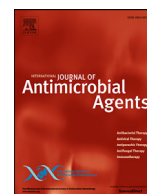




Contents lists available at ScienceDirect

International Journal of Antimicrobial Agents

journal homepage: www.elsevier.com/locate/ijantimicag

Structural and molecular modelling studies reveal a new mechanism of action of chloroquine and hydroxychloroquine against SARS-CoV-2 infection

Jacques Fantini^{a,b,*}, Coralie Di Scala^c, Henri Chahinian^{a,b}, Nouara Yahi^{a,b}^aINSERM UMR_S 1072, Marseille, France^bDepartment of Biology, Aix-Marseille Université, Marseille, France^cINMED, INSERM U1249, Parc Scientifique de Luminy, Marseille, France

ARTICLE INFO

Article history:

Available online xxx

Editor: Jean-Marc Rolain

Keywords:

Coronavirus

Pandemic

SARS-CoV-2

Ganglioside

Spike

Chloroquine

ABSTRACT

The recent emergence of the novel pathogenic SARS-coronavirus 2 (SARS-CoV-2) is responsible for a worldwide pandemic. Given the global health emergency, drug repositioning is the most reliable option to design an efficient therapy for infected patients without delay. The first step of the viral replication cycle [i.e. attachment to the surface of respiratory cells, mediated by the spike (S) viral protein] offers several potential therapeutic targets. The S protein uses the angiotensin-converting enzyme-2 (ACE-2) receptor for entry, but also sialic acids linked to host cell surface gangliosides. Using a combination of structural and molecular modelling approaches, this study showed that chloroquine (CLQ), one of the drugs currently under investigation for SARS-CoV-2 treatment, binds sialic acids and gangliosides with high affinity. A new type of ganglioside-binding domain at the tip of the N-terminal domain of the SARS-CoV-2 S protein was identified. This domain (111–158), which is fully conserved among clinical isolates worldwide, may improve attachment of the virus to lipid rafts and facilitate contact with the ACE-2 receptor. This study showed that, in the presence of CLQ [or its more active derivative, hydroxychloroquine (CLQ-OH)], the viral S protein is no longer able to bind gangliosides. The identification of this new mechanism of action of CLQ and CLQ-OH supports the use of these repositioned drugs to cure patients infected with SARS-CoV-2. The in-silico approaches used in this study might also be used to assess the efficiency of a broad range of repositioned and/or innovative drug candidates before clinical evaluation.

© 2020 Elsevier B.V. and International Society of Chemotherapy. All rights reserved.

1. Introduction

The recent emergence of the novel pathogenic SARS-coronavirus 2 (SARS-CoV-2) is responsible for a global pandemic [1], and there is an urgent need to identify active antiviral agents. Given the global health emergency, drug repositioning is the most reliable option to design an efficient therapy for infected patients without delay [2,3]. Several drugs have already been tested, among which chloroquine (CLQ), a well-known antimalarial drug, is one of the most promising as it has shown apparent efficacy in the treatment of COVID-19-associated pneumonia in recent clinical studies [4]. However, the mechanism of action of CLQ against SARS-CoV-2 is unclear as the drug seems to exert a broad range of potential antiviral effects [5]. Therefore, although CLQ is classically considered as an inhibitor of endocytic pathways through elevation of endoso-

mal pH [6], its detailed molecular mechanism of action as an antiviral compound remains unclear [3,5]. Interestingly, CLQ has been shown to interfere with the terminal glycosylation of angiotensin-converting enzyme-2 (ACE-2) [7], which acts as a plasma membrane receptor for both SARS-CoV [8] and SARS-CoV-2 [9], and CLQ could act at several steps of the coronavirus replication cycle [7]. These data suggest the interesting and mostly unexplored possibility that CLQ could prevent viral attachment through a direct effect on host cell surface molecules.

One important characteristic of human coronaviruses is that besides their protein membrane receptor, they also depend upon sialic-acid-containing glycoproteins and gangliosides that act as primary attachment factors along the respiratory tract [10]. The present study used a combination of structural and molecular modelling approaches [11] to investigate the potential interaction between CLQ and sialic acids. A ganglioside-binding site in the N-terminal domain (NTD) of the spike (S) glycoprotein of SARS-CoV-2 was identified, and CLQ was shown to be a potential blocker of the S-ganglioside interaction which occurs in the first step of the vi-

* Corresponding author. Address: INSERM UMR_S 1072, Boulevard Pierre Drame, 13015 Marseille, France.

E-mail address: jacques.fantini@univ-amu.fr (J. Fantini).

<https://doi.org/10.1016/j.ijantimicag.2020.105960>

0924-8579/© 2020 Elsevier B.V. and International Society of Chemotherapy. All rights reserved.

ral replication cycle (i.e. attachment to the surface of respiratory cells, mediated by the S protein). In addition, the antiviral potential of CLQ and its derivative hydroxychloroquine (CLQ-OH) against SARS-CoV-2 were compared. Overall, this study found that CLQ and CLQ-OH may be used to fight pathogenic human coronaviruses [3–6], including SARS-CoV-2 [12].

2. Materials and methods

In-silico analyses were performed using Hyperchem and Molegro Molecular viewer as described [11,13,14]. The initial coordinates of GM1 were obtained from CHARMM-GUI Glycolipid Modeler (<http://www.charmmgui.org/?doc=input/glycolipid>; [15]), which uses the internal coordinate information of common glycosidic torsion angle values, orients the ganglioside perpendicular to the membrane, and performs Langevin dynamics with a cylindrical restraint potential to keep the whole GM1 molecule cylindrical, especially the membrane-embedded ceramide part. In the next step, the ganglioside was included in a periodic box solvated with 1128 water molecules. The system was energy-minimized six times, switching between runs using the steepest descent gradients and runs using Polak–Ribière conjugate gradients until convergence to machine precision. The ganglioside was subsequently merged with the NTD domain of SARS-CoV-2 S protein as obtained from pdb file # 6VSB [16]. Initial conditions corresponded to minimized structures obtained with the Polak–Ribière algorithm. Docked complexes were subsequently submitted to iterative cycles of molecular dynamics using the CHARMM36 force field optimized for carbohydrates [17]. Interaction energies were calculated from stable complexes using the Ligand Energy Inