

## Article

# SARS-CoV-2 and MERS-CoV Spike Protein Binding Studies Support Stable Mimic of Bound 9-O-Acetylated Sialic Acids

Lisa Oh <sup>1</sup>, Ajit Varki <sup>2</sup>, Xi Chen <sup>1</sup> and Lee-Ping Wang <sup>1,\*</sup><sup>1</sup> Department of Chemistry, University of California, Davis, CA 95616, USA<sup>2</sup> Glycobiology Research and Training Center, Departments of Medicine and Cellular and Molecular Medicine, University of California, San Diego, CA 92093, USA

\* Correspondence: leeping@ucdavis.edu

**Abstract:** Many disease-causing viruses target sialic acids (Sias), a class of nine-carbon sugars known to coat the surface of many cells, including those in the lungs. Human beta coronaviridae, known for causing respiratory tract diseases, often bind Sias, and some preferentially bind to those with 9-O-Ac modification. Currently, co-binding of SARS-CoV-2, a beta coronavirus responsible for the COVID-19 pandemic, to human Sias has been reported and its preference towards  $\alpha$ 2-3-linked Neu5Ac has been shown. Nevertheless, O-acetylated Sias-protein binding studies are difficult to perform, due to the ester lability. We studied the binding free energy differences between Neu5,9Ac2 $\alpha$ 2-3Gal $\beta$ pNP and its more stable 9-NAc mimic binding to SARS-CoV-2 spike protein using molecular dynamics and alchemical free energy simulations. We identified multiple Sia-binding pockets, including two novel sites, with similar binding affinities to those of MERS-CoV, a known co-binder of sialic acid. In our binding poses, 9-NAc and 9-OAc Sias bind similarly, suggesting an experimentally reasonable mimic to probe viral mechanisms.

**Keywords:** SARS-CoV-2; MERS-CoV; CoV S protein; sialic acid; MM-PBSA; SOMD; binding free energy simulations; molecular dynamics



**Citation:** Oh, L.; Varki, A.; Chen, X.; Wang, L.-P. SARS-CoV-2 and MERS-CoV Spike Protein Binding Studies Support Stable Mimic of Bound 9-O-Acetylated Sialic Acids. *Molecules* **2022**, *27*, 5322. <https://doi.org/10.3390/molecules27165322>

Academic Editor: Irini Doytchinova

Received: 29 July 2022

Accepted: 16 August 2022

Published: 20 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

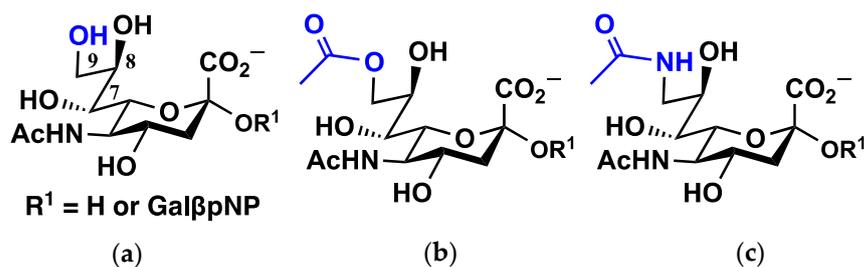
Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the beta coronavirus responsible for the COVID-19 pandemic, is structurally highly similar to SARS-CoV-1 (73% sequence identity between spike proteins), yet is remarkably more infective [1]. Both SARS-CoV-1 and SARS-CoV-2 share the same primary human cellular receptor angiotensin-converting enzyme-2 (ACE-2), but this commonality in binding cannot explain the difference in infectivity. One possible mechanism for how SARS-CoV-2 achieves its high infectivity is by binding to sialic acids (Sias), which coat many cell surfaces, especially in the lungs, and are targeted by many disease-causing viruses. In the airways the first encounter of the virus would be with soluble sialomucins, where the very high density of Sias would provide high avidity. On the other hand, the action of cilia are constantly pushing soluble mucins towards the upper airways. Some particles may escape by binding to membrane-bound mucins and sulfated glycosaminoglycans, then are eventually handed off to the high affinity ACE-2 receptors much closer to the plasma membrane.

Sias are part of a large family of over 50 derivatives of the 9-carbon sugar neuraminic acid, where N-acetylneuraminic acid (Neu5Ac) is most common [2]. The beta coronavirus Middle East respiratory syndrome coronavirus (MERS-CoV) is known to co-bind to Neu5Ac in addition to its DPP4 primary receptor in a two-step binding mechanism, and depletion of Neu5Ac was found to inhibit MERS-CoV entry into human airway cells [3]. Additionally, MERS-CoV binds preferentially to  $\alpha$ 2-3-linked Sias over  $\alpha$ 2-6-linked ones [3]. In another example, beta coronaviruses OC43 and HKUI are known to bind to 9-O-acetylated (9-OAc) Sias [1,4,5], where O-acetylation is one of the most common Sia modifications found in nature.

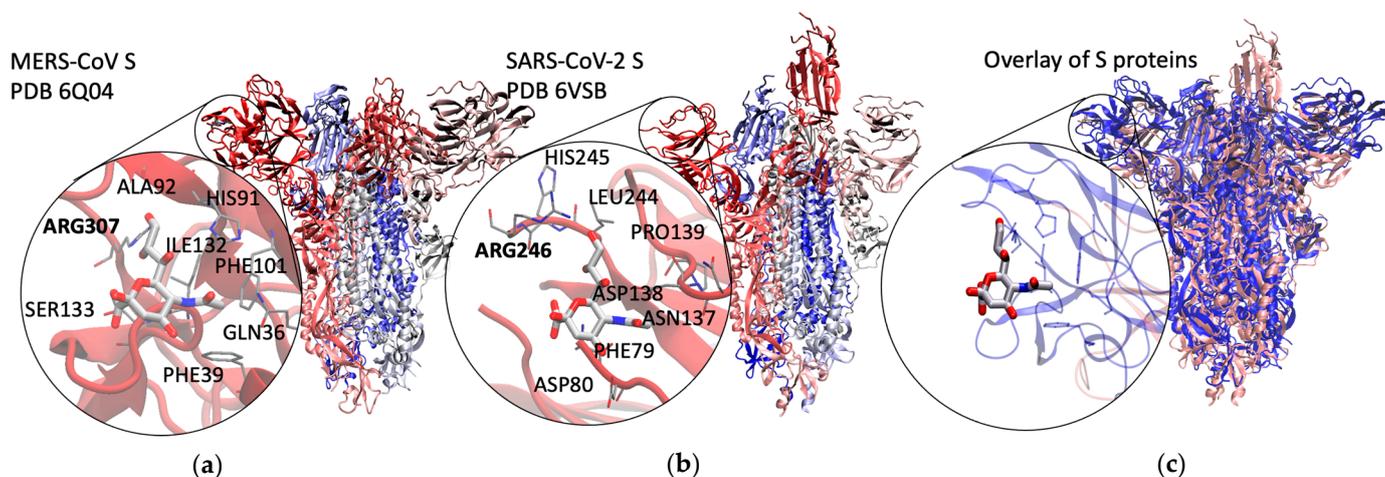
While SARS-CoV-1 is not known for its binding to Sias, recent studies have shown Sia binding by the SARS-CoV-2 spike (S) protein, and suggest that sialylated glycans can facilitate viral entry [6,7]. In a lateral flow test, the SARS-CoV-2 S protein bound to both Neu5Ac and  $\alpha$ 2-3/ $\alpha$ 2-6-linked Sias, where glycans are each attached to gold nanoparticles (auNP) [8]. While stronger binding under experimental conditions was observed with Neu5Ac-auNP over  $\alpha$ 2-3/ $\alpha$ 2-6-linked Sia-auNPs, the report did not preclude involvement of  $\alpha$ 2-3/ $\alpha$ 2-6-linked Sias in SARS-CoV-2 S protein binding [8]. Binding studies using sialylated glycans indicate low affinities to the SARS-CoV-2 S protein ( $\sim$ 10 kcal/mol) [9] and specifically along the ACE-2 receptor binding domain (RBD), with mono-sialylated ganglioside glycan affinities of 100–200  $\mu$ M and multi-sialylated glycan affinities approaching 900  $\mu$ M, using catch-and-release ESI-MS (CaR-ESI-MS) ( $\sim$ 20 kcal/mol) [6]. A few potential Sia-binding domains have been proposed in addition to the ACE-2 RBD, especially along the flat region of the SARS-CoV-2 S protein *N*-terminal domain (NTD), but specific sites have not been confirmed experimentally [10–17]. The S protein is made up of two subunits: S1 for cell recognition, containing both NTD and ACE-2 RBD, and S2, which is responsible for viral cell membrane fusion [18]. Given the density of Sias along cell surfaces, and the role of S1 in cell recognition, Sias could increase viral binding affinity by acting as an intermediate target or co-binder. We attempt to provide a quick overview of current work on Sias-SARS-CoV-2 S protein binding, where a more thorough study may be found in the recent review by Sun [19].

Experimental studies that test S protein binding to Neu5Ac and *O*-acetylated Sias using glycan microarrays are informative towards understanding the disease and preferential binding [9]. Even so, experimental testing is difficult with *O*-acetylated Sias due to the instability of the ester with respect to migration and cleavage, which depends on pH, temperature, and the presence of esterases [20–23]. *N*-Acetylated (NAc) Sias have been proposed to be stable synthetic mimics, as they are chemically and structurally similar, as seen in experimental and computational NMR studies [21]. Determining the similarity in S protein binding of 9-NAc to 9-OAc Sias would be valuable when performing binding array studies, and understanding the binding sites of sialic acids to SARS-CoV-2 S protein is important when considering potential druggable sites.

To this end, we computed the binding free energies of modified Neu5Ac monosaccharides and sialyloligosaccharides to SARS-CoV-2 S protein using molecular dynamics (MD) simulations, starting with a binding pose based on MERS-CoV (Figures 1 and 2). Our simulations revealed new possible Sia-binding sites ranging along the S1 unit, along the NTD and ACE-2 RBD, with some approaching the S2 domain. Each binding site contains a salt bridge connecting a conserved arginine residue to the carboxylate group of Sia, a known motif in Sia-lectin binding [24]; while the binding to individual sites is predicted to be weak, multiple binding to cell surface Sias could strengthen binding overall [1,25]. In addition, we calculated relative binding free energies of ligands that differ in the chemical modification at C9 (9-OH, 9-OAc, and 9-NAc), showing that the synthetic 9-NAc analogues are excellent structural mimics of their naturally occurring 9-OAc counterparts. These insights into Sia-S protein binding could lead to the design of therapeutics that inhibit the binding of S protein to Sias on the cell surface, thereby limiting SARS-CoV-2 transmission.



**Figure 1.** Structures of *N*-acetylneuraminic acid (Neu5Ac) and analogues, the simulated ligands in this study. (a) Neu5Ac; (b) Neu5,9Ac<sub>2</sub>; and (c) Neu5Ac9NAc are simulated ligands, corresponding to the unsubstituted Neu5Ac and its 9-OAc and 9-NAc forms. Two substitutions at R<sup>1</sup> were simulated, corresponding to the Sia monosaccharide or an  $\alpha$ 2-3-linked sialoside containing the terminal disaccharide of the GM3 ganglioside, commonly found on cell membranes, followed by *para*-nitrophenol (*p*NP), a molecule used in quantifying Sia binding and cleavage in sialidase activity assays.



**Figure 2.** Binding pose in SARS-CoV-2 S protein based on MERS-CoV S protein. (a) MERS-CoV S protein with Neu5Ac bound, cryo-EM structure; (b) initial binding pose of Neu5Ac in SARS-CoV-2 S protein based on MERS-CoV S protein, used as a starting point for MD simulations; (c) structural overlay of pink SARS-CoV-2 S protein (PDB ID: 6VSB) with blue MERS-CoV S protein in complex with Neu5Ac, highlighting binding residues from MERS-CoV S protein (PDB ID: 6Q04).

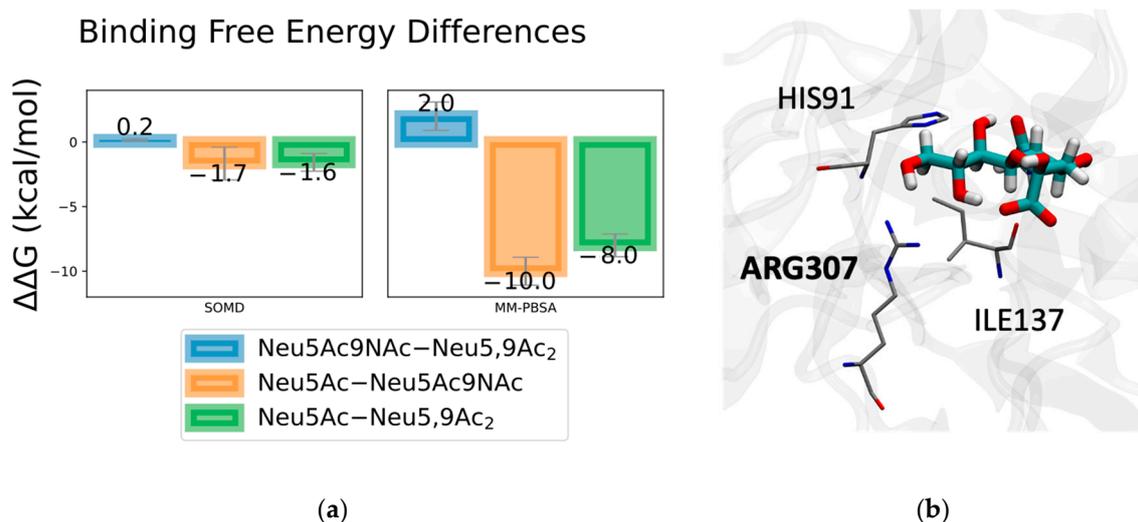
## 2. Results and Discussion

### 2.1. Method Validation with Sia-MERS-CoV S Protein Binding

Modern free energy simulation methods are estimated to be accurate to within 1–2 kcal/mol for well-behaved [26] protein/ligand systems, but this system presents additional challenges due to the exceptional flexibility of both the SARS-CoV-2 S protein and the ligands, the latter of which we have studied both computationally [20] and using NMR [21]. Based on a recent SAMPL6 challenge that evaluates binding free energy prediction methods [27], we used two methods in this study—an inexpensive implicit solvent approach known as molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) [28] to estimate absolute binding free energies, and a more rigorous approach based on alchemical intermediates known as SIRE-OpenMM molecular dynamics (SOMD) [29,30] to compute relative binding free energies of ligands that differ only in the chemical modification at the Sia C9.

Given the exploratory nature of Sia-SARS-CoV-2 S protein binding, we first validated our methods with Sias bound to the RBD in MERS-CoV S protein, known to preferentially bind to Neu5Ac-containing sialosides over Neu5,9Ac<sub>2</sub>- and Neu5Ac9NAc-containing sialosides. For over 1250 ns of combined simulation time, Neu5Ac was observed to stay within the experimentally determined binding site (Neu5Ac remained bound after over 700 ns in a single simulation). Both MM-PBSA binding energy and SOMD alchemical free energy

differences show that Neu5Ac binds stronger than either modified variant, consistent with experimental findings (Figure 3). For ~100 ns of combined simulation time, Neu5,9Ac<sub>2</sub> remained bound, with each simulation unbinding before reaching 40 ns. For ~260 ns of combined simulation time, Neu5,9Ac<sub>2</sub> remained bound, with over 200 ns in a single simulation. Interestingly, after 70 ns of binding, Neu5Ac9NAc unbound and returned the binding pocket within 16 ns.



**Figure 3.** Sia-binding free energy differences in MERS-CoV S protein. **(a)** SOMD and MM-PBSA binding free energy differences for Neu5Ac, Neu5Ac9NAc and Neu5,9Ac<sub>2</sub> bound in MERS-CoV S protein. Neu5Ac binds stronger than Neu5,9Ac<sub>2</sub> and Neu5Ac9NAc, where the Neu5Ac9NAc and Neu5,9Ac<sub>2</sub> result in nominal binding energy differences.; **(b)** Neu5Ac in MERS-CoV S protein binding domain. Top three contributing residues for Neu5Ac are annotated with top contributor bolded. See Figure S3 for energy decomposition results.

