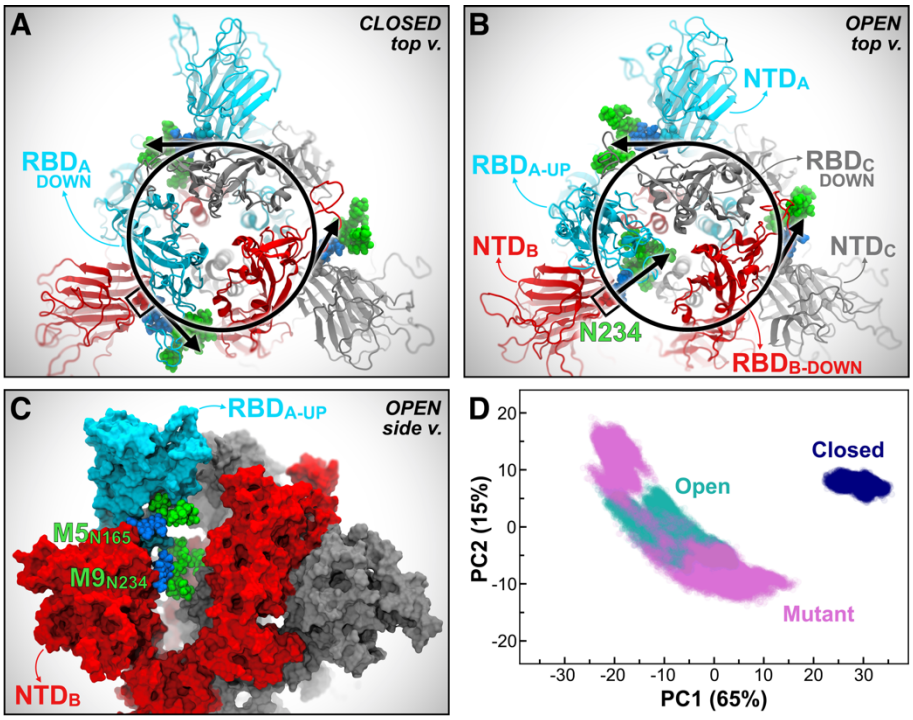


Figure 2. N234A and N165A mutations show increased instability of RBD-A in the "up" state. (**A-B**) Top view of the S protein as in the Closed (**A**) and Open (**B**) systems. Protein is represented with cartoons, colored in cyan, red, and gray for chains A, B and C, respectively. Oligomannose N-glycans at position N234 from all chains are depicted with VdW spheres, where GlcNAc is colored in blue and Man in green. In Closed (**A**), all the N-glycans at N234 are tangential to a hypothetical circle going through N234. In Open (**B**), the N-glycan at N234 of chain B moves inward, filling in the vacancy under RBD-A in the "up" conformation. (**C**) Side view of the S protein (surface representation) in Open, where the RBD of chain A (RBD-A, cyan) is stabilized by N-glycans at N165 and N234 in the "up" conformation. Same color scheme as panels A and B is applied. (**D**) PCA plot showing PC1 vs. PC2 of RBD-A (residues 330–530) in Closed, Open, and Mutant in blue, teal, and magenta, respectively. The amount (%) of variance accounted by each PC is shown between parentheses.

To obtain further insights into the RBD-A dynamic behavior and examine the differences in the explored conformational spaces revealed by PCA, we monitored the variations along the dynamics of two angles describing RBD-A fluctuations, hereafter referred to as the "lateral-angle" and "axial-angle." The lateral-angle reports an in-plane motion of the RBD along a hypothetical circle centered on the central helices of the spike (**Figure 3A**). This angle is described by three points corresponding to the (i) center of mass (COM) of RBD-A core β-sheets at frame 0, (ii - vertex) COM of the top section of the CH, and (iii) COM of RBD-A core β-sheets at frame n. The axial-angle identifies a tilting motion of the RBD either away from or toward the CH of the spike (**Figure 3B)**. This angle is defined by three points corresponding to (i) the COM of RBD-A core β-sheets, (ii - vertex) the COM of the CH, and (iii) the COM of the top section of the CH. The distributions of the lateral and axial-angle fluctuations along the trajectories of each system are shown in **Figures 3C** and **3D.** Angle variations were calculated with respect to their initial values at frame 0 of the trajectories. Full details on angle definitions and calculations are outlined in the Material and Methods section in the SI.

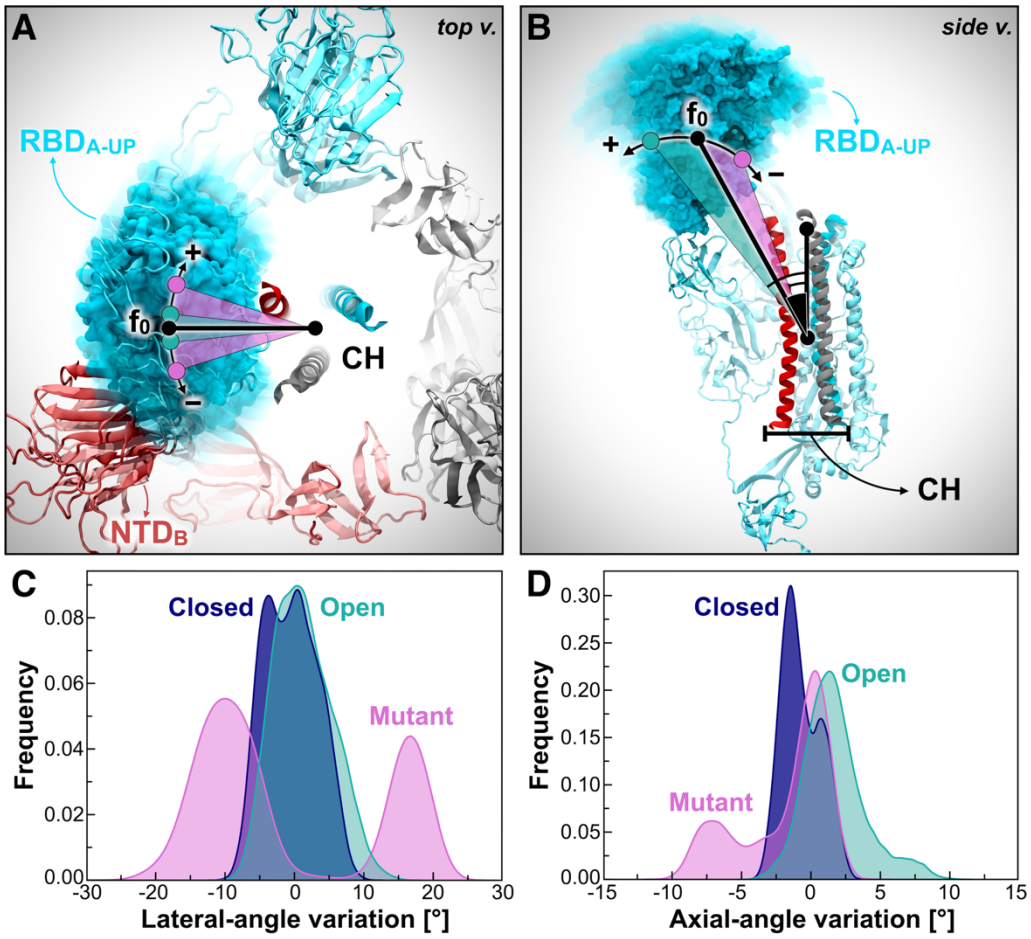


Figure 3. RBD-A lateral and axial-angle fluctuations. (**A, B**) RBD-A lateral-angle (**A**) and axial-angle (**B**), where chains A, B, and C of the spike are represented as ribbons colored in cyan, red, and gray, respectively. Positive and negative variations with respect to the initial frame (0) are indicated with the "+" and "–" symbols, respectively. N-/O-Glycans and some structural domains of the spike are omitted for clarity. (**C, D**) Distributions of RBD-A lateral-angle (**C**) and axial-angle (**D**) fluctuations along the trajectories (across all replicas) in Closed (blue), Open (teal), and Mutant (magenta). Angle variations were calculated with respect to their value at frame 0. Frequencies have been normalized within the respective data sets.

