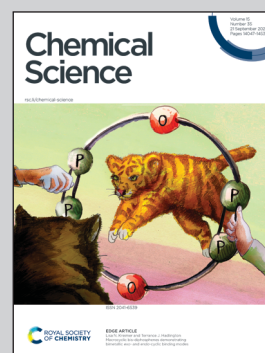**Showcasing research from Professor Jean-Philip Piquemal's group, Department of Chemistry, Sorbonne Université, Paris, France, and from Qubit Pharmaceuticals, Paris, France.**

Water–glycan interactions drive the SARS-CoV-2 spike dynamics: insights into glycan-gate control and camouflage mechanisms

To elucidate the molecular mechanisms of SARS-CoV-2 spike dynamics, we performed μs-long allatom AMOEBA high-resolution polarizable adaptive sampling molecular dynamics simulations, zooming into the interaction interface in its both open and closed states. Our findings revealed a sophisticated protein-solvent-glycan polarization network essential for stabilizing the open state. Additionally, we demonstrated that the glycan shield consistently preserves viral camouflage across both conformations, effectively protecting the virus from immune detection. These insights not only deepen our understanding of viral behavior but also pave the way for further exploration of viral evasion strategies.

As featured in:

See Louis Lagardère, Jean-Philip Piquemal *et al.*, *Chem. Sci.*, 2024, **15**, 14177.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemical-science

Check for updates

Water–glycan interactions drive the SARS-CoV-2 spike dynamics: insights into glycan-gate control and camouflage mechanisms†

Marharyta Blazhynska, [ID] [a] Louis Lagardère,*[a] Chengwen Liu,[bc] Olivier Adjoua,[a] Pengyu Ren[b] and Jean-Philip Piquemal [ID] *[a]

To develop therapeutic strategies against COVID-19, we introduce a high-resolution all-atom polarizable model capturing many-body effects of protein, glycan, solvent, and membrane components in SARS-CoV-2 spike protein open and closed states. Employing μs-long molecular dynamics simulations powered by high-performance cloud-computing and unsupervised density-driven adaptive sampling, we investigated the differences in bulk-solvent–glycan and protein–solvent–glycan interfaces between these states. We unraveled a sophisticated solvent–glycan polarization interaction network involving the N165/N343 glycan-gate patterns that provide structural support for the open state and identified key water molecules that could potentially be targeted to destabilize this configuration. In the closed state, the reduced solvent polarization diminishes the overall N165/N343 dipoles, yet internal interactions and a reorganized sugar coat stabilize this state. Despite variations, our glycan–solvent accessibility analysis reveals the glycan shield capability to conserve constant interactions with the solvent, effectively camouflaging the virus from immune detection in both states. The presented insights advance our comprehension of viral pathogenesis at an atomic level, offering potential to combat COVID-19.

1   Introduction

The emergence of the COVID-19 pandemic[1] has underscored the critical importance of comprehensively understanding the biochemical nature of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Despite widespread collaborative efforts within the scientific community to develop treatments and vaccines,[2,3] many aspects of the virus behavior remain poorly understood. Notably, elucidating the structural dynamics of the spike protein[4–7] is of paramount importance for discerning its functional properties and identifying viable targets for the design of new therapeutic strategies.

The coronavirus uses spike envelope glycoproteins (S-proteins comprising subunits S1 and S2) for receptor recognition in order to bind with the host, leading to membrane fusion, as well as to the entry into the host cell to initiate the viral infection.[5,6,8–10] Structurally, the spike protein features a trimeric configuration, with each protomer composed of an N-terminal

domain (NTD), a receptor-binding domain (RBD), a central helix (CH), a fusion peptide (FP), a connecting domain (CD), a heptad repeat (HR), a transmembrane domain (TM), and cytoplasmic tail (CT) (Fig. 1A and B). It is important to point out that the spike protein exhibits distinct prefusion conformations, namely the closed and open states.[4] In the closed state, the spike protein is characterized by concealed RBDs (Fig. 1C), therefore limiting their interaction with the host-cell receptors. Conversely, the open state features one (Fig. 1D) or multiple exposed RBDs, accessible for the binding to host cells. The transition from the closed to the open state is triggered by the binding of the spike protein to host cell receptors, such as the angiotensin-converting enzyme 2 (ACE2), allowing the viral genome to initiate its viral replication and, therefore, the infection.[9,10] Besides RBDs, glycans surround the SARS-CoV-2 spike protein and play a crucial role in modulating its structure and function, influencing various aspects of viral pathogenesis.[5,6,8,11–13] While not being directly encoded in the viral genetic sequence,[14] glycans act as a protective barrier, camouflaging the underlying protein structure from the host immune system, thus promoting the virus infection.[12,15,16] Among experimental studies on elucidating the role of spike–glycan interactions,[9,10,12,13,15–24] the work of Huang *et al.* demonstrated that enzymatic removal of glycans from the SARS-CoV-2 spike protein enhances immune responses and protection in animal models, further highlighting the significance of glycans in viral pathogenesis and vaccine development.[22]

[a]Laboratoire de Chimie Théorique, Sorbonne Université, UMR 7616 CNRS, 75005 Paris, France. E-mail: louis.lagardere@sorbonne-universite.fr; jean-philip.piquemal@sorbonne-universite.fr

[b]Department of Biomedical Engineering, The University of Texas at Austin, Texas 78712, USA

[c]Qubit Pharmaceuticals, 75014 Paris, France

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4sc04364b
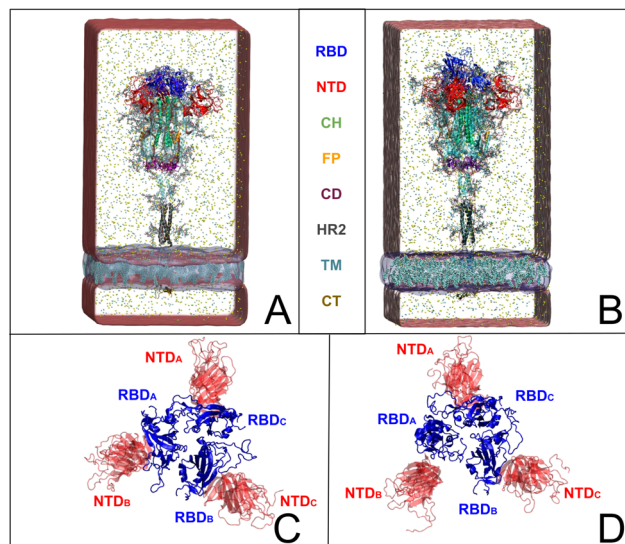
Fig. 1 SARS-CoV-2 full spike protein structure shown in (A and C) closed and (B and D) open (one RBD$_A$ up-conformation) states embedded in a water box and the viral membrane with cholesterol molecules (depicted as transparent red and blue rectangles, respectively). The receptor binding domain (RBD), the N-terminal domain (NTD), the central helix (CH), the fusion peptide (FP), the connector domain (CD), the heptad repeat 2 (HR2), the transmembrane domain (TM) and the cytoplasmic tail (CT) are represented in blue, red, green, orange, mauve, gray, cyan, and ochre, respectively. Top views (C and D) illustrate the structural reorganization of immune recognition sites on the RBD and NTD between closed (C) and open (D) states.

