Two Receptor Binding Strategy of SARS-CoV-2 is Mediated by Both the N-Terminal and Receptor-Binding Spike Domain

Michele Monti*,[*,†,‡,∇] Edoardo Milanetti,[¶,‡,∇] Myrthe T. Frans,[§,∇] Mattia Miotto,[‡,∇] Lorenzo Di Rienzo,[‡,∇] Maksim V. Baranov,[§] Giorgio Gosti,[‖,‡] Arun Kumar Somavarapu,[⊥] Madhu Nagaraj,[⊥] Thaddeus W. Golbek,[#] Emiel Rossing,[@] Sam J. Moons,[@] Thomas J. Boltje,[@] Geert van den Bogaart,[§] Tobias Weidner,[#] Daniel E. Otzen,[⊥] Gian Gaetano Tartaglia,[†] Giancarlo Ruocco,[¶,‡] and Steven J. Roeters*[,#,△]

†*RNA System Biology Lab, Department of Neuroscience and Brain Technologies, Istituto Italiano di Tecnologia, Via Morego 30, 16163, Genoa, Italy*

‡*Center for Life Nanoscience, Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161, Rome, Italy*

¶*Department of Physics, Sapienza University, Piazzale Aldo Moro 5, 00185, Rome, Italy*

§*Molecular Immunology – Groningen Biomolecular Sciences and Biotechnology, Nijenborgh 7, 9747 AG, Groningen, The Netherlands*

‖*DHILab, Istituto di Scienze del Patrimonio Culturale, Sede di Roma, Consiglio Nazionale delle Ricerche, Via Salaria km 29300, Rome, 00010, Italy*

⊥*Interdisciplinary Nanoscience Center (iNANO), Aarhus University, Gustav Wieds Vej 14, 8000 Aarhus C, Denmark*

#*Department of Chemistry, Aarhus University, Langelandsgade 140, 8000, Aarhus C, Denmark*

@*Synthetic Organic Chemistry, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands*

△*Amsterdam UMC, Vrije Universiteit, Department of Anatomy and Neurosciences, De Boelelaan 1108, 1081 HZ, Amsterdam, The Netherlands*

∇*These authors contributed equally to the presented work.*

E-mail: michele.monti@iit.it; s.j.roeters@amsterdamumc.nl

Abstract

It is not well understood why SARS-CoV-2 spreads much faster than other β-coronaviruses such as SARS-CoV and MERS-CoV. In a previous publication, we predicted the binding of the N-terminal domain (NTD) of SARS-CoV-2 spike to sialic acids. Here, we experimentally validate this interaction, and present simulations that reveal a secondary possible interaction between sialic acids and the spike protein via a binding site located in the the Receptor Binding Domain (RBD). The predictions from the 2D-Zernike binding-site recognition approach and molecular-dynamics simulations were validated through flow-induced dispersion analysis (FIDA), which reveals the capability of the SARS-CoV-2 spike to bind to sialic-acid containing (glyco)lipid vesicles, and flow-cytometry measurements, which show that spike binding is strongly decreased upon inhibition of sialic-acid expression on the membranes of HEK-ACE2 cells. Our analyses reveal that the sialic-acid binding of the NTD and RBD strongly enhances the infection-inducing ACE2 binding. Altogether, our work provides *in-silico*, *in-vitro* and cellular evidence that the SARS-CoV-2 virus utilizes a two-receptor (sialic acid and ACE2) strategy. This allows the SARS-CoV-2 spike to use sialic-acid moieties on the cell membrane as a binding anchor, which increases the residence time of the virus on the cell surface, and aids binding of the main receptor, ACE2, via 2D-diffusion.

Introduction

In the past years, the exceptionally fast transmission of SARS-CoV-2 virus has led to millions of deaths all over the world.[1] Although the symptoms of COVID-19 are less severe as compared to some other β-coronaviruses, its fast viral diffusion still causes the virus to clog public health systems across the world, in spite of effective vaccines. Seven coronavirus strains have been shown to infect humans.[2,3] In the past twenty years, along with SARS-CoV-2, two other β-coronavirus have caused three of the most severe epidemics reported in the world: SARS-CoV that causes the severe acute respiratory syndrome (SARS),[4,5] and

MERS-CoV[6] that causes the Middle-East respiratory syndrome (MERS). These three viruses share a close evolutionary history,[7] but the infection and the diffusion rates are different, which indicates that the SARS-CoV-2 virus has developed a different infection mechanism and/or pathology with respect to the other species. The initial binding partners that SARS-CoV and MERS interact with for the infection of cells are different. For SARS-CoV, the receptor is the angiotensin converting enzyme-2 (ACE2). Instead, for MERS-CoV, the infection is mediated by an initial binding to nine-carbon acidic monosaccharides typically found at a terminal position of glycan chains present in the cell membrane, such as sialic acid (SA), which serve as attachment factors before binding with the main entry receptor of the virus, dipeptidyl peptidase 4 (DPP4).[2,8] These different receptors partially explain the distinct spreading efficiency of the two viruses.[9] Herein, we would like to stress that we use the term "receptor" with regards to sialic acid solely to indicate that it facilitates the initial binding, and that it thereby relays the signal to be incorporated into the biochemical pathway that follows from subsequent ACE2 binding, following the definition of receptor used in refs. 10–15 . The interaction between these viruses and human-cell receptors are being extensively studied to understand the diffusion mechanism and to explain the differences in mortality/spreading rate .[16–18] Even though the molecular and genomic differences between SARS-CoV-2 and the other β-coronaviruses are small, they have resulted in catastrophic effects to our societies in very different degrees.

In our previous work,[19] we identified how the spike of SARS-CoV-2 has evolved an N-terminal domain (NTD) sialic-acid binding pocket that is similar to the MERS-CoV one in the same genomic region, which has recently been experimentally shown to interact with sialic-acid receptors.[20,21]

SAs are a class of α-keto acid sugars with a nine-carbon backbone. This small molecular moiety is omnipresent on all cell membranes of vertebrates and some invertebrates, mostly attached to the outermost ends of lipids and proteins that constitute the surface of the cell.[22] These glycans are present in cells forming the external respiratory airways,[13] so viruses like

influenza-A[23] use it as receptor while others, like MERS-CoV[20] exploit it as an initial, less specific, but ubiquitous attachment factor to promote their infections, by reducing the three-dimensional search for the main receptor to a two-dimensional plane. The evolutionary advantage of this reduction in dimensionality relates directly to the findings obtained in our previous computational analysis based on the recently developed '2D-Zernike polynomial' formalism, which allows the characterization of the shape of portions of the molecular surface of proteins. With this method, we investigated the structural area around the SA binding site of MERS-CoV, confirming that SARS-CoV-2 has evolved a similar SA-binding pocket in the NTD. This computational analysis shows the high similarity between the SA binding patch of the MERS-CoV spike, which was selected from the X-ray structure of the complex, and the corresponding patch of the SARS-CoV-2 spike located in the same NTD region. In contrast, the same analysis performed on the SARS-CoV spike shows its inability to bind SA, since no patches with similar geometric properties were found in comparison to the MERS binding site. Importantly, the site interacting with SA is hypervariable and is mainly composed of disordered regions. The binding site belonging to the MERS spike has much more extensive disordered regions than the corresponding SARS-CoV region. Interestingly, the SARS-CoV-2 spike region has loops with an intermediate length between the MERS spike and the SARS-CoV spike. This suggests that the length of the SARS-CoV-2 loops is sufficient to bind with SA. Yet, the small differences between the MERS and SARS-CoV-2 regions could lead to a distinct binding affinity and thus to a different role of the SA receptor.

Previous studies have revealed that also the receptor binding domain (RBD) of the SARS-CoV-2 spike can bind to SA.[24] Here, we characterize and compare, both computationally and experimentally, the nature of the binding and the relative affinities of SA and the RBD and the NTD as, until now, this has not been analyzed in detail. We present an extensive theoretical and experimental study of the binding between the SARS-CoV-2 spike protein and SA. Using different experimental and computational techniques, we were able to cover the mechanism of the binding and to demonstrate its biological relevance. From

the experimental perspective, we performed experiments at different scales. We used two different techniques to characterize the strength and the biological effect of the binding between SA and spike proteins. On one hand we have used flow-induced dispersion analysis (FIDA)[25] that allows the quantification of the strength of protein–glycolipid vesicle binding, on the other hand we have used flow cytometry (FCM) to study the binding to (sialic-acid inhibitor incubated) HEK-ACE2 cells by different domains of the spike protein, to obtain a more detailed understanding of the molecular details of this biochemical process.

To quantify the binding dynamics we performed extensive MD simulations of the system. This computational analysis aims to reveal and compare the nature of the binding among the different domains at the atomic level. We ran the simulations on various spike segments, investigating the interaction of sialic acid with the whole S1 chain of the spike, and with the NTD and RBD domains.

From these results, we infer that both the NTD and the RBD are able to bind sialic-acid moieties in a way that is beneficial for ACE2 binding. The MD simulations indicate that the NTD-SA binding is slightly stronger than the RBD-SA binding, the FIDA measurements experimentally validate the importance of the NTD-SA binding, and the FCM experiments reveal that the RBD-SA binding does not result in sequestering of its ACE2 binding site, but instead leads to an increased ACE2 binding, given the strong sialic-acid ánd ACE2 dependence of the RBD binding to ACE2 transgenic HEK cells. The binding to sialic-acid moieties, omnipresent on the membranes of cells in the respiratory tract,[13] allows the virus to diffuse on the surface of the cells. The subsequent binding to its main receptor, ACE2, is strongly enhanced by this two-receptor strategy.

Methods

Molecular dynamics (MD) simulations

The molecular dynamics (MD) simulations are performed using GROMACS 2019.3.[26] The topologies of the system are built using the CHARMM-27 force field.[27] The selection of the CHARMM-27 force field for simulating protein-small molecule binding follows from its well-established reputation and compatibility with biomolecular systems.[28] CHARMM-27 is specifically parameterized for proteins and nucleic acids, ensuring accurate representation of their structural and dynamic behavior. Crucially, this force field includes comprehensive parameters for small molecules, providing a robust foundation for studying protein-small molecule interactions. Furthermore, the transferability of CHARMM force fields is advantageous, allowing for consistent and reliable simulations across various systems,[29] e.g., systems containing (only, or combinations of) lipids, proteins, nucleic acids, small molecules, et cetera. The extensive benchmarking and validation performed on CHARMM-27 parameters[30] provide confidence in the accuracy of the simulations. Finally, the force field is widely embraced in the molecular dynamics simulation community, presenting the advantage of yielding easily interpretable and comparable results.

The protein is placed in a dodecahedric simulative box, with periodic boundary conditions, filled with TIP3P water molecules.[31] For all simulated systems, we check that each atom of the proteins was at least at a distance of 1.1 nm from the box borders. Each system is then minimized with the steepest descent algorithm. Next, a relaxation of water molecules and thermalization of the system is run in NVT and NPT environments each for 0.1 ns at 2 fs time-step. The temperature is kept constant at 300 K with v-rescale algorithm;[32] the final pressure is fixed at 1 bar with the Parrinello-Rahman algorithm.[33] The LINCS algorithm[34] is used to constraint hydrogen bonds. A cut-off of 12 Å is imposed for the evaluation of short-range non-bonded interactions and the Particle Mesh Ewald method[35] is used for the long-range electrostatic interactions. The described procedure is used for all the performed

simulations. As for NTD, the simulation was carried out for 1.75 μs with a time step of 2 fs.

As the SARS-CoV-2 spike has 22 glycosylation sites, which play a crucial role in shaping the virus's ability to infect specific cell types,[36,37] we have checked if glycosylation might affect our 2D-Zernike and MD-simulation results. However, none of the specific residues identified in our study as being involved in contact with sialic acid are found in close proximity to the glycosylation sites. Consequently, it is reasonable to leave out the glycan groups in our analyses, as the identified binding regions should retain capacity to interact with sialic acid, irrespective of glycosylation.

The full-length SARS-CoV-2 spike protein was simulated starting from the X-ray structure of the complex (PDB id:6M17). We perform a 100 ns long simulation with time step 2 fs. The system is rendered electroneutral by adding 24 sodium counter ions. The water density was set to 998 kg/m^3.

To probe the interaction with sialic acid, the SARS-CoV-2 spike N-terminal domain (residue range: 16-290) is simulated in the presence of one molecule of sialic acid (Neu5Ac) in solution. We select the domain ranging from residue 16 to 290 of the A chain of the trimeric complex. The simulation is carried out for 1.75 μs with time steps of 2 fs. Likewise, the SARS-CoV-2 spike receptor binding domain (RBD, residue range: 331-524) is simulated in the presence of one molecule of sialic acid (Neu5Ac) in solution. We select this domain from residue 331-524 to match the sequence used in our experiments. The same procedure is followed for the S1 domain (residue range: 1-700) of SARS-CoV-2 spike. The simulation has been carried out for 1.75 μs with a time step of 2 fs.

Energy calculation

From the NTD simulation in complex with the SA (3 μs) we extracted a frame every 1 ns for a total of 3001 frames. On these frames, we calculated the energy using the Autodock software and the distance between the centroid of the Zernike pocket and the centroid of the SA. Then placing ourselves in an interval that goes from the nearest integer less than the

lowest energy found and the nearest integer greater than the maximum energy found, we grouped the energies at intervals of 0.5 kcal/mol and finally calculated the energy vs distance boxplot of Figure 2.

The insert on the MERS NTD binding energy distribution is instead calculated on the first 200 ns (201 frames, 1 ns step) of the MERS NTD dynamics in complex with the SA.

In addition, to estimate the binding free energy between the spike protein and the SA molecule, we use the fastDHR tool, which is an open-access web server capable of predicting the binding free energy through methods based on MM/PB(GB)SA.[38] More specifically, two binding poses extracted from the simulation of the NTD and sialic acid were selected: the first corresponding to the minimum energy configuration as calculated through the Autodock method, and the second corresponding to the pose where SA is closest to the center of the pocket predicted by the computational method based on the Zernike formalism.[39] In this procedure, we selected the ff99SB force field (with TIP3P water model) for the receptor (spike protein) and the GAFF2 force field for the ligand (SA molecule).

Flow-induced dispersion analysis (FIDA)

Flow-induced dispersion analysis (FIDA) experiments are based on the fact that the hydrodynamic radius of particles determines their laminar flow profile.[25,40] By flowing mixtures of glycolipid-containing vesicles and spike proteins through e.g. 75-μm diameter sized capillaries, one will observe different flow profiles with vs. without binding, because the bound complex has a larger hydrodynamic radius than the two unbound species. This effect is detected through fluorescently labeling of the spike proteins, recording the fluorescence at a given point along the tubing as a function of time (see the FIDA figure in the results section for a graphical representation hereof), and plotting this in a so-called Taylorgram.[41]

Sample preparation

The S1 segment (Gln14-Arg685) of the SARS-CoV-2 spike is expressed in modified human embryonic kidney (HEK) 293 cells by GenScript Biotech (NJ, USA), and the sequence and purity were checked by mass spectrometry. The S1 segments of MERS-CoV (Tyr18-Pro747) and SARS-CoV (Ser14-Leu666) are also expressed in HEK293 cells, but by Bio-Techne Ltd. (UK). We dissolve the proteins in phosphate-buffered saline (PBS, Sigma-Aldrich, MO, USA) buffer (pH 7.4) and subsequently Alexa–488 label it in a non-specific manner, via amine coupling to the exposed lysine side chains (with a 50% efficiency). A stock solution of 2.4 μM S1-Alexa488 is prepared, which is subsequently diluted to a 100 nM concentration for the FIDA experiments.

Sialic acid (N-Acetylneuraminic acid) with a >98% purity is purchased from CarboSynth (Compton, UK), and dissolved in PBS buffer and used at a 10 μM concentration during the experiment.

The lipid vesicles are prepared according to the protocol described in ref. 42. The 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) and gangliosides GM1 and GM3 (extracted from ovine brains) lipids are purchased from Avanti Polar Lipids, Inc. (AL, USA), dissolved in chloroform and 50/50 vol% chloroform/methanol, respectively, dried under an N_2 stream, and kept under light vacuum overnight. The next day the lipids were re-suspended in a PBS buffered H_2O (milli-Q) solution, both as pure DPPC and at a GM1/GM3:DPPC molar ratio of 1:9, with a total concentrations of 1.3 mM. The solutions are then placed in a 60 °C water bath for 1 hour, and subsequently extruded 21 times through a 30 nm polycarbonate membrane with a mini-extruder system (both from Avanti Polar Lipids, Inc.) placed on a 60 °C heating block, which resulted in the formation of vesicles with a ∼50 nm diameter, as ascertained by dynamic light scattering (DLS), with a PDI of ∼0.15. The vesicles are then diluted to a 50-200 μM lipid concentration for the FIDA experiments. In the concentration range 50-200 µM (corresponding to roughly 35-140 µg/ml), the volume fraction of lipids remains well below 0.1% and we expect no major changes in solution viscosity over this

range.

FIDA experiment

To characterize the binding of the S1 segment of the SARS-CoV, MERS-CoV, and SARS-CoV-2 spikes to glycolipid-containing SUVs, we perform flow-induced dispersion analysis (FIDA) experiments using a FIDAlyzer instrument (Fida Biosystems ApS, Copenhagen, Denmark), with laserinduced fluorescence (LIF) detection (ZETALIF Evolution, Picometrics, Labege, France) with an excitation wavelength of 488 nm (Melles Griot Diode laser, Picometrics), in connection with an optical high-pass filter. The sample is flowed through a standard fused silica capillary (Fida Biosystems ApS) with an inner diameter of 75 μm, an outer diameter of 375 μm, a total length of 100 cm (with 84 cm between the sample reservoirs and the detection window. The capillary temperature is set to 25°C inside the FIDAlyzer instrument, excluding the minor part connected to the LIF detector. Also the capillary inlet and sample temperatures are heated to 25°C.

The experimental protocol is then performed as follows (similar to ref. 40): first the capillary is rinsed and equilibrated prior to each sample analysis with 1 M NaOH and PBS buffer, at 3500 mbar for 45 seconds and 2 minutes, respectively. Then, the analyte sample (the vesicles) are injected at 3500 mbar for 20 seconds, after which the indicator sample (S1-Alexa488, mixed with the vesicles and the sialic-acid solution, with a pre-injection incubation time of >10 min) is injected at 50 mbar for 10 seconds (39 nL, corresponding to 1% of the capillary volume). Finally, the injected indicator sample is then flowed towards the detection point with the vesicle sample at 50 mbar for 20 minutes.

All samples were performed in duplicate, and the Taylorgrams were processed using the FIDA data analysis software (Fida Biosystems ApS, Copenhagen, Denmark).

K_D derivation

Because the FIDA data of the S1 spikes to the 1:9 GM3:DPPC vesicles does not have a sigmoidal shape, we estimate the dissociation constants (K_D) by fitting the data with the normalized binding model, according to:

$$\text{Percentage of S1 bound} = 100 * [\text{Lipid}]/(K_D + [\text{Lipid}]) \quad (1)$$

This model assumes (1) that at higher concentrations there will be a 100% occupancy, and (2) that the dose response curve has a standard slope, equal to a Hill slope (or slope factor) of 1.0. All data could satisfactorily be fitted to a simple binding isotherm; there was no statistical basis for more complicated models, e.g. to include cooperativity.

Flow cytometry (FCM)

Flow cytometry (FCM) was employed to measure the binding of the various Alexa-488-labeled spike proteins to the ACE2 receptor. The HEK293 cell line stably expressing ACE2 was cultured in the presence of the Alexa-488-labeled spike proteins as described below, forming complexes upon binding. The fluorescence emitted from these complexes was then detected as the cells passed through a laser beam in a flow cytometer.

Gating strategy

For the identification of the HEK-ACE2 cells, we first use the forward scatter area (FSC-A) vs. the side scatter area (SSC-A) signals. The FSC-A and SSC-A areas refer to the measurement of the total light intensity in the forward and side scatter detectors, providing information about the sizes of the cells as they pass through the laser beam (see flow cytometry figure in the results section, panel (a)).

After this, we use the FSC-A plotted against the forward scatter height (FSC-H; the peak intensity) signals to select all the single cells within that HEK-ACE2 cell ensemble.

This gating strategy facilitates the detection of single cells and discern them for example from events with multiple cells or cell debris (see flow cytometry figure in the results section, panel (b)).

The Alexa-488 fluorescence signal, which we record using the FL1-A channel for each event, is then directly proportional to the amount of labeled spike protein bound to each cell (see flow cytometry figure in the results section, panel (c)). The geometric mean of this fluorescence intensity (the gMFI) can thus be used to compare the binding under the various experimental circumstances.

Sialostatin experiment

The human embryonic kidney cells (HEK-293T) expressing human angiotensin-converting enzyme 2, HEK-293T-hACE2 Cell Line, NR-52511, were obtained through BEI Resources, NIAID, NIH. The transformed cells were seeded at 50,000 cells/well in a 96-wells U-bottom suspension plate (CELLSTAR), and cultured in DMEM medium supplemented with 10% FBS, 1% L-Glutamine, 1% antibiotic-antimytotic and 1% L-pyruvate for 5 days with addition of either 300 µM sialostatin or an equal volume of DMSO. At day 5, the cells were transferred to a 96-wells V-bottom plate and washed twice with PBS. Cells were subsequently stained in 25 µl/well of PBS supplemented with 0.5% BSA and 10 µg/ml of one of the following SARS-CoV-2 derived proteins or protein segments; spike glycoprotein (S), spike glycoprotein receptor binding domain (RBD) or spike glycoprotein domain S1 (S1) for 30 minutes at 4 °C. The S, S1, and RBD protein (segments) were produced under HHSN272201400008C and obtained through BEI Resources, NIAID, NIH. The full length S protein from SARS-Related Coronavirus 2 (Wuhan-Hu-1) has a C-Terminal Histidine Tag, and was recombinantly expressed in HEK293F cells. It lacks the signal sequence and contains 1196 residues (the ectodomain) of the SARS-CoV-2 spike glycoprotein; the recombinant protein was modified to remove the polybasic S1S2 cleavage site (RRAR to A; residues 682 to 685), stabilized with a pair of mutations (K986P and V987P, wild-type

numbering; GenPept: YP-009724390) and includes a thrombin cleavage site, T4 foldon trimerization domain and C-terminal hexa-histidine tag. The S1 segment was produced by transfection in human embryonic kidney HEK293 cells and purified. The S1 segment also lacks the signal sequence, contains 670 residues of the SARS-CoV-2 spike glycoprotein (amino acid residues V16 to R685), and features a C-terminal poly-histidine tag. Finally, the RBD segment contains 223 residues (R319 to F541) of the SARS-CoV-2 spike, features a C-terminal hexa-histidine tag, and this protein segment has also been recombinantly expressed in HEK293F Cells. Sialostatin, or methyl 5-(ethylcarbamado)-2,4,7,8,9-penta-O-acetyl-3,5-dideoxy3-fluoro-D-glycero-β-galacto-non-2-ulopyranosonate is an efficient metabolic inhibitor of sialyltransferases and was synthesized from commercially available sialic acid as previously reported (see ref. 43 for the most recently reported EC_{50} values and updated synthesis; it was first reported in ref. 44).

After incubation, the cells were washed twice with PBS supplemented with 0.5% BSA, followed by a primary antibody or probe staining for 30 minutes at 4 °C with either Rabbit anti-His tag antibody (Abcam, ab14923, 1:90), goat anti-ACE2 antibody (R&D Systems, AF933, 1:400) or SiaFind Pan-Specific Lectenz (LectenzBio, 1:400). Subsequently, two washes with PBS and 0.5% BSA were followed by the addition of PBS with 0.5% BSA and either Donkey-anti-Rabbit IgG (H&L) Alexa 488 (Thermo Scientific, A21206, 1:400), Donkey-anti-goat-IgG (H&L) Alexa 488 (Thermo Scientific, A11055, 1:400) or Streptavidin-488 (Thermo Scientific, S32354, 1:1000). The secondary antibody staining was incubated for 10 minutes at 4 °C. Prior to analysis on a Cytoflex S flow cytometer (Beckman Coulter), the cells were washed twice with PBS with 0.5% BSA and resuspended in PBS.

HSase experiment

The binding of the SARS-CoV-2 S, S1 and RBD to HEK-ACE2 cells upon addition of HSase was assessed using flow cytometry. In short, 100,000 HEK-ACE2 cells were seeded into a 96-wells plate and incubated with or without the addition of HSase mix (2.5 mU/mL HSase

I, 2.5 mU/mL HSase II, and 5 mU/mL HSase III; obtained from IBEX Pharmaceuticals, Inc.). The cells were subsequently stained with 20 µg/ml of the His-tagged SARS-CoV-2 S, S1 and RBD proteins (vide supra) in PBS with 0.5% BSA. After washing the cells, the cells were stained with either Rabbit anti-His tag antibody (Abcam, ab14923, 1:90) or HS4C3 anti HS antibody (used 1:10 from aliquots). The secondary antibody staining was done using Donkey-anti-Rabbit IgG (H&L) Alexa 488 and P5D4 mouse anti-histag antibody (used 1:10 from aliquots). As a tertiary antibody, Goat- anti mouse IgG – Alexa-fluor488 (Fisher Scientific catalog #: 10256302) was used. Before analysis on a Cytoflex-S flow cytometer (Beckman Coulter), the cells were washed twice and resuspended in PBS.

Results

Quantification and prediction of the binding between SARS-CoV-2 spike and sialic acid using MD simulations

In an earlier computational investigation,[19] we compared the structures of the N-terminal domain of MERS-CoV, SARS-CoV, and SARS-CoV-2 spike proteins by evaluating their similarity using a computational method that relies on so-called 2D-Zernike polynomials. This methodology allows a structural comparison of different molecular regions, and highlighting of conformational differences and similarities. Based on our findings we identified a binding region in the NTD of the SARS-CoV-2 spike for SA molecules. In that analysis, the SARS-CoV NTD was employed as a negative control, as this virus lacks SA-binding capabilities.[7] In Figure 1a-c), we show the surface of the NTD region in the three different spikes, with the region that has a high similarity with the MERS-CoV region colored in red. Notably, the finding that SARS-CoV-2 has SA-binding capabilities was subsequently experimentally validated,[21,45] thus confirming the predictive power of the 2D-Zernike method to recognize binding pockets (see Figure 1d-f).

Here, we extend our analysis to the whole S1 region of the spike, which contains the NTD

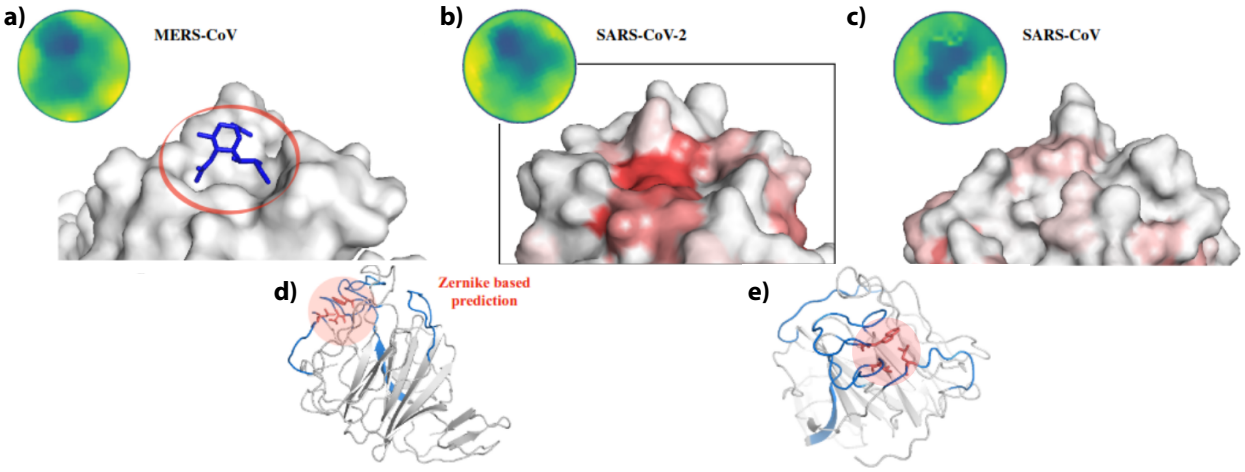


Figure 1: Characterization of a SARS-CoV-2 spike region very similar to the SA binding site on MERS-CoV spike. (a) Molecular surface representation of the N-terminal region of MERS-CoV spike bound to sialic acid, along with the corresponding representative 2D-patch disk. (b) Molecular surface representation of the N-terminal region of SARS-CoV-2, colored from white to red according to their shape similarity with the MERS-CoV binding region, along with the corresponding representative 2D-patch disk. (c) Same as in panel b, but for SARS-CoV. (d) Cartoon representation of the SARS-CoV-2 N-terminal domain with the predicted binding region highlighted in red. (e) As in d, but rotated.

and the receptor-binding domain (RBD) using MD simulations, allowing us to compare the sialic-acid binding capabilities of different spike domains. Specifically, we performed MD simulations on the full S1 segment, the RBD, and the NTD, in each case together with one molecule of SA in explicit water. No additional potential was included, so that the sialic acid molecule was free to diffuse in the aqueous environment and interact with the protein domain.

Recent cryo-electron microscopy experiments have shown that the structure of the separate domains of the spike protein are conserved when they are isolated from the full trimeric complex.[46] In order to determine if also the dynamics of the isolated domains are the same as in the full-length trimeric spike, and to avoid simulating the entire spike trimeric complex, thereby speeding up the simulations, we have first performed a simulation of the various segments and the full-length and trimeric SARS-CoV-2 spike. These simulations confirm that the RBDs, NTDs and S1 segments have a similar structure and dynamics in their isolated and trimer-incorporated form (see Supporting Information Figure S1). The analysis indicates

that the behavior of each NTD, RBD and S1 segment of the trimer system is comparable with the domain simulated alone in solution, both in terms of root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) descriptors. In particular, the comparison between the RMSF of the domains considered alone vs. in the trimer shows that the most and least fluctuating regions are conserved. The average Pearson correlation between the trend of the RMSF of the three domains of the trimer with the RMSF of the NTD of chain A alone is 0.63. Taken together, these results allow us to simulate only the NTD, RBD, and S1 of the SARS-CoV-2 spike protein, thus significantly reducing the computational cost of our simulations.

Furthermore, simulating S1 allowed us to analyze the correlated motions among residues of both RDB and NTD of the spike protein. The analysis of correlated motions among residues, both considering covariance (to assess the average fluctuation of each residue) and considering the Pearson coefficient (which is independent of the fluctuation of individual residues), provides insight into communication among different regions that may play a key role in function and structural dynamics.[47–49] We examine both Pearson correlation and covariance between residues (averaging over x, y, and z correlation values for each residue pair). Specifically, we focused on the cross-correlation between the NTD and RBD domains of the S1, studying the correlated movements both within the domain (i.e., among residues belonging to the same domains; NTD-NTD and RBD-RBD) and between domains (i.e., among residues belonging to two different domains, NTD-RBD).

Interestingly, we observe a positive Pearson correlation for residue pairs within the same domain, both NTD and RBD, with average values of 0.056 and 0.043 for residues belonging to the NTD and RBD, respectively (see Supporting Information Figure S2). Conversely, for residues belonging to the two different domains (NTD and RBD), we note a significant anticorrelation (negative correlation) with a Pearson coefficient of -0.042. Although analyses of correlated motions are widely used for analyzing molecular simulation data, the interpretation of average intra- and inter-domain correlations is not straightforward. The

positive correlation between nearby residues within the same domain may likely determine the conservation of binding regions, both directly and indirectly. Conversely, the correlation (in this case negative) between residues belonging to the two domains could quantify how the two domains communicate with each other, probably to maintain (or alternate) binding propensity.

To better understand the mechanisms of synchronized motion between residues of the two domains (NTD and RBD), we selected the set of residues that maximizes the anti-correlation between the two domains (see also Figure S2). A set of residue-residue pairs was chosen, which exhibited anti-correlation values ranging from -0.33 to -0.38. Specifically, the residues range from 182 to 187 in the NTD and from 386 to 391 in the RBD. In both cases, the residues belong to loop regions. Interestingly, we identify two regions that do not overlap with the sialic acid binding regions, but which could play an important role in the dynamic-structural relationship between the two domains.

Next, we performed MD simulations of the NTD, RBD, and S1 domains of the SARS-CoV-2 spike in the presence of a sialic-acid molecule to observe events of interaction. To study the interaction between the protein and SA, we first defined a dynamic binding propensity score for each residue of the simulated domains, which is based on the time each residue spends in interaction with sialic acid during the evolution of the simulation. For each residue, we calculated the fraction of frames in which the centroid atom of the sialic-acid molecule is located at a distance less than 6 Å to at least one of the heavy atoms of the residue (see Figure 2.a). This allows us to compare the sialic-acid contact probability per residue obtained from MD simulation of the various SARS-CoV-2 spike domains with the results from the Zernike shape complementarity evaluation,[19] through the binding propensity derived from the MD trajectory.

To avoid spurious contacts, i.e. contacts due to the limited size of the simulation box, we considered only configurations having favorable energies. To this end, we firstly estimated the binding energy between the NTD of MERS-CoV spike and sialic acid performing a

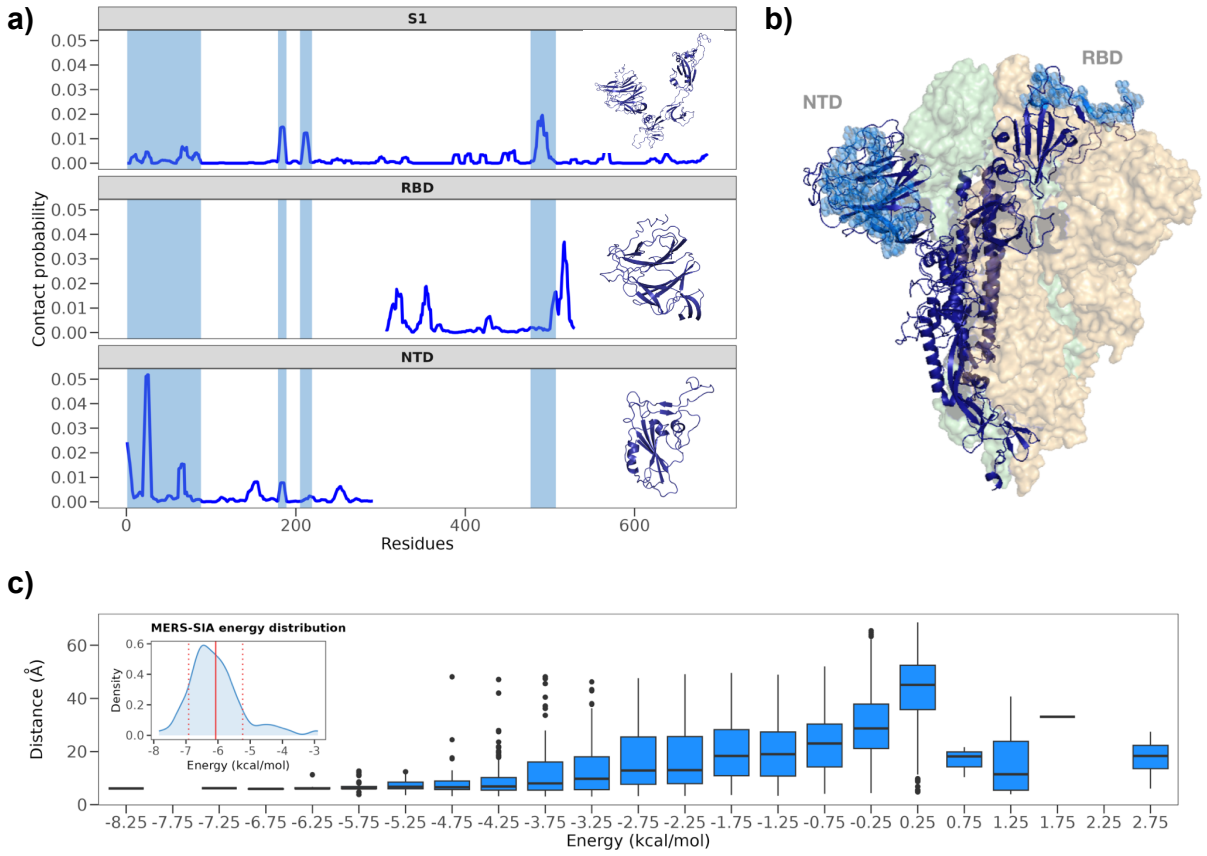


Figure 2: Contact probability as determined with MD simulations of the S1, RBD, and NTD domains of the SARS-CoV-2 spike. (a) Contact probability between each residue and the sialic-acid molecule is shown for the three molecular systems: the S1 segment, receptor-binding domain (RBD) and N-terminal domain (NTD) of the spike protein. (b) Molecular structure of the spike in its trimeric form: one chain is depicted as a blue-colored cartoon, the molecular surfaces are highlighted for the remaining two chains. The residues with the highest probability of interaction, as calculated from the MD trajectory of the largest system, S1, are represented in cyan, and are located both in the NTD and in the RBD. (c) Box-and-whisker plot of the minimum distance of the sialic-acid molecule to the nearest protein residue during the SARS-CoV-2 NTD MD simulation as a function of the binding energies. The inset displays the distribution of binding energies of the sialic-acid molecule to the MERS-CoV NTD during the simulation, with the thick red line indicating the mean value, and the thin red lines the standard deviation.

molecular dynamics simulation starting from the experimental pose[20] (see Methods). More specifically, sampling configurations of the system at the equilibrium (see Methods for detail), we calculated the inter-molecular energy contribution using the AutoDock[50] for each selected pose. The distribution of these energies is reported in the inset in Figure 2c. Despite the differences between these two systems, the energy distribution of the interaction between MERS-CoV spike protein and the sialic-acid molecule allows us to energetically relate the interaction between sialic-acid molecules and the SARS-CoV-2 spike with what we would expect experimentally, under very similar conditions.

We then extracted the interacting conformations and calculated the inter-molecular energy using the AutoDock algorithm for each of them: we thus selected only the poses characterized by an inter-molecular energy lower than -4 kcal/mol. As further validation, we studied the intermolecular interaction between sialic acid and SARS-CoV-2 spike protein as a function of the distance between sialic acid and the putative binding region predicted with the Zernike-based approach. As is evident from Figure 2.c, when the sialic-acid molecule is close to the predicted region its interaction energies are favourable, while higher energies characterizes the poses where the sialic-acid molecule is far from the predicted binding site. This result confirms that the preferred spike region of interaction is the one with the highest binding compatibility, as calculated by our Zernike method.

To further test the revealed binding region, we sought for an evaluation of binding free energy via fastDRH[38] (freely available as web server), which provides a MM/PB(GB)SA-based free energy estimation given a certain binding pose. In particular, we selected two binding configurations between the spike protein (S1) of SARS-CoV-2 and the sialic acid molecule. The first configuration corresponds to the complex for which the binding energy , estimated with Autodock,[50] is minimal compared to all other frames (-8.2 kcal/mol). In this configuration, the sialic-acid molecule is located at a distance of approximately 6 $\mathring{A}$ from the center of the binding pocket predicted computationally by the Zernike polynomial-based method. The second selected configuration is the one in which the intermolecular energy

between the spike and the sialic-acid molecule is minimal among the frames in which sialic acid binds to the pocket predicted by the Zernike-based method. In this way, we selected the structure with the minimal energy of the spike-SA complex in which sialic acid binds to the predicted pocket. Calculating the free energy with MM/PB(GB)SA methods for both selected configurations, we confirmed the estimated energy value for the first configuration (obtaining a value of -8.22 kcal/mol with Autodock, and an average value of -8.61 kcal/mol with MM/PB(GB)SA), on the other hand, we got a considerably more favorable (negative) free energy value for the binding between SA and the pocket predicted with Zernike when MM/PB(GB)SA methods are used. Specifically, for this configuration, the energy calculated with Autodock is -4.99 kcal/mol (indicating a good interaction energy, among the best explored blindly in the simulation), while the energy calculated with MM/PB(GB)SA is -12.10 kcal/mol. This shows that the pocket predicted with the Zernike method could bind to the SA molecule even more favorably than the possible neighboring binding sites. The comparison between the estimates obtained with Autodock and those obtained with fastDRH is reported in the Table S1 in the Supporting Information.

Interestingly, looking at Figure 2a), not all the peaks in the S1 simulation match with single-domain simulations. This suggests that some binding region of the single domains could be in non-physical regions; in other words, in a region of the sequence that faces the inner part of the fully folded spike trimer that cannot interact with other molecules in the biological context. To elucidate this, in Figure 2b) the full trimer is depicted with one S1 chain highlighted in blue, and the solvent-accessible regions of the electrostatic surface of the RBD and NTD that have a high probability of interaction with sialic acid are marked in cyan. To compute the regions that are solvent exposed in the trimer, we made use of the short MD simulation of the full trimer (vide supra). For the RBD this excludes the region from residue 320 to 380. During the SA-RBD simulation, 23.3% of the SA binding time occurs to RBD residues that are not accessible to sialic acid molecules in the fully-folded spike protein. The identification of non-biologically relevant (i.e. regions that are not

solvent exposed in the full-length, folded protein) binding sites, indicates that the RBD is not a proper model protein to study SA binding by the S1 segment or full-length spike.

To quantify which binding is stronger and more physically significant, we computed the relative time that the SA molecule spends in the NTD and in the RBD binding pocket in the S1 MD simulations. By integrating over the contact probability and determining the ratio among these times we can estimate which is the stronger binder among the two domains. We find that the NTD spends 0.211 ms bound, against 0.167 ms for the RBD. Since the binding probability is intrinsically proportional to the time spent bound,[51] and because these two residence times are computed from the same simulation, we can speculate that the SA binding strength of the NTD is around 25% stronger than the SA binding strength of the RBD. From this, we conclude that NTD and RBD binding to sialic-acid moieties have an approximately equally strong, positive effect on the probability of infection of SARS-CoV-2.

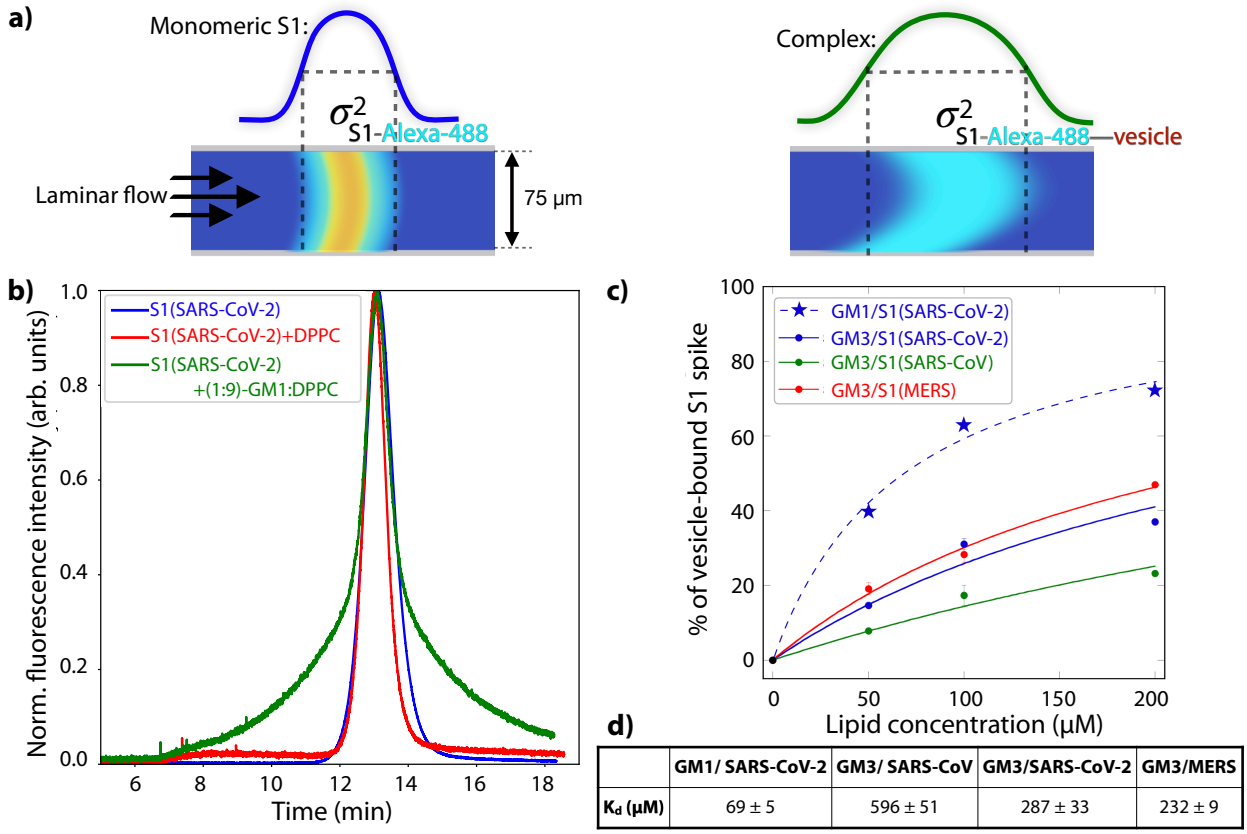
Flow-induced dispersion analysis (FIDA) shows that SARS-CoV-2 spike binds to sialic acid *in vitro*

Our *in silico* findings predict that the S1 segment of the SARS-CoV-2 spike protein is able to bind to SA moieties.[19] To experimentally confirm these predictions we perform flow-induced dispersion analysis (FIDA) measurements[25,40] with the S1 segment of the protein (residues Val-16 to Arg-685, see GenBank accession number GenBank: QHD43416.1[52,53] for the full sequence), recombinantly expressed in human cells (see Methods). The FIDA technique relies on the fact that particles with a larger hydrodynamic radius (R_h) will be dispersed more strongly in a laminar flow than particles with a small R_h (see Figure 3a). By labeling the S1 spike segment non-specifically with a fluorescent dye and flowing it through 75-μm diameter capillaries in presence and absence of vesicles partly composed of sialic-acid containing glycolipids, the observed R_h is expected to strongly increase upon S1 – glycolipid binding. The S/N of the FIDA setup is not sufficient to detect complex formation between the S1 spike and free SA, but the binding to glycolipid-containing vesicles can be characterized

well. The mechanistic nature of this experimental setup allows us to determine if SA can play a physical role in SARS-CoV-2 infection in a cellular context. Indeed, if the SARS-CoV-2 S1 segment can bind SA-containing vesicles, the physical strength of this binding can be considered relevant in a cell–virion context.

When we compare the Taylorgrams of a pure S1 solution with those of either pure DPPC ∼50 nm small-unilaminar vesicles (SUVs), or with 1:9 GM1/DPPC composed SUVs, we observe the R_h of monomeric S1 if there are no glycolipids present in the system, but if there are glycolipid-containing vesicles, the increased R_h indicates complex formation. This result is corroborated by previously published findings for sialic-acid molecules that were connected with 50-mer PHEA polymer chains to Au nanoparticles, which also exhibited SARS-CoV-2 spike binding.[45] Our experiments are performed in the presence of 10 μM free sialic acid, to prevent nonspecific binding between the spike S1 protein and DPPC, as evinced by sharp peaks in the Taylorgrams (see Figure S3). While pure monomeric S1, and S1 in the presence of SUVs composed solely of DPPC give a unimodal Gaussian lineshape of the Taylorgrams, the S1 binding to 1:9 GM1:DPPC vesicles results in a two-modal Gaussian lineshape, which (through fitting, see ref. [25]), indicates that approximately 50% of the S1 spikes are bound to the vesicles, while the other ∼50% remain unbound.

To investigate if the Zernike analysis also correctly predicts the absence of this sialic-acid binding pocket in the S1 segment of SARS-CoV ,[19] while it is also present in MERS, and to determine if the binding is dependent on the type of glycolipid, we repeated the experiment by incubating 100 nM of each the three S1 segments, but now in the presence of varying amounts of 1:9 GM3:DPPC lipid vesicles. The glycolipid is varied from GM1 to GM3 to investigate if the exact position of the sialic-acid group within the sugar chain in the lipid headgroup[16] affects the binding. The FIDA measurements show (1) that the SARS-CoV-2 and MERS-CoV S1 segments bind ∼twice as strong than the SARS-CoV S1 segment, as evinced by the ∼twice as high percentage of vesicle-bound S1 at a given lipid concentration, and (2) that the SARS-CoV-2 spike binds stronger to GM1 than to GM3 (Figure 3c and Figure S4 for the



	GM1/ SARS-CoV-2	GM3/ SARS-CoV	GM3/SARS-CoV-2	GM3/MERS
K_d (µM)	69 ± 5	596 ± 51	287 ± 33	232 ± 9

Figure 3: FIDA analysis of SARS-CoV-2 spike S1 segment binding to sialic-acid containing glycolipids. (a) Schematic of how complex formation by fluorescently-labeled S1 will affect Taylorgrams of FIDA measurements. A single species that includes a fluorescent label will give rise to a single Gaussian lineshape, with a Gaussian width σ that increases with an increasing hydrodynamic radius. (b) Taylorgrams recorded for pure 100 nM spike S1 protein, with 200 μM DPPC vesicles, and with 1:9 GM1:DPPC vesicles, partly composed of the sialic-acid containing GM1 glycolipid. In the case of the glycolipid-containing vesicles, a double Gaussian lineshape is observed, indicating the presence of both free S1 species and S1-vesicle complexes, as opposed to the single Gaussian lineshapes observed for the pure S1 for in the presence of the DPPC vesicles that do not contain glycolipids, indicating the absence of S1-binding in that case. (c) The relative number of of bound species for the S1 segments of SARS-CoV, SARS-CoV-2, and MERS-CoV bound to (1:9)-GM1:DPPC and -GM3:DPPC lipid vesicles. (d) Estimated dissociation constants.

associated Taylorgrams). In order to obtain an estimation of the dissociation constant k_D, even though no saturation is obtained before vesicle clustering occurred during the FIDA measurement, we assume that at higher lipid concentrations 100% occupancy would have been reached. Under this assumption, we find apparent k_Ds that are (1) ∼twice as small for SARS-CoV-2 and MERS as compared to SARS-CoV-1, and (2) ∼four times as small for SARS-CoV-2 spike binding to GM1 as compared to GM3 (see Figure 3d). The latter observation indicates that the extra sugar groups present in the headgroup of GM1 (absent in the GM3 headgroup) mediate the spike-sialic acid binding. Furthermore, we think that the non-zero SARS-CoV S1 binding to the glycolipids is due to non-specific interactions, and note that – given the omnipresence of sialic-acid moieties on cell membranes – even the relatively weak sialic-acid binding revealed here may suffice to greatly enhance the infection efficiency.

Flow cytometry (FCM) shows that SARS-CoV-2 spike binds to SA moieties on cell membranes

The *in silico* 2D-Zernike analyses and MD simulations predict that the spike S1 can bind SA through different domains, while the FIDA measurements show the physical relevance of the structural variation amongst various β-coronavirus spike NTDs in a lipid-membrane context. In order to see if these indications also stand in the spike interaction with cells we perform flow-cytometry (FCM) measurements. We do this by determining the binding of three different spike segments to transgenic HEK cells that are transfected to express a large number of ACE2 receptors. First, we investigate the full-length spike protein (Cys-15–Pro-1213) – given that this has the closest similarity to the spike protein attached to the SARS-CoV-2 virus, the S1 segment (Val-16–Arg-685) – given that it contains both SA-binding regions of the SARS-CoV-2 spike protein, and the RBD region (Arg-319–Phe-541) – given the unexpected result obtained in the MD simulations that the ACE2 binding site (which gives the RBD its name) also has SA-binding capabilities. To reveal the influence of

cell-membrane associated sialic-acid groups on the binding, we add a sialic-acid expression inhibitor called sialostatin (see *Methods* for details). When we gate the FCM such that the Alexa-488 fluorescence reflects the spike-binding events to the HEK-ACE2 cells (see Figure 4a-c), a significant decrease in the binding of each of the three investigated domains (the full-length spike, the S1 domain and the receptor-binding domain; see Figure 4(d)) occurs when the HEK-ACE2 cells are incubated with sialostatin. The fact that the binding of all of the investigated domains decreases is consistent with the MD simulations that indicated sialic-acid binding capabilities in the three segments.

As can be seen from the SiaFind Pan-Specific Lectenz binding, which reports on the level of sialic acid on the cell membrane, the degree of sialostain-induced signal decrease is similar for the three spike domains as for Lectenz, while the ACE2 levels are not affected. This indicates that the availability of sialic-acid groups on the membranes of HEK-ACE2 cells very strongly affects the ability of the three spike domains to bind, and that there is no significant contribution to the spike-cell binding through direct ACE2 binding.

To further disentangle the relative roles of the SA and ACE2 receptors in the spike (segment) binding, we also performed FCM measurements with wild-type HEK cells (Figure S5). Instead of removing SA moieties from the cell membrane as in the experiment shown in Figure 4, we thus observe the effect of removing ACE2 groups. Comparison of the spike binding to transgenic ACE2-HEK cells and to wt-HEK cells reveals a strongly decreased binding to the latter. The two FCM experiments thus show that the cellular binding of the spike segments is dependent on the presence of both SA and ACE2 receptors. We can thus derive that the initial spike–SA binding, which is of vital importance to the initial spike–cell binding according to the results depicted in Figure 4 – does not suffice by itself to maintain the spike-cell binding at a similar level as for the HEK cells that have a higher ACE2 expression. We think that the origin of the decreased binding as a function of decreased SA expression (Figure 4) instead lies in a decreased initial molecular recognition between the spike and SARS-CoV-2 host cell. Potentially, if larger decreases in sialic-acid

expression could be obtained in future studies, the relative importance of the two receptors could be explored in more detail. As we now see maximal inhibition of the binding (1:1) for the achieved inhibition of sialic acid expression, and also a very strong decrease in the spike binding without enhanced ACE2 expression (Figure S 5), it is probable that the importance of both receptors for successful binding is comparable in magnitude.

As we have incubated the cells with a single concentration of spike (domain) proteins, we cannot determine a binding strength, but the FCM measurements nonetheless clearly confirm the biological relevance of the most important computational predictions, namely that SARS-CoV-2 exhibits a two-receptor strategy, and that the RBD is also strongly involved in sialic-acid binding, which has a positive contribution to eventual ACE2 binding.

Overall, the FCM experiments demonstrate the biological relevance of the SA – spike (segment) binding observed with the *in silico* and FIDA methods, at a cellular level.

Discussion: A molecular model of SARS-CoV-2 binding to sialic acid

By closely integrating molecular simulations and specific experimental binding assays, we have we investigated the interaction between the spike protein of SARS-CoV-2 and SA. In this research, we have used computational methods to first discover sialic acid as a previously unknown binding partner for the SARS-CoV-2 spike protein (with the Zernike approach[19]), after which we have identified two binding regions on the spike (with the MD simulations – see Figures 1 and 2). To validate these findings, we have first confirmed the specific SARS-CoV-2 – sialic-acid binding in the well-defined model system of sialic-acid containing lipid vesicles and the spike S1 segment that contains both sialic-acid binding sites (with FIDA – see Figure 3), and subsequently confirmed a similarly specific binding to ACE2-expressing cells for various spike segments, to show that also in a biologically relevant and complex setting this is a relevant process (with FCM – see Figure 4).

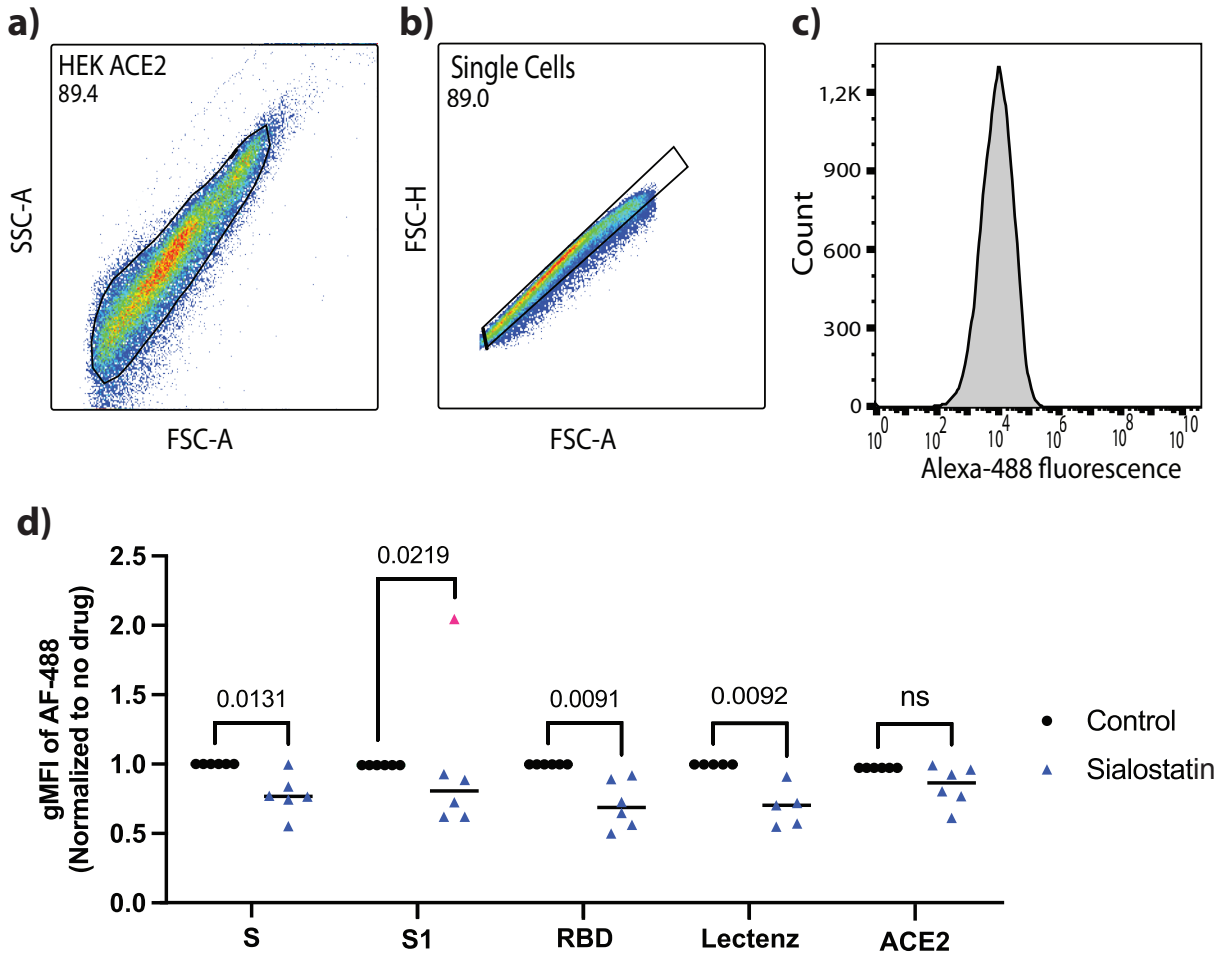


Figure 4: Decreased binding of the SARS-CoV-2 full-length spike, S1 and RBD proteins to HEK-ACE2 cells upon inhibition of sialic-acid expression. (a-c) Example of applied flow cytometry gating strategy to assess full-length spike protein binding to single HEK-ACE2 cells, and the Alexa-488 fluorescence resulting from spike binding to the gated ensemble. d) The normalized geometric mean of the Alexa-488 fluorescence intensity (gMFI) that indicates the relative binding strength of various segments of the SARS-CoV-2 spike protein to HEK-ACE2 cells, the Lectenz binding (which reports on the presence of sialic-acid groups on the cellular membrane), and ACE2 levels, as determined by flow cytometry, with and without addition of sialostatin (n=6, 2-sided paired t-test). Statistics per experiment: SARS-CoV-2 Spike S1 domain: n=5, 2-sided paired t-test), with one outlier (marked pink) was excluded based on a GRUBS test (alpha 0.05); SARS-CoV-2 Spike RBD domain: n=6, 2-sided paired t-test; SiaFind Pan-Specific Lectenz with and without addition of sialostatin: n=5, 2-sided paired T-test; ACE2 expression level of HEK-ACE2 cells: (n=6, 2-sided paired T-test). ns: not significant.

The main finding from this this combination of methods is the observation that within the S1 segment of the spike protein there are two regions that interact with sialic acid groups, namely the N-terminal domain (NTD) and the receptor-binding domain (RBD). So far there is not a well-known biological role of the NTD for the entry of the viral RNA inside the cell, while the RBD is known to bind the principal SARS-CoV-2 receptor ACE2, which leads to virion internalization.[54] For MERS-CoV, the binding pocket of SA is known to be an anchor for its spike to bind the cell membrane as a first step in the infection mechanism.[20] Here, we show that the NTD plays a similarly critical role in the infection mechanism for SARS-CoV-2. Based on the experimental and *in silico* evidence presented in this work, it appears that the SARS-CoV-2 spike first binds SA moieties on the cell membrane, after which the spike-SA complex diffuses on the membrane to find the ACE2 receptor.[55] Combination of the 3D diffusion, to reach the cell, and 2D diffusion to search across the membrane, strongly increases the residence time on the membrane during which it can meet its principal receptor. We note that a quantification is difficult because of the large number of physical constants (relative concentrations, exact *in vivo* binding constants, etc.). Building a molecular model that fully describes all the relevant interactions can not be achieved by experiments or simulations alone.

We therefore performed experiments at different biologically-relevant levels and complementary MD simulations to obtain a molecular-level interpretation. Indeed, on one hand based on flow-induced dispersion analysis (FIDA), we are able to show that at an *in vitro* scale, the S1 segment of the SARS-CoV-2 spike is able to bind lipid vesicles that are partly composed of sialic-acid containing glycolipids. We observe that this binding was much stronger for the MERS-CoV and SARS-CoV-2 S1 proteins, whose NTD has been shown to have a similarly shaped SA binding pocket,[19,56] as compared to SARS-CoV, for which such a similarly-shaped region was found to be absent in the Zernike analysis. This substantiates the importance of the role of the NTD in the SA – SARS-CoV-2 spike binding, and corroborates recent experiments that directly reveal SARS-CoV-2 spike NTD – SA bind-

ing.[57,58] On the other hand, we also performed flow cytometry (FCM) measurements to show that the spike-SA interaction is essential in the cellular binding probability, by incubating the cells with the sialic-acid inhibitor sialostatin. By assessing this binding for various spike segments, we find a similar SA inhibition effect on the cellular binding of the RBD, the S1 and the full-length spike, consistent with the MD findings that each of these segments contains SA binding sites.

The MD simulations also provide a molecular-level image of what happens during the binding experiments. By analyzing the structure of the RBD, NTD and S1 segment, either as a part of the fully folded spike protein, or as a separate domain, we were able to reproduce the NTD binding pocket that was previously found with the Zernike analysis,[19] and show that the experimentally-observed sialic-acid binding of the isolated RBD is to a certain degree biologically irrelevant as $\sim$25% of the SA-binding residues are not accessible to sialic acid in the fully-folded spike protein. From the sialic-acid – S1 MD trajectory, we have determined the respective sialic-acid affinities to the NTD and RBD domains. Using a simple dynamical model, deriving physical constants from the MD simulations, we are able to estimate the effects of the two domains on sialic-acid binding, which – as a secondary receptor – helps the viral infection. This indicated an additional $\sim$25% stronger binding of SA to the NTD than to the RBD.

Interestingly, we observe SA-, but not heparan-sulfate (HS) binding in both the FIDA (for the S1) and (for neither the RBD, S1, and full-length S segments) FCM experiments (see Figure S6), while a previous FCM study has reported RBD and S binding to e.g. *Vero E6* and *A549* cells,[16] a fluorescent microarray and surface-plasmon resonance study has reported RBD and S binding to HS, but not to SA,[61] and an NMR study has reported that S, but not RBD binds to α2,3- and α22,6-sialyl N-acetyllactosamine), i.e., two sialic-acid containing trisaccharides.[57] In these FCM experiments, performed with HSases (instead of sialostatin), we observed strongly decreased HS levels on the HEK-ACE2 cells, but not any decrease in spike (segment) binding, as opposed to earlier findings for different cell lines, by Clausen

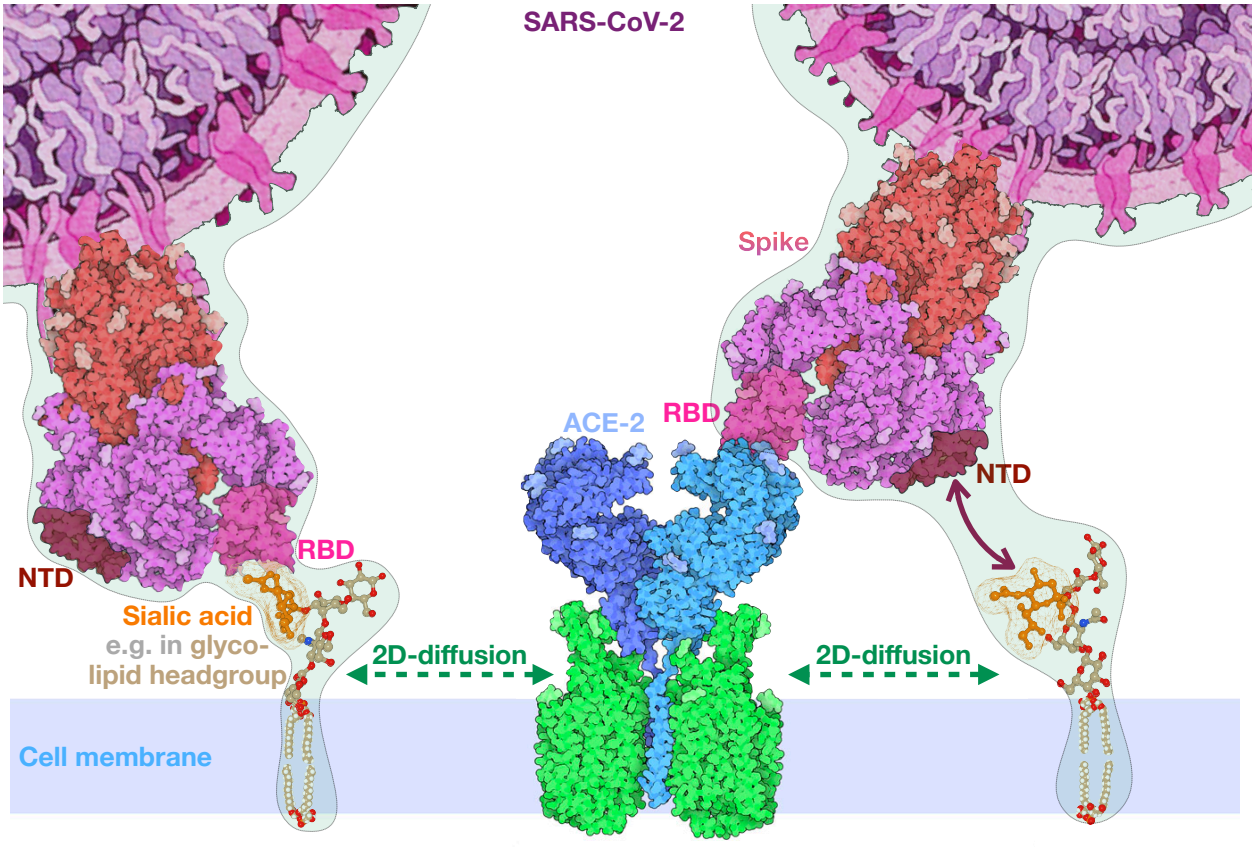


Figure 5: The two-receptor binding strategy of SARS-CoV-2 is mediated by both the N-terminal and receptor-binding spike domain. Computations and experiments reveal that both the NTD and RBD can bind to sialic-acid groups at the cell membrane. Both binding events have a positive effect on subsequent ACE2 binding. After a 3D diffusion process in which the spike proteins can interact with the omnipresent sialic-acid moieties on cellular membranes, sialic-acid binding will reduce the dimensionality of the search for the ACE2 to a 2D-diffusion process. As the latter binding triggers virion internalization,[54] the two-receptor and two binding-domain strategy of SARS-CoV-2 strongly enhances its infection rate. Protein and virion renderings are reproduced from ref.[59] under the CC-BY-4.0 license, lipid rendering is created in *UCSF Chimera*.[60]

et al.[16] And in FIDA experiments, we observe no RBD or S1 binding to heparin of various molecular weights (6 – 30 kDa) nor to the HS-chain containing proteoglycan syndecan-2 (data not shown), while this was observed for immobilized HS ,[62,63] and – in a typically unspecific manner – for heparin.[64–67] Taken together, this indicates that the exact choice of molecules, cells, and techniques significantly influences the outcome of these experiments.

Conclusions

The observed sialic-acid (SA) binding by the SARS-CoV-2 spike could have profound consequences in the context of the human pandemic situation. Indeed, closely-related species of β-coronaviruses, such as MERS-CoV and SARS-CoV, have developed different strategies to infect the host cells. SARS-CoV utilizes the ACE2 angiotensin converting enzyme 2 (ACE2) receptor with a high affinity, and binding ACE2 with its spike protein starts the chain of chemical reactions that leads to the insertion of the viral RNA. In a different fashion, the MERS coronavirus uses sialic-acid moieties on the cell membrane as the main attachment receptor, after which it searches for its main receptor, dipeptidyl peptidase 4 (DPP4), on the membrane. Here, we show that SARS-CoV-2 has developed a different infection mechanism compared to these other coronaviruses (see Figure 5). On the one hand – as is the case for SARS-CoV infection – the SARS-CoV-2 virion internalizes after binding to the ACE2 receptor.[54] On the other hand – as is the case for MERS-CoV – it also has an affinity for SA, which is used as a first anchor to engage the virus to the host cell. Two-dimensional diffusion on the cell membrane after SA binding strongly accelerates the search for SARS-COV-2's internalization receptor, ACE2. In light of these observations, we propose the possibility that SARS-CoV-2 uses a combined strategy. Such a double-receptor strategy, in which two spike domains are both able to bind the sialic-acid moieties on the cellular membrane in the initial step, could explain the very fast spread of this virus.

One of the main problems with regards to the spreading of this virus has been the very

high number of asymptomatic patients that have spread the virus without being aware of having it. It is possible that this observation can be linked to the here-described SA binding. Indeed, SA has been found to be omnipresent on the membranes of cells of the external respiratory airways.[68] So possibly, the virus binds to these cells without going through the lung and bronchus, where the ACE2 receptor is expressed much more than in external respiratory airway cells.[69] Trapping the virus at the external respiratory cells could thus result in a lower infection probability, but a high virion concentration in the respiratory tract might increase the probability of spreading the virus.

Acknowledgement

We would like to thank Leonardo Bò for his help in the early stages of the presented work. We acknowledge Henk Dekker and Gertjan Kramer for assessing the purity of the protein sample with mass spectrometry. We thank Dr. Henrik Jensen, FIDAbio, for outstanding support in the use of FIDA technology. We acknowledge James D. Pickering for proof reading the manuscript. M Monti acknowledges funding through the H2020 project IN-FORE 825080. A.K.S. is supported by the Lundbeck Foundation (grant R287-2018-1836). TWG thanks the Lundbeck Foundation (Fellowship Grant, R322-2019-2461). SJR acknowledges the Lundbeck Foundation for funding through fellowship grant R303-2018-3495. TW thanks the Novo Nordisk Foundation (Facility Grant NanoScat, NNF18OC0032628) for support of this work. DEO and MN gratefully acknowledge support from the Innovation Foundation Denmark (grant no. 0211-00066B). GR and M Miotto acknowledge support by European Research Council Synergy grant ASTRA (n. 855923). The following reagents were obtained through BEI Resources, NIAID, NIH: Human Embryonic Kidney Cells (HEK-293T) Expressing Human Angiotensin-Converting Enzyme 2, HEK-293T-hACE2 Cell Line, NR-52511. Spike Glycoprotein (Stabilized) from SARS-Related Coronavirus 2, Wuhan-Hu-1 with C-Terminal Histidine Tag, Recombinant from HEK293F Cells, NR-52397. Spike Glycopro-

tein S1 Domain from SARS-Related Coronavirus 2, Wuhan-Hu-1 with C-Terminal Histidine Tag, Recombinant from HEK293 Cells, NR-53798. Spike Glycoprotein Receptor Binding Domain (RBD) from SARS-Related Coronavirus 2, Wuhan-Hu-1 with C-Terminal Histidine Tag, Recombinant from HEK293F Cells, NR-52366.

Supporting Information Available

The supporting information contains additional computational (molecular dynamics) and experimental (FIDA and flow cytometry) results.

References

(1) Wang, H.; Paulson, K. R.; Pease, S. A.; Watson, S.; Comfort, H.; Zheng, P.; Aravkin, A. Y.; Bisignano, C.; Barber, R. M.; Alam, T. et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *The Lancet* **2022**, *399*, 1513–1536.

(2) Liu, D. X.; Liang, J. Q.; Fung, T. S. Human Coronavirus-229E, -OC43, -NL63, and -HKU1 (Coronaviridae). *Encyclopedia of Virology* **2021**, 428–440.

(3) Mingaleeva, R. N.; Nigmatulina, N. A.; Sharafetdinova, L. M.; Romozanova, A. M.; Gabdoulkhakova, A. G.; Filina, Y. V.; Shavaliyev, R. F.; Rizvanov, A. A.; Miftakhova, R. R. Biology of the SARS-CoV-2 Coronavirus. *Biochemistry (Moscow)* **2022**, *87*, 1662–1678.

(4) Drosten, C.; Günther, S.; Preiser, W.; Van Der Werf, S.; Brodt, H.-R.; Becker, S.; Rabenau, H.; Panning, M.; Kolesnikova, L.; Fouchier, R. A. et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *New England journal of medicine* **2003**, *348*, 1967–1976.

(5) Ksiazek, T. G.; Erdman, D.; Goldsmith, C. S.; Zaki, S. R.; Peret, T.; Emery, S.; Tong, S.; Urbani, C.; Comer, J. A.; Lim, W. et al. A novel coronavirus associated with severe acute respiratory syndrome. *New England journal of medicine* **2003**, *348*, 1953–1966.

(6) Zaki, A. M.; Van Boheemen, S.; Bestebroer, T. M.; Osterhaus, A. D.; Fouchier, R. A. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *New England Journal of Medicine* **2012**, *367*, 1814–1820.

(7) Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, 1–4.

(8) Wielgat, P.; Rogowski, K.; Godlewska, K.; Car, H. Coronaviruses: Is Sialic Acid a Gate to the Eye of Cytokine Storm? From the Entry to the Effects. *Cells* **2020**, *9*.

(9) Wu, Z.; Harrich, D.; Li, Z.; Hu, D.; Li, D. The unique features of SARS-CoV-2 transmission: Comparison with SARS-CoV, MERS-CoV and 2009 H1N1 pandemic influenza virus. *Reviews in Medical Virology* **2021**, *31*, e2171.

(10) Matrosovich, M.; Herrler, G.; Klenk, H. D. In *SialoGlyco Chemistry and Biology II: Tools and Techniques to Identify and Capture Sialoglycans*; Gerardy-Schahn, R., Delannoy, P., von Itzstein, M., Eds.; Springer International Publishing: Cham, 2015; pp 1–28.

(11) Alberts, B.; Bray, D.; Hopkin, K.; Johnson, A. D.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Essential cell biology*; Garland Science, 2015.

(12) Tortorici, M. A.; Walls, A. C.; Lang, Y.; Wang, C.; Li, Z.; Koerhuis, D.; Boons, G.-J.; Bosch, B.-J.; Rey, F. A.; de Groot, R. J. et al. Structural basis for human coronavirus attachment to sialic acid receptors. *Nature Structural & Molecular Biology* **2019**, *26*, 481–489.

(13) Nicholls, J. M.; Bourne, A. J.; Chen, H.; Guan, Y.; Peiris, J. Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses. *Respiratory research* **2007**, *8*, 1–10.

(14) Huang, X.; Dong, W.; Milewska, A.; Golda, A.; Qi, Y.; Zhu, Q. K.; Marasco, W. A.; Baric, R. S.; Sims, A. C.; Pyrc, K. et al. Human coronavirus HKU1 spike protein uses O-acetylated sialic acid as an attachment receptor determinant and employs hemagglutinin-esterase protein as a receptor-destroying enzyme. *Journal of virology* **2015**, *89*, 7202–7213.

(15) Schwegmann-Weßels, C.; Herrler, G. Sialic acids as receptor determinants for coronaviruses. *Glycoconjugate Journal* **2006**, *23*, 51–58.

(16) Clausen, T. M.; Sandoval, D. R.; Spliid, C. B.; Pihl, J.; Perrett, H. R.; Painter, C. D.; Narayanan, A.; Majowicz, S. A.; Kwong, E. M.; McVicar, R. N. et al. SARS-CoV-2 infection depends on cellular heparan sulfate and ACE2. *Cell* **2020**, *183*, 1043–1057.

(17) Miotto, M.; Rienzo, L. D.; Gosti, G.; Bo', L.; Parisi, G.; Piacentini, R.; Boffi, A.; Ruocco, G.; Milanetti, E. Inferring the stabilization effects of SARS-CoV-2 variants on the binding with ACE2 receptor. *Communications Biology* **2022**, *5*, 20221.

(18) Miotto, M.; Di Rienzo, L.; Grassmann, G.; Desantis, F.; Cidonio, G.; Gosti, G.; Leonetti, M.; Ruocco, G.; Milanetti, E. Differences in the organization of interface residues tunes the stability of the SARS-CoV-2 spike-ACE2 complex. *Frontiers in Molecular Biosciences* **2023**, *10*.

(19) Milanetti, E.; Miotto, M.; Di Rienzo, L.; Nagaraj, M.; Monti, M.; Golbek, T. W.; Gosti, G.; Roeters, S. J.; Weidner, T.; Otzen, D. E. et al. In-silico evidence for a two receptor based strategy of SARS-CoV-2. *Frontiers in molecular biosciences* **2021**, 509.

(20) Tortorici, M. A.; Walls, A. C.; Lang, Y.; Wang, C.; Li, Z.; Koerhuis, D.; Boons, G.-J.; Bosch, B.-J.; Rey, F. A.; de Groot, R. J. et al. Structural basis for human coronavirus

attachment to sialic acid receptors. *Nature structural & molecular biology* **2019**, *26*, 481–489.

(21) Guo, H.; Li, A.; Lin, H.-F.; Liu, M.-Q.; Chen, J.; Jiang, T.-T.; Li, B.; Wang, Y.; Letko, M. C.; Peng, W. et al. The Glycan-Binding Trait of the Sarbecovirus Spike N-Terminal Domain Reveals an Evolutionary Footprint. *Journal of Virology* **2022**, *96*.

(22) Varki, A. Sialic acids in human health and disease. *Trends in Molecular Medicine* **2008**, *14*, 351–360.

(23) García-Sastre, A. Influenza Virus Receptor Specificity. *The American Journal of Pathology* **2010**, *176*, 1584–1585.

(24) Li, B.; Wang, L.; Ge, H.; Zhang, X.; Ren, P.; Guo, Y.; Chen, W.; Li, J.; Zhu, W.; Chen, W. et al. Identification of potential binding sites of sialic acids on the RBD domain of SARS-CoV-2 spike protein. *Frontiers in Chemistry* **2021**, 349.

