Prediction of Single Point Mutations in Ganglioside-Binding Domain of SARS-CoV-2 S and Their Effects on Binding of 9-O-Acetylated Sialic Acid and Hidroxychloroquine

Petar M. Mitrasinovic*

Center for Biophysical and Chemical Research, Belgrade Institute of Science and Technology, 11060 Belgrade, Serbia;

*E-mail: pmitrasinovic.ist-belgrade.edu.rs@tech-center.com

Abstract

The infectious disease CoViD-19 is caused by a new severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), also referred to as hCoV-19. A possible infection mechanism includes dual host receptor recognitions by the SARS-CoV-2 transmembrane spike (S) glycoproteins. SARS-CoV-2 S contains two different domains, the receptor-binding domain (RBD) and the N-terminal domain (NTD), which interact with the angiotensin-converting enzyme 2 (ACE2) and the ganglioside-rich domain of the plasma membrane at the surface of respiratory cell, respectively. The NTD amino acid residues (111-162) form a functional ganglioside-binding domain (GBD) that is conserved in all clinical isolates. Herein, the single point mutations (SPMs) of the GBD residues to which the virus is prone during genetic adaptation are predicted using an *in silico* protein engineering approach. Consequently, their effects on the attachment of SARS-CoV-2 S to the ganglioside-linked 9-O-acetylated sialic acid (9-O-Ac-Sia) are explored using molecular docking simulations. Val120Tyr and Asn122Trp are found to be the most likely SPMs in the GBD of SARS-CoV-2 S being involved in very specific recognition with 9-O-Ac-Sia through electrostatic interactions. Val120Tyr and Asn122Trp are also found to be the most likely SPMs in the GBD of SARS-CoV-2 S that is involved in conspicuously hydrophobic recognition with hidroxychloroquine (Hcq), thereby indicating the ability of Hcq to competitively inhibit GBD interactions with lipid rafts. However, the considerably non-specific binding of Hcq and the micromolar range of the dissociation constants of the SARS-CoV-2 S/Hcq complexes do not support the proposal of treating Hcq as a drug candidate. Maintaining a clear resemblance of the structure of a potential drug candidate to a natural substrate, accompanied by essential functional group modifications, may be a usable guideline for the structure-based design of anti-CoViD-19 drugs.

Key words: coronavirus, CoViD-19, hidroxychloroquine, mutation, pandemic, SARS-CoV-2

1. Introduction

Coronaviruses (CoVs) are enveloped, single-stranded RNA (+) viruses that are divided into four genera (alpha, beta, gamma and delta) and involved in respiratory, enteric, hepatic and neuronal infectious diseases in animals and in humans. The beta genus comprises bCoV, hCoV-OC43, MHV, SARS-CoV and MERS-CoV.[1] The new virus SARS-CoV-2 caused an infectious disease CoViD-19 that was first identified in 2019 in Wuhan, China.[2] CoViD-19 has since spread through interpersonal contacts globally, so that a world 2019-20 CoV pandemic was declared by the World Health Organization (WHO) on March 11, 2020.[3] The possibility of having the second upsurge of it is a matter of debate worldwide. Substantial efforts are being invested through current research and therapeutic product development, with aim to mitigate the far-reaching consequences of the pandemic on human lives. Pharmaceutical industry expectations are to have a vaccine on the market in the next 12 to 18 months.[4] The United States (US) Food and Drug Administration (FDA) is testing coronavirus treatments, including hidroxychloroquine and chloroquine, by looking at widespread clinical trials of the drugs.[5] Hcq was suggested to be more preferred as an initial therapy than Cq.[6] It was indicated that Hcq, as a less toxic derivative of Cq, is effective in inhibiting the SARS-CoV-2 infection *in vitro*.[7] The present communication is imagined to contribute a little bit to dissecting these standpoints at the molecular scale.

Infections are initiated by the SARS-CoV-2 S proteins that recognize the ACE2 receptors at cell surface. Biophysical and structural studies revealed that SARS-CoV-2 S binds ACE2 with higher affinity than does SARS-CoV S.[8] Molecular modeling observations showed that the SARS-CoV-2 S protein contains two distinct domains, the receptor-binding domain (RBD) that interacts with ACE2 and the N-terminal domain

(NTD) that interacts with the ganglioside-rich domain of the plasma membrane (Figure 1, left).[6] Amino acid residues (111-162) of the NTD of SARS-CoV-2 S form a functional ganglioside-binding domain (GBD) (Figure 1, right) being both involved in facilitating the ACE2 recognition by RBD and completely conserved in clinical isolates worldwide. Hcq was found to bind sialic acids (linked to gangliosides) with high affinity, indicating its potential to block the interactions of GBD with lipid rafts.[6] The plan of this work is to place specificity issue, underlying the recognition of ganglioside by GBD, on a more rational ground.

Based on the structural knowledge of SARS-CoV-2 S,[8] herein, the recognition modes of the GBD of SARS-CoV-2 S by 9-O-Ac-Sia and Hcq are investigated, in terms of affinity and specificity, using molecular docking simulations. Recent *in vitro* data[9] highlighted the observation that SARS-CoV-2 mutations, including several ones in the spike glycoprotein,[9,10] are not an infrequent event and that certain SARS-CoV-2 mutations confer a greater pathogenic effect than others.[9] In this article, the likely SPMs in the GBD of SARS-CoV-2 S are predicted using an *in silico* protein engineering approach[11-14] in a systematic fashion. Consequently, a question, how the most likely SPMs affect the affinity and specificity of 9-O-Ac-Sia and Hcq to the GBD of SARS-CoV-2 S, is explored.

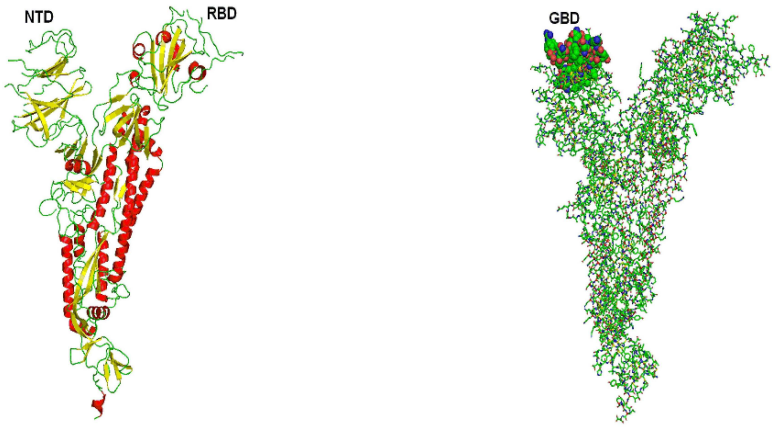

Figure 1. (**left**) The SARS-CoV-2 transmembrane spike (S) glycoprotein (PDB ID: 6VSB) has two distinct domains, the receptor-binding domain (RBD) that interacts with ACE2 and the N-terminal domain (NTD) that interacts with the ganglioside-rich domain of the plasma membrane; (**right**) the NTD amino acid residues being in between 111 and 162 (denoted by spheres) in the protein sequence form a functional ganglioside-binding domain (GBD) that is conserved in all clinical isolates.