Additionally, significant computational efforts have been dedicated to uncover intricate interactions between the spike protein and surrounding glycans.[5,6,8,25] Thus, employing molecular dynamics (MD) simulations, Amaro et al.[8] showed that during the RBD opening, the glycan present at N234 undergoes an inward rotation, filling the resultant void and contributing to the stabilization of the open configuration. Subsequently, their recent MD investigation, complemented by cryo-electron microscopy and biolayer interferometry experiments[25] revealed that the glycans positioned at N343, N234, and N165 play a pivotal role in spike opening by sequentially interacting with multiple residues within the RBD, a mechanism referred to as "glycan gating". It should be noted that in their studies, Amaro et al.[8,25] utilized classical non-polarizable MD simulations tuning the conformational sampling to get insight into spike opening and offering valuable insights into spike-glycan dynamics. Despite being all-atom, the non-polarizable classical MD-based approaches may not fully capture some important mechanisms of charge redistribution in complex molecular systems; therefore, potential gaps in understanding the complete spectrum of spike protein behavior should be considered. Particularly, classical MD simulations may inadequately depict weak intermolecular interactions involving many-body effects,[26] requiring the use of polarizable force fields[27] for more nuanced exploration. In this context, the use of High-Performance Computing (HPC) via a distributed cloud infrastructure in conjunction with advanced sampling

algorithms offers a promising avenue for the high-resolution description of intricate dynamics, especially for large and convoluted biological systems, such as the SARS-CoV-2 spike protein.

To address these questions, we leveraged the Tinker-HP simulation package[28] employing its GPU-accelerated implementation,[29] allowing us to conduct a series of all-atom MD simulations using the AMOEBA polarizable force field.[27,30] Furthermore, our simulations rely on extensive density-driven adaptive conformational sampling[31] of both open and closed configurations provided by Amaro et al.[8] It is worth noting that our study marks the first realization of a parametrization of all the components of the full-length spike protein including glycans, water solvent, counter-ions, and a membrane bilayer with a polarizable force field, thus allowing us to access a more comprehensive and accurate representation of the spike complex interactions.

Our findings revealed subtle differences in protein–glycan interactions between the spike protein closed and open states. Notably, while the overall integrity of the glycan shield remained remarkably consistent in both forms, the mechanisms governing local interactions varied. In the open state, highly polarizable water molecules created dynamic hydrogen bonding that mediated specific protein–glycan interactions, particularly at residues N343 and N165, likely contributing to stabilization of the spike opening. In contrast, the closed state features loosely packed, low-polarization water molecules at the protein–glycan interface, suggesting that the stability of the protein structure in this state is maintained due to the inner reorganization of the protein and stabilization arising from the interacting glycans. The presented insights highlight the intricate balance between the SARS-CoV-2 state and the solvent interactions, indicating a sophisticated viral adaptation toward evading immune detection. Overall, our study provides a foundation for the development of targeted therapeutics aimed at disrupting viral infectiousness and enhancing host immune recognition, therefore offering promising avenues for fighting COVID-19 and other viral diseases.

2 Computational assays

2.1 Structural data and parametrization

The starting structure files of open and closed states (6VSB[8,32] and 6VXX[8,33] Protein Data Bank (PDB) ids, respectively) were taken from the Amaro Lab site,[34] comprising the full-length spike protein (58 634 atoms) enveloped with 70 glycans (14 236 atoms), the membrane of POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine), POPI (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoinositol), POPS (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylserine), POPE (1-palmitoyl-2-oleoyl-phosphatidyl-ethanolamine), cholesterol (165 743 atoms in total), water molecules (1 451 508 atoms), and sodium (1647) and chloride ions (1366) creating a physiological concentration of 150 mM. The total rectangular box size was $204.7 \times 199.4 \times 408.5$ Å^3. The protein, water, and counterions were parameterized by using the AMOEBA polarizable force field by means of PDBXYZ utility available in the TINKER software package. More guidelines

about the TINKER source code can be found in ref. 35. The AMOEBA parameters for the rest of the system (*i.e.*, lipids and glycans) were generated using Poltype2, a tool designed for automated parameterization of small molecules. The final parameters for the glycans used in our study have been included in the parameter files for both the open and closed states, available in ESI data files S3 and S4.† For detailed parameterization steps and outcomes specific to these glycans, refer to Data S2 in the ESI.† Although these glycan parameters are not yet part of the general AMOEBA force field, users can generate similar parameters using Poltype2 if needed. Additional information about Poltype2 is available in the public GitHub repository[36] and in ref. 37 and 38.

2.2 Advanced molecular dynamics simulation details

To conduct an extensive unsupervised adaptive sampling simulation of open and closed states, we initiated our study with 10 ns conventional MD simulations employing the AMOEBA polarizable force field. Each system underwent an initial minimization process using a limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm until achieving a Root Mean Square (RMS) gradient of 1 kcal mol^{-1}. Satisfied with the resultant models for both states, we proceeded with the extensive adaptive sampling.

For each MD simulation, we resorted to the BAOAB-RESPA1 integrator with an outer timestep of 10 fs,[39] as implemented in the Tinker-HP software package,[39] which is available for public usage in the GitHub repository[40] or in the Tinker-HP web site.[41] The integration scheme incorporated three time step sizes: a short time step of 1 fs, an intermediate time step of 4 fs and the largest one of 10 fs. The short time step was used to accurately capture the fastest motions within the system, while the intermediate time step addressed medium-frequency motions and interactions.

Our sampling technique, elaborately described in ref. 31, is a multi-iterative approach tailored for execution on large distributed computing resources such as supercomputers equipped with hundreds of GPU cards. Initially, we conducted nine independent 10 ns simulations, each starting from the same conformations for both states but with varied initial velocities. Subsequently, the resulting structures from these simulations were aligned with those obtained from the study by Casalino *et al.*[8] Further fully automated structure selections were made by means of principal component analysis (PCA) based on their density in a low-dimensional space, prioritizing exploration of less-explored regions. Multiple 1.5 ns simulations were then launched from these initial structures, with each simulation state recorded at regular intervals. Cumulatively, 340 micro-iterations were performed to obtain a total simulation time of 510 ns for each state, allowing us to capture a significantly broader conformational landscape compared to traditional single trajectory sampling methods. It should be noted that for each iteration we computed a debiasing score to remove the statistical influence of the selection strategy described above, and therefore allowing us to compute accurate and unbiased observables. Table 1 provides detailed

information about the adaptive sampling process, including the number of micro-iterations, the duration of each micro-iteration, and the associated speed of the simulations.

2.3 Adaptive sampling on the cloud

In recent years, cloud computing has emerged as an invaluable resource for handling computing demands, with providers such as Amazon Web Services (AWS), Google Cloud, and Microsoft Azure leading the way in offering flexible and secure services. Unlike traditional supercomputing clusters, where resources are allocated on an annual or bi-annual basis following rigorous scientific planning, cloud computing offers greater flexibility in resource allocation. Despite potentially higher associated costs, cloud computing provides researchers with an alternative that is well-suited to dynamic project requirements. With the support of AWS specialists, we successfully deployed a cloud-based supercomputer, utilizing Slurm for job submission management, as depicted in Fig. 2. The architecture of this supercomputer consisted of a Slurm head node based on a C5 large instance, with compute nodes utilizing spot or on-demand Nvidia GPUs (P3 or P4D instances). The P3 instances, configured with 4 V100 GPUs per node, enabled 51 micro-iterations to run in parallel, utilizing a total of 408 V100 GPUs simultaneously (*i.e.*, for both open and closed states). Although P4D instances equipped with 8 A100 GPUs per node were available, they were not employed for the AS. The setup also included customized Amazon Machine Images (AMIs) with essential pre-installed libraries. Additionally, a dedicated post-processing server equipped to execute Python-based scripts was included. Notably, all instances were interconnected through a shared Elastic File System (EFS). The distributed nature of the adaptive sampling strategy aligns seamlessly with cloud computing, as it operates without the need for synchronization during molecular dynamics trajectories. Such inherent adaptability renders the strategy independent of the specific computing resources it utilizes, making it well-suited for deployment within cloud computing environments.

3 Principal component analysis of the SARS-CoV-2 spike protein

To gain insights into the structural dynamics governing the open and closed states of the SARS-CoV-2 spike protein, we employed the principal component analysis (PCA). Leveraging the solute-only trajectories shared by the Amaro laboratory[34] and following their PCA calculation strategy discussed in detail in ref. 8, we replicated their PCA spaces for both open and closed conformations. Subsequently, we projected our extensively sampled simulations onto such spaces (Fig. 3). It should be noted that in the work presented by Casalino *et al.*,[8] the Amaro laboratory implemented an enhanced sampling strategy solely for the open state, whereas exploration of the closed state was carried out through production simulations. Despite the limited conformational sampling in the latter, our analysis revealed an emergence of multiple intersecting clusters in their simulations, indicating the conformational heterogeneity

Table 1 Details of the adaptive sampling process

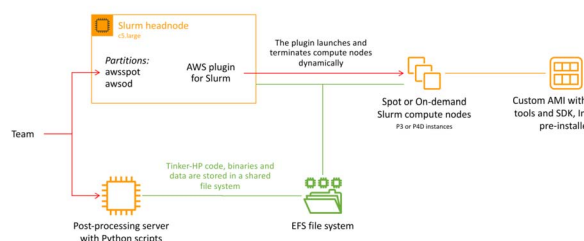| # of macro-iterations	# of micro-iterations	Length (ns/micro-iteration)	Speed (ns per day)
1	10	1.5	1.6
2	25		
3	25		
4	50		
5	50		
6	45		
7	45		
8	45		
9	45		

Fig. 2 Schematic representation of the cloud-based virtual cluster used for this study, enabling the use of AWSSpot and AWSod (on demand) partitions and leveraging a shared elastic file system.
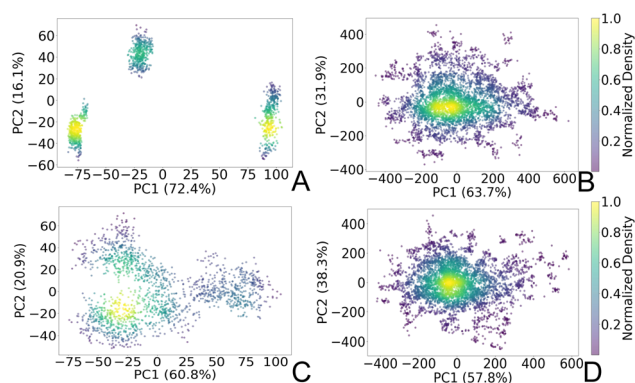

Fig. 3 PCA plots representing two principal components (PC1 and PC2) of the SARS-CoV-2 spike C_α atoms in (A) open state reproduced from the Amaro laboratory enhanced sampling simulations from three replicas, (B) open state projected to Amaro laboratory PCA space (this work), (C) closed state reproduced from Amaro laboratory production simulations from three replicas, and (D) closed state projected to Amaro laboratory PCA space (this work). The points are colored based on the normalized probability densities obtained from kernel density estimation. The amount of variance of each PC is shown in %. Images are generated with the Matplotlib Python library.[42]

within this state (Fig. 3C). Furthermore, our extensive sampling strategy applied to the closed state (Fig. 3D) enabled the delineation of distinct clusters corresponding to both highly visited regions of high probability densities (0.6–1.0), alongside capturing rare conformations of lower densities (0–0.4) within

this state. In contrast, a significant difference in the PCA spaces was observed in the comparison with open-state simulations. While Amaro's data showed distinct, non-interacting clusters for this state (see Fig. 3A), our analysis depicted more intertwined clusters (Fig. 3B). Our analysis unveiled stable structural states of the protein (Fig. 3B and D), providing valuable insights into the spike dynamic behavior. It should be noted that similar improvements in configuration space sampling over conventional MD have been observed previously by our group (see ref. 31), thus underscoring the effectiveness of the implemented adaptive sampling strategy coupled with a polarizable force field.

4 Analysis of protein and glycan dynamics

To validate our models against those documented by Casalino *et al.*,[8] we complemented our analysis by examining the root mean square fluctuations (RMSFs) of protein C_α atoms which elucidate the protein dynamic behavior in open and closed states (Fig. 4A). Our results revealed that the RMSF profiles derived from simulations utilizing the polarizable force field exhibit diminished fluctuations compared to the classical MD data.[8] The observed variation likely arises from differences in how electrostatic interactions within the protein are captured, particularly in how they impact the arrangement and stability of charged groups near the C_α atoms. Indeed, classical force fields, which typically employ fixed charges assigned to atoms, tend to oversimplify these interactions since they do not consider
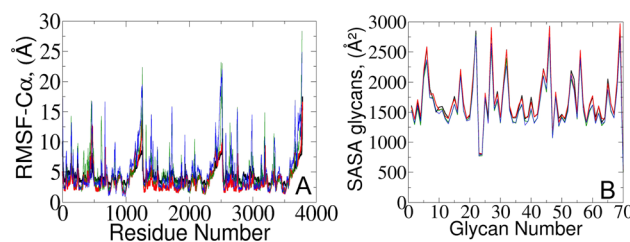

Fig. 4 (A) RMSF of protein C_α atoms and (B) averaged SASA of 70 surrounding glycans obtained from the current study (represented in red and black) and from simulations reported by the Amaro lab[8] (green and blue) for open and closed states, respectively.

many-body polarization effects. The application of the polarizable force field provides a route to include these effects *via* a more accurate representation.[27,30,43]

Additional hydrogen-bonding and salt-bridge analysis (see Fig. S1–S6 and Data S1 in the ESI†) provides further protein structural details. From our polarizable simulations, we identified only 27 common stable inner hydrogen bonds present in both the open and closed states. The closed state exhibited 77 unique stable hydrogen bonds and 4 unique salt bridges, while the open state had 45 unique stable hydrogen bonds and only 2 unique salt bridges. Such disparity indicates that the inner protein structure of the closed state is more extensively stabilized by interactions between its residues compared to the open state, suggesting a more rigid structure, which may contribute to its structural integrity until the fusion event unfolds. Conversely, the open state, with fewer stabilizing interactions, appears more plastic and adaptable, facilitating processes such as receptor binding and viral entry.

To gain further insights into potential triggers modulating SARS-CoV-2 protein conformational changes, we investigated the average solvent-accessible surface area (SASA) of the glycans covering the protein surface. The total SASA for each glycan was accumulated across all simulation frames, and the average value was determined by dividing the total SASA by the total number of frames (Fig. 4B). Remarkably, our analysis of surrounding glycans showed similar average SASA profiles in both classical and polarizable simulations. Regardless of the conformational state and the protein structural changes along the trajectories, the glycans maintained a stable level of exposure to the surrounding solvent molecules. Such stability suggests the formation of a consistent and robust protective barrier, effectively camouflaging[17–19,23] the underlying protein structure from the external environment. Prior literature underscores the crucial role of glycan camouflage in CoV spike proteins (*i.e.*, SARS-CoV-2), impacting host attachment, immune responses, and virion.[11,16–19,44,45] However, it was previously unknown that the glycan covering persists unchanged across different conformational states of the spike protein, suggesting a sophisticated strategy employed by the virus to limit accessibility to the spike and potentially escape detection by the immune system.

It should be noted that despite similar glycan SASA profiles, the interaction patterns at the protein–glycan interface differ between the two states. To gain insights into glycan spatial dynamics around the spike protein, we conducted analyses of glycan root mean square deviation (RMSD) and their radial positioning relative to the protein central axis (Fig. 5A and B). The RMSD analysis revealed lower values for glycans in the closed state, indicating a more rigid and tighter packing around the protein, which likely enhances its structural integrity. Glycan radial distance analysis (Fig. 5C), along with protein–glycan dynamic cross-correlation (Fig. S7†) and relative distance evaluation (Fig. S9†), supported a closer association between glycans and the protein surface at immune recognition sites (*i.e.*, RBD and NTD) in the closed state.

Given the prevalence and role of α-D-mannose residues in viral attachment,[8,23,46–48] we also evaluated their contribution to
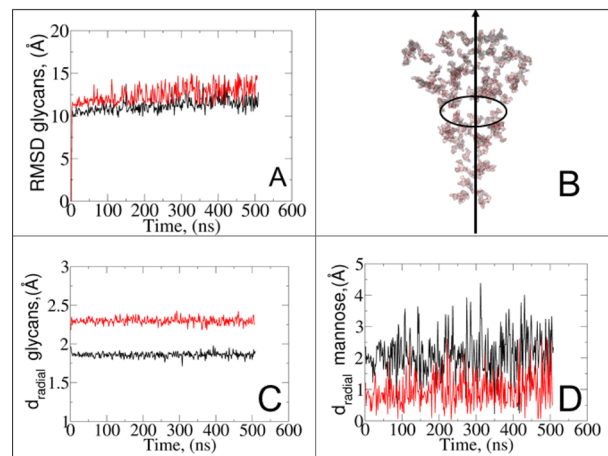

Fig. 5 (A) RMSD of all surrounding spike glycans excluding hydrogen atoms. (B) Schematic representation of the radial distance of glycans from the z-axis of the protein, which is used for the graphs in panels (C) and (D). (C) Radial distances of all glycans relative to the protein. (D) Contribution of α-D-mannose residues in the total glycan radial distances. Red and black colors indicate the open and closed states, respectively.

the glycan radial motion. In the closed state, the radial distances of these residues closely matched the total glycan radial distances (Fig. 5D), indicating their crucial role in the stabilization of the closed form. Conversely, in the open state, the smaller contribution of α-D-mannose residues (∼1 Å) to the total radial motion compared to that of all glycans (∼2.4 Å) suggests their lesser involvement in overall glycan dynamics. However, their closer radial arrangement in the open state may indicate their reorganization near the protein surface, potentially influencing viral infectivity, immune recognition, and evasion mechanisms. Our observation aligns with recent studies[49–59] that explored lectin-based inhibitors targeting high-mannose residues on coronaviruses, suggesting a viable strategy for preventing viral entry and enhancing immune recognition.

5 Solvent interactions at the glycan–protein interface through radial distribution function analysis

To explore the dynamics at the glycan–protein interface, we investigated its interactions with the solvent using a radial distribution function (RDF) analysis. Initially, we computed two distinct RDFs: one between water and protein oxygen atoms, and another between water and glycan oxygen atoms. While the total intensity values of $g(r)$ were small, our focus remained on qualitative aspects. Remarkably, the RDF profiles exhibited similarities between the open and closed states (Fig. 6A and B), mirroring the consistent glycan SASA patterns (Fig. 4B). We further refined the analysis by splitting the RDF between water and glycan oxygen atoms (Fig. 6B) into two parts: glycan oxygens adjacent to the protein interface (Fig. 6C) and those closer to the bulk solvent (Fig. 6D).
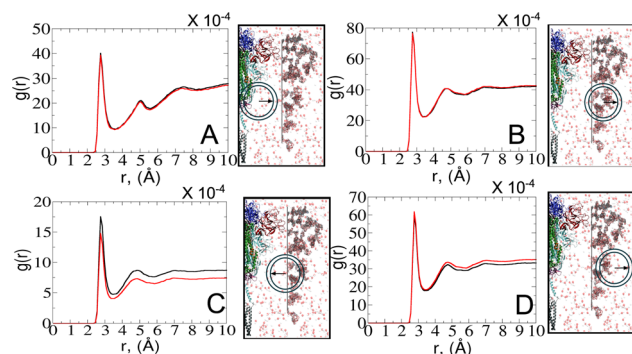
Fig. 6 RDFs depict the interactions between water–protein and water–glycan oxygen atoms in the closed (black) and open (red) states. Each RDF plot is accompanied by a schematic representation on the right side of the graph, denoting interfaces from left to right: protein, interfacial solvent, glycan, and bulk solvent. The positions of the radial circles where the RDF calculations were performed are shown. (A) RDF between water and protein oxygen atoms. (B) RDF between water and glycan oxygen atoms. (C) RDF between water and glycan oxygen atoms adjacent to the protein interface. (D) RDF between water and glycan oxygen atoms closer to the bulk solvent.

An interesting observation emerged from the RDF of water–glycan oxygen pairs oriented towards the protein interface: the closed state exhibited a more pronounced RDF compared to the open state, indicating a potentially more permissive arrangement of interfacial water molecules. Such a finding suggests that the closed conformation, characterized by a predominance of glycans near the protein surface (see Fig. S7–S9†), may create a distinct microenvironment that influences the dynamics of water molecules differently compared to the open state. Conversely, the denser packing of water molecules in the open state tends to show a tighter interaction of the solvent with the protein–glycan interface, impacting its stability and functionality. The observed water behavior is further compensated for by the local bulk solvent reorganization at the glycan interface (Fig. 6D).

6 Role of polarizable water in mediating protein–glycan interactions

The disparities in water behavior at the protein–glycan interface across distinct conformational states prompted an in-depth investigation into the potential presence of solvent polarization that influences their interactions. It is important to acknowledge that only a limited number of experimental studies have explored the polarizing effect of the environment on viral activity. While investigations into enveloped viruses like SARS-CoV-2, vaccinia,[60] and influenza,[61–63] have shed light on the role of polarization in facilitating viral fusion processes and mediating proton transport, capturing the polarizable nature of water molecules in experiments poses significant challenges. Water dipoles, which are pivotal for polarization effects, are small in magnitude and transient, rendering their detection difficult.[64–67] Consequently, computational investigations into solvent polarizability become essential. Non-polarizable

classical force fields struggle to accurately capture the dynamic behavior of water molecules,[27,68] making the use of polarizable force fields crucial for accurately representing the behavior of water molecules and elucidating their role in viral processes.

Initially, we conducted a thorough examination to identify water molecules exhibiting elevated dipole moments (*i.e.*, surpassing 2.8 D (ref. 69)) and a high occupancy (*i.e.*, exceeding 30%) proximal to the protein–glycan interface. Our analysis revealed the presence of highly polarizable water molecules at only two locations: adjacent to N343 (N343$_{RBD-A}$) located in the RBD in the up-conformation with attached glycan G10, and adjacent to N165 (N165$_{NTD-B}$) located in the NTD with glycan G30, exclusively in the open state (Fig. 7). Notably, these protein residues (*e.g.*, N165 (ref. 5, 8 and 20) and N343 (ref. 5, 20 and 21)) have been established as pivotal players in the spike opening. These findings are particularly significant in the context of the critical discovery of glycan gating by Sztain *et al.*[25] While their work emphasized the importance of glycans at N343 in initiating and N234 and N165 in supporting the RBD opening, our study underscores the multifaceted nature of spike protein dynamics, integrating both glycan-mediated mechanisms and solvent interactions. Surprisingly, for N234, we observed only the reorganization of the attached high-mannose glycan G31, as depicted in contact maps comparing the open and closed states (see Fig. S9 in the ESI†), which contrasts with the solvent-mediated interactions involving N343$_{RBD-A}$ and N165$_{NTD-B}$ patterns.