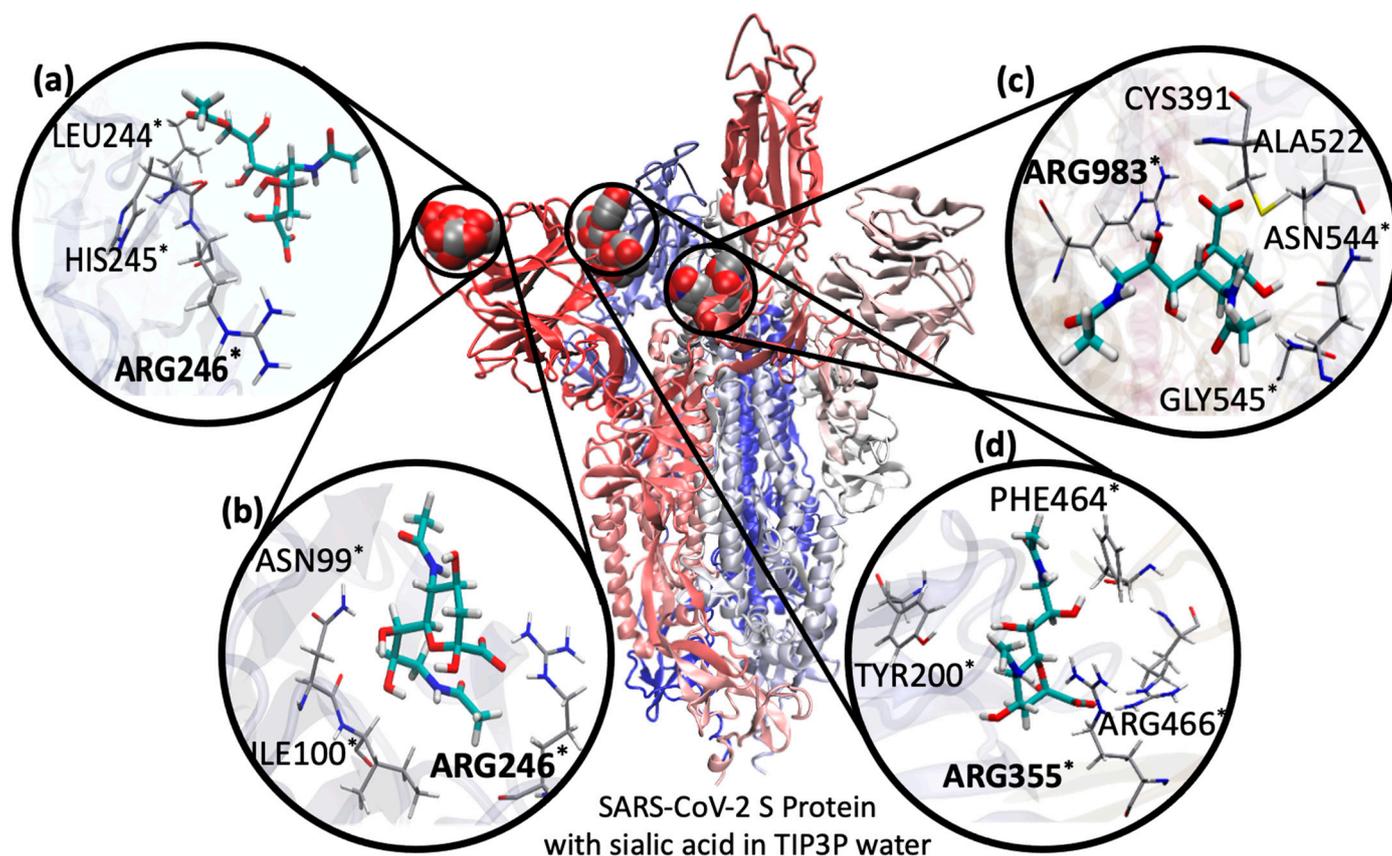
In the case of SOMD, Neu5Ac binds stronger than Neu5Ac9NAc and Neu5,9Ac<sub>2</sub> by 1.6–1.7 kcal/mol. Both Neu5Ac9NAc and Neu5,9Ac<sub>2</sub> bind very similarly, with a nominal difference of 0.2 kcal/mol, within the margin of error. While the magnitudes of the binding free energy differences are larger in the MM-PBSA results, they follow the same trend as our SOMD results, where Neu5Ac binds stronger than Neu5Ac9NAc and Neu5,9Ac<sub>2</sub> (by 8–10 kcal/mol), with the modified Sias binding very similarly. Given the shallow nature of the binding pocket, we did not perform binding free energy simulations with the larger sialosides, as the modified Neu5Ac test case is sufficient for validating our methods. As expected from shallow binding and experimental reports for weak binding affinity [31], MM-PBSA binding energy results indicated overall weak binding (−20 kcal/mol for Neu5Ac). Based on energy decomposition analysis, one of the key residues contributing to this difference in binding is ARG307, which binds stronger to Neu5Ac (Figure 3).

Given that the MERS-CoV simulation predictions largely aligned with the experiment, we proceeded toward computational discovery of Sia-binding sites in SARS-CoV-2 S protein and estimation of binding free energies using the same computational approach.

## 2.2. Discovery and Analysis of Sias-SARS-CoV-2 S Protein Binding

We performed MD simulations of Neu5Ac, Neu5,9Ac<sub>2</sub> and Neu5Ac9NAc bound in SARS-CoV-2 S protein, starting with a docked pose based on the Neu5Ac-MERS-CoV S protein complex (Figure 2, see Methods for details). After initially setting up docked ligands in the cryo-EM structure of SARS-CoV-2 S protein (PDB ID: 6VSB), we found a large RMSD in protein structure after ~200 ns of simulation time indicating a conformational change of the S protein (Figure S1). We observed the single ACE-2 RBD-up moves downwards, and each NTD shrinks inward towards the trimer core, indicating the flexibility of the S1 subunit, specifically the NTD and ACE-2 RBD. Dynamic cross correlation maps support

the structural dependence between these regions, specifically between the NTD and ACE-2 RBD-down and NTD and ACE-2 RBD-up regions (Figure S2) [32]. During the final equilibration phase with no restraints, each Sia (Neu5Ac, Neu5,9Ac<sub>2</sub>, and Neu5Ac9NAc) along the NTD of SARS-CoV-2 S protein became unbound. Interestingly, Neu5,9Ac<sub>2</sub> and Neu5Ac9NAc were observed to return to the original RBD and remained there for ~3 ps before unbinding again. All three Sias subsequently sampled temporary binding events to many regions of the S protein. From the simulation trajectories, we initially selected six poses for further investigation, four of which are in Figure 4.



**Figure 4.** Sia-binding poses with the SARS-CoV-2 S protein. (a) Binding pose based on MERS-CoV with Neu5,9Ac<sub>2</sub> highlighted; (b–d) binding poses based on initial MD, with Neu5Ac9NAc highlighted. Nearby protein residues are annotated, with top contributor bolded, based on MM-PBSA energy decomposition analysis. \* Residues conserved across omicron (PDB ID: 7TB4), kappa (PDB ID: 7VXB), delta (PDB ID: 7W92), gamma (PDB ID: 7M8K), and original (PDB ID: 6VSB) variants of SARS-CoV-2.

All of the proposed sites from MD involve arginine forming a salt bridge with the Sia's carboxylate, a known motif in Sia–lectin binding [24]. In other Sia–S protein observations without an interacting arginine, the Sia quickly dissociated in generally less than 10 ns. While we observed many other binding sites in the MD simulations, such as along the flat top of the NTD within the S1 domain (quick dissociation observed), along the S2 domain and between the S1 and S2 domains (deeper within the S protein trimer), we only considered regions with longer association times and that may be accessible by Sias (see SI video S1 for an overlay of example MD simulations of Sia unbinding and binding events).

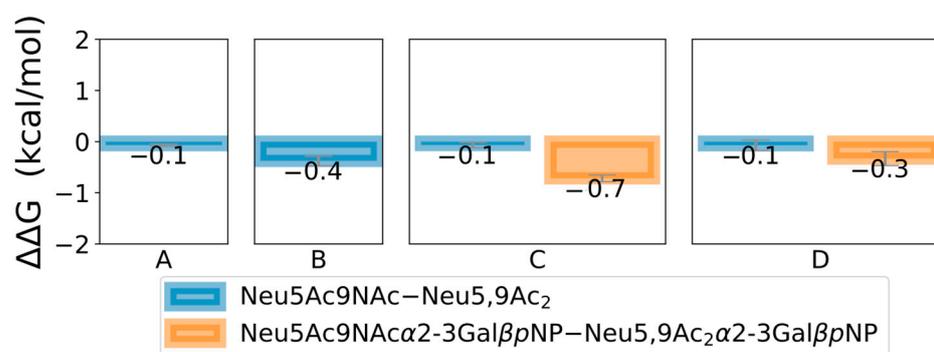
Including the initial pose based on MERS-CoV, we chose four binding poses to investigate in closer detail (Figure 4). Here, we describe how each pose was found in our first set of simulations. Structures of Neu5Ac, Neu5,9Ac<sub>2</sub>, Neu5Ac9NAc were docked into the initial binding pocket along the NTD front (based on MERS-CoV, Figure 4a). Given the shallow pocket, each sialic acid unbound and sampled different regions along the S protein. Neu5Ac and Neu5,9Ac<sub>2</sub> sampled the region behind this initial pose (Figure 4b). This pose

is close to pose a, potentially allowing for more accessible Sia-binding domains along the S protein NTD. Neu5Ac9NAc sampled a deeper binding pose between chains A (NTD) and C (ACE-2 RBD) in the initial simulations (Figure 4c). Neu5Ac sampled the region between the ACE-2 RBD and S2 domain and between chains A & B (Figure 4d). This region is accessible for gangliosides to attach to the S protein and direct or inhibit the down to up state of the ACE-2 RBD. Indeed, a previous study indicated that residues along this region may function as a binding pose for glycosaminoglycans including heparan sulfate [33–35], highlighting the importance of this region in cellular recognition. Both of these deeper binding domains (Figure 4c,d) overlap with key regions in stabilizing the S protein and affecting its ability to transition the ACE-2 RBD between “down” and “up” states [36]. Sia binding in these regions may shift the S protein flexibility and ability to transition between states, altering ACE-2 binding.