2. Methods

To obtain the initial coordinates of atoms, the experimental structure of the trimeric SARS-CoV-2 S (PDB ID: 6VSB) was retrieved from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB).[8]

The FoldX molecular design toolkit version 5.0[15-23] was employed to systematically perform single point mutations of each residue in the GBD of SARS-CoV-2 S. FoldX provides a fast quantitative estimation of the importance of the interactions contributing to the stability of proteins and protein complexes. The algorithm uses a full atomic description of the structure of the proteins, while different energy terms are weighted using empirical data obtained from protein engineering experiments. The code mutates one amino acid to the other 24, including phosphorylated Tyr, Ser and Thr, as well as hydroxyl Proline in addition to 20 standard amino acids, and repairs the neighbor residues. The way it functions is: it mutates the selected position to Ala and annotates all

neighbor residues; it mutates the wild-type (wt) residue to itself, and then the neighbors to themselves followed each time by the wt residue to itself. In this way it ensures that, when mutating, any residue that has not been moved in the wt reference will not move. Once this is done, the new wt reference is mutated at the selected position to the target amino acids. To prevent problems, neighbor side chains are only optimized when creating the wt reference after self mutation, but not when making the individual mutants unless a new rotamer for the neighbor is selected.[15-23] The predictive power was tested on a very large set of point mutants (1088) that comprise most of the structural environments found in proteins.[24] A training database of 339 mutants in nine different proteins was initially considered and the set of parameters and weighting factors that best accounted for the changes in stability of the mutants was optimized. The predictive power was then tested using a blind test mutant database of 667 mutants, as well as a database of 82 protein-protein complex mutants. The global correlation for 95 % of the entire mutant database (1030 mutants) was 0.83, with a standard deviation of 0.81 kcal mol^{-1} and a slope of 0.76.[24] In the present paper, single point mutations of the amino acid residues in the GBD of the SARS-CoV-2 S protein were predicted at pH 8, which was used for the experimental determination of the starting structure (PDB ID: 6VSB).[8] It was shown that pH acidification of the medium, such as the one occurring in the endosomal compartment, does not trigger fusogenic conformational changes of some beta hCoV S proteins.[25]

Docking calculations were performed using the AScore/ShapeDock protocol from the ArgusLab 4.0.1 suite of programs.[26] AScore is based on the decomposition of the total protein-ligand binding free energy into the following contributions: the van der Waals interaction between the ligand and the protein, the hydrophobic effect, the

hydrogen bonding between the ligand and the protein, the hydrogen bonding that involves charged donor and/or acceptor groups, the deformation effect, and the effects of the translational and rotational entropy loss in binding process, respectively. Flexible ligand docking was done by describing the ligand as a torsion tree. Groups of bonded atoms that do not have rotatable bonds are nodes, while torsions are connections between the nodes. Topology of a torsion tree is a determinative factor influencing efficient docking. The AScore/ShapeDock protocol is fast, reproducible and formally explores all energy minima.[26] This particular protocol was shown to be consistent for docking ligands in the crystal structures of viral proteins,[27-31] while the calculated binding free energies were well-correlated with the experimental inhibitory concentrations.[29] This approach predicted a dissociation constant of 48.7 μM for the hCoV-OC43 S/9-O-Ac-Sia complex,[14] a value that is in agreement with an experimental observation of 49.7 ± 10.7 μM.[25]

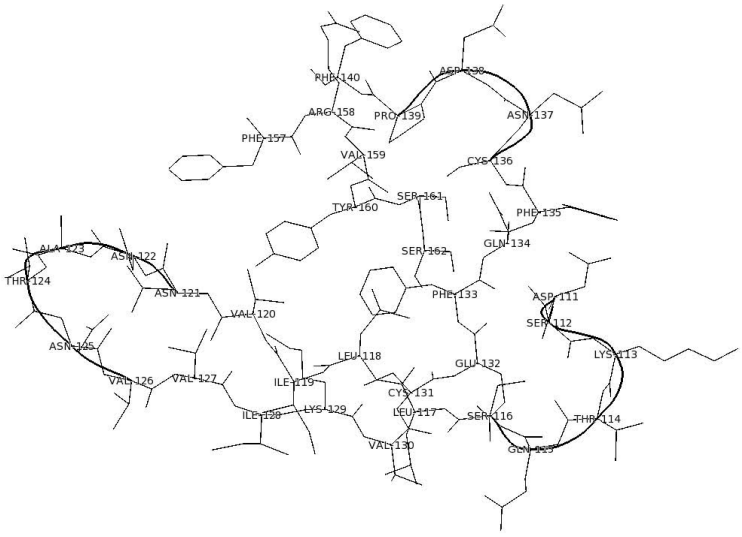

Figure 2. Amino acid residues (111-162) of the NTD of SARS-CoV-2 S form GBD that consists of three loops (denoted by ribbon), L1 (111-Asp-Ser-Lys-Thr-Glu-Ser-116), L2 (121-Asn-Asn-Ala-Thr-Asn-Val-126) and L3 (136-Cys-Asn-Asp-Pro-139), and five hydrophobic pockets, P1 (Leu117, Leu118, Ile119 and Val120), P2 (Val126, Val127, Ile128, Lys129, Val130 and Cys131), P3 (Phe133, Val159 and Tyr160), P4 (Phe135 and Cys136) and P5 (Phe140, Leu141 and Phe157).