Focusing on N343$_{RBD-A}$ and N165$_{NTD-B}$, we delved into the dynamics of their water-mediated interactions with glycans (*i.e.*, G10 and G30, respectively). Across all iterations of the open-state simulations, we noted clear variations in water behavior at the declared locations. Benefiting from extensive conformational sampling, we hypothesized that these dynamic water fluctuations could play an integral role in the mechanism of spike opening. We identified four key phases of water behavior, based on the average dipole moments and occupancies of the polarizable water molecules across all iterations of open-state simulations. These phases encompass bridging, clustering, replacement, and relaxation, as demonstrated on the interaction interfaces of N165-G30 and N343-G10, depicted in Fig. 8.
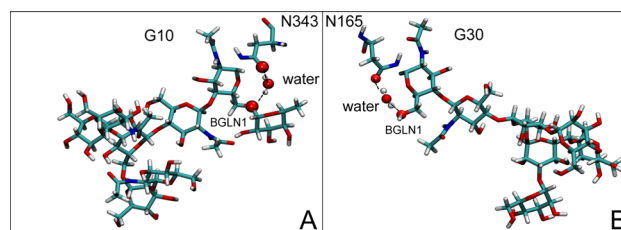

Fig. 7 Polarizable water molecule creates bridging between (A) N343$_{RBD-A}$ in the up-conformation and β-*N*-acetyl-D-glucosamine (BGLN1) of glycan G10, and (B) N165$_{NTD-B}$ and BGLN1 of glycan G30 in the open state. In β-D-glucosamine, the oxygen atoms involved in this water-bridging interaction correspond to the ether oxygen (A) or hydroxyl group (B). In asparagine residues, the oxygen atoms involved in this water-bridging interaction are located in the carboxamide group. Interacting patterns are shown in van der Waals representation.
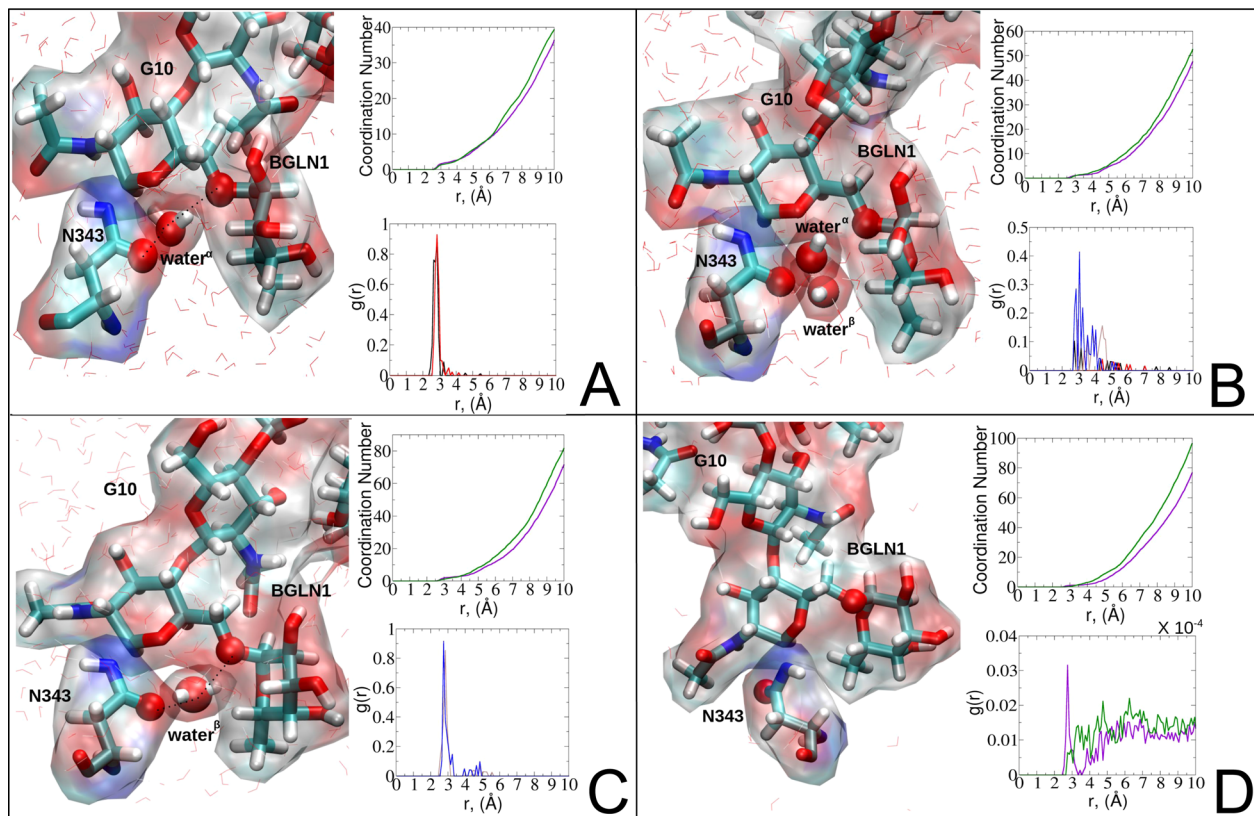
Fig. 8 Phases of dynamic polarizable water molecule bridging formation between oxygen of N343 of RBD-A in the up-conformation and oxygen of BGLN1 of glycan G10 with the corresponding pair distribution functions (PDF), $g(r)$, (black – N343(O)−water$^{\alpha}$(O), red – G10(O)−water$^{\alpha}$(O), brown – N343(O)−water$^{\beta}$(O), and blue – G10(O)−water$^{\beta}$(O)), and coordination number corresponding to the interfacial water located between N343 and BGLN1 of G10 glycan residues (violet – N343(O)−water(O) and green – G10(O)−water(O)). (A) Bridging through a water molecule, (B) regrouping, (C) replacing, and (D) relaxation. Interacting patterns are shown in van der Waals representation. To visualize the molecular surfaces of the involved moieties we employed the MSIS (molecular surface) representation[70] available *via* the Visual Molecular Dynamics (VMD) environment.[71]

In the bridging phase, a specific water molecule (water$^{\alpha}$) actively forms hydrogen-bond interaction between the protein and glycan oxygen atoms (Fig. 8A). To be classified in this phase, water$^{\alpha}$ exhibits a dipole moment that exceeds 2.8 D and an occupancy of more than 30%. To strengthen our findings, we examined the total number of water molecules proximal to the protein–glycan atoms engaged in bridging, alongside analyzing the pair distribution functions (PDFs) of water$^{\alpha}$ with these atoms. The relatively high intensity of the PDF peak, nearing unity, indicates a significant concentration of water$^{\alpha}$ at a distance of approximately 2.8 Å from the N343 and G10 oxygen atoms, corresponding to the first coordination number, underscoring a strong correlation between the number of water molecules present at that distance and the peak intensity of the PDF.

Conversely, during the clustering phase, water$^{\alpha}$ experiences a reduction in dipole moment, falling below the threshold of 2.8 D, and a decrease in occupancy, dipping below 30%, as it conglomerates with other polarizable water molecules, thereby diminishing its individual contribution to interactions. Furthermore, within this phase, another water molecule, water$^{\beta}$, exhibits a heightened dipole moment (more than 2.8 D),

positioning to potentially replace water$^{\alpha}$ (Fig. 8B). Herein, the PDFs of water$^{\alpha}$ and water$^{\beta}$ oxygen atoms appear notably noisy and less intensive compared to the bridging phase, indicating a competition between these water atoms.

During the replacement phase, water$^{\alpha}$ is finally substituted by a highly polarizable water molecule, water$^{\beta}$ that exhibits a dipole moment exceeding 2.8 D and an occupancy of more than 30%, ensuring the continuity of interactions (Fig. 8C). During this phase, water$^{\beta}$ undergoes an initial stabilization period, reflected in the PDF displaying peaks of low intensity of 4–5 Å.

Subsequently, upon achieving stability, water$^{\beta}$ facilitates a bridging interaction between the protein and glycan residues, leading to a PDF closely resembling that observed during the bridging phase, characterized by high intensity.

Finally, the relaxation phase signifies the absence of significant water interactions at the interface (Fig. 8D). Notably, comprehensive RDFs for all water molecules proximal to the protein–glycan interface affirm the lack of pronounced interactions during this phase, with the closest interfacial water oxygen observed to be ∼3.5 Å. The loss of water mediation is attributed to the reorganization of the involved protein

interacting sites, leading to changes in conformation and potentially disrupting the sites of interaction with water molecules.