Notably, and to the best of our knowledge, these two Sia-binding poses (Figure 4c,d) have not been previously proposed for Sia binding. One possible explanation is that a number of simulations have been performed using a single NTD or subunit rather than the full S protein trimer, which is needed to describe Sias binding between multiple trimer units (such as between chains A & B in pose d).

In all initial MD simulations using the cryo-EM structure, Sia-unbinding events occurred rapidly, given the high flexibility of the SARS-CoV-2 S protein and the resolution of the cryo-EM structure. This flexibility was indicated by relatively fast increase in RMSD at the start of simulations that used the cryo-EM structure. As such, we ran simulations with the SARS-CoV-2 S protein in water to equilibrate the S protein, ideally to a more stable form. We extracted several SARS-CoV-2 S protein structures from the equilibrated trajectories, and used these for docking of the described four poses for production runs, which resulted in longer simulation lengths prior to Sia unbinding. Results from these simulations are in Figure 5.

### SOMD Binding Free Energy Differences in 4 Poses



**Figure 5.** SOMD binding free energy differences for Sias bound in SARS-CoV-2 S protein. Results for transformations between Neu5,9Ac<sub>2</sub> and Neu5Ac9NAc shown for poses (A–D), and between Neu5,9Ac<sub>2</sub>α<sub>2</sub>-3GalβpNP and Neu5Ac9NAcα<sub>2</sub>-3GalβpNP shown for poses c and d. Error bars are plotted from standard error of means across 4 simulations, when available. See Figure S4 for all SOMD results, including Neu5Ac and Neu5Acα<sub>2</sub>-3GalβpNP, and Figure S5 for MM-PBSA energies and energy decomposition results.

Figure 5 shows that in each binding pose, Neu5,9Ac<sub>2</sub> and Neu5Ac9NAc bind very similarly, with differences and errors much less than 1 kcal/mol. Each pose contains a conserved arginine motif, which are conserved across most SARS-CoV-2 variants (omicron (7TB4), kappa (7VXB), delta (7W92), gamma (7M8K), original (6VSB)). Binding residues conserved across these variants are annotated in Figure 4. Neu5Ac binds weaker than either modified Sia in pose a, with slightly weaker binding to LEU244 (Figures S4 and S5, but each Sia binding strongest to ARG246). In pose b, ARG246 is again the tightest binding residue to the Sias, and Neu5Ac binds more strongly than either Neu5,9Ac<sub>2</sub> or Neu5Ac9NAc, likely

due to its size fitting better in the shallow pocket (Figure S4). In pose C, ARG983 was the tightest binding residue, and ARG355 and ARG466 for pose d. Neu5Ac binds slightly stronger than either modified variant, especially to these key arginine residues (Figure S4). Due to their size, Neu5Ac $\alpha$ 2-3Gal $\beta$ pNP, Neu5,9Ac $_2$  $\alpha$ 2-3Gal $\beta$ pNP, and Neu5Ac9NAc $\alpha$ 2-3Gal $\beta$ pNP in pose c were slightly displaced from the deep ARG983, resulting in the top binding residue as LYS529 (Figure S4). While both Neu5,9Ac $_2$  and Neu5Ac9NAc bound weaker than the Neu5Ac, these differences were not large (~1 kcal/mol). In binding pose d, while the Neu5Ac $\alpha$ 2-3Gal $\beta$ pNP binds stronger than Neu5,9Ac $_2$  $\alpha$ 2-3Gal $\beta$ pNP, the standard error of means is relatively large. Larger error bars calculated for binding poses c and d are due to the range of flexibility observed in the S protein, Sias, and the position and orientation of Sia in the binding pocket.

Based on MM-PBSA binding free energies, Sia binding in each pose of the SARS-CoV-2 S protein is weak (<25 kcal/mol), but is consistent with experimental results [9], and is comparable to that of Sia-MERS-CoV S protein binding (see Figures S4 and S5 for MM-PBSA energies and decomposition analysis). Of all the binding poses, binding pose d resulted in tighter binding of  $\alpha$ 2-3-linked Sias over their monomer counterparts in its initial equilibrated pose used in SOMD. This highlights the importance of pose d in tighter SARS-CoV-2 S protein co-binding to cell-surface Sias, influencing cellular recognition and ACE-2 binding.

A number of other possible binding sites were observed in the MD simulations, but we did not investigate them further due to the very short dwell times in the trajectories. Weak binding was also observed with ARG158 and ARG237 along the NTD, but the binding pocket was shallow, and dissociation quickly occurred during the equilibration of the Sias-S protein complex. Binding was also observed to ARG328 along the outside of the S below the ACE-2 RBD, but further simulations also showed the site as shallow and a weak binder.

### 3. Methods

#### 3.1. Choice of Protein Structures, Sialic Acids, and Binding Poses

The initial protein-ligand complex for molecular dynamics simulations was chosen based on available structures from the Protein Databank. At the start of this study, there were no experimental or computationally modeled SARS-CoV-2 S protein structures with sialic acids bound, to the best of our knowledge. We therefore searched for other human beta coronavirus S protein structures available on the Protein Databank with a bound sialic acid. The MERS-CoV S protein (PDB ID: 6Q04 [3]) has a 35% identity with SARS-CoV-2 S protein, and has an available structure with a bound sialic acid (Neu5Ac). There is no known Sia binding or bound complex with SARS-CoV [1], although it shares a higher sequence identity of 73% with the SARS-CoV-2 S protein (PDB ID: 6VSB [37]).

To generate an initial sialic acid binding pose with the SARS-CoV-2 S protein, we performed sequence alignment with both S proteins (6VSB and 6Q04) using MultiSeq, [38] an extension of the Multiple Alignment tool available in VMD [39], a freely available structural graphics program for visualization and analysis. From the aligned structures and conserved binding residues, notably arginine, we placed the Sia directly into the overlaid SARS-CoV-2 S protein N-terminal domain and used the resulting complex as our initial structure (Figure 2).

For additional ligand-protein complexes in both S proteins, we docked Sias (Figure 1) into binding pockets using the OEDocking tool [40–43]. We docked Neu5Ac, Neu5,9Ac $_2$ , and Neu5Ac9NAc into each S protein, based on evidence that other beta coronaviruses (OC43 and HKUI [1,4]) prefer to bind Neu5,9Ac $_2$  over Neu5Ac, and that N-acetylated Sias have shown to be chemically and structurally reasonable mimics to O-acetylated counterparts. Given that MERS-CoV S protein preferentially binds to  $\alpha$ 2-3 over  $\alpha$ 2-6-linked sialic acids, we docked  $\alpha$ 2-3-linked sialic acids (Neu5Ac $\alpha$ 2-3Gal $\beta$ pNP, Neu5,9Ac $_2$  $\alpha$ 2-3Gal $\beta$ pNP, and Neu5Ac9NAc $\alpha$ 2-3Gal $\beta$ pNP) into deeper binding pockets of the SARS-CoV-2 S protein, where these new poses were inferred from the molecular dynamics simulations in which a free sialic acid would associate to various residues on the S protein.