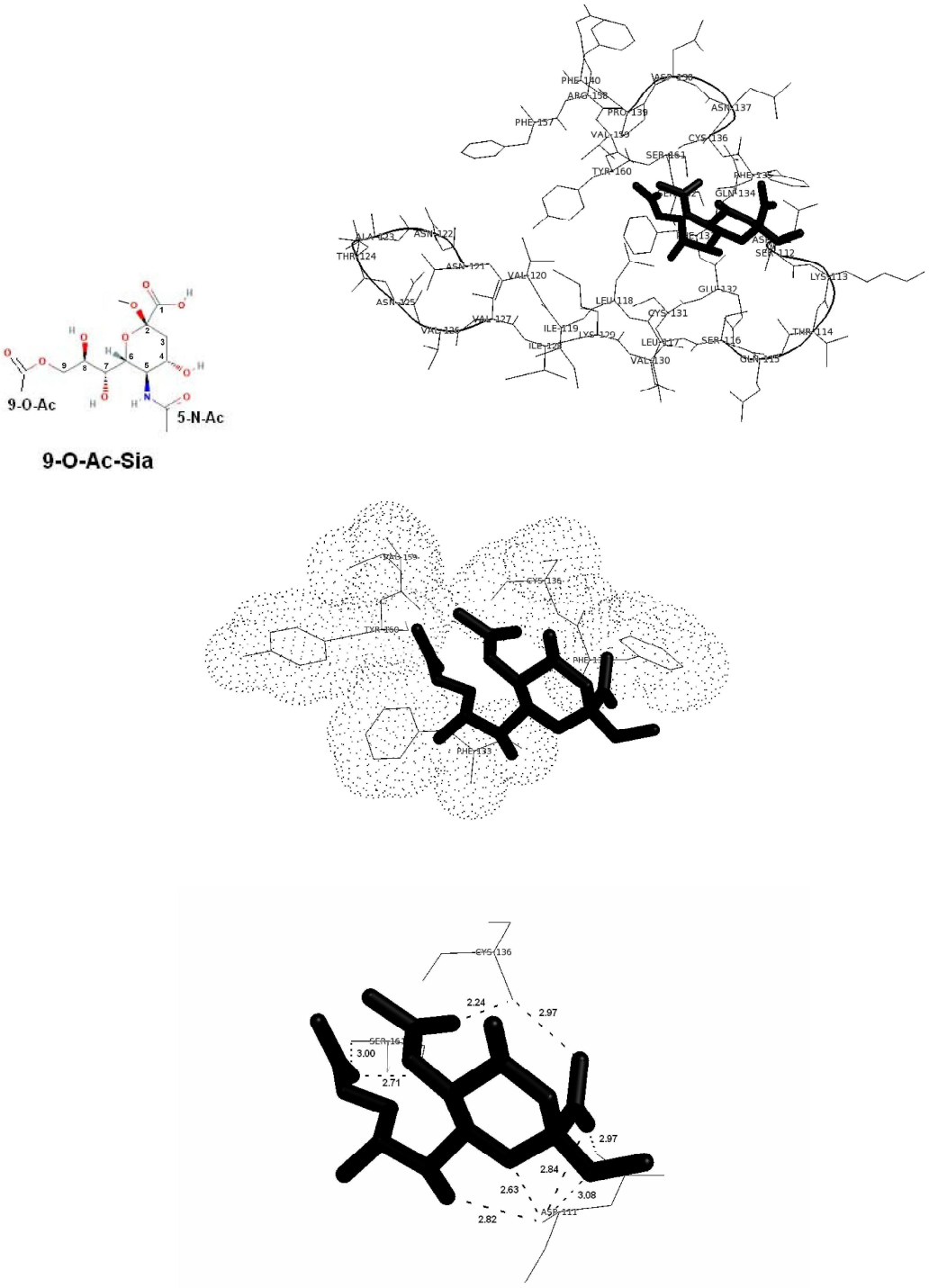

Figure 3. (**top**, **left**) The chemical structure of 9-O-Ac-Sia; (**top**, **right**) 9-O-Ac-Sia (bold sticks) docked in the GBD of SARS-CoV-2 S (PDB ID: 6VSB); (**center**) Val159 (dots) from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket accommodates the 5-N-Ac methyl, while Tyr160 (dots) from the same hydrophobic pocket accommodates the 9-O-Ac methyl; (**bottom**) 9 electrostatic contacts of 9-O-Ac-Sia with the GBD residues – 5 with Asp111, 2 with Cys136 and 2 with Ser161, respectively.

3. Results and Discussion

A possible infection mechanism is based on dual host receptor recognition - the recognition of ACE2 by the RBD of SARS-CoV-2 S that is simultaneously supported by the recognition of the ganglioside-rich domain of the plasma membrane by the GBD of SARS-CoV-2 S. The GBD subdomain of the NTD of SARS-CoV-2 S consists of amino acids being located between 111 and 162 in the protein sequence and forming a functional domain that is entirely conserved in all clinical isolates.[6] The particular residues are mainly arranged in three loops, L1 (111-Asp-Ser-Lys-Thr-Glu-Ser-116), L2 (121-Asn-Asn-Ala-Thr-Asn-Val-126) and L3 (136-Cys-Asn-Asp-Pro-139), and in five hydrophobic pockets, P1 (Leu117, Leu118, Ile119 and Val120), P2 (Val126, Val127, Ile128, Lys129, Val130 and Cys131), P3 (Phe133, Val159 and Tyr160), P4 (Phe135 and Cys136) and P5 (Phe140, Leu141 and Phe157), as shown in Figure 2. To mimic the interaction between the SARS-CoV-2 transmembrane S glycoprotein and the ganglioside-rich domain, the interaction between the GBD and 9-O-Ac-Sia was simulated by means of molecular docking (Figure 3, top, right). The complex formed illustrates that Val159 from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket accommodates the 5-N-Ac methyl, Tyr160 from the same hydrophobic pocket accommodates the 9-O-Ac methyl, while Phe135 from the P4 (Phe135 and Cys136) hydrophobic pocket accommodates the 2-O methyl (Figure 3, center). The complex is also stabilized by 9 electrostatic contacts of 9-O-Ac-Sia with the GBD residues (Figure 3, bottom) – 2 between the 1-O hydroxyl and Asp111 (2.97 and 2.84 Å), 1 between the 2-O and Asp111 (3.08 Å), 1 between the O6 and Asp111 (2.63 Å), 1 between the 7-O and Asp111 (2.82 Å), 1 between the 1-carboxylate and Cys136 (2.97 Å), 1 between the 5-N-Ac and Cys136 (2.24 Å) and 2 between the 9-O-Ac and Ser161 (2.71 and 3.00 Å), respectively. Thus, the

particular sequence AspAspAspAspAspCysCysSerSer is the signature of the high specificity of 9-O-Ac-Sia to the GBD. Asp111, Phe135, Cys136, Tyr160 and Ser161 were previously identified using a molecular dynamics (MD) approach as the sites of SARS-CoV-2 S in contact with GM1 (a ganglioside with sialic acid), of which Asp111 contributed most to the stability of the NTD/9-O-Ac-Sia complex in a more realistic mode of MD simulation.[6] Evaluated strength of the GBD/9-O-Ac-Sia interaction, seen through a dissociation constant of circa 17.5 μM (Table 1), indicates potent S-mediated virion attachment, especially in high-density receptor environments. E.g., the GBD of SARS-CoV-2 S binds 9-O-Ac-Sia ($K_d \approx 17.5$ μM) with about 3 times greater potency than does the interactive site of hCoV-OC43 S ($K_d \approx 48.7$ μM).[14,25]

Each of the GBD residues is mutated to the other 24 (20 standard amino acids, phosphorylated Tyr, Ser and Thr, as well as hydroxyl Proline). All the single point mutants are energetically evaluated with reference to the original receptor (PDB ID: 6VSB). SPMs that stabilize the wild-type receptor structure for more than 2 kcal mol^{-1} are treated as likely ones (Table 1). An average level of thermochemical accuracy of 2 kcal mol^{-1} is acceptable for the structure-based drug (or ligand) design purposes.[32-43]

Table 1. The Binding Free Energies Obtained by Docking 9-O-Ac-Sia and Hcq in the Single Point Mutants of SARS-CoV-2 S (PDB ID: 6VSB)

Ligand: 9-O-Ace-Sia			Ligand: Hcq		
Receptor[(a)]	**ΔG_{bind} (kcal mol^{-1})**	**K_d (μM)[(b)]**	**Receptor[(a)]**	**ΔG_{bind} (kcal mol^{-1})**	**K_d (μM)[(b)]**
wt SARS-CoV-2 S	-6.53	17.50	wt SARS-CoV-2 S	-7.96	1.60
Leu117Met	-6.60	15.56	Phe157Met	-7.98	1.54
Glu132Ala	-6.62	15.05	Asp138Met	-7.99	1,51
Asp138Arg	-6.64	14.55	Phe135Lys	-8.03	1.41
Phe157Met	-6.64	14.55	Thr114Ptr	-8.17	1.12
Thr124Arg	-6.69	13.38	Asn121Leu	-8.17	1.12
Thr114Ptr	-6.69	13.38	Asn121Met	-8.17	1.12
Asn121Phe	-6.69	13.38	Asn121Phe	-8.17	1.12
Asn122Met	-6.69	13.38	Asn125Ile	-8.18	1.10
Ala123Gln	-6.69	13.38	Asn125Met	-8.18	1.10
Ala123Glu	-6.69	13.38	Asp111Gly	-8.19	1.08
Ala123Lys	-6.69	13.38	Asn137Pro	-8.19	1.08
Ser161Tyr	-6.88	9.73	Val159Trp	-8.19	1.08
Asp138Leu	-6.91	9.25	Asn121Ile	-8.20	1.06
Asp138Met	-6.91	9.25	Ala123Arg	-8.20	1.06
Asp138Trp	-6.97	8.36	Ala123His	-8.20	1.06
Phe140Ser	-6.98	8.23	Ala123Gln	-8.20	1.06
Phe135Ala	-7.06	7.19	Ala123Glu	-8.20	1.06
Val159Trp	-7.11	6.61	Ala123Lys	-8.20	1.06
Val130Lys	-7.13	6.40	Ala123Ptr	-8.20	1.06
Ile119Met	-7.17	5.98	Ala123Tyr	-8.20	1.06
Asn121Leu	-7.17	5.98	Thr124Arg	-8.20	1.06
Asn121Ile	-7.17	5.98	Ile119Met	-8.21	1.05
Asn121Met	-7.17	5.98	Asn122Ile	-8.21	1.05
Ala123Arg	-7.17	5.98	Asn122Met	-8.21	1.05
Ala123His	-7.17	5.98	Asn122Sep	-8.21	1.05
Ala123Ptr	-7.17	5.98	Asn137Hyp	-8.21	1.05
Asn125Ile	-7.17	5.98	Asp138Arg	-8.23	1.01
Asn125Met	-7.17	5.98	Asp138Ptr	-8.24	0.99
Asn122Ile	-7.18	5.88	Phe140Ser	-8.39	0.77
Asn122Leu	-7.18	5.88	Asp138Trp	-8.49	0.65
Asn137Hyp	-7.18	5.88	Asp138Leu	-8.51	0.63
Asp111Gly	-7.20	5.69	Leu117Met	-8.78	0.40
Asn122Sep	-7.20	5.69	Asn122Leu	-8.84	0.36
Asp138Ptr	-7.21	5.59	Val130Lys	-8.88	0.34
Val127Glu	-7.29	4.89	Phe135Ala	-8.90	0.33
Ala123Tyr	-7.31	4.73	Val127Glu	-8.94	0.31
Asn137Pro	-7.31	4.73	Glu132Ala	-9.10	0.23
Phe135Lys	-7.38	4.20	Ser161Tyr	-9.24	0.19
Asn122Trp	-7.40	4.07	Asn122Trp	-9.67	0.09
Val120Tyr	-7.42	3.93	Val120Tyr	-9.76	0.08