To evaluate variations in the local environment and interactions experienced by the water molecules, we performed an in-depth analysis of the average dipole moments observed during each phase, along with their respective duration, across all iterations (Table 2). To compare, we also estimated the average dipole moments relative to the same group of oxygen atoms in the closed state, where no water-mediated interactions were observed. Additionally, we conducted a distribution analysis of the dipole moments for both the open and closed states separately (see Fig. S11–S14 in the ESI†). The water molecules involved in bridging interactions demonstrated higher polarity compared to other interfacial water molecules located in the proximity, thereby enhancing the stability of these interactions. As it can be noticed from Table 2, substantial variances are observed in the dipole moments of protein residues across different conformational states, indicating significant conformational changes occurring throughout the simulations, reflective of dynamic structural transitions within the protein.

Besides, a net contrast arises from the comparison between the absolute values of the dipole moments of the oxygen of the carboxamide group of glycan-gating residues N343 and N165 in the open state, which could be attributed to the differences in their local environments at the protein–solvent–glycan interface and functional roles within the protein structure. Specifically, N343 experiences a more pronounced interaction with surrounding water molecules, which is also reflected in the cumulative duration of the bridging and replacement phases (exceeding 70%), leading to higher dipole moments compared to that of N165. In contrast, the absolute values of dipole moments for both N343 and N165 oxygen atoms in the closed state remain similar, suggesting that these residues may experience comparable degrees of involvement in stabilizing the protein structure in this particular state. Moreover, in the case of G30, the oxygen atom involved in water bridging is located in the hydroxyl group. The absence of high solvent polarizing effects in the closed state may result in fewer interactions contributing to the overall dipole moment, consequently yielding a lower value compared to the open state where such interactions are more prevalent. It should be noted that the dipole moments related to the oxygen atoms of glycans G10 and

Table 2 Comparison of average dipole moments and duration of water-mediated interactions in open and closed states for N343$_{RBD-A}$ and N165$_{NTD-B}$ residues[a]

	N343$_{RBD-A}$ open state					
Phases	$\vec{\mu}$(N343), D[a]	$\vec{\mu}$(W$^\alpha$), D	$\vec{\mu}$(W$^\beta$), D	$\vec{\mu}$(W^{int}), D[b]	$\vec{\mu}$(G10), D[c]	Duration, %
Bridging	615.4 ± 8.9	2.9 ± 0.1	—	2.7 ± 0.0	49.2 ± 0.0	17
Clustering	620.0 ± 16.5	2.7 ± 0.1	2.8 ± 0.1	2.7 ± 0.0	49.5 ± 1.4	19
Replacement	616.6 ± 13.9	—	2.9 ± 0.1	2.7 ± 0.0	49.4 ± 1.2	55
Relaxation	616.3 ± 8.9	—	—	2.7 ± 0.0	49.0 ± 0.9	10

	N343$_{RBD-A}$ closed state				
$\vec{\mu}$(N343), D	$\vec{\mu}$(W$^\alpha$), D	$\vec{\mu}$(W$^\beta$), D	$\vec{\mu}$(W^{int}), D	$\vec{\mu}$(G10), D	Duration, %
568.0 ± 7.3	—	—	2.7 ± 0.1	46.6 ± 0.7	100

	N165$_{NTD-B}$ open state					
Phases	$\vec{\mu}$(N165), D[d]	$\vec{\mu}$(W$^\alpha$), D	$\vec{\mu}$(W$^\beta$), D	$\vec{\mu}$(W^{int}), D	$\vec{\mu}$(G30), D[e]	Duration, %
Bridging	651.0 ± 5.2	2.9 ± 0.1	—	2.8 ± 0.1	52.3 ± 0.4	7
Clustering	650.2 ± 8.6	2.7 ± 0.1	2.9 ± 0.1	2.8 ± 0.1	52.3 ± 0.6	8
Replacement	654.5 ± 9.3	—	2.9 ± 0.1	2.8 ± 0.1	52.6 ± 0.7	17
Relaxation	650.6 ± 8.3	—	—	2.7 ± 0.0	52.4 ± 0.6	68

	N165$_{NTD-B}$ closed state				
$\vec{\mu}$(N165), D	$\vec{\mu}$(W$^\alpha$), D	$\vec{\mu}$(W$^\beta$), D	$\vec{\mu}$(W^{int}), D	$\vec{\mu}$(G30), D	Duration, %
563.9 ± 10.5	—	—	2.7 ± 0.1	45.1 ± 0.8	100

[a] [a] and [d] correspond to the averaged atomic dipole moment in Debye of the oxygen atom of N343 and N165 residues, respectively, involved in hydrogen bonding with a polarizable water molecule, [b] refers to the interfacial water located at the proximity of 5 Å to the defined oxygen atoms of protein and glycan residues, and [c] and [e] correspond to the averaged atomic dipole moment.

G30 exhibit relatively low variances, underscoring the consistent role of glycans in shielding protein interfaces in both states, contributing to the overall stability and function of the spike.

7 Conclusion

In this study, we performed high performance molecular dynamics simulations of the SARS-CoV-2 spike protein (1 693 134 atom system) using the AMOEBA polarizable force field coupled to extensive density-driven conformational sampling. Aggregating 1.2 µs of simulation data, we extensively explored the structural dynamics of the open and closed prefusion states. Our analysis reveals that, regardless of the spike protein conformational state, the glycan shield remains stable. This stability, despite the inherent variability of glycans—both across species and within individual cells—suggests a sophisticated strategy employed by the virus, tending to indicate the mechanism of viral adaptiveness at evading immune detection, enabling its persistence and spread within the host.

Contrary to the uniform glycan behavior, distinct structural properties were observed in the spike protein. The closed state maintained greater stability through robust hydrogen bonds and salt bridges, along with a tighter association between α-D-mannose-rich glycans and key immune recognition sites. In contrast, the open state experienced a reorganization of the glycans on the protein surface, hinting at their potential role in modulating viral infectivity.

Since the virus may fine-tune its immune evasion strategies, we further zoomed into how the solvent dynamics influences protein–glycan interactions. A detailed examination of the interfacial water reorganization at the protein–glycan interface showed a more dispersed arrangement of water molecules around glycans in the closed state, compared to the open one. Based on these differences, we further investigated how the polarizable nature of water molecules influences the stability and dynamics of protein–glycan interactions. We uncovered that highly polarizable interfacial water, exclusively present in the open state, plays a pivotal role in mediating interactions at glycan-gating patterns between N343 and N165 protein residues and their corresponding glycans. Additionally, the identification of distinct phases in open-state trajectories, namely bridging, clustering, replacement, and relaxation, elucidates water-mediated mechanisms at the protein–glycan interface. In contrast, the stability of the closed state may primarily result from internal protein and glycan interface reorganization rather than solvent-driven influences.

Relying on a robust computational approach, our study underscores the adaptability of the SARS-CoV-2 spike protein and the crucial role of the glycan shield in viral survival. Disrupting this shield could enhance immune detection, offering new therapeutic avenues. Additionally, understanding solvent dynamics provides a deeper insight into the spike's opening mechanism, presenting opportunities to prevent viral entry by targeting the water-mediated interactions at the glycan-gating sites that could potentially destabilize the open state and inhibit ACE2 binding.

Data availability

All the data used in the analyses are available in the paper and/or the ESI.† Further information and requests should be directed to and will be fulfilled by the lead contact, JPP (jean-philip.piquemal@sorbonne-universite.fr).

Author contributions

Conceptualization: MB, LL, and JPP. Methodology: OA, CL, LL, PR, and JPP. Investigation: MB and LL. Visualization: MB. Supervision: LL and JPP. Writing—original draft: MB. Writing—review & editing: LL, CL, PR, and JPP.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 V. J. McGowan and C. Bambra, *Lancet Public Health*, 2022, **7**, e966–e975.

2 Y. Yuan, B. Jiao, L. Qu, D. Yang and R. Liu, *Front. Immunol.*, 2023, **14**, 1125246.

3 T. I. Ng, I. Correia, J. Seagal, D. A. DeGoey, M. R. Schrimpf, D. J. Hardee, E. L. Noey and W. M. Kati, *Viruses*, 2022, **14**, 961.

4 Y. Cai, J. Zhang, T. Xiao, H. Peng, S. M. Sterling, J. Walsh, M. Richard, S. Rawson, S. Rits-Volloch and B. Chen, *Science*, 2020, **369**, 1586–1592.

5 Y. T. Pang, A. Acharya, D. L. Lynch, A. Pavlova and J. C. Gumbart, *Commun. Biol.*, 2022, **5**, 1170.

6 M. I. Zimmerman, J. R. Porter, M. D. Ward, S. Singh, N. Vithani, A. Meller, U. L. Mallimadugula, C. E. Kuhn, J. H. Borowsky, R. P. Wiewiora, M. F. D. Hurley, A. M. Harbison, C. A. Fogarty, J. E. Coffland, E. Fadda, V. A. Voelz, J. D. Chodera and G. R. Bowman, *Nat. Chem.*, 2021, **13**, 651–659.

7 M. Wieczór, V. Genna, J. Aranda, R. M. Badia, J. L. Gelpí, V. Gapsys, B. L. de Groot, E. Lindahl, M. Municoy, A. Hospital and M. Orozco, *Wiley Interdisc. Rev. Comput. Mol. Sci.*, 2023, **13**, e1622.

8 L. Casalino, Z. Gaieb, J. A. Goldsmith, C. K. Hjorth, A. C. Dommer, A. M. Harbison, C. A. Fogarty, E. P. Barros, B. C. Taylor, J. S. McLellan, E. Fadda and R. E. Amaro, *ACS Cent. Sci.*, 2020, **6**, 1722–1734.

9 I. Berger and C. Schaffitzel, *Cell Res.*, 2020, **30**, 1059–1060.

10 M. Lu, P. D. Uchil, W. Li, D. Zheng, D. S. Terry, J. Gorman, W. Shi, B. Zhang, T. Zhou, S. Ding, R. Gasser, J. Prévost, G. Beaudoin-Bussières, S. Anand, A. Laumaea, J. R. Grover, L. Liu, D. D. Ho, J. R. Mascola, A. Finzi and W. Mothes, *Cell Host Microbe*, 2020, **28**, 880–891.

11 A. C. Walls, M. A. Tortorici, B. Frenz, *et al.*, *Nat. Struct. Mol. Biol.*, 2016, **23**, 899–905.

12 W. T. Harvey, A. M. Carabelli, B. Jackson, R. K. Gupta, E. C. Thomson, E. M. Harrison, C. Ludden, R. Reeve, A. Rambaut, COVID-19 Genomics UK (COG-UK) Consortium, S. J. Peacock and D. L. Robertson, *Nat. Rev. Microbiol.*, 2021, **19**, 409–424.

13 M. L. Newby, C. A. Fogarty, J. D. Allen, J. Butler, E. Fadda and M. Crispin, *J. Mol. Biol.*, 2023, **435**, 167928.

14 D. Campos, M. Girgis and M. Sanda, *Proteomics*, 2022, **22**, e2100322.

15 D. Chmielewski, E. A. Wilson, G. Pintilie, P. Zhao, M. Chen, M. F. Schmid, G. Simmons, L. Wells, J. Jin, A. Singharoy and W. Chiu, *Nat. Commun.*, 2023, **14**, 7175.

16 Y. Watanabe, Z. T. Berndsen, J. Raghwani, G. E. Seabright, J. D. Allen, O. G. Pybus, J. S. McLellan, I. A. Wilson, T. A. Bowden, A. B. Ward and M. Crispin, *Nat. Commun.*, 2020, **11**, 2688.

17 T. J. Yang, Y. C. Chang, T. P. Ko, P. Draczkowski, Y. C. Chien, Y. C. Chang, K. P. Wu, K. H. Khoo, H. W. Chang and S. D. Hsu, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 1438–1446.

18 Z. Sun, K. Ren, X. Zhang, J. Chen, Z. Jiang, J. Jiang, F. Ji, X. Ouyang and L. Li, *Engineering*, 2021, **7**, 1441–1451.

19 Z. Sun, X. Zheng, F. Ji, M. Zhou, X. Su, K. Ren and L. Li, *Infect. Microbes Dis.*, 2021, **3**, 149–157.

20 Y. Watanabe, J. D. Allen, D. Wrapp, J. S. McLellan and M. Crispin, *Science*, 2020, **369**, 330–333.

21 L. Zheng, Y. Ma, M. Chen, G. Wu, C. Yan and X.-E. Zhang, *Biochem. Biophys. Res. Commun.*, 2021, **579**, 69–75.

22 H.-Y. Huang, H.-Y. Liao, X. Chen, S.-W. Wang, C.-W. Cheng, M. Shahed-Al-Mahmud, Y.-M. Liu, A. Mohapatra, T.-H. Chen, J. M. Lo, Y.-M. Wu, H.-H. Ma, Y.-H. Chang, H.-Y. Tsai, Y.-C. Chou, Y.-P. Hsueh, C.-Y. Tsai, P.-Y. Huang, S.-Y. Chang, T.-L. Chao, H.-C. Kao, Y.-M. Tsai, Y.-H. Chen, C.-Y. Wu, J.-T. Jan, T.-J. R. Cheng, K.-I. Lin, C. Ma and C.-H. Wong, *Sci. Transl. Med.*, 2022, **14**, eabm0899.

23 Y.-P. Hsu, M. Frank, D. Mukherjee, V. Shchurik, A. Makarov and B. F. Mann, *Glycobiol*, 2023, **33**, 126–137.

24 A. Q. Stiving, D. J. Foreman, Z. L. VanAernum, E. Durr, S. Wang, J. Vlasak, J. Galli, J. O. Kafader, T. Tsukidate, X. Li, H. A. Schuessler and D. D. Richardson, *J. Am. Soc. Mass Spectrom.*, 2024, **35**, 62–73.

25 T. Sztain, S.-H. Ahn, A. T. Bogetti, L. Casalino, J. A. Goldsmith, E. Seitz, R. S. McCool, F. L. Kearns, F. Acosta-Reyes, S. Maji, G. Mashayekhi, J. A. McCammon, A. Ourmazd, J. Frank, J. S. McLellan, L. T. Chong and R. E. Amaro, *Nat. Chem.*, 2021, **13**, 963–968.

26 G. Klesse, S. Rao, S. J. Tucker and M. S. Sansom, *J. Am. Chem. Soc.*, 2020, **142**, 9415–9427.

27 J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson and T. Head-Gordon, *J. Phys. Chem. B*, 2010, **114**, 2549–2564.

28 L. Lagardère, L.-H. Jolly, F. Lipparini, F. Aviat, B. Stamm, Z. F. Jing, M. Harger, H. Torabifard, G. A. Cisneros, M. J. Schnieders, N. Gresh, Y. Maday, P. Y. Ren, J. W. Ponder and J.-P. Piquemal, *Chem. Sci.*, 2018, **9**, 956–972.

29 O. Adjoua, L. Lagardère, L.-H. Jolly, A. Durocher, T. Very, I. Dupays, Z. Wang, T. J. Inizan, F. Célerse, P. Ren, J. W. Ponder and J.-P. Piquemal, *J. Chem. Theory Comput.*, 2021, **17**, 2034–2053.

30 C. Zhang, C. Lu, Z. Jing, C. Wu, J.-P. Piquemal, J. W. Ponder and P. Ren, *J. Chem. Theory Comput.*, 2018, **14**, 2084–2108.

31 T. Jaffrelot Inizan, F. Célerse, O. Adjoua, D. El Ahdab, L.-H. Jolly, C. Liu, P. Ren, M. Montes, N. Lagarde, L. Lagardère, P. Monmarché and J.-P. Piquemal, *Chem. Sci.*, 2021, **12**, 4889–4907.