(25) Pedersen, M. E.; Østergaard, J.; Jensen, H. *Clinical Applications of Capillary Electrophoresis*; Springer, 2019; pp 109–123.

(26) Spoel, D. V. D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry* **2005**, *26*, 1701–1718.

(27) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. et al. CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry* **2009**, *30*, 1545–1614.

(28) Lin, F.-Y.; MacKerell, A. D. In *Biomolecular Simulations: Methods and Protocols*; Bonomi, M., Camilloni, C., Eds.; Springer New York: New York, NY, 2019; pp 21–54.

(29) Zhu, X.; Lopes, P. E. M.; MacKerell Jr, A. D. Recent developments and applications of the CHARMM force fields. *WIREs Computational Molecular Science* **2012**, *2*, 167–185.

(30) Rabet, S.; Raabe, G. Comparison of the GAFF, OPLSAA and CHARMM27 force field for the reproduction of the thermodynamics properties of furfural, 2-methylfuran, 2,5-dimethylfuran and 5-hydroxymethylfurfural. *Fluid Phase Equilibria* **2022**, *554*, 113331.

(31) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935.

(32) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007**, *126*, 014101.

(33) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Physical Review Letters* **1980**, *45*, 1196–1199.

(34) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **1997**, *18*, 1463–1472.

(35) Cheatham, T. E. I.; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. *Journal of the American Chemical Society* **1995**, *117*, 4193–4194.

(36) 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. et al. Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nature communications* **2020**, *11*, 2688.

(37) Breiman, A.; Ruvën-Clouet, N.; Le Pendu, J. Harnessing the natural anti-glycan immune response to limit the transmission of enveloped viruses such as SARS-CoV-2. *PLOS Pathogens* **2020**, *16*, 1–4.

(38) Wang, Z.; Pan, H.; Sun, H.; Kang, Y.; Liu, H.; Cao, D.; Hou, T. fastDRH: a webserver to predict and analyze protein–ligand complexes based on molecular docking and MM/PB (GB) SA computation. *Briefings in Bioinformatics* **2022**, *23*, bbac201.

(39) Milanetti, E.; Miotto, M.; Di Rienzo, L.; Monti, M.; Gosti, G.; Ruocco, G. 2D Zernike polynomial expansion: Finding the protein-protein binding regions. *Computational and structural biotechnology journal* **2021**, *19*, 29–36.

(40) Cholak, E.; Bugge, K.; Khondker, A.; Gauger, K.; Pedraz-Cuesta, E.; Pedersen, M. E.; Bucciarelli, S.; Vestergaard, B.; Pedersen, S. F.; Rheinstädter, M. C. et al. Avidity within the N-terminal anchor drives α-synuclein membrane interaction and insertion. *The FASEB Journal* **2020**,

(41) Taylor, G. I. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1953**, *219*, 186–203.

(42) Bolean, M.; Simão, A.; Favarin, B.; Millán, J.; Ciancaglini, P. Thermodynamic properties and characterization of proteoliposomes rich in microdomains carrying alkaline phosphatase. *Biophysical Chemistry* **2011**, *158*, 111 – 118.

(43) Moons, S. J.; Rossing, E.; Janssen, M. A.; Heise, T.; Büll, C.; Adema, G. J.; Boltje, T. J. Structure–Activity Relationship of Metabolic Sialic Acid Inhibitors and Labeling Reagents. *ACS chemical biology* **2022**, *17*, 590–597.

(44) Heise, T.; Pijnenborg, J. F.; Büll, C.; van Hilten, N.; Kers-Rebel, E. D.; Balneger, N.; Elferink, H.; Adema, G. J.; Boltje, T. J. Potent metabolic sialylation inhibitors based on C-5-modified fluorinated sialic acids. *Journal of medicinal chemistry* **2018**, *62*, 1014–1021.

(45) Baker, A. N.; Richards, S.-J.; Guy, C. S.; Congdon, T. R.; Hasan, M.; Zwetsloot, A. J.; Gallo, A.; Lewandowski, J. R.; Stansfeld, P. J.; Straube, A. et al. The SARS-COV-2

spike protein binds sialic acids and enables rapid detection in a lateral flow point of care diagnostic device. *ACS central science* **2020**, *6*, 2046–2052.

(46) Zhang, J.; Xiao, T.; Cai, Y.; Chen, B. Structure of SARS-CoV-2 spike protein. *Current Opinion in Virology* **2021**, *50*, 173–182.

(47) Amadei, A.; Linssen, A. B.; Berendsen, H. J. Essential dynamics of proteins. *Proteins: Structure, Function, and Bioinformatics* **1993**, *17*, 412–425.

(48) Xu, D.; Meisburger, S. P.; Ando, N. Correlated motions in structural biology. *Biochemistry* **2021**, *60*, 2331–2340.

(49) Miotto, M.; Di Rienzo, L.; Gosti, G.; Bo', L.; Parisi, G.; Piacentini, R.; Boffi, A.; Ruocco, G.; Milanetti, E. Inferring the stabilization effects of SARS-CoV-2 variants on the binding with ACE2 receptor. *Communications biology* **2022**, *5*, 20221.

(50) Huey, R.; Morris, G. M.; Olson, A. J.; Goodsell, D. S. A semiempirical free energy force field with charge-based desolvation. *Journal of computational chemistry* **2007**, *28*, 1145–1152.

(51) Gilabert, J. F.; Gracia Carmona, O.; Hogner, A.; Guallar, V. Combining Monte Carlo and molecular dynamics simulations for enhanced binding free energy estimation through Markov state models. *Journal of Chemical Information and Modeling* **2020**, *60*, 5529–5539.

(52) Clark, K.; Karsch-Mizrachi, I.; Lipman, D.; Ostell, J.; Sayers, E. GenBank. *Nucleic Acids Res.* **2016**, D67–72.

(53) Wu, F.; Zhao, S.; Yu, B.; Chen, Y.-M.; Wang, W.; Song, Z.-G.; Hu, Y.; Tao, Z.-W.; Tian, J.-H.; Pei, Y.-Y. et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265–269.

(54) 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* **2020**, *367*, 1444–1448.

(55) Sun, X.-L. The role of cell surface sialic acids for SARS-CoV-2 infection. *Glycobiology* **2021**, *31*, 1245–1253.

(56) Bò, L.; Miotto, M.; Rienzo, L. D.; Milanetti, E.; Ruocco, G. Exploring the Association Between Sialic Acid and SARS-CoV-2 Spike Protein Through a Molecular Dynamics-Based Approach. *Frontiers in Medical Technology* **2021**, *2*.

(57) Unione, L.; Moure, M. J.; Lenza, M. P.; Oyenarte, I.; Ereño-Orbea, J.; Ardá, A.; Jiménez-Barbero, J. The SARS-CoV-2 Spike Glycoprotein Directly Binds Exogeneous Sialic Acids: A NMR View. *Angewandte Chemie International Edition* **2022**, *61*, e202201432.

(58) Guo, H.; Li, A.; Lin, H.-F.; Liu, M.-Q.; Chen, J.; Jiang, T.-T.; Li, B.; Wang, Y.; Letko, M. C.; Peng, W. et al. The glycan-binding trait of the sarbecovirus spike N-terminal domain reveals an evolutionary footprint. *Journal of Virology* **2022**, *96*, e00958–22.

(59) Goodsell, D. S.; Dutta, S.; Zardecki, C.; Voigt, M.; Berman, H. M.; Burley, S. K. The RCSB PDB "molecule of the month": inspiring a molecular view of biology. *PLoS biology* **2015**, *13*, e1002140.

(60) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry* **2004**, *25*, 1605–1612.

(61) Hao, W.; Ma, B.; Li, Z.; Wang, X.; Gao, X.; Li, Y.; Qin, B.; Shang, S.; Cui, S.; Tan, Z. Binding of the SARS-CoV-2 spike protein to glycans. *Science Bulletin* **2021**, *66*, 1205–1214.

(62) Yue, J.; Jin, W.; Yang, H.; Faulkner, J.; Song, X.; Qiu, H.; Teng, M.; Azadi, P.; Zhang, F.; Linhardt, R. J. et al. Heparan Sulfate Facilitates Spike Protein-Mediated SARS-CoV-2 Host Cell Invasion and Contributes to Increased Infection of SARS-CoV-2 G614 Mutant and in Lung Cancer. *Frontiers in Molecular Biosciences* **2021**, *8*.

(63) Liu, L.; Chopra, P.; Li, X.; Bouwman, K. M.; Tompkins, S. M.; Wolfert, M. A.; de Vries, R. P.; Boons, G.-J. Heparan Sulfate Proteoglycans as Attachment Factor for SARS-CoV-2. *ACS Central Science* **2021**, *7*, 1009–1018.

(64) Kwon, P. S.; Oh, H.; Kwon, S.-J.; Jin, W.; Zhang, F.; Fraser, K.; Hong, J. J.; Linhardt, R. J.; Dordick, J. S. Sulfated polysaccharides effectively inhibit SARS-CoV-2 in vitro. *Cell discovery* **2020**, *6*, 50.

(65) Mycroft-West, C. J.; Su, D.; Pagani, I.; Rudd, T. R.; Elli, S.; Gandhi, N. S.; Guimond, S. E.; Miller, G. J.; Meneghetti, M. C. Z.; Nader, H. B. et al. Heparin Inhibits Cellular Invasion by SARS-CoV-2: Structural Dependence of the Interaction of the Spike S1 Receptor-Binding Domain with Heparin. *Thromb Haemost* **2020**, *120*, 1700–1715.

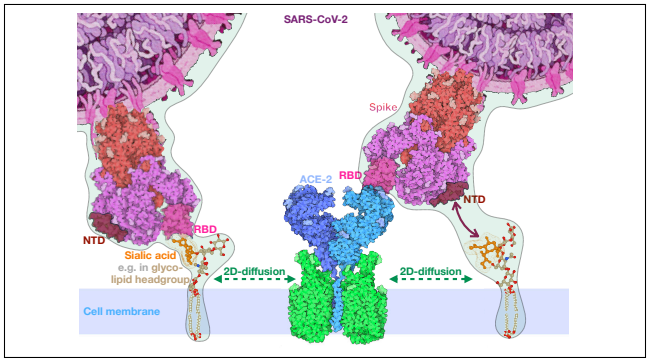
(66) Partridge, L. J.; Urwin, L.; Nicklin, M. J.; James, D. C.; Green, L. R.; Monk, P. N. ACE2-independent interaction of SARS-CoV-2 spike protein with human epithelial cells is inhibited by unfractionated heparin. *Cells* **2021**, *10*, 1419.

(67) Paiardi, G.; Richter, S.; Oreste, P.; Urbinati, C.; Rusnati, M.; Wade, R. C. The binding of heparin to spike glycoprotein inhibits SARS-CoV-2 infection by three mechanisms. *Journal of Biological Chemistry* **2022**, *298*.

(68) Varki, A. Sialic acids in human health and disease. *Trends in molecular medicine* **2008**, *14*, 351–360.

(69) Hikmet, F.; Méar, L.; Edvinsson, Å.; Micke, P.; Uhlén, M.; Lindskog, C. The protein expression profile of ACE2 in human tissues. *Molecular systems biology* **2020**, *16*, e9610.

TOC Graphic


SARS-CoV-2
Spike
ACE-2
RBD
NTD
RBD
NTD
Sialic acid
e.g. in glyco-
lipid headgroup
2D-diffusion
2D-diffusion
Cell membrane