In agreement with the dynamic behavior revealed by PCA (**Figure 2D)**, the RBD-A lateral-angle displays a well-defined single distribution centered on zero in Open, whereas it exhibits a bimodal population in Mutant (**Figure 3C**). This analysis shows that the presence of the NTD-B N-glycans at N165 and N234 is crucial to stabilize RBD-A "up" conformation, preventing it from undergoing disordered motions. Interestingly, the axial-angle analysis results display a similar behavior of RBD-A in Open and Mutant, showing a long tail distribution in both systems (**Figure 3D**). However, Mutant exhibits a significant negative trend, whereas Open shows a positive trend. This suggests that the N-glycans at N165 and N234 may not only play a critical role in stabilizing the RBD "up" conformation, but also in modulating RBD opening and closing, thus potentially affecting binding to ACE2. Indeed, whereas the RBD in the "down" state overlies on the CH of the spike trimer, thus not being accessible to ACE2, when it transitions to the "up" state it detaches from the central axis, protruding into the solvent and becoming available for binding.

To confirm the hypothesis unveiled by our simulations, we sought to experimentally assess the effects of N165 and N234 deletions on the RBD state. To quantify RBD accessibility in

solution, we used biolayer interferometry to measure ACE2 binding to S protein containing either N165A or N234A substitutions. As a result, ablation of these N-linked sites reduces the binding responses for N165A or N234A to ~90% and ~60% (p=0.0051 and p=0.0002, Student's t-test) as high as the parent S-2P variant, respectively (**Figure 4**). We remark that the S-2P variant of the S protein bears two consecutive proline substitutions within the S2 subunit that are introduced to stabilize the pre-fusion conformation.[32] Importantly, a negative control spike (HexaPro), engineered with S383C/D985C mutations to lock all three RBDs in the closed state through a disulfide bond,[57] shows no binding to ACE2. These experiments corroborate the hypothesis that N165 and N234 glycans stabilize the RBD "up" conformation, therefore facilitating binding to ACE2, with the N234 glycan playing a larger role than N165.

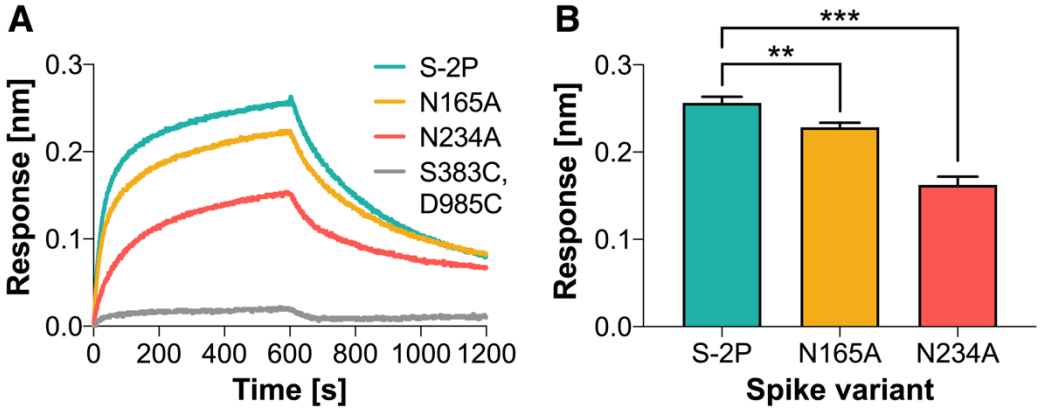


Figure 4. N234A and N165A mutations reduce RBD binding to ACE2. (**A**) Representative biolayer interferometry sensorgrams showing binding of ACE2 to spike variants. (**B**) Binding responses for biolayer interferometry measurements of ACE2 binding to spike variants. Data are shown as mean ± S.D. from three independent experiments for each variant. Asterisks represent statistical significance (Student's t test; *0.01<p<0.05, **0.001>p>0.01, ***0.0001<p<0.001).

To characterize how the N-glycans at N165 and N234 stabilize the RBD in the "up" state, we examined their interaction with the S protein through hydrogen bond analysis. Man9 linked to N234 on NTD-B deeply extends into the large pocket created by the opening of the RBD-A (**Figure 5C**). Specifically, it largely interconnects with the lower part of RBD-A (H519 in particular), propping it up from underneath, makes stable hydrogen bonds with D198 of NTD-B, and interacts as deep as R983, D985, E988 located within the CH of chain B (**Movie S3**). All of these hydrogen bonds are stable for more than 40% of the 4.2 µs trajectory of the Open system, with the majority of hydrogen bonds formed with the CH of chain B and RBD-A (see **Figure 5A**). The N-glycan at N165 is more exposed to the solvent relative to the Man9 glycan at N234; nevertheless, it extensively engages in interactions with the "up" conformation of RBD-A (~90% frequency) with a variable hydrogen bonding pattern across simulation replicas (**Figures 5B** and **5D, Figure S8, Movie S3**). The per-replica hydrogen bond networks observed for the N-glycans at N234 and N165 are shown in **Figure S9**.

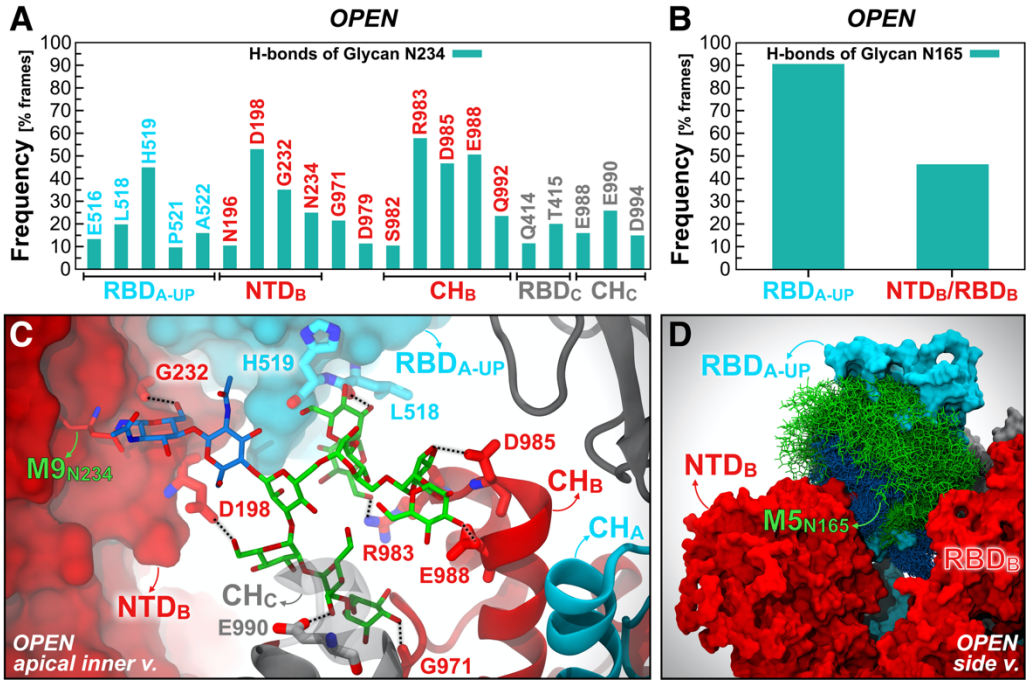


Figure 5. Hydrogen bond interactions of N-glycans at N234 and N165. The main hydrogen bond interactions of N-glycans at N234 (**A**) and N165 (**B**) within the Open system are shown as occupancy across all replicas (% frames). (**C**) A snapshot capturing Man9 glycan at N234 (licorice representation) within NTD-B establishing multiple hydrogen bonds with S protein residues (thicker licorice representation) belonging to RBD-A (cyan surface), NTD-B (red surface), CH-B (red cartoons), and CH-C (gray cartoons). GlcNAc and Man carbons are colored in blue and green, respectively. (**D**) Molecular representation of Man5 glycan at N165 within NTD-B interacting with RBD-A. Multiple (1000) equally interspersed configurations (selected across all replicas) of the glycan at N165 from NTD-B are simultaneously shown. The glycan is represented as colored licorices (GlcNAc in blue, Man in green), whereas RBD-A and NTD-B are represented as cyan and red surfaces, respectively.

Altogether, the heterogeneous conformational dynamics disclosed by PCA and angle analysis, together with the reduced ACE2-binding responses revealed by biolayer interferometry experiments, show that the conformational plasticity of the RBD is affected by the absence of N-glycans at N165 and N234. Notably, two flexible linker loops, connecting the RBD with the receptor-binding C-terminal domain 2 (CTD2) located underneath, prime the RBD to undergo large conformational rearrangements, such as the ACE2-induced hinge motion.[58] In this scenario, the subtle "up"/"down" equilibrium of the RBD can be dramatically altered by introducing single point mutations within the RBD or nearby regions,[57] or also by varying pH conditions.[59] Our work reveals that, similarly to MERS spike's RBD,[60] the SARS-CoV-2 spike's RBD "up" state is metastable within its conformational ensemble. We find that the RBD necessitates two N-glycans located on the adjacent NTD, namely N165 and N234, to "load-and-lock" the "up" conformation and to successfully bind to ACE2 receptors. Although our simulations do not show a full "up-to-down" transition of the RBD, the investigated glycan-deleting mutations cause the RBD in the "up" state to experience a larger conformational freedom and to undergo wider tilting motions. As a consequence, this conformational instability might increase the vulnerability of the virus, as cryptic epitopes located on the RBD might become more exposed to antibody recognition,[61,62] and, most likely, might decrease the virus infectivity as a result of a possible conformational shift of the RBD towards the closed state. Here, through biolayer interferometry experiments, we show that S protein binding to ACE2 is remarkably reduced for the N234A mutant, and slightly impaired

for the N165A variant, whereas is completely abolished in the case of an engineered S protein with all three RBDs locked in the "down" state. As such, these experiments, confirming the computational predictions, advocate for a conformational shift of the RBD population towards the "down" state as a result of the depletion of N234 and N165 glycans. This is consistent with the dynamics of the Man9 glycan linked to N234, which we observe crawling towards the apical portion of the spike's central axis, filling in the space left vacant by the RBD in the "up" conformation, and engaging in persistent interactions both with the same RBD "up" and the CH. Hence, it is not surprising that this glycan plays a larger role in stabilizing the RBD "up" than the glycan at N165, which is more solvent exposed and whose interaction pattern is less defined. However, we show that N165 steadily interacts with the RBD, therefore also contributing to loading and locking the "up" state, although to a lesser extent.

The SARS-CoV-2 spike's marked sensitivity to mutations affecting its structural dynamics represents an opportunity for vaccine design,[32,63] especially when it entails the alteration of the RBD "up"/"down" equilibrium.[57] In this context, our findings pinpoint the possibility of controlling the RBD conformational plasticity by introducing N165A and N234A mutations. These N-linked sites – N234 in particular – are essential for locking the RBD in the open state, thus contributing to priming the SARS-CoV-2 virus to recognize ACE2 receptors and invade host cells. Excising these glycans might result in less infectious, or "weakened", viruses bearing S proteins with predominantly "down" RBDs. This scenario would not prevent a humoral immune response targeting the upper side of the RBD,[64] including the RBM, since the RBDs would still be able to transition to the "up" state, although to a lesser extent. Instead, it might stimulate a larger production of antibodies recognizing other known epitopes located either on the core of the RBD,[65,66] or on the NTD,[67] or other S2 regions,[68] which might be engaged even in the closed state.

Glycan Shield of the Full-Length SARS-CoV-2 S Protein

In the context of vaccine design, it is critical to consider all the strategies developed by viruses to evade the host immune response. Within this framework, many viruses use a glycan shield to mask the immunogenic epitopes targeted by neutralizing antibodies.[11,15] SARS-CoV-2 S protein glycosylation is extensive and capable of thwarting the host humoral recognition.[11,15,69] In addition to the spike head and RBD,[61,70] another possible attractive target for antibodies is the stalk, as also shown in influenza virus studies.[71] Whereas 19 N-glycans essentially camouflage the spike head region, only 3 N-glycosylation sites (N1158, N1174, and N1194) are present on the stalk. Exploiting the extensive sampling derived from all-atom simulations of our full-length model of SARS-CoV-2 S glycoprotein, we provide an unprecedented end-to-end overview of the spike's glycan shield (**Figure 6**). In **Figure 6A**, each blue bush-like structure represents an ensemble of equally interspersed conformations of a single glycan sampled along a 1 μs trajectory for a total of 300 superimposed poses. Considering the different timescales of the dynamics of these N-glycans (nanoseconds), relative to the antibody/spike binding process (microseconds to milliseconds), this representation provides a realistic view of the shielding that molecules targeting the spike might encounter before binding. To quantify this shielding effect, we calculated the S protein's accessible surface area (ASA) that is covered by glycans on both the head and stalk regions of the Open system. We also screened a wide range of probe radii, from 1.4 to 15 Å, to measure the effectiveness of the shield with respect to molecules of different sizes. A value of 1.4 Å is commonly set to approximate the radius of a water molecule, whereas larger radii can be used to describe larger moieties, ranging from small molecules at 2–5 Å to larger peptide- and antibody-sized molecules at 10–15 Å.[69,72]

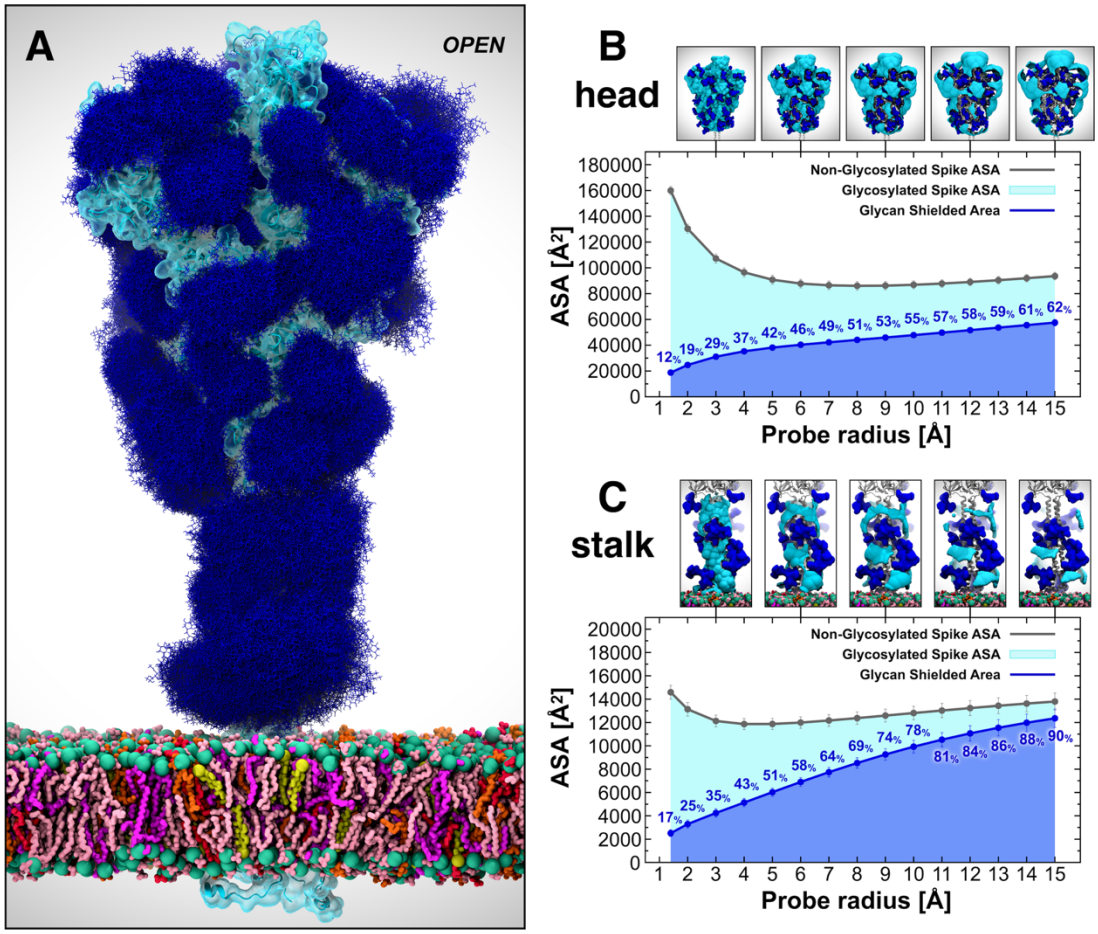


Figure 6. Glycan shield of the SARS-CoV-2 S protein. (**A**) Molecular representation of the Open system. Glycans at several frames (namely, 300 frames, one every 30 ns from one replica) are represented with blue lines, whereas the protein is shown with cartoons and highlighted with a cyan transparent surface. Color code used for lipid tails: POPC (pink), POPE (purple), POPI (orange), POPS (red), cholesterol (yellow). P atoms of the lipid heads are shown with green spheres. Cholesterol's O3 atoms are shown with yellow spheres. (**B-C**) Accessible surface area of the head (**B**) and stalk (**C**) and the area shielded by glycans at multiple probe radii from 1.4 (water molecule) to 15 Å (antibody-sized molecule). The values have been calculated and averaged across all replicas of Open and are reported with standard deviation. The area shielded by the glycans is presented in blue (rounded % values are reported), whereas the gray line represents the accessible area of the protein in the absence of glycans. Highlighted in cyan is the area that remains accessible in the presence of glycans, which is also graphically depicted on the structure in the panels located above the plots.