32 D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham and J. S. McLellan, *Science*, 2020, **367**, 1260–1263.

33 A. C. Walls, Y.-J. Park, M. A. Tortorici, A. Wall, A. T. McGuire and D. Veesler, *Cell*, 2020, **181**, 281–292.

34 R. E. Amaro, Amaro Lab – COVID-19, **https://amarolab.ucsd.edu/covid19.php**.

35 J. Podler, Jay Podler Lab – Tinker Molecular Modeling Package, **https://dasher.wustl.edu/tinker/distribution/doc/tinker-guide.pdf**.

36 C. Liu, Poltype 2: Automated Parameterization and Free Energy Prediction for AMOEBA , **https://github.com/TinkerTools/poltype2**.

37 B. Walker, C. Liu, E. Wait and P. Ren, *J. Comput. Chem.*, 2022, **43**, 1530.

38 J. J. Wu, G. Chattree and P. Ren, *Theor. Chem. Acc.*, 2012, **131**, 1138.

39 L. Lagardère, F. Aviat and J.-P. Piquemal, *J. Phys. Chem. Lett.*, 2019, **10**, 2593–2599.

40 O. Adjoua, Tinker-HP: High-Performance Massively Parallel Evolution of Tinker on CPUs & GPUs, **https://github.com/TinkerTools/Tinker-HP**.

41 J.-P. Piquemal, L. Lagardère, O. Adjoua and L.-H. Jolly, Tinker-HP: High-Performance Massively Parallel Evolution of Tinker on CPUs & GPUs, **https://tinker-hp.org/**.

42 J. D. Hunter, *Comput. Sci. Eng.*, 2007, **9**, 90–95.

43 Y. Shi, Z. Xia, J. Zhang, R. Best, C. Wu, J. W. Ponder and P. Ren, *J. Chem. Theory Comput.*, 2013, **9**, 4046–4063.

44 D. Chang and J. Zaia, *Mol. Cell. Proteomics*, 2019, **18**, 2348–2358.

45 R. de Groot, *Glycoconjugate J.*, 2006, **23**, 59–72.

46 V. Calvaresi, A. G. Wrobel, J. Toporowska, D. Hammerschmid, K. J. Doores, R. T. Bradshaw, R. B. Parsons, D. J. Benton, C. Roustan, E. Reading, M. H. Malim, S. J. Gamblin and A. Politis, *Nat. Commun.*, 2023, **14**, 1421.

47 Y. Zhang, W. Zhao, Y. Mao, Y. Chen, S. Wang, Y. Zhong, T. Su, M. Gong, D. Du, X. Lu, J. Cheng and H. Yang, *Mol. Cell. Proteomics*, 2021, **20**, 100058.

48 D. Zhou, X. Tian, R. Qi, C. Peng and W. Zhang, *Glycobiology*, 2021, **31**, 69–80.

49 A. Barre, E. J. M. Van Damme, M. Simplicien, S. Le Poder, B. Klonjkowski, H. Benoist, D. Peyrade and P. Rougé, *Cells*, 2021, **10**, 1619.

50 A. Nangarlia, F. F. Hassen, G. Canziani, P. Bandi, C. Talukder, F. Zhang, D. Krauth, E. N. Gary, D. B. Weiner, P. Bieniasz, S. Navas-Martin, B. R. O'Keefe, C. G. Ang and I. Chaiken, *Biochemistry*, 2023, **62**, 2115–2127.

51 K. Tabynov, M. Solomadin, N. Turebekov, M. Babayeva, G. Fomin, G. Yadagiri, S. Renu, T. Yerubayev, N. Petrovsky, G. J. Renukaradhya and K. Tabynov, *Sci. Rep.*, 2023, **13**, 12115.

52 V. R. Grosche, L. P. F. Souza, G. M. Ferreira, M. Guevara-Vega, T. Carvalho, R. R. D. S. Silva, K. L. R. Batista, R. P. F. Abuna, J. S. Silva, M. F. Calmon, P. Rahal, L. C. N. da Silva, B. S. Andrade, C. S. Teixeira, R. Sabino-Silva and A. C. G. Jardim, *Viruses*, 2023, **15**, 1886.

53 T. Nazmul, B. M. Lawal-Ayinde, T. Morita, R. Yoshimoto, A. Higashiura, A. Yamamoto, T. Nomura, Y. Nakano, M. Hirayama, H. Kurokawa, Y. Kitamura, K. Hori and T. Sakaguchi, *Microbiol. Immunol.*, 2023, **67**, 334–344.

54 Y. Li, S. Xu, Q. Ye, H. Chi, Z. Guo, J. Chen, M. Wu, B. Fan, B. Li, C.-F. Qin and Z. Liu, *Adv. Sci.*, 2023, **10**, e2202689.

55 M. Stravalaci, I. Pagani, H. Zhong, M. Sironi, S. Bondesan, F. Barzaghi, P. Carrera, F. Izzo, E. Zoia, B. Bottazzi, R. Asselta, G. Casari, A. Aiuti, A. Mantovani and C. Garlanda, *Immunobiology*, 2023, **228**, 152566.

56 A. Sutta, B. González-García, L. Pérez-Alós, A. Rosbjerg, P. Garred and R. Bayarri-Olmos, *Immunobiology*, 2023, **228**, 152592.

57 G. Albertini, L. Pisani, A. Bolzoni, M. Annunziata, C. Porta and L. Pastorelli, *J. Crohns Colitis*, 2024, **18**, i436.

58 J. Cramer, A. Lakkaichi, B. Aliu, R. P. Jakob, S. Klein, I. Cattaneo, X. Jiang, S. Rabbani, O. Schwardt, G. Zimmer, M. Ciancaglini, T. Abreu Mota, T. Maier and B. Ernst, *J. Am. Chem. Soc.*, 2021, **143**, 17465–17478.

59 A. Gupta, K. Yadav, A. Yadav, R. Ahmad, A. Srivastava, D. Kumar, M. A. Khan and U. N. Dwivedi, *Glycoconjugate J.*, 2024, **41**(1), 1–33.

60 R. D. M. Gray, D. Albrecht, C. Beerli, M. Huttunen, G. H. Cohen, I. J. White, J. J. Burden, R. Henriques and J. Mercer, *Nat. Microbiol.*, 2019, **4**, 1636–1644.

61 J. L. Thomaston, R. A. Woldeyes, T. Nakane, A. Yamashita, T. Tanaka, K. Koiwai, A. S. Brewster, B. A. Barad, Y. Chen, T. Lemmin, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 2017, **114**, 13357–13362.

62 J. K. Williams, D. Tietze, M. Lee, J. Wang and M. Hong, *J. Am. Chem. Soc.*, 2016, **138**, 8143–8155.

63 S. O. Shepelenko, A. S. Salnikov, S. V. Rak, E. P. Goncharova and A. B. Ryzhikov, *J. Phys.: Conf. Ser.*, 2009, **177**, 012008.

64 D. I. Svergun, S. Richard, M. H. J. Koch, Z. Sayers, S. Kuprin and G. Zaccai, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 2267–2272.

65 B. Halle and L. Nilsson, *J. Phys. Chem. B*, 2009, **113**, 8210–8213.

66 S. Mukherjee, S. Mondal and B. Bagchi, *J. Chem. Phys.*, 2017, **147**, 024901.

67 S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 20749–20752.

68 R. Fuentes-Azcatl and J. Alejandre, *J. Phys. Chem. B*, 2014, **118**, 1263–1272.

69 P. Y. Ren and J. W. Ponder, *J. Phys. Chem.*, 2003, **107**, 5933–5947.

70 M. F. Sanner, A. J. Olson and J.-C. Spehner, *Biopolymers*, 1996, **38**, 305–320.

71 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.