[(a)] wild-type (wt), Hyp - hydroxyl Pro; Ptr – phosphorylated Thr, Tpo - phosphorylated Tyr; Sep - phosphorylated Ser

[(b)] ΔG_{bind} – the binding free energy, K_d – the dissociation constant, $\Delta G_{bind} = RT \ln(K_d)$, R – the gas constant ($1.9872 \times 10^{-3}$ kcal K^{-1} mol^{-1}), T – the absolute temperature (300 K), 1 μM = 10^{-6} M

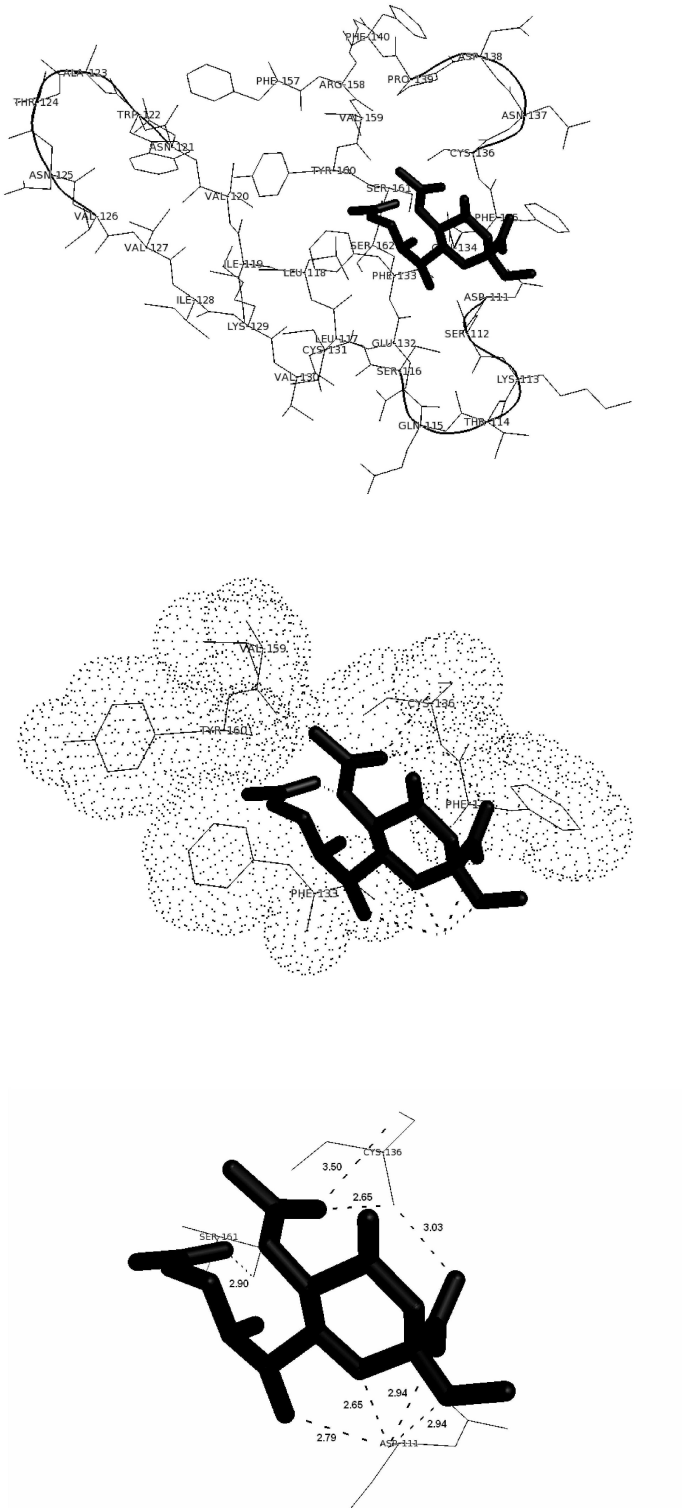

Figure 4. (**top**) 9-O-Ac-Sia (bold sticks) docked in the GBD of the Asn122Trp mutant of SARS-CoV-2 S (PDB ID: 6VSB); (**center**) Val159 (dots) from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket accommodates the 5-N-Ac methyl, while Phe133 and Tyr160 (dots) from the same hydrophobic pocket accommodate the 9-O-Ac methyl; (**bottom**) 8 electrostatic contacts of 9-O-Ac-Sia with the GBD residues – 4 with Asp111, 3 with Cys136 and 1 with Ser161, respectively.

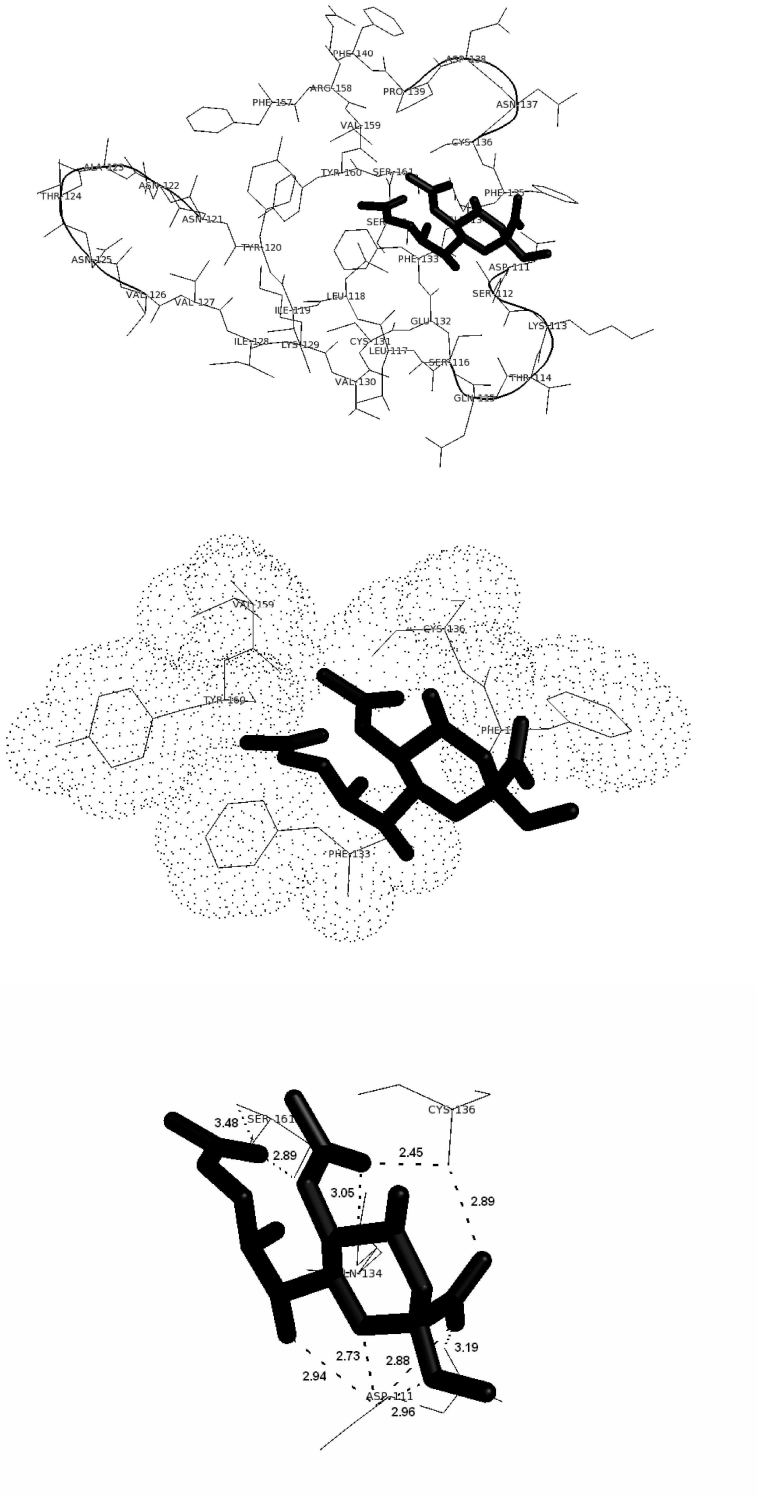

Figure 5. (**top**) 9-O-Ac-Sia (bold sticks) docked in the GBD of the Val120Tyr mutant of SARS-CoV-2 S (PDB ID: 6VSB); (**center**) Val159 (dots) from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket accommodates the 5-N-Ac methyl, while Phe133 and Tyr160 (dots) from the same hydrophobic pocket accommodate the 9-O-Ac methyl; (**bottom**) 10 electrostatic contacts of 9-O-Ac-Sia with the GBD residues – 5 with Asp111, 1 with Gln134, 2 with Cys136 and 2 with Ser161, respectively.

An inspection of the values in Table 1 shows that Asn122Trp and Val120Tyr stabilize the wt/9-O-Ac-Sia complex to a largest extent. In comparison to the original complex (Figure 3, top, right), these SPMs do not change very much the overall spatial orientation of 9-O-Ac-Sia with respect to the receptors (Figures 4 and 5, top), but rather cause subtle conformational changes that are hard to be noted by looking at the binding modes superficially.

Asn122Trp and Val120Tyr increase the hydrophobic nature of the GBD/9-O-Ac-Sia interaction. In both cases, Val159 from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket accommodates the 5-N-Ac methyl, while Phe135 from the P4 (Phe135 and Cys136) hydrophobic pocket accommodates the 2-O methyl (Figures 4 and 5, center), as similarly displayed by the wt/9-O-Ac-Sia interaction interface (Figure 3, center). The increased hydrophobicity is reflected through the interaction of the 9-O-Ac methyl not only with Tyr160 (Figure 3, center), but with both Phe133 and Tyr160 (Figures 4 and 5, center). These are expected consequences of the particular SPMs, considering the physicochemical features of the original amino acids and their substituents. Asn122 - a small and non-hydrophobic amino acid is replaced by Trp122 – a large and very hydrophobic amino acid having two aromatic rings fused in its side chain. Val120 - a small, aliphatic and hydrophobic amino acid is replaced by Tyr120 - a large and hydrophobic amino acid having an aromatic ring in its side chain. As hydrophobic amino acids do not like to reside in an aqueous environment, the SPMs in the GBD increase the hydrophobic effect and the release of protein-bound water molecules, thereby causing to have the substituent (Trp122 or Tyr120) buried within the hydrophobic core of the NTD of SARS-CoV-2 S, or within the lipid portion of the membrane.

The introduction of Trp122, instead of Asn122, in the GBD is associated with 8 electrostatic contacts of 9-O-Ac-Sia with the GBD residues (Figure 4, bottom) – 1 between the 1-O hydroxyl and Asp111 (2.94 Å), 1 between the 2-O and Asp111 (2.94 Å), 1 between the O6 and Asp111 (2.65 Å), 1 between the 7-O and Asp111 (2.79 Å), 1 between the 1-carboxylate and Cys136 (3.03 Å), 2 between the 5-N-Ac and Cys136 (2.65 and 3.50 Å) and 1 between the 9-O-Ac and Ser161 (2.90 Å), respectively. Thus, Asn122Trp changes the specificity pattern, AspAspAspAspAspCysCysSerSer, of the wt/9-O-Ac-Sia complex to AspAspAspAspCysCysCysSer. The specificity differences, the loss of an Asp and the substitution of a Ser by a Cys, due to the Asn122Trp mutation can be seen as consequences at the expense of the increased hydrophobic effect. The introduction of Tyr120, instead of Val120, in the GBD is associated with 10 electrostatic contacts of 9-O-Ac-Sia with the GBD residues (Figure 5, bottom) – 2 between the 1-O hydroxyl and Asp111 (3.19 and 2.88 Å), 1 between the 2-O and Asp111 (2.96 Å), 1 between the O6 and Asp111 (2.73 Å), 1 between the 7-O and Asp111 (2.94 Å), 1 between the 5-N-Ac and Gln134 (3.05 Å), 1 between the 1-carboxylate and Cys136 (2.89 Å), 1 between the 5-N-Ac and Cys136 (2.45 Å) and 2 between the 9-O-Ac and Ser161 (2.89 and 3.48 Å), respectively. Thus, Val120Tyr changes the specificity pattern, AspAspAspAspAspCysCysSerSer, of the wt/9-O-Ac-Sia complex to AspAspAspAspAspGlnCysCysSerSer. The specificity order, extended by a Gln due to the Val120Tyr mutation, is simultaneously followed by the increased hydrophobic effect.