### 3.2. Molecular Dynamics Simulation Setup

Molecular dynamics simulations were performed using AMBER18 software suite, using *tleap* for setup and *pmemd.cuda* for dynamics, running on servers equipped with Intel Xeon CPUs and Nvidia GTX 980 Ti or 1080 Ti GPUs [44–48]. We used the ff14SB protein force field, GLYCAM06 carbohydrate force field, the GAFF small molecule force field for portions of the sialosides, and the TIP3P water model. Protein or protein–ligand complexes were solved in a truncated octahedron box with 12.0 Å padding between the biomolecule and simulation cell edge, resulting in around 120,000 water molecules, and Na<sup>+</sup> ions were added for a net neutral charge. Simulations were run with periodic boundary conditions: a 2.0 fs time step, Langevin thermostat set to 298.15 K, and a collision frequency of 5.0 ps<sup>−1</sup>. The particle mesh Ewald method was used to treat long-range electrostatics with a real-space cutoff of 9.0 Å, and the SHAKE algorithm used to constrain all bonds involving hydrogen. With restraints on complex heavy atoms with a restraint weight of 10 kcal mol<sup>−1</sup> Å<sup>−2</sup>, 1000 minimization steps were performed, followed by an additional 500 steps with no constraints. We subsequently heated our system from 0.1 K to 298.15 K in a single simulation with equally spaced temperatures across 500 ps, at constant pressure using a Berendsen barostat set to 1.0 atm and a compressibility of 4.5 × 10<sup>−5</sup> bar<sup>−1</sup>, and harmonic restraints on the protein complex heavy atoms with a force constant of 10 kcal mol<sup>−1</sup> Å<sup>−2</sup>. Subsequent equilibration was performed at constant pressure using 5 × 100 ns simulations in which the harmonic restraints were set to 10.0, 1.0, 0.1, 0.01, and 0 kcal mol<sup>−1</sup> Å<sup>−2</sup>, respectively. Three production simulations were performed with the NVT ensemble using the simulation cell volume taken from the final structure of the equilibration run.

Trajectory analyses were carried out using the *cpptraj* [49] simulation analysis package and free energy differences plots generated with *matplotlib* plotting software package in python. Molecular structures were visualized with VMD [39].

The structures used in molecular dynamics simulations were 6VSB and 6Q04, each without sialylation.

### 3.3. Computation of Binding Free Energies Using MM-PBSA and Alchemical Simulation Approaches

This study used a combination of two approaches for computing binding free energies: a relatively approximate and inexpensive Molecular Mechanics Poisson–Boltzmann Surface Area (MM-PBSA) approach to estimate binding free energies and a relatively accurate and computationally expensive approach using alchemical intermediates to estimate binding free energy differences between chemical modifications on the sialic acid ligand.

The MM-PBSA calculations used structures taken from the first set of molecular dynamics simulations, production runs, and the equilibration runs for the alchemical free energy simulations, and results were averaged across all sampled trajectories [50–52]. MM-PBSA can be used to calculate free energy differences, where the linearized Poisson–Boltzmann equation is used to compute the solvation free energy components (charge distribution and solvation free energy), and the LCPO method implemented in *sander* is used to generate an empirical term for hydrophobic contributions based on surface area. The binding free energy is the difference between the sum of ligand–protein complex binding and its solvation and the sum of the solvation energies of the free ligand and protein (Figure S6). Implicit solvation in the thermodynamic cycle calculations reduces computational time compared to the explicit water molecules used for solvation in the above alchemical free energy approach. Key interacting protein residues are identified by energy decomposition analysis (EDA). Calculations were performed using MMPBSA.py [30] AMBER software, from molecular dynamics simulation trajectories. Simulations ran for up to 800 ns, and were terminated earlier if the Sia was observed to leave the binding pocket, resulting in an average of 300 ns trajectories for Sias in the SARS-CoV-2 S protein. This was deemed sufficient time for relative binding free energies, as our total simulation lengths sampled in our MM-PBSA calculations are greater than the 1–10 ns range tested as

sufficient in a paper studying MM-PBSA protein–ligand binding free energy convergences from crystal structures [50].

### 3.4. Alchemical Free Energy Simulations: Setup and Procedure

Alchemical free energy simulations calculate the free energy difference between two chemical states (i.e., ligand bound vs. unbound) by employing fictitious “alchemical” intermediate states to improve thermodynamic overlap. These simulations were setup in a similar fashion, but employed the automated Sire FESetup [53] tool to prepare systems for alchemical free energy simulations using the Sire and OpenMM software packages (SOMD) [32,54]. Due to differences in the software, there were some differences in the simulation details: we used a rectangular box with a 12.0 Å padding, performed minimization over 300 steps, heated the system from 5.0 K to 298.15 K, and ran equilibration simulations with decreasing constraint force constants in order of 8.0, 6.0, 4.0, 2.0, and 0 kcal mol<sup>-1</sup> Å<sup>-2</sup>; the first equilibration ran for 100 ps, and subsequent simulations ran for 100 ns, as in the above case. We again used FESetup to generate the simulation inputs for the alchemical intermediates describing the transformation occurring (such as Neu5Ac to Neu5,9Ac<sub>2</sub>). These intermediate systems used dummy atoms and interpolated values of parameters, as well as soft-core potentials to describe partial electrostatic and van der Waals interactions according to the following, Equations (1) and (2) [55]:

$$r_{\text{LJ}} = (2\sigma_{ij}\lambda + r_{ij}^2)^{1/2}, \quad (1)$$

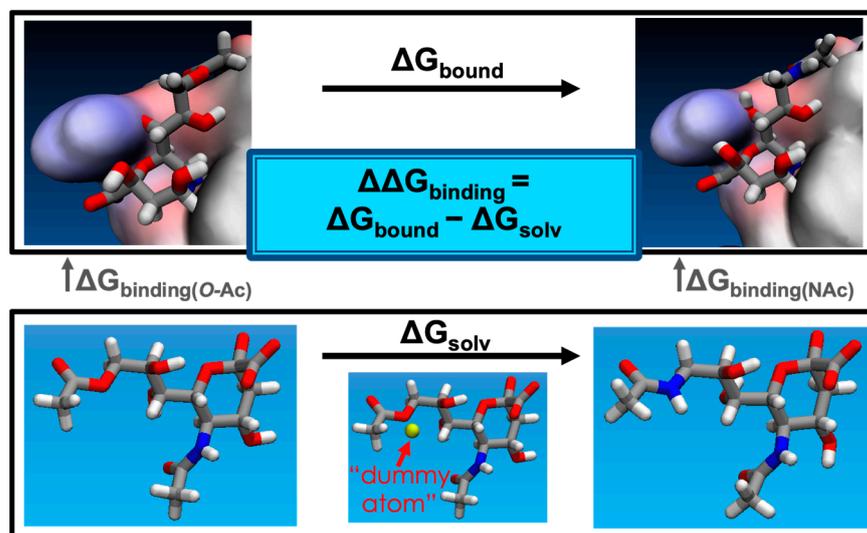
$$r_{\text{Coul}} = (\lambda + r_{ij}^2)^{1/2}, \quad (2)$$

Here,  $r_{\text{LJ}}$  and  $r_{\text{Coul}}$  are “renormalized” distances that are input into the Lennard-Jones and Coulomb potentials, respectively, ensuring that the potential never reaches the singularity at  $r = 0$ , and  $r_{ij}$  represents the distance between disappearing/appearing atoms  $i/j$ . The value of  $\lambda$  ranges from 0.0 to 1.0, in which  $\lambda = 0.0$  represents the full interaction with the original ligand’s transforming/disappearing atoms, and  $\lambda = 1.0$  represents no interaction with these atoms. For the appearing atoms,  $\lambda$  is replaced with  $1 - \lambda$ , such that interaction is fully turned on at  $\lambda = 1.0$  (while full interaction is turned off for the disappearing atoms).

Ligands were transformed over 13–21 linearly-spaced  $\lambda$  windows, based on a previous study that used SOMD and other alchemical free energy methods in an assessment of binding affinities [56], and following the scheme outlined in the SI of Loeffler et al. [55]. In general, the first set of simulations for each transformation was run with 13 windows, and the second set of simulations with 21 for greater certainty. Each simulation was minimized over 1000 steps prior to 5 ns production run in the NPT ensemble, where free energies were computed from the last 4 ns, based on the assessments from Kuhn et al. and Mey et al. [56,57]. Each transformation was computed in the reverse direction, and simulations were repeated in the forward and reverse directions, starting from the same equilibrated starting complex. For example, Neu5,9Ac<sub>2</sub> was transformed to Neu5Ac9NAc by slowly turning off the interactions of Neu5,9Ac<sub>2</sub> and slowly turning them on for Neu5Ac9NAc, over a set of discrete steps, as described the SI of Loeffler et al. [55]. We used *analyse\_frenerg* to calculate free energy differences at 298 K using multistate Bennett acceptance ratio (MBAR) and thermodynamic integration (TI) as a reference estimate [58–60].