Our results indicate that the head is overall less shielded by the glycans than the stalk at all probe radii (**Figures 6B** and **6C**). Interestingly, the stalk is almost completely inaccessible to large molecules such as antibodies, with glycan coverage equal to 90% of the protein accessible area for 15-Å-radius probe. On the contrary, the head represents an easier target as its glycan camouflaging is insufficient (62%) to cover its larger surface area, which includes one RBD in the "up" conformation. When smaller probes are screened (1.4–3 Å radius), the stalk and the head result to be similarly shielded at an average of 26% and 20%, respectively, suggesting that small molecules can equally penetrate either areas. ASA average and standard deviation values for the head and stalk in Open are provided in **Tables S6** and **S7**, respectively. Overall, taking glycosylation into account, the stalk appears as a potentially more difficult therapeutic target than the head, despite

being a highly conserved domain among betacoronaviruses. Its smaller surface is well-protected by large sialylated and fucosylated tetrantennary glycans, which are found to be almost 100% effective in shielding large molecules. Interestingly, glycans at N1174 and N1198 have been found to always be tetrantennary, 100% fucosylated (N1174 and N1198), and 100% sialylated (N1198).[11,15] However, small drugs might still interfere with the fusion process by binding to the HR2 domain of the stalk. In contrast, the head shows a higher vulnerability, which can be harnessed for vaccine development.

Glycan Shield of the Receptor Binding Domain

As discussed in the previous section, the glycan shield plays a critical role in hiding the S protein surface from molecular recognition. However, to effectively function, the spike needs to recognize and bind to ACE2 receptors as the primary host cell infection route. For this reason, the RBM must become fully exposed and accessible.[73] In this scenario, the glycan shield works in concert with a large conformational change that allows the RBD to emerge above the N-glycan coverage. Here, we quantify the ASA of the RBM within RBD-A, corresponding to the RBD/ACE2-interacting region (residues 400–508), at various probe radii in both the Open and Closed systems (**Figures 7A** and **7D,** full data in **Tables S8-S10**.). As expected, the ASA plots show a significant difference between the "down" (Closed) and "up" (Open) RBD conformations, with the RBM area covered by glycans being remarkably larger in the former. When RBD-A is in the "up" conformation, its RBM shows an average (across all radii) of only ~9% surface area covered by glycans, compared with ~35% in the Closed system (**Figures 7A** and **7D**). This difference is further amplified when considering a larger probe radius of 15 Å, with a maximum of 11% and 46% for Open and Closed, respectively. Interestingly, for smaller probes (1.4–3 Å) the shielding becomes weak in both systems, with an average of 6% and 16% for Open and Closed, respectively.

Note that the RBD region that does not directly interact with ACE2 remains shielded by the glycans in both "up" and "down" conformations (**Figure S10**). This region is equally protected regardless of the RBD conformation mostly owing to the presence of N-glycans bound to the RBD itself at N331 and N343 and to the N-glycans at N165 and N234 (**Figure S11**). Ultimately, this analysis shows that the RBM is always accessible when RBD is "up", whereas it is very well-camouflaged when "down" (**Figure 7**). This suggests that the glycan shield of this critical domain is effectively paired with its "down-to-up" conformational change, allowing the RBM to transiently emerge from the glycan shield and bind to ACE2 receptors. Furthermore, while antibody targeting the RBD might be ineffective when the RBD is "down", small molecules could more easily circumvent the glycan coverage. This is in agreement with structural data reporting the "up" conformation as a requirement for RBD neutralization by host antibodies.[74] In this respect, several SARS-CoV-2 antibodies targeting the S protein have been identified (**Table S11**).[61,62,64–68,75–79] The majority of these antibodies recognizes epitopes on the RBD, while only a few have been shown to address other antigenic regions within the NTD and CD (**Figure S12**). Among the RBD antibodies, B38 interacts with the RBM at the RBD/ACE2 interface,[64] whereas S309 and CR3022 target the side/bottom part of the RBD.[61,62,65] In addition, 4A8 and 1A9 have been found to engage with the NTD and CD, respectively.[67,68] Analysis of the glycan shield effective on these epitopes is provided in **Section 3** of SI.

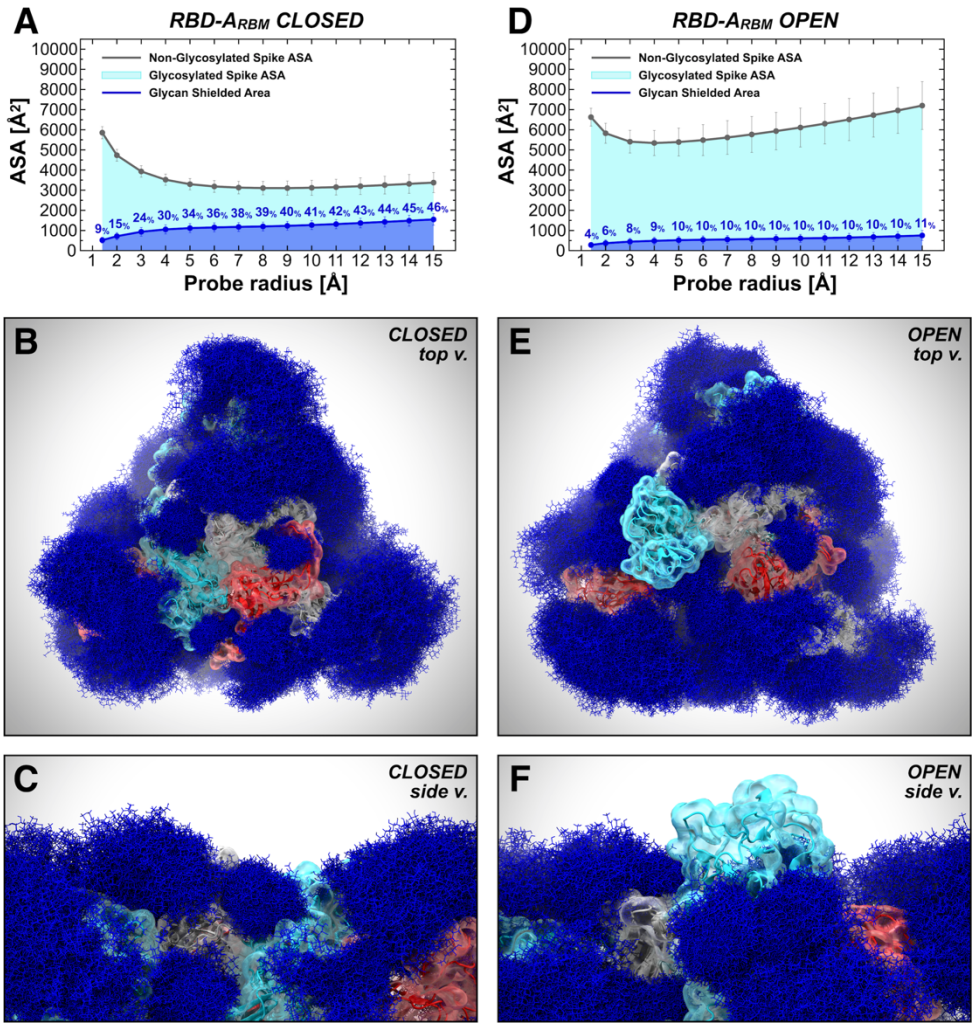


Figure 7. Glycan shield of the RBD ACE2-interacting region. The accessible surface area of the RBM-A and the area shielded by neighboring glycans in the Closed (**A**) and Open (**D**) systems are plotted at multiple probe radii from 1.4 (water molecule) to 15 Å (antibody-sized molecule). The values have been averaged across replicas and are reported with standard deviation. In blue is the area of the RBM-A covered by the glycans (rounded % values are reported), whereas the gray line is the accessible area in the absence of glycans. Highlighted in cyan is the RBM-A area that remains accessible in the presence of glycans, which is also graphically depicted on the structure in the panels located below the plots. (**B-F**) Molecular representation of Closed and Open systems from top (**B** and **E**, respectively) and side (**C** and **F**, respectively) views. Glycans (blue lines) are represented at several frames equally interspersed along the trajectories (300 frames along 0.55 ns for Closed and 1.0 µs for Open), while the protein is shown with colored cartoons and transparent surface (cyan, red and gray for chains A, B and C, respectively). Importantly, in panel **E** and **F**, RBD within chain A (cyan) is in the 'up' conformation and emerges from the glycan shield.