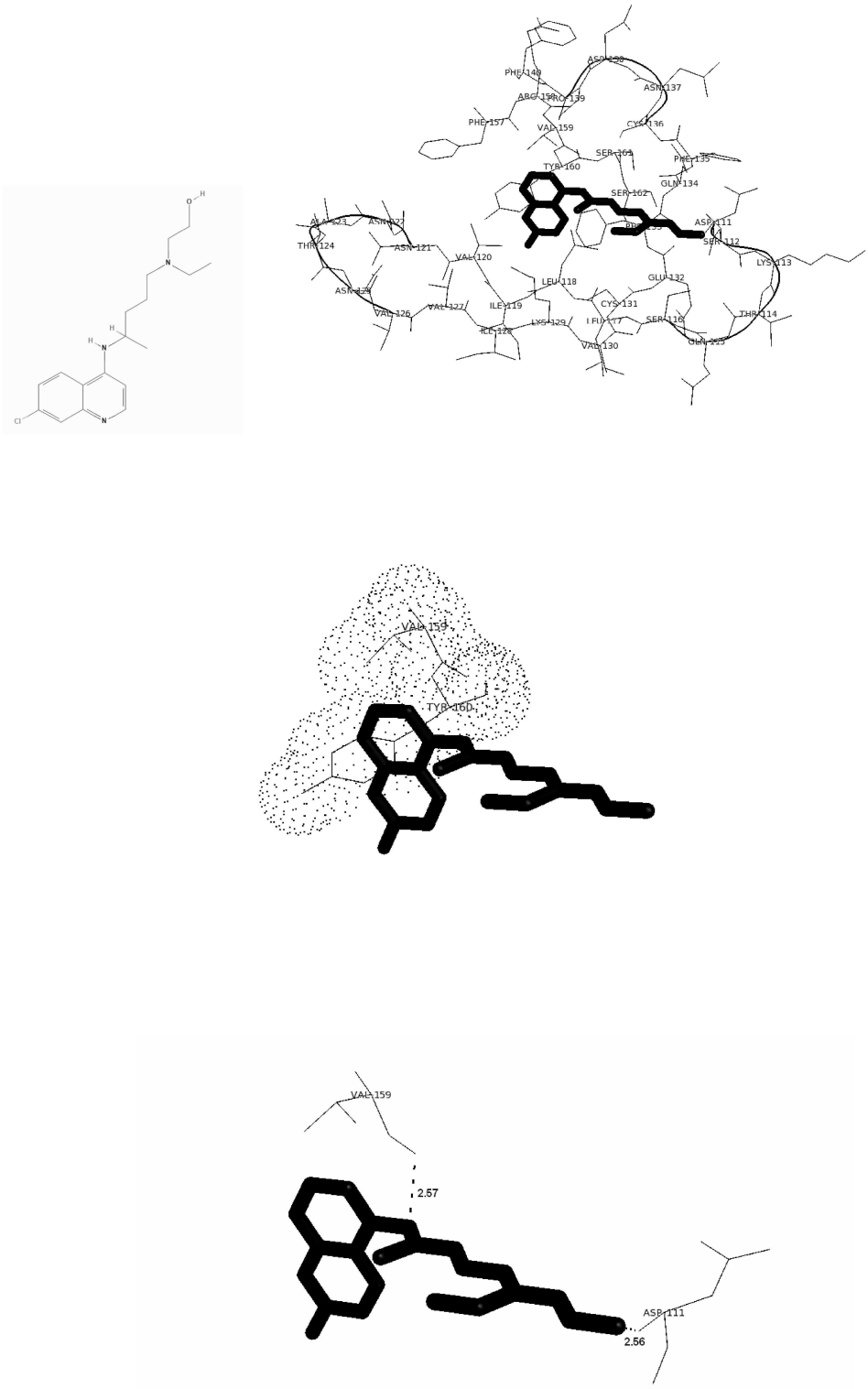

Figure 6. (**top**, **left**) The chemical structure of Hcq; (**top**, **right**) Hcq (bold sticks) docked in the GBD of SARS-CoV-2 S (PDB ID: 6VSB); (**center**) Val159 and Tyr160 (dots) from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket accommodate the aromatic core scaffold of Hcq; (**bottom**) 2 electrostatic contacts of the Hcq side chain with the GBD residues – 1 with Asp111 and 1 with Val159, respectively.

The chemical structure of Hcq (Figure 6, top, left) contains an aromatic core to which both a Cl atom and a large side chain are bound, indicating a pronounced hydrophobic character of the ligand. The binding mode obtained by docking Hcq in the GBD of SARS-CoV-2 S is illustrated in Figure 6 (top, right). Val159 and Tyr160 from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket accommodate the aromatic core scaffold of Hcq (Figure 6, center). Two electrostatic contacts between the Hcq side chain and the GBD residues are also involved in formation of the complex - 1 with Asp111 (2.56 Å) and 1 with Val159 (2.57 Å), respectively (Figure 6, bottom).

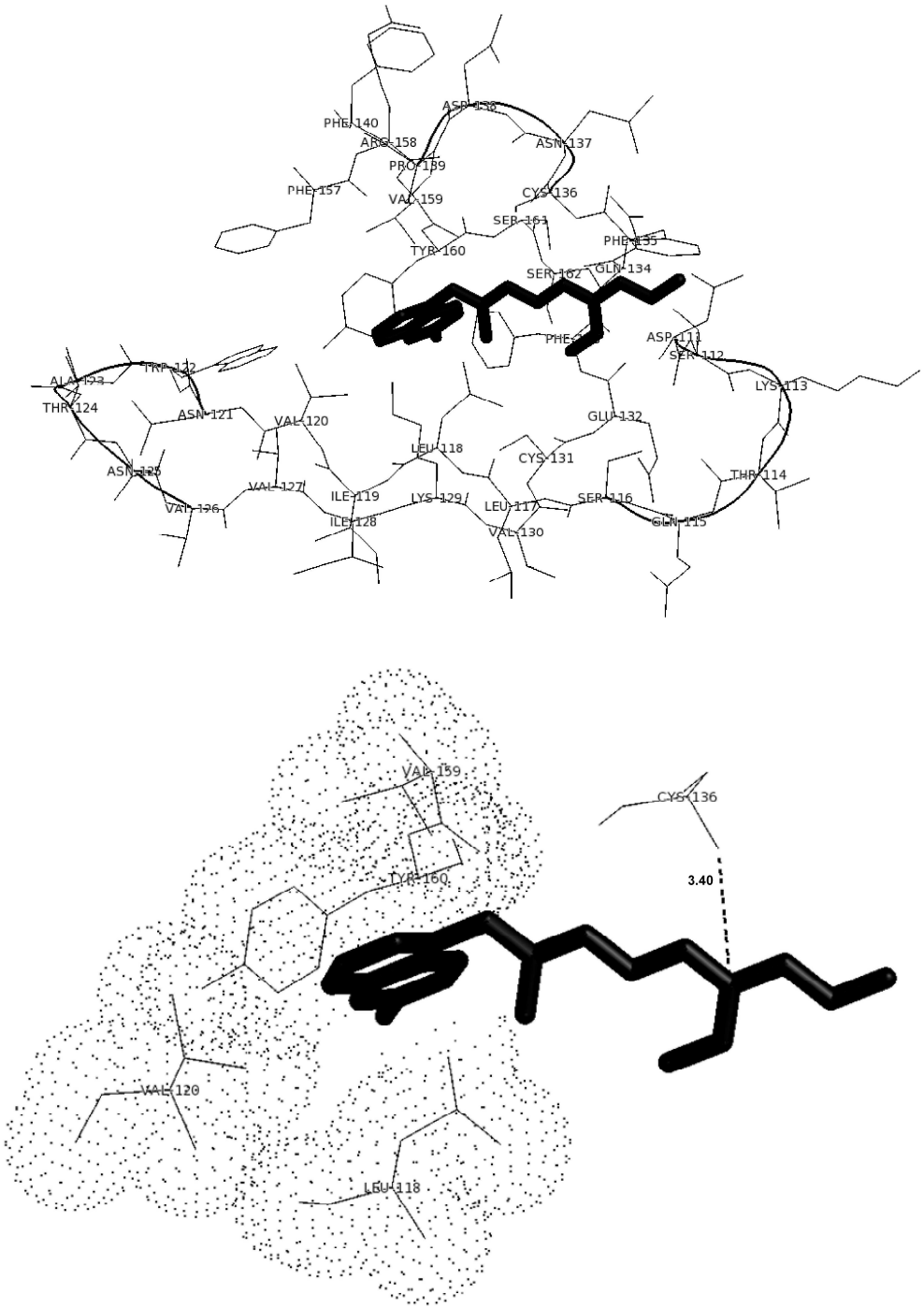

Figure 7. (**top**) Hcq (bold sticks) docked in the GBD of the Asn122Trp mutant of SARS-CoV-2 S (PDB ID: 6VSB); (**bottom**) Leu118 and Val120 (dots) from the P1 (Leu117, Leu118, Ile119 and Val120) hydrophobic pocket as well as Val159 and Tyr160 (dots) from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket accommodate the aromatic core scaffold of Hcq; 1 electrostatic contact of the Hcq side chain with Cys136.

The introduction of Trp122 (Figure 7, top), instead of Asn122 (Figure 6, top, right), in the GBD causes the vertical conformational switch (rotation by around 270º clockwise) of Hcq, making the aromatic core scaffold of Hcq interact with both Leu118 and Val120 from the P1 (Leu117, Leu118, Ile119 and Val120) hydrophobic pocket and Val159 and Tyr160 from the P3 (Phe133, Val159 and Tyr160) hydrophobic pocket (Figure 7, bottom). The orientation of the Cl atom being right toward Tyr160 additionally contributes to the much more hydrophobic nature of the Asn122Trp/Hcq interaction compared to the wt/Hcq interaction. One electrostatic contact between the Hcq side chain and Cys136 (3.40 Å) is also involved in formation of the complex (Figure 7, bottom), indicating the very low specificity of Hcq binding.