The relative free energy of binding is computed as the free energy difference of the alchemical transformation of the ligand bound to the protein and the ligand in solution, i.e., water (Figure 6). The reliability of results was determined based on the similarity between MBAR and TI, and from overlap matrices, where the first off-diagonal element is at least 0.03 [57], but ideally close to 0.10 to ensure enough thermodynamic overlap between simulation windows (Figure S7). Some test cases of longer simulations were performed, but the thermodynamic overlap did not significantly improve compared to more windows. The high degree of flexibility known for the SARS-CoV-2 S protein may contribute to this observation. Thus, we increased the number of windows when necessary,

and duplicate simulations were run for higher certainty and increased sampling. Solvated ligand simulations each took about 45 min, and the solvated ligand-protein complex calculations took about 1 day each, running on NVIDIA GTX 980 Ti or 1080 Ti hardware.



**Figure 6.** Visual representation of  $\Delta\Delta G_{\text{binding}}$  of Neu5Ac9NAc versus Neu5,9Ac<sub>2</sub> in the SARS-CoV-2 S protein, where  $\Delta\Delta G_{\text{binding}}$  is the difference of two alchemical transformations in explicit solvent (boxed). The upper box is the transformation from Neu5,9Ac<sub>2</sub> to Neu5Ac9NAc bound in the protein and solvated in explicit water. The lower box is the transformation in explicit water. Additional atoms required in transformations are represented by “dummy atoms” whose interactions are gradually turned on as one progresses over the alchemical intermediates.

### 3.5. Sialic Acid Parametrization Procedure

We parametrized *p*NP with the same approach we used for Neu5,9Ac<sub>2</sub> and Neu5Ac9NAc in our previous study, which was designed to maximize compatibility with the GLYCAM format [20,61]. Here, we briefly describe the procedure. A set of atomic partial charges was derived for *p*NP. The procedure involves averaging over an ensemble of structures, sampled using a MD simulation of Gal $\beta$ *p*NP in TIP3P water. We used a GLYCAM06 force field for Gal and GAFF [62] for the para-nitrophenol (*p*NP) aglycon. The system was equilibrated for 1 ns at 298.15 K, 1.0 atm (NPT), followed by a 100 ns production run at 298.15 K (NVT). One hundred structures of Gal $\beta$ *p*NP were saved at 2 ns intervals and energy-minimized at the HF/6-31G\* level of theory, with all exocyclic dihedral angles constrained to their MD-sampled values. For each of the 100 constrained energy-minimized structures, electrostatic potential (ESP) calculations were computed using Gaussian16 software [63] and ESP data used in a single-stage restrained ESP fitting (RESP) calculation, with a restraint weight of 0.01 applied to all atoms using the resp program from AmberTools. An arithmetic average over the 100 sets of fitted charges yielded the final set of charges in the model.

The GLYCAM and GAFF force fields lacked torsional parameters to describe the torsional energy profiles about the *p*NP functional group and Gal $\beta$ -*p*NP linkage; these parameters were derived by fitting to reproduce torsional profiles from DFT at the HF/6-31G\* level of theory. We used the TeraChem quantum chemistry software for the energy minimizations and the *torsiondrive* software to scan over the dihedral angles recursively [64–68]. The optimized geometries were used for single-point energies and atomistic forces calculated at the  $\omega$ B97X-D3/6-31++G(2d,2p) level of theory [69]. The parameters were optimized by fitting to the quantum chemical energies using the ForceBalance optimization software [68,70]. Bond stretching and angle bending parameters for the Gal $\beta$ *p*NP linkage were copied from analogous parameters available in GLYCAM06. The values of optimized parameters and simulation-ready parameter files are provided in Table S1.

#### 4. Conclusions

We have identified multiple weak Sia-binding sites on the SARS-CoV-2 S protein, with two novel Sia-binding poses. Each novel pose, between the NTD and ACE-2 RBD (chains A and C) and between the ACE-2 RBD and S2 domain (chains A and B), is accessible to gangliosides for S protein attachment, and overlaps with regions known in stabilizing the S protein. The binding in each pose is predicted to be weak, in overall agreement with glycan microarray experiments [6,9]. We validated our methods using MERS-CoV-S protein, confirming the experimental result that Neu5Ac binds stronger than that with either 9-OAc or 9-NAc modification, and binding energies comparable to Sia-SARS-CoV-2 S protein binding. Each Sia-binding pose in SARS-CoV-2 S protein contains an arginine residue that is conserved across SARS-CoV-2 variants. The multiple Sia-binding sites on SARS-CoV-2 S protein may lead to increased binding affinity to multiple Sias collocated on the cell surface, and the existence of multiple binding sites to the S protein may be validated experimentally. Neu5Ac in SARS-CoV-2 S protein binding sites tend to have stronger binding than their 9-OAc or 9-NAc modifications, but the differences are within the expected margin of error. Binding free energies support 9-NAc Sias as close structural and chemical mimics of 9-OAc Sias in SARS-CoV-2 and MERS-CoV S proteins, given the small energy differences (all <1 kcal/mol with SOMD). This supports 9-NAc Sias as an experimentally stable mimic of 9-OAc counterparts to probe Sia-virus binding. Future experimental studies can validate our differential binding free energy results and understanding the binding of modified Sias can further elucidate the role of Sias in cellular recognition and the high transmissibility of SARS-CoV-2.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/molecules27165322/s1>, Figure S1: Conformational change of the SARS-CoV-2 S protein; Figure S2: Dynamic cross correlation heat map supporting conformational change in the SARS-CoV-2 S protein; Figure S3: MM-PBSA energies and decomposition analysis for Neu5,9Ac2, Neu5Ac9NAc and Neu5Ac in MERS-CoV S protein; Figure S4: SOMD relative binding free energy differences for Sias binding poses with the SARS-CoV-2 S protein; Figure S5: MM-PBSA energies and decomposition analysis for Neu5,9Ac2, Neu5Ac9NAc, Neu5Ac, Neu5,9Ac2 $\alpha$ 2-3Gal $\beta$ pNP, Neu5Ac9NAc $\alpha$ 2-3Gal $\beta$ pNP, and Neu5Ac $\alpha$ 2-3Gal $\beta$ pNP in all binding poses (a-d) of SARS-CoV-2 S protein; Figure S6: Thermodynamic cycle to estimate binding free energies; Figure S7: Representative binding free energy difference of Neu5Ac—Neu5,9Ac2 in the SARS-CoV-2 S protein; Table S1: Simulation-ready parameter files for the pNP residue in AMBER; Video S1: Overlay of example MD simulations of Sia unbinding and binding events in the SARS-CoV-2 S protein (6VSB).

**Author Contributions:** Conceptualization, L.O. and L.-P.W.; methodology, L.O.; software, L.O.; validation, L.O.; formal analysis, L.O.; investigation, L.O.; resources, L.-P.W., X.C. and A.V.; data curation, L.O.; writing—original draft preparation, L.O.; writing—review and editing, L.O., L.-P.W., X.C. and A.V.; visualization, L.O.; supervision, L.-P.W., X.C. and A.V.; project administration, L.-P.W., X.C. and A.V.; funding acquisition, L.-P.W., X.C. and A.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the United States National Institutes of Health, grant no. R01AI130684 and R01AI130684-03S1.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available in this article.

**Acknowledgments:** We are grateful for the insightful discussions with Caleb Oh at the UC Davis Medical Center.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Sample Availability:** Not applicable.