Conclusions

Our work presents multiple microsecond-long, all-atom, explicitly solvated MD simulations of the full-length model of the glycosylated SARS-CoV-2 S protein embedded in a viral membrane. We show how the time-averaged glycan shield covers a vast amount of the S protein surface area and how it changes depending on the conformational state of the protein between open and closed states. Interestingly, our simulated models reveal a role – beyond shielding – for N-

glycans at positions N165 and N234 as modulators of the RBD conformational plasticity. Simulations of the N165A/N234A and of N234A mutants of the S protein highlight their critical structural role in stabilizing the RBD "up" conformation. To confirm our simulation results, biolayer interferometry experiments conducted on N165A and N234A variants show a reduced ACE2 binding when these glycans are removed, revealing an RBD conformational shift towards the "down" state, with a larger effect for N234A. Overall, our work sheds new light on the full structure of this critical target and points to opportunities and challenges for small molecules and vaccine design. We make available through the NSF MolSSI and BioExcel COVID-19 Molecular Structure and Therapeutics site (https://covid.molssi.org/) our models to enable other groups to use and explore this dynamic system in atomic detail.[80]

Supporting Information

The Supporting Information file (.pdf) contains the following: Material and methods, antibody accessibility analysis, additional simulations of the spike's head model, supplementary figures and tables, captions of Movies S1-S3.

Corresponding Author

*Rommie E. Amaro (ramaro@ucsd.edu).

Author Contributions

[+]L.C and Z.G. equally contributed to this work. R.E.A. and E.F. designed and oversaw the research project. L.C, Z.G., built the full-length spike models and performed the MD simulations. A.C.D. built the membrane bilayer. L.C., Z.G., R.E.A performed the simulation analyses. A.C.D. performed the analyses of the membrane bilayer. L.C., Z.G. created the figures. L.C. made movies S1 and S3. A.M.H., C.A.F., E.F. built the spike's head model and performed analyses discussed in Section 3 of SI. A.M.H., C.A.F., E.F. made movie S2. J.S.M. designed and oversaw biolayer interferometry experiments. J.A.G. and C.K.H. performed the experiments and wrote the corresponding results and methods sections. L.C., Z.G., R.E.A., E.F. wrote the paper with contributions from all authors.

Funding Sources

This work was supported by NIH GM132826, NSF RAPID MCB-2032054, an award from the RCSA Research Corp., a UC San Diego Moore's Cancer Center 2020 SARS-COV-2 seed grant, and the Irish Research Council. LC is funded by a Visible Molecular Cell Consortium fellowship. This work was supported in part by NIH grant R01-AI127521 to J.S.M.

Acknowledgements

We are grateful for the efforts of the Texas Advanced Computing Center (TACC) Frontera team and for the compute time made available through a Director's Discretionary Allocation (made possible by the National Science Foundation award OAC-1818253), and also to the Irish Centre for High-End Computing (ICHEC) for computational resources. We thank Prof. Michael Feig and Dr. Lim Heo (Michigan State University), Prof. Syma Khalid (University of Southampton), Prof. Carlos Simmerling and his research group (SUNY Stony Brook), Prof. Ben Neuman (Texas A&M University), Prof. Greg Voth and Dr. Viviana Monje-Galvan (University of Chicago), Prof. Adrian Mulholland and his research group (University of Bristol), Prof. Julien Michel (University of

Edinburgh) and his research group, Prof. Jean-Philip Piquemal (Sorbonne University), and Dr. Reda Rawi (NIH Vaccine Research Center) for system structure checks and helpful discussions.

References

(1) Chan, J. F. W.; Yuan, S.; Kok, K. H.; To, K. K. W.; Chu, H.; Yang, J.; Xing, F.; Liu, J.; Yip, C. C. Y.; Poon, R. W. S.; Tsoi, H. W.; Lo, S. K. F.; Chan, K. H.; Poon, V. K. M.; Chan, W. M.; Ip, J. D.; Cai, J. P.; Cheng, V. C. C.; Chen, H.; Hui, C. K. M.; Yuen, K. Y. A Familial Cluster of Pneumonia Associated with the 2019 Novel Coronavirus Indicating Person-to-Person Transmission: A Study of a Family Cluster. *Lancet* **2020**, *395* (10223), 514–523. https://doi.org/10.1016/S0140-6736(20)30154-9.

(2) Chan, J. F.-W.; Kok, K.-H.; Zhu, Z.; Chu, H.; To, K. K.-W.; Yuan, S.; Yuen, K.-Y. Genomic Characterization of the 2019 Novel Human-Pathogenic Coronavirus Isolated from a Patient with Atypical Pneumonia after Visiting Wuhan. *Emerg. Microbes Infect.* **2020**, *9* (1), 221–236. https://doi.org/10.1080/22221751.2020.1719902.

(3) Cheng, V. C. C.; Lau, S. K. P.; Woo, P. C. Y.; Yuen, K. Y. Severe Acute Respiratory Syndrome Coronavirus as an Agent of Emerging and Reemerging Infection. *Clin. Microbiol. Rev.* **2007**, *20* (4), 660 LP – 694. https://doi.org/10.1128/CMR.00023-07.

(4) Cui, J.; Li, F.; Shi, Z. L. Origin and Evolution of Pathogenic Coronaviruses. *Nature Reviews Microbiology*. Nature Publishing Group March 1, 2019, pp 181–192. https://doi.org/10.1038/s41579-018-0118-9.

(5) Remembering the Nearly 100,000 Lives Lost to Coronavirus in America - The New York Times https://www.nytimes.com/interactive/2020/05/24/us/us-coronavirus-deaths-100000.html (accessed May 25, 2020).

(6) McKee, M.; Stuckler, D. If the World Fails to Protect the Economy, COVID-19 Will Damage Health Not Just Now but Also in the Future. *Nature Medicine*. Nature Research April 9, 2020, pp 640–642. https://doi.org/10.1038/s41591-020-0863-y.

(7) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N. H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181* (2), 271-280.e8. https://doi.org/10.1016/j.cell.2020.02.052.

(8) Davidson, A. D.; Williamson, M. K.; Lewis, S.; Shoemark, D.; Carroll, M. W.; Heesom, K. J.; Zambon, M.; Ellis, J.; Lewis, P. A.; Hiscox, J. A.; Matthews, D. A. Characterisation of the Transcriptome and Proteome of SARS-CoV-2 Reveals a Cell Passage Induced in-Frame Deletion of the Furin-like Cleavage Site from the Spike Glycoprotein. *Genome Med.* **2020**, *12* (1), 68. https://doi.org/10.1186/s13073-020-00763-0.

(9) Belouzard, S.; Chu, V. C.; Whittaker, G. R. Activation of the SARS Coronavirus Spike Protein via Sequential Proteolytic Cleavage at Two Distinct Sites. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (14), 5871–5876. https://doi.org/10.1073/pnas.0809524106.

(10) Apellániz, B.; Huarte, N.; Largo, E.; Nieva, J. L. The Three Lives of Viral Fusion Peptides. *Chemistry and Physics of Lipids*. Elsevier Ireland Ltd 2014, pp 40–55. https://doi.org/10.1016/j.chemphyslip.2014.03.003.

(11) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-Specific Glycan Analysis of the SARS-CoV-2 Spike. *Science* **2020**. https://doi.org/10.1126/science.abb9983.

(12) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, 1–12. https://doi.org/10.1016/j.cell.2020.02.058.

(13) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, 1–11. https://doi.org/10.1016/j.cell.2020.03.045.

(14) Alejandra Tortorici, M.; Walls, A. C.; Lang, Y.; Wang, C.; Li, Z.; Koerhuis, D.; Boons, G. J.; Bosch, B. J.; Rey, F. A.; de Groot, R. J.; Veesler, D. Structural Basis for Human Coronavirus Attachment to Sialic Acid Receptors. *Nat. Struct. Mol. Biol.* **2019**, *26* (6), 481–489. https://doi.org/10.1038/s41594-019-0233-y.