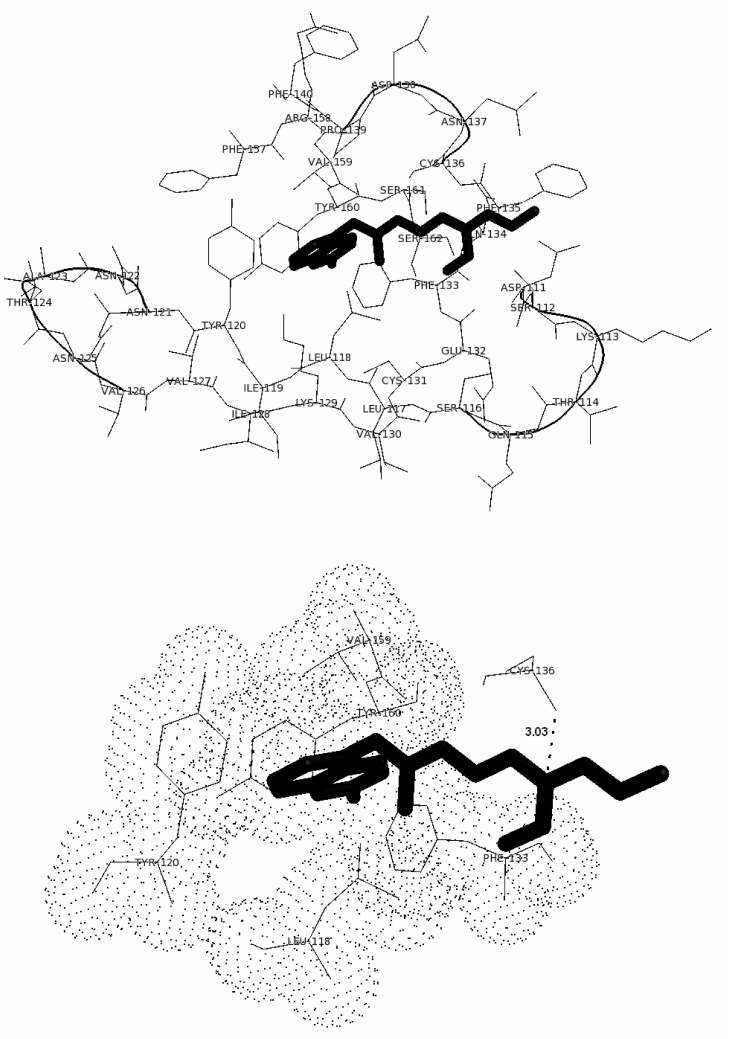

Figure 8. (**top**) Hcq (bold sticks) docked in the GBD of the Val120Tyr mutant of SARS-CoV-2 S (PDB ID: 6VSB); (**bottom**) Leu118 and Tyr120 (dots) from the P1 (Leu117, Leu118, Ile119 and Tyr120) hydrophobic pocket as well as the entire P3 (Phe133, Val159 and Tyr160) hydrophobic pocket (dots) accommodate the aromatic core scaffold of Hcq; 1 electrostatic contact of the Hcq side chain with Cys136.

The values in Table 1 indicate that Val120Tyr is the most likely SPM for the binding of Hcq. The hydrophobic nature of the Val120Tyr/Hcq interaction mode (Figure 8, top) is similar to that of the Asn122Trp/Hcq interaction mode (Figure 7, top) qualitatively. However, the Val120Tyr/Hcq interaction (Figure 8, bottom) is more hydrophobic than the Asn122Trp/Hcq interaction (Figure 7, bottom), primarily due to having the whole hydrophobic pocket P3 (Phe133, Val159 and Tyr160) involved in

accommodating both the aromatic core scaffold of Hcq and its Cl atom. One electrostatic contact between the Hcq side chain and Cys136 (3.03 Å, Figure 8, bottom) underlines the remarkably low specificity of Hcq binding.

The binding free energies (Table 1) of the complexes, formed by docking 9-O-Ac-Sia and Hcq to the same receptor, show that Hcq has higher affinity to SARS-CoV-2 S than does 9-O-Ac-Sia. This standpoint can be rationalized by the considerably hydrophobic and almost non-specific recognition of SARS-CoV-2 S by Hcq, thereby speaking in favor of the potential of Hcq to competitively inhibit the interactions of GBD with lipid rafts. However, rapid kinetics underlying S-mediated virion attachment to 9-O-Ac-Sia is associated with recognition that is very specific. It is, therefore, convenient to contrast the chemical structures of 9-O-Ac-Sia and Hcq by estimating their respective entropy losses upon binding. Flexible ligand docking is based on active torsions in ligand structure, conceivable as particular sp^3 bonds that are directly involved in finding the lowest energy receptor/ligand conformations. Entropy loss upon ligand binding is related to the loss of its degrees of freedom. The torsional potential indirectly takes care of the particular entropy amount by being proportional to the number of active torsions in ligand structure. An active torsion has been energetically estimated to cost circa 0.3 kcal mol^{-1}.[44,45] It means that the structure of 9-O-Ac-Sia with eleven active torsions experiences a negative entropy change of -3.3 kcal mol^{-1}, while the structure of Hcq with one active torsion experiences a tiny entropic decrease of -0.3 kcal mol^{-1}. Finally, the dissociation constants of the SARS-CoV-2 S/Hcq complexes are in the micromolar (10^{-6} M) range (Table 1), which is acceptable for hit ligand molecule but not for drug candidate.

4. Summary

Val120Tyr and Asn122Trp have been predicted as the most likely SPMs in the GBD of SARS-CoV-2 S involved in very specific recognition with 9-O-Ac-Sia through electrostatic interactions. The same SPMs have been found to be critical in the GBD of SARS-CoV-2 S involved in very hydrophobic recognition with Hcq. The considerably non-specific interaction between SARS-CoV-2 S and Hcq as well as the micromolar range of the dissociation constants of the SARS-CoV-2 S/Hcq complexes may be some of the key glitches in considering the potential of Hcq to become a drug candidate. Keeping a clear structural similarity of a potential drug candidate with a natural substrate, combined with essential functional group modifications, may be a reasonable guideline for the successful structure-based design of anti-CoViD-19 drugs. A relevant example may be zanamivir (Relenza), an anti-influenza drug that was directly derived from sialic acid with minimal functional modifications.[46,47]

5. References

1. J. Reguera, G. Mudgal, C. Santiago, J. M. Casasnovas, *Virus Res.* **2014**, *194*, 3-15. **DOI**:10.1016/j.virusres.2014.10.005

2. D. S. Hui, E. I. Azhar, T. A. Madani, C. Drosten, A. Zumla, E. Petersen, *Int. J. Infect. Dis.* **2020**, *91*, 264-266. **DOI**:10.1016/j.ijid.2020.01.009

3. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (accessed on 14 June 2020)

4. https://medicalxpress.com/news/2020-03-pharma-chiefs-coronavirus-vaccine-.html (accessed on 14 June 2020)

5. https://techcrunch.com/2020/03/19/fda-testing-coronavirus-treatments-including-chloroquine-plasma-from-recovered-covid-19-patients/ (accessed on 14 June 2020)

6. J. Fantini, C. Di Scala, H. Chahinian, N. Yahi, *Int. J. Antimicrob. Agents* **2020**, *55(5)*, 105960. **DOI**:10.1016/j.ijantimicag.2020.105960

7. J. Liu, R. Cao, M. Xu, X. Wang, H. Zhang, H. Hu, Y. Li, Z. Hu, W. Zhong, M. Wang, *Cell Discovery* **2020**, *6*, 16. **DOI**: 10.1038/s41421-020-0156-0