## References

1. 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. *Nat. Struct. Mol. Biol.* **2019**, *26*, 481–489. [[CrossRef](#)] [[PubMed](#)]
2. Varki, A.; Schauer, R. Sialic Acids. In *Essentials of Glycobiology*; Varki, A., Cummings, R.D., Esko, J.D., Freeze, H.H., Stanley, P., Bertozzi, C.R., Hart, G.W., Etzler, M.E., Eds.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2009; ISBN 978-0-87969-770-9.
3. Park, Y.-J.; Walls, A.C.; Wang, Z.; Sauer, M.M.; Li, W.; Tortorici, M.A.; Bosch, B.-J.; DiMaio, F.; Velesler, D. Structures of MERS-CoV Spike Glycoprotein in Complex with Sialoside Attachment Receptors. *Nat. Struct. Mol. Biol.* **2019**, *26*, 1151–1157. [[CrossRef](#)] [[PubMed](#)]
4. Hulswit, R.J.G.; Lang, Y.; Bakkers, M.J.G.; Li, W.; Li, Z.; Schouten, A.; Ophorst, B.; van Kuppeveld, F.J.M.; Boons, G.-J.; Bosch, B.-J.; et al. Human Coronaviruses OC43 and HKU1 Bind to 9-O-Acetylated Sialic Acids via a Conserved Receptor-Binding Site in Spike Protein Domain A. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 2681–2690. [[CrossRef](#)] [[PubMed](#)]
5. Li, Z.; Lang, Y.; Liu, L.; Bunyatov, M.I.; Sarmiento, A.I.; de Groot, R.J.; Boons, G.-J. Synthetic O-Acetylated Sialosides Facilitate Functional Receptor Identification for Human Respiratory Viruses. *Nat. Chem.* **2021**, *13*, 496–503. [[CrossRef](#)] [[PubMed](#)]
6. Nguyen, L.; McCord, K.A.; Bui, D.T.; Bouwman, K.M.; Kitova, E.N.; Elaish, M.; Kumawat, D.; Daskhan, G.C.; Tomris, I.; Han, L.; et al. Sialic Acid-Containing Glycolipids Mediate Binding and Viral Entry of SARS-CoV-2. *Nat. Chem. Biol.* **2022**, *18*, 81–90. [[CrossRef](#)]
7. Nguyen, K.; Chakraborty, S.; Mansbach, R.A.; Korber, B.; Gnanakaran, S. Exploring the Role of Glycans in the Interaction of SARS-CoV-2 RBD and Human Receptor ACE2. *Viruses* **2021**, *13*, 927. [[CrossRef](#)]
8. 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 Cent. Sci.* **2020**, *6*, 2046–2052. [[CrossRef](#)]
9. Dhar, C.; Sasmal, A.; Diaz, S.; Verhagen, A.; Yu, H.; Li, W.; Chen, X.; Varki, A. Are Sialic Acids Involved in COVID-19 Pathogenesis? *Glycobiology* **2021**, *31*, 1068–1071. [[CrossRef](#)]
10. 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. *Front. Chem.* **2021**, *9*, 659764. [[CrossRef](#)]
11. Wang, C.-W.; Lee, O.K.; Fischer, W.B.; Wang, C.-W.; Lee, O.K.; Fischer, W.B. Screening Coronavirus and Human Proteins for Sialic Acid Binding Sites Using a Docking Approach. *AIMS Biophys.* **2021**, *8*, 248–263. [[CrossRef](#)]
12. 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. *Front. Mol. Biosci.* **2021**, *8*, 509. [[CrossRef](#)] [[PubMed](#)]
13. Fantini, J.; Di Scala, C.; Chahinian, H.; Yahi, N. Structural and Molecular Modelling Studies Reveal a New Mechanism of Action of Chloroquine and Hydroxychloroquine against SARS-CoV-2 Infection. *Int. J. Antimicrob. Agents* **2020**, *55*, 105960. [[CrossRef](#)] [[PubMed](#)]
14. Zamorano Cuervo, N.; Grandvaux, N. ACE2: Evidence of Role as Entry Receptor for SARS-CoV-2 and Implications in Comorbidities. *eLife* **2020**, *9*, e61390. [[CrossRef](#)] [[PubMed](#)]
15. Vandelli, A.; Monti, M.; Milanetti, E.; Armaos, A.; Rupert, J.; Zacco, E.; Bechara, E.; Delli Ponti, R.; Tartaglia, G.G. Structural Analysis of SARS-CoV-2 Genome and Predictions of the Human Interactome. *Nucleic Acids Res.* **2020**, *48*, 11270–11283. [[CrossRef](#)]
16. Seyran, M.; Takayama, K.; Uversky, V.N.; Lundstrom, K.; Palù, G.; Sherchan, S.P.; Attrish, D.; Rezaei, N.; Aljabali, A.A.A.; Ghosh, S.; et al. The Structural Basis of Accelerated Host Cell Entry by SARS-CoV-2. *FEBS J.* **2021**, *288*, 5010–5020. [[CrossRef](#)]
17. Awasthi, M.; Gulati, S.; Sarkar, D.P.; Tiwari, S.; Kateriya, S.; Ranjan, P.; Verma, S.K. The Sialoside-Binding Pocket of SARS-CoV-2 Spike Glycoprotein Structurally Resembles MERS-CoV. *Viruses* **2020**, *12*, 909. [[CrossRef](#)]
18. Huang, Y.; Yang, C.; Xu, X.; Xu, W.; Liu, S. Structural and Functional Properties of SARS-CoV-2 Spike Protein: Potential Antivirus Drug Development for COVID-19. *Acta Pharmacol. Sin.* **2020**, *41*, 1141–1149. [[CrossRef](#)]
19. Sun, X.-L. The Role of Cell Surface Sialic Acids for SARS-CoV-2 Infection. *Glycobiology* **2021**, *31*, 1245–1253. [[CrossRef](#)]
20. Khedri, Z.; Xiao, A.; Yu, H.; Landig, C.S.; Li, W.; Diaz, S.; Wasik, B.R.; Parrish, C.R.; Wang, L.-P.; Varki, A.; et al. A Chemical Biology Solution to Problems with Studying Biologically Important but Unstable 9-O-Acetyl Sialic Acids. *ACS Chem. Biol.* **2017**, *12*, 214–224. [[CrossRef](#)]
21. Li, W.; Battistel, M.D.; Reeves, H.; Oh, L.; Yu, H.; Chen, X.; Wang, L.-P.; Freedberg, D.I. A Combined NMR, MD and DFT Conformational Analysis of 9-O-Acetyl Sialic Acid-Containing GM3 Ganglioside Glycan and Its 9-N-Acetyl Mimic. *Glycobiology* **2020**, *30*, 787–801. [[CrossRef](#)]
22. Ji, Y.; Sasmal, A.; Li, W.; Oh, L.; Srivastava, S.; Hargett, A.A.; Wasik, B.R.; Yu, H.; Diaz, S.; Choudhury, B.; et al. Reversible O-Acetyl Migration within the Sialic Acid Side Chain and Its Influence on Protein Recognition. *ACS Chem. Biol.* **2021**, *16*, 1951–1960. [[CrossRef](#)] [[PubMed](#)]
23. Oh, L.; Ji, Y.; Li, W.; Varki, A.; Chen, X.; Wang, L.-P. O-Acetyl Migration within the Sialic Acid Side Chain: A Mechanistic Study by the Ab Initio Nanoreactor. Department of Chemistry, University of California, Davis, CA, USA. 2022, *submitted*.
24. Varki, A.; Angata, T. Siglecs—The Major Subfamily of I-Type Lectins. *Glycobiology* **2006**, *16*, 1R–27R. [[CrossRef](#)] [[PubMed](#)]
25. Marsh, L. Strong Ligand-Protein Interactions Derived from Diffuse Ligand Interactions with Loose Binding Sites. *BioMed Res. Int.* **2015**, *2015*, e746980. [[CrossRef](#)]

26. Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M.K.; Greenwood, J.; et al. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J. Am. Chem. Soc.* **2015**, *137*, 2695–2703. [[CrossRef](#)]
27. Rizzi, A.; Murkli, S.; McNeill, J.N.; Yao, W.; Sullivan, M.; Gilson, M.K.; Chiu, M.W.; Isaacs, L.; Gibb, B.C.; Mobley, D.L.; et al. Overview of the SAMPL6 Host–Guest Binding Affinity Prediction Challenge. *J. Comput. Aided Mol. Des.* **2018**, *32*, 937–963. [[CrossRef](#)] [[PubMed](#)]
28. Miller, B.R.; McGee, T.D.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321. [[CrossRef](#)]
29. Eastman, P.; Swails, J.; Chodera, J.D.; McGibbon, R.T.; Zhao, Y.; Beauchamp, K.A.; Wang, L.-P.; Simmonett, A.C.; Harrigan, M.P.; Stern, C.D.; et al. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLoS Comput. Biol.* **2017**, *13*, e1005659. [[CrossRef](#)]
30. Woods, C.J.; Mey, A.S.; Calabro, G.; Julien, M. Sire Molecular Simulation Framework. Version 2020.1.0. Available online: <http://Siremol.Org.2020.1.0> (accessed on 19 May 2020).
31. Li, W.; Hulswit, R.J.G.; Widjaja, I.; Raj, V.S.; McBride, R.; Peng, W.; Widagdo, W.; Tortorici, M.A.; van Dieren, B.; Lang, Y.; et al. Identification of Sialic Acid-Binding Function for the Middle East Respiratory Syndrome Coronavirus Spike Glycoprotein. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8508–E8517. [[CrossRef](#)]
32. Brown, D.K.; Penkler, D.L.; Sheik Amamuddy, O.; Ross, C.; Atilgan, A.R.; Atilgan, C.; Tastan Bishop, Ö. MD-TASK: A Software Suite for Analyzing Molecular Dynamics Trajectories. *Bioinformatics* **2017**, *33*, 2768–2771. [[CrossRef](#)]
33. 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. [[CrossRef](#)]
34. Kim, S.H.; Kearns, F.L.; Rosenfeld, M.A.; Casalino, L.; Papanikolas, M.J.; Simmerling, C.; Amaro, R.E.; Freeman, R. GlycoGrip: Cell Surface-Inspired Universal Sensor for Betacoronaviruses. *ACS Cent. Sci.* **2022**, *8*, 22–42. [[CrossRef](#)] [[PubMed](#)]
35. 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. [[CrossRef](#)] [[PubMed](#)]
36. Henderson, R.; Edwards, R.J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S.M.C.; Kopp, M.; Li, D.; Parks, R.; Hsu, A.L.; et al. Controlling the SARS-CoV-2 Spike Glycoprotein Conformation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 925–933. [[CrossRef](#)]
37. Wrapp, D.; Wang, N.; Corbett, K.S.; Goldsmith, J.A.; Hsieh, C.-L.; Abiona, O.; Graham, B.S.; McLellan, J.S. Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation. *Science* **2020**, *367*, 1260–1263. [[CrossRef](#)]
38. Roberts, E.; Eargle, J.; Wright, D.; Luthey-Schulten, Z. MultiSeq: Unifying Sequence and Structure Data for Evolutionary Analysis. *BMC Bioinform.* **2006**, *7*, 382. [[CrossRef](#)] [[PubMed](#)]
39. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [[CrossRef](#)]
40. McGann, M. FRED and HYBRID Docking Performance on Standardized Datasets. *J. Comput. Aided Mol. Des.* **2012**, *26*, 897–906. [[CrossRef](#)]
41. McGann, M. FRED Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.* **2011**, *51*, 578–596. [[CrossRef](#)]
42. Kelley, B.P.; Brown, S.P.; Warren, G.L.; Muchmore, S.W. POSIT: Flexible Shape-Guided Docking For Pose Prediction. *J. Chem. Inf. Model.* **2015**, *55*, 1771–1780. [[CrossRef](#)]
43. OEDOCKING 3.3.1: OpenEye Scientific Software, Inc., Santa Fe, NM. Available online: <http://www.eyesopen.com> (accessed on 28 May 2019).
44. Götz, A.W.; Williamson, M.J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R.C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542–1555. [[CrossRef](#)]
45. Le Grand, S.; Götz, A.W.; Walker, R.C. SPFP: Speed without Compromise—A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374–380. [[CrossRef](#)]
46. Salomon-Ferrer, R.; Gotz, A.W.; Poole, D.; Le Grand, S.; Walker, R.C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888. [[CrossRef](#)] [[PubMed](#)]
47. Salomon-Ferrer, R.; Case, D.A.; Walker, R.C. An Overview of the Amber Biomolecular Simulation Package. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198–210. [[CrossRef](#)]
48. Case, D.A.; Ben-Shalom, I.Y.; Brozell, S.R.; Cerutti, D.S.; Cheatham, T.E., III; Cruzeiro, V.W.D.; Darden, T.A.; Duke, R.E.; Ghoreishi, D.; Gilson, M.K.; et al. *AMBER 2018*; University of California: San Francisco, CA, USA, 2018.
49. Roe, D.R.; Cheatham, T.E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095. [[CrossRef](#)] [[PubMed](#)]
50. Wang, C.; Nguyen, P.H.; Pham, K.; Huynh, D.; Le, T.-B.N.; Wang, H.; Ren, P.; Luo, R. Calculating Protein-Ligand Binding Affinities with MMPBSA: Method and Error Analysis: Calculating Protein-Ligand Binding Affinities. *J. Comput. Chem.* **2016**, *37*, 2436–2446. [[CrossRef](#)] [[PubMed](#)]
51. Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889–897. [[CrossRef](#)]

52. Srinivasan, J.; Miller, J.; Kollman, P.A.; Case, D.A. Continuum Solvent Studies of the Stability of RNA Hairpin Loops and Helices. *J. Biomol. Struct. Dyn.* **1998**, *16*, 671–682. [[CrossRef](#)]
53. Woods, C. FESetup: Automating Setup for Alchemical Free Energy Simulations. *J. Chem. Inf. Modeling* **2015**, *55*, 2485–2490.
54. Mey, A.S.J.S.; Jiménez, J.J.; Michel, J. Impact of Domain Knowledge on Blinded Predictions of Binding Energies by Alchemical Free Energy Calculations. *J. Comput. Aided Mol. Des.* **2018**, *32*, 199–210. [[CrossRef](#)]
55. Loeffler, H.H.; Bosisio, S.; Duarte Ramos Matos, G.; Suh, D.; Roux, B.; Mobley, D.L.; Michel, J. Reproducibility of Free Energy Calculations across Different Molecular Simulation Software Packages. *J. Chem. Theory Comput.* **2018**, *14*, 5567–5582. [[CrossRef](#)]
56. Kuhn, M.; Firth-Clark, S.; Tosco, P.; Mey, A.S.J.S.; Mackey, M.; Michel, J. Automated Assessment of Binding Affinity via Alchemical Free Energy Calculations. *J. Chem. Inf. Model.* **2020**, *60*, 3120–3130. [[CrossRef](#)] [[PubMed](#)]
57. Mey, A.S.J.S.; Allen, B.K.; Macdonald, H.E.B.; Chodera, J.D.; Hahn, D.F.; Kuhn, M.; Michel, J.; Mobley, D.L.; Naden, L.N.; Prasad, S.; et al. Best Practices for Alchemical Free Energy Calculations [Article v1.0]. *Living J. Comput. Mol. Sci.* **2020**, *2*, 18378. [[CrossRef](#)] [[PubMed](#)]
58. Shirts, M.R.; Chodera, J.D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105. [[CrossRef](#)] [[PubMed](#)]
59. Kirkwood, J.G. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313. [[CrossRef](#)]
60. Merz, K.M.; Kollman, P.A. Free Energy Perturbation Simulations of the Inhibition of Thermolysin: Prediction of the Free Energy of Binding of a New Inhibitor. *J. Am. Chem. Soc.* **1989**, *111*, 5649–5658. [[CrossRef](#)]
61. Kirschner, K.N.; Yongye, A.B.; Tschampel, S.M.; González-Outeiriño, J.; Daniels, C.R.; Foley, B.L.; Woods, R.J. GLYCAM06: A Generalizable Biomolecular Force Field. *Carbohydrates. J. Comput. Chem.* **2008**, *29*, 622–655. [[CrossRef](#)]
62. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [[CrossRef](#)]
63. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Petersson, G.A.; Nakatsuji, H.; et al. *Gaussian 16, Revision C.01*; Gaussian, Inc.: Wallingford, CT, USA, 2016.
64. Ufimtsev, I.S.; Martínez, T.J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619–2628. [[CrossRef](#)]
65. Ufimtsev, I.S.; Martínez, T.J. Quantum Chemistry on Graphical Processing Units. 1. Strategies for Two-Electron Integral Evaluation. *J. Chem. Theory Comput.* **2008**, *4*, 222–231. [[CrossRef](#)]
66. Ufimtsev, I.S.; Martínez, T.J. Quantum Chemistry on Graphical Processing Units. 2. Direct Self-Consistent-Field Implementation. *J. Chem. Theory Comput.* **2009**, *5*, 1004–1015. [[CrossRef](#)]
67. Wang, L.-P.; Song, C. Geometry Optimization Made Simple with Translation and Rotation Coordinates. *J. Chem. Phys.* **2016**, *144*, 214108. [[CrossRef](#)] [[PubMed](#)]
68. Wang, L.-P.; McKiernan, K.A.; Gomes, J.; Beauchamp, K.A.; Head-Gordon, T.; Rice, J.E.; Swope, W.C.; Martínez, T.J.; Pande, V.S. Building a More Predictive Protein Force Field: A Systematic and Reproducible Route to AMBER-FB15. *J. Phys. Chem. B* **2017**, *121*, 4023–4039. [[CrossRef](#)] [[PubMed](#)]
69. Kesharwani, M.K.; Karton, A.; Martin, J.M.L. Benchmark Ab Initio Conformational Energies for the Proteinogenic Amino Acids through Explicitly Correlated Methods. Assessment of Density Functional Methods. *J. Chem. Theory Comput.* **2016**, *12*, 444–454. [[CrossRef](#)] [[PubMed](#)]
70. Wang, L.-P.; Martínez, T.J.; Pande, V.S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891. [[CrossRef](#)]