(15) 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.; Crispin, M. Vulnerabilities in Coronavirus Glycan Shields despite Extensive Glycosylation. *Nat. Commun.* **2020**, *11* (1), 2688. https://doi.org/10.1038/s41467-020-16567-0.

(16) Watanabe, Y.; Bowden, T. A.; Wilson, I. A.; Crispin, M. Exploitation of Glycosylation in Enveloped Virus Pathobiology. *Biochimica et Biophysica Acta - General Subjects*. Elsevier B.V. October 1, 2019, pp 1480–1497. https://doi.org/10.1016/j.bbagen.2019.05.012.

(17) Raman, R.; Tharakaraman, K.; Sasisekharan, V.; Sasisekharan, R. Glycan–Protein Interactions in Viral Pathogenesis. *Current Opinion in Structural Biology*. Elsevier Ltd October 1, 2016, pp 153–162. https://doi.org/10.1016/j.sbi.2016.10.003.

(18) Doores, K. J.; Bonomelli, C.; Harvey, D. J.; Vasiljevic, S.; Dwek, R. A.; Burton, D. R.; Crispin, M.; Scanlan, C. N. Envelope Glycans of Immunodeficiency Virions Are Almost Entirely Oligomannose Antigens. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (31), 13800–13805. https://doi.org/10.1073/pnas.1006498107.

(19) Crispin, M.; Ward, A. B.; Wilson, I. A. Structure and Immune Recognition of the HIV Glycan Shield. *Annu. Rev. Biophys.* **2018**, *47* (1), 499–523. https://doi.org/10.1146/annurev-biophys-060414-034156.

(20) Stewart-Jones, G. B. E.; Soto, C.; Lemmin, T.; Chuang, G. Y.; Druz, A.; Kong, R.; Thomas, P. V.; Wagh, K.; Zhou, T.; Behrens, A. J.; Bylund, T.; Choi, C. W.; Davison, J. R.; Georgiev, I. S.; Joyce, M. G.; Kwon, Y. Do; Pancera, M.; Taft, J.; Yang, Y.; Zhang, B.; Shivatare, S. S.; Shivatare, V. S.; Lee, C. C. D.; Wu, C. Y.; Bewley, C. A.; Burton, D. R.; Koff, W. C.; Connors, M.; Crispin, M.; Baxa, U.; Korber, B. T.; Wong, C. H.; Mascola, J. R.; Kwong, P. D. Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. *Cell* **2016**, *165* (4), 813–826. https://doi.org/10.1016/j.cell.2016.04.010.

(21) Yang, M.; Huang, J.; Simon, R.; Wang, L. X.; MacKerell, A. D. Conformational Heterogeneity of the HIV Envelope Glycan Shield. *Sci. Rep.* **2017**, *7* (1), 1–15. https://doi.org/10.1038/s41598-017-04532-9.

(22) Wei, X.; Decker, J. M.; Wang, S.; Hui, H.; Kappes, J. C.; Wu, X.; Salazar-Gonzalez, J. F.; Salazar, M. G.; Kilby, J. M.; Saag, M. S.; Komarova, N. L.; Nowak, M. A.; Hahn, B. H.; Kwong, P. D.; Shaw, G. M. Antibody Neutralization and Escape by HIV-1. *Nature* **2003**, *422* (6929), 307–312. https://doi.org/10.1038/nature01470.

(23) Go, E. P.; Herschhorn, A.; Gu, C.; Castillo-Menendez, L.; Zhang, S.; Mao, Y.; Chen, H.; Ding, H.; Wakefield, J. K.; Hua, D.; Liao, H.-X.; Kappes, J. C.; Sodroski, J.; Desaire, H. Comparative Analysis of the Glycosylation Profiles of Membrane-Anchored HIV-1 Envelope Glycoprotein Trimers and Soluble Gp140. *J. Virol.* **2015**, *89* (16), 8245–8257. https://doi.org/10.1128/jvi.00628-15.

(24) Cunningham, A. L.; Harman, A. N.; Donaghy, H. DC-SIGN "AIDS" HIV Immune Evasion and Infection. *Nat. Immunol.* **2007**, *8* (6), 556–558. https://doi.org/10.1038/ni0607-556.
(25) Goncalves, A.-R.; Moraz, M.-L.; Pasquato, A.; Helenius, A.; Lozach, P.-Y.; Kunz, S. Role of DC-SIGN in Lassa Virus Entry into Human Dendritic Cells. *J. Virol.* **2013**, *87* (21), 11504–11515. https://doi.org/10.1128/jvi.01893-13.
(26) Seitz, C.; Casalino, L.; Konecny, R.; Huber, G.; Amaro, R. E.; McCammon, J. A. Multiscale Simulations Examining Glycan Shield Effects on Drug Binding to Influenza Neuraminidase. *bioRxiv* **2020**, 2020.08.12.248690. https://doi.org/10.1101/2020.08.12.248690.
(27) Bonomelli, C.; Doores, K. J.; Dunlop, D. C.; Thaney, V.; Dwek, R. A.; Burton, D. R.; Crispin, M.; Scanlan, C. N. The Glycan Shield of HIV Is Predominantly Oligomannose Independently of Production System or Viral Clade. *PLoS One* **2011**, *6* (8). https://doi.org/10.1371/journal.pone.0023521.
(28) Altman, M. O.; Angel, M.; Košík, I.; Trovão, N. S.; Zost, S. J.; Gibbs, J. S.; Casalino, L.; Amaro, R. E.; Hensley, S. E.; Nelson, M. I.; Yewdell, J. W. Human Influenza a Virus Hemagglutinin Glycan Evolution Follows a Temporal Pattern to a Glycan Limit. *MBio* **2019**. https://doi.org/10.1128/mBio.00204-19.
(29) Behrens, A. J.; Crispin, M. Structural Principles Controlling HIV Envelope Glycosylation. *Current Opinion in Structural Biology*. Elsevier Ltd June 1, 2017, pp 125–133. https://doi.org/10.1016/j.sbi.2017.03.008.
(30) Geijtenbeek, T. B. H.; Kwon, D. S.; Torensma, R.; Van Vliet, S. J.; Van Duijnhoven, G. C. F.; Middel, J.; Cornelissen, I. L. M. H. A.; Nottet, H. S. L. M.; KewalRamani, V. N.; Littman, D. R.; Figdor, C. G.; Van Kooyk, Y. DC-SIGN, a Dendritic Cell-Specific HIV-1-Binding Protein That Enhances Trans-Infection of T Cells. *Cell* **2000**, *100* (5), 587–597. https://doi.org/10.1016/S0092-8674(00)80694-7.
(31) Shajahan, A.; Supekar, N. T.; Gleinich, A. S.; Azadi, P. Deducing the N-and O-Glycosylation Profile of the Spike Protein of Novel Coronavirus SARS-CoV-2. *Glycobiology 2020*, 1–8. https://doi.org/10.1093/glycob/cwaa042.
(32) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM Structure of the 2019-NCoV Spike in the Prefusion Conformation. *Science* **2020**, *1263* (March), 1260–1263. https://doi.org/10.1126/science.abb2507.
(33) Šali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234* (3), 779–815. https://doi.org/10.1006/jmbi.1993.1626.
(34) Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nat. Protoc.* **2010**, *5* (4), 725–738. https://doi.org/10.1038/nprot.2010.5.
(35) Shulla, A.; Gallagher, T. Role of Spike Protein Endodomains in Regulating Coronavirus Entry. *J. Biol. Chem.* **2009**, *284* (47), 32725–32734. https://doi.org/10.1074/jbc.M109.043547.
(36) Thorp, E. B.; Boscarino, J. A.; Logan, H. L.; Goletz, J. T.; Gallagher, T. M. Palmitoylations on Murine Coronavirus Spike Proteins Are Essential for Virion Assembly and Infectivity. *J. Virol.* **2006**, *80* (3), 1280–1289. https://doi.org/10.1128/jvi.80.3.1280-1289.2006.
(37) Park, S. J.; Lee, J.; Qi, Y.; Kern, N. R.; Lee, H. S.; Jo, S.; Joung, I.; Joo, K.; Lee, J.; Im,

W. CHARMM-GUI Glycan Modeler for Modeling and Simulation of Carbohydrates and Glycoconjugates. *Glycobiology* **2019**. https://doi.org/10.1093/glycob/cwz003.

(38) Jo, S.; Song, K. C.; Desaire, H.; MacKerell, A. D.; Im, W. Glycan Reader: Automated Sugar Identification and Simulation Preparation for Carbohydrates and Glycoproteins. *J. Comput. Chem.* **2011**. https://doi.org/10.1002/jcc.21886.

(39) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 2967–2970. https://doi.org/10.1002/jcc.

(40) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder toward Realistic Biological Membrane Simulations. *Journal of Computational Chemistry*. 2014. https://doi.org/10.1002/jcc.23702.

(41) Huang, J.; Mackerell, A. D. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**. https://doi.org/10.1002/jcc.23354.

(42) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5* (9), 2353–2370. https://doi.org/10.1021/ct900242e.

(43) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kal??, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *Journal of Computational Chemistry*. 2005, pp 1781–1802. https://doi.org/10.1002/jcc.20289.

(44) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**. https://doi.org/10.1016/0263-7855(96)00018-5.

(45) McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, *109* (8), 1528–1532. https://doi.org/10.1016/j.bpj.2015.08.015.

(46) Ghysels, A.; Krämer, A.; Venable, R. M.; Teague, W. E.; Lyman, E.; Gawrisch, K.; Pastor, R. W. Permeability of Membranes in the Liquid Ordered and Liquid Disordered Phases. *Nat. Commun.* **2019**, *10* (1), 1–12. https://doi.org/10.1038/s41467-019-13432-7.

(47) Shahane, G.; Ding, W.; Palaiokostas, M.; Orsi, M. Physical Properties of Model Biological Lipid Bilayers: Insights from All-Atom Molecular Dynamics Simulations. *J. Mol. Model.* **2019**, *25* (3), 1–13. https://doi.org/10.1007/s00894-019-3964-0.

(48) Saito, H.; Morishita, T.; Mizukami, T.; Nishiyama, K.-I.; Kawaguchi, K.; Nagao, H. Molecular Dynamics Study of Binary POPC Bilayers: Molecular Condensing Effects on Membrane Structure and Dynamics. *J. Phys* **2018**, 12022. https://doi.org/10.1088/1742-6596/1136/1/012022.

(49) Venable, R. M.; Brown, F. L. H.; Pastor, R. W. Mechanical Properties of Lipid Bilayers from Molecular Dynamics Simulation. *Chemistry and Physics of Lipids*. Elsevier Ireland Ltd November 1, 2015, pp 60–74. https://doi.org/10.1016/j.chemphyslip.2015.07.014.

(50) Janosi, L.; Gorfe, A. A. Simulating POPC and POPC/POPG Bilayers: Conserved Packing and Altered Surface Reactivity. *J. Chem. Theory Comput.* **2010**, *6* (10), 3267–3273. https://doi.org/10.1021/ct100381g.

(51) Berkowitz, M. L. Detailed Molecular Dynamics Simulations of Model Biological Membranes Containing Cholesterol. *Biochimica et Biophysica Acta - Biomembranes*. Elsevier January 1, 2009, pp 86–96. https://doi.org/10.1016/j.bbamem.2008.09.009.

(52) Róg, T.; Pasenkiewicz-Gierula, M. Cholesterol Effects on a Mixed-Chain Phosphatidylcholine Bilayer: A Molecular Dynamics Simulation Study. *Biochimie* **2006**, *88* (5), 449–460. https://doi.org/10.1016/j.biochi.2005.10.005.
(53) Harbison, A. M.; Brosnan, L. P.; Fenlon, K.; Fadda, E. Sequence-to-Structure Dependence of Isolated IgG Fc Complex Biantennary N-Glycans: A Molecular Dynamics Study. *Glycobiology* **2019**, *29* (1), 94–103. https://doi.org/10.1093/glycob/cwy097.
(54) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.
(55) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29* (4), 622–655. https://doi.org/10.1002/jcc.20820.
(56) Walls, A. C.; Park, Y. J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181* (2), 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058.
(57) Henderson, R.; Edwards, R. J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S. M. C.; Kopp, M.; Li, D.; Parks, R.; Hsu, A. L.; Borgnia, M. J.; Haynes, B. F.; Acharya, P. Controlling the SARS-CoV-2 Spike Glycoprotein Conformation. *Nat. Struct. Mol. Biol.* **2020**, 1–9. https://doi.org/10.1038/s41594-020-0479-4.
(58) Song, W.; Gui, M.; Wang, X.; Xiang, Y. Cryo-EM Structure of the SARS Coronavirus Spike Glycoprotein in Complex with Its Host Cell Receptor ACE2. *PLOS Pathog.* **2018**, *14* (8), e1007236.
(59) Zhou, T.; Tsybovsky, Y.; Olia, A. S.; Gorman, J.; Rapp, M. A.; Cerutti, G.; Katsamba, P. S.; Nazzari, A.; Schon, A.; Wang, P. D.; Bimela, J.; Shi, W.; Teng, I.-T.; Zhang, B.; Boyington, J. C.; Chuang, G.-Y.; Sampson, J. M.; Sastry, M.; Stephens, T.; Stuckey, J.; Wang, S.; Friesner, R. A.; Ho, D. D.; Mascola, J. R.; Shapiro, L.; Kwong, P. D. A PH-Dependent Switch Mediates Conformational Masking of SARS-CoV-2 Spike. *bioRxiv Prepr. Serv. Biol.* **2020**, 2020.07.04.187989. https://doi.org/10.1101/2020.07.04.187989.
(60) 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.; McLellan, J. S. Immunogenicity and Structures of a Rationally Designed Prefusion MERS-CoV Spike Antigen. *Proc. Natl. Acad. Sci. U. S. A.* **2017**. https://doi.org/10.1073/pnas.1707304114.
(61) Yuan, M.; Wu, N. C.; Zhu, X.; Lee, C.-C. D.; So, R. T. Y.; Lv, H.; Mok, C. K. P.; Wilson, I. A. A Highly Conserved Cryptic Epitope in the Receptor-Binding Domains of SARS-CoV-2 and SARS-CoV. *Science (80-.).* **2020**. https://doi.org/10.1126/science.abb7269.
(62) 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.; Stuart, D. I. Neutralization of SARS-CoV-2 by Destruction of the Prefusion Spike. *Cell Host Microbe* **2020**. https://doi.org/10.1016/j.chom.2020.06.010.
(63) 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.; McLellan, J. S. Structure-Based Design of Prefusion-Stabilized SARS-CoV-2 Spikes. *Science (80-.).* **2020**, eabd0826. https://doi.org/10.1126/science.abd0826.

(64) Wu, Y.; Wang, F.; Shen, C.; Peng, W.; Li, D.; Zhao, C.; Li, Z.; Li, S.; Bi, Y.; Yang, Y.; Gong, Y.; Xiao, H.; Fan, Z.; Tan, S.; Wu, G.; Tan, W.; Lu, X.; Fan, C.; Wang, Q.; Liu, Y.; Zhang, C.; Qi, J.; Gao, G. F.; Gao, F.; Liu, L. A Noncompeting Pair of Human Neutralizing Antibodies Block COVID-19 Virus Binding to Its Receptor ACE2. *Science (80-.).* **2020**, eabc2241. https://doi.org/10.1126/science.abc2241.

(65) Pinto, D.; Park, Y.; Beltramello, M.; Walls, A. C.; Tortorici, M. A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; Marco, A. De; 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.; Corti, D. Cross-Neutralization of SARS-CoV-2 by a Human Monoclonal SARS-CoV Antibody. *Nature* **2020**, 1–10. https://doi.org/10.1038/s41586-020-2349-y.

(66) Wang, C.; Li, W.; Drabek, D.; Okba, N. M. A.; van Haperen, R.; Osterhaus, A. D. M. E.; van Kuppeveld, F. J. M.; Haagmans, B. L.; Grosveld, F.; Bosch, B. J. A Human Monoclonal Antibody Blocking SARS-CoV-2 Infection. *Nat. Commun.* **2020**, *11* (1), 1–6. https://doi.org/10.1038/s41467-020-16256-y.

(67) Chi, X.; Yan, R.; Zhang, J.; Zhang, G.; Zhang, Y.; Hao, M.; Zhang, Z.; Fan, P.; Dong, Y.; Yang, Y.; Chen, Z.; Guo, Y.; Zhang, J.; Li, Y.; Song, X.; Chen, Y.; Xia, L.; Fu, L.; Hou, L.; Xu, J.; Yu, C.; Li, J.; Zhou, Q.; Chen, W. A Neutralizing Human Antibody Binds to the N-Terminal Domain of the Spike Protein of SARS-CoV-2. *Science (80-.).* **2020**, eabc6952. https://doi.org/10.1126/science.abc6952.

(68) Zheng, Z.; Monteil, V. M.; Maurer-Stroh, S.; Yew, C. W.; Leong, C.; Mohd-Ismail, N. K.; Cheyyatraivendran Arularasu, S.; Chow, V. T. K.; Lin, R. T. P.; Mirazimi, A.; Hong, W.; Tan, Y.-J. Monoclonal Antibodies for the S2 Subunit of Spike of SARS-CoV-1 Cross-React with the Newly-Emerged SARS-CoV-2. *Eurosurveillance* **2020**, *25* (28), 2000291. https://doi.org/10.2807/1560-7917.ES.2020.25.28.2000291.

(69) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 Spike Protein Glycan Shield: Implications for Immune Recognition. https://doi.org/10.1101/2020.04.07.030445.

(70) Jiang, S.; Hillyer, C.; Du, L. Neutralizing Antibodies against SARS-CoV-2 and Other Human Coronaviruses. *Trends in Immunology*. 2020. https://doi.org/10.1016/j.it.2020.03.007.

(71) Ng, S.; Nachbagauer, R.; Balmaseda, A.; Stadlbauer, D.; Ojeda, S.; Patel, M.; Rajabhathor, A.; Lopez, R.; Guglia, A. F.; Sanchez, N.; Amanat, F.; Gresh, L.; Kuan, G.; Krammer, F.; Gordon, A. Novel Correlates of Protection against Pandemic H1N1 Influenza A Virus Infection. *Nat. Med.* **2019**. https://doi.org/10.1038/s41591-019-0463-x.

(72) Urbanowicz, R. A.; Wang, R.; Schiel, J. E.; Keck, Z.; Kerzic, M. C.; Lau, P.; Rangarajan, S.; Garagusi, K. J.; Tan, L.; Guest, J. D.; Ball, J. K.; Pierce, B. G.; Mariuzza, R. A.; Foung, S. K. H.; Fuerst, T. R. Antigenicity and Immunogenicity of Differentially Glycosylated Hepatitis C Virus E2 Envelope Proteins Expressed in Mammalian and Insect Cells. *J. Virol.* **2019**, *93* (7). https://doi.org/10.1128/jvi.01403-18.

(73) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural Basis for the Recognition

of SARS-CoV-2 by Full-Length Human ACE2. *Science (80-.).* **2020**, *367* (6485), 1444–1448. https://doi.org/10.1126/science.abb2762.

(74) Gui, M.; Song, W.; Zhou, H.; Xu, J.; Chen, S.; Xiang, Y.; Wang, X. Cryo-Electron Microscopy Structures of the SARS-CoV Spike Glycoprotein Reveal a Prerequisite Conformational State for Receptor Binding. *Cell Res.* **2017**. https://doi.org/10.1038/cr.2016.152.

(75) 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.; van Gils, M. J. Potent Neutralizing Antibodies from COVID-19 Patients Define Multiple Targets of Vulnerability. *Science (80-.).* **2020**, eabc5902. https://doi.org/10.1126/science.abc5902.

(76) Zost, S. J.; Gilchuk, P.; Case, J. B.; Binshtein, E.; Chen, R. E.; Nkolola, J. P.; Schäfer, A.; Reidy, J. X.; Trivette, A.; Nargi, R. S.; Sutton, R. E.; Suryadevara, N.; Martinez, D. R.; Williamson, L. E.; Chen, E. C.; Jones, T.; Day, S.; Myers, L.; Hassan, A. O.; Kafai, N. M.; Winkler, E. S.; Fox, J. M.; Shrihari, S.; Mueller, B. K.; Meiler, J.; Chandrashekar, A.; Mercado, N. B.; Steinhardt, J. J.; Ren, K.; Loo, Y. M.; Kallewaard, N. L.; McCune, B. T.; Keeler, S. P.; Holtzman, M. J.; Barouch, D. H.; Gralinski, L. E.; Baric, R. S.; Thackray, L. B.; Diamond, M. S.; Carnahan, R. H.; Crowe, J. E. Potently Neutralizing and Protective Human Antibodies against SARS-CoV-2. *Nature* **2020**, 1–7. https://doi.org/10.1038/s41586-020-2548-6.

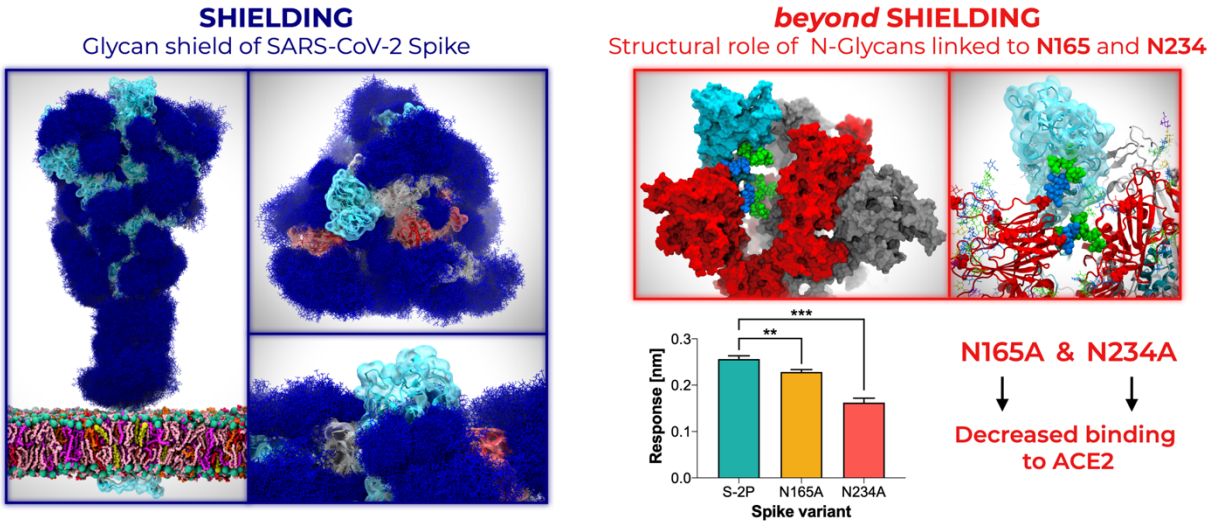
(77) Tai, W.; Zhang, X.; He, Y.; Jiang, S.; Du, L. Identification of SARS-CoV RBD-Targeting Monoclonal Antibodies with Cross-Reactive or Neutralizing Activity against SARS-CoV-2. *Antiviral Res.* **2020**, *179*, 104820. https://doi.org/10.1016/j.antiviral.2020.104820.

(78) Wrapp, D.; De Vlieger, D.; Corbett, K. S.; Schepens, B.; Saelens, X.; Mclellan, J. S. Article Structural Basis for Potent Neutralization of Betacoronaviruses by Single-Domain Camelid Antibodies. *Cell* **2020**, *181*, 1004-1015.e15. https://doi.org/10.1016/j.cell.2020.04.031.

(79) Ejemel, M.; Li, Q.; Hou, S.; Schiller, Z. A.; Wallace, A. L.; Amcheslavsky, A.; Yilmaz, N. K.; Toomey, J. R.; Schneider, R.; Close, B. J.; Chen, D.-Y.; Conway, H. L.; Saeed, M.; Cavacini, L. A.; Klempner, M. S.; Schiffer, C. A.; Wang, Y. IgA MAb Blocks SARS-CoV-2 Spike-ACE2 Interaction Providing Mucosal Immunity. *bioRxiv* **2020**, 2020.05.15.096719. https://doi.org/10.1101/2020.05.15.096719.

(80) Amaro, R. E.; Mulholland, A. J. A Community Letter Regarding Sharing Biomolecular Simulation Data for COVID-19. *J. Chem. Inf. Model.* **2020**, 0–6. https://doi.org/10.1021/acs.jcim.0c00319.

For Table of Contents Use Only



Synopsis

The glycan shield is a sugary barrier that helps the viral SARS-CoV-2 spikes to evade the immune system. Beyond shielding, two of the spike's glycans are discovered to prime the virus for infection.