8. D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C. L. Hsieh, O. Abiona, B. S. Graham, J. S. McLellan, *Science* **2020**, *367(6483)*, 1260-1263. **DOI**:10.1126/science.abb2507

9. H. Yao, X. Lu, Q. Chen, K. Xu, Y. Chen, L. Cheng, F. Liu, Z. Wu, H. Wu, C. Jin, M. Zheng, N. Wu, C. Jiang, L. Li, *medRxiv* **2020**. **DOI**:10.1101/2020.04.14.20060160

10. B. Korber, W. M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, B. Foley, E. E. Giorgi, T. Bhattacharya, M. D. Parker, D. G. Partridge, C. M. Evans, T. M. Freeman, T. I. de Silva, A. Angyal, R. L. Brown, L. Carrilero, L. R. Green, D. C. Groves, K. J. Johnson, A. J. Keeley, B. B. Lindsey, P. J. Parsons, M. Raza, S. Rowland-Jones, N. Smith, R. M. Tucker, D. Wang, M. D. Wyles, C. C. LaBranche, D. C. Montefiori, *bioRxiv* **2020**. **DOI**:10.1101/2020.04.29.069054

11. P. M. Mitrasinovic, *Curr. Drug Targets* **2013**, *14*, 817-829. **DOI**:10.2174/1389450111314070009

12. P. M. Mitrasinovic, *Med. Chem.* **2014**, *10*, 252-270. **DOI**:10.2174/15734064100314030414344 2

13. P. M. Mitrasinovic, *Med. Chem.* **2014**, *10*, 46-58. **DOI**:10.2174/15734064100113122612212 4

14. P. M. Mitrasinovic, *Acta Chim. Slov.* **2020**, *67*, 1-8. **DOI**:10.17344/acsi.2020.6009

15. P. Vanhee, E. Verschueren, L. Baeten, F. Stricher, L. Serrano, F. Rousseau, J. Schymkowitz, *Nucleic Acids Res.* **2011**, *39*, D435-D442. **DOI**:10.1093/nar/gkq972

16. M. Petukhov, D. Cregut, C. M. Soares, L. Serrano, *Protein Sci.* **1999**, *8*, 1982-1989. **DOI**:10.1110/ps.8.10.1982

17. V.Muñoz, L. Serrano, *Proteins* **1994**, *20*, 301-311. **DOI**:10.1002/prot.340200403

18. R. Abagyan, M. Totrov, *J. Mol. Biol.* **1994**, *235*, 983-1002. **DOI**:10.1006/jmbi.1994.1052

19. M. Vijayakumar, K. Y. Wong, G. Schreiber, A. R. Fersht, A. Szabo, H. X. Zhou, *J. Mol. Biol.* **1998**, *278*, 1015-1024. **DOI**:10.1006/jmbi.1998.1747

20. J. W. H. Schymkowitz, F. Rousseau, I. C. Martins, J. Ferkinghoff-Borg, F. Stricher, L. Serrano, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10147-10152. **DOI**:10.1073/pnas.0501980102

21. J. Van Durme, J. Delgado, F. Stricher, L. Serrano, J. Schymkowitz, F. Rousseau, *Bioinformatics* **2011**, *27*, 1711-1712. **DOI**:10.1093/bioinformatics/btr254

22. J. Schymkowitz, J. Borg, F. Stricher, R Nys, F. Rousseau, L. Serrano, *Nucleic Acids Res.* **2005**, *33*, W382-W388. **DOI**:10.1093/nar/gki387

23. L. Baeten, J. Reumers, V. Tur, F. Stricher, T. Lenaerts, L. Serrano, F. Rousseau, J. Schymkowitz, *PLoS Comput. Biol.* **2008**, *4*, e1000083. **DOI**:10.1371/journal.pcbi.1000083

24. R. Guerois, J. E. Nielsen, L. Serrano, *J. Mol. Biol.* **2002**, *320(2)*, 369-387. **DOI**:10.1016/S0022-2836(02)00442-4

25. M. Alejandra Tortorici, A. C. Walls, Y. Lang, C. Wang, Z. Li, D. Koerhuis, G.-J. Boons, B.-J. Bosch, F. A. Rey, R. J. de Groot, D. Veesler, *Nat. Struct. & Mol. Biol.* **2019**, *26*, 481-489. **DOI**:10.1038/s41594-019-0233-y

26. ArgusLab 4.0.1, M. A. Thompson, Planaria Software, LLC, Seatle WA, **2004**.

27. P. M. Mitrasinovic, *Biophys. Chem.* **2009**, *140*, 35-38.

DOI:10.1016/j.bpc.2008.11.004

28. P. M. Mitrasinovic, *Curr. Drug Targets,* **2010***, 11,* 315-326.

DOI:10.2174/138945010790711932

29. M. L. Mihajlovic, P. M. Mitrasinovic, *Mol. Simulation* **2009**, *35*, 311-324.

DOI: 10.1080/08927020802430752

30. M. L. Mihajlovic, P. M. Mitrasinovic, *Biophys. Chem.* **2008**, *136*, 152-158.

DOI:10.1016/j.bpc.2008.06.003

31. M. L. Mihajlovic, P. M. Mitrasinovic, *J. Serb. Chem. Soc.* **2009**, *74*, 1-13.

DOI:10.2298/JSC0901001M

32. A. D. Becke, *J. Chem. Phys.* **2002**, *117*, 6935-6938. **DOI**:10.1063/1.1503772

33. S. Cosconati, S. Forli, A. L. Perryman, R. Harris, D. S. Goodsell, A. J. Olson, *Expert Opin. Drug Discov.* **2010**, *5(6)*, 597-607. **DOI**:10.1517/17460441.2010.484460

34. N. Deng, L. Wickstrom, P. Cieplak, C. Lin, D. Yang, *J. Phys. Chem. B* **2017**, *121(46)*, 10484-10497. **DOI**:10.1021/acs.jpcb.7b09406

35. P. M. Mitrasinovic, *Chem. Phys.* **2003**, *286*, 1-13. **DOI**:10.1016/S0301-0104(02)00902-3

36. P. M. Mitrasinovic, *Acta Chim. Slov.* **2020**, *67*, 386-395.

DOI:10.17344/acsi.2019.5105

37. P. M. Mitrasinovic, *Acta Chim. Slov.* **2020**, *67*, 1-18. **DOI**: 10.17344/acsi.2020.5823

38. P. M. Mitrasinovic, *Acta Chim. Slov.* **2020**, *67*, 1-11. **DOI**:10.17344/acsi.2020.5847

39. P. M. Mitrasinovic, *Croat. Chem. Acta* **2019**, *92(1)*, 43-57. **DOI**:10.5562/cca3456

40. P. M. Mitrasinovic, *J. Biomol. Struct. Dyn.* **2018**, *36(9)*, 2292-2302.

DOI:10.1080/07391102.2017.1358670

41. P. M. Mitrasinovic, *J. Chem. Inf. Model.* **2015**, *55(2)*, 421-433.

DOI:10.1021/ci5006965

42. A. Pavlov, P. M. Mitrasinovic, *Curr. Org. Chem.* **2010**, *14*, 129-138. **DOI**:10.2174/138527210790069866

43. P. M. Mitrasinovic, *Curr. Org. Chem.* **2010**, *14*, 198-211.

DOI:10.2174/138527210790069857

44. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639-1662. **DOI**:10.1002/(SICI)1096-987X(19981115)19:14%3c1639::AID-JCC10%3e3.0.CO;2-B

45. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785–2791. **DOI**:10.1002/jcc.21256

46. M. Von Itzstein, *Nat. Rev. Drug Discov.* **2007**, *6*, 967-974. **DOI**:10.1038/nrd2400

47. P. M. Mitrasinovic, in: P. M. Mitrasinovic (Ed): Global View of the Fight against Influenza, Nova Science Publishers, New York, U.S.A., **2009**, pp. 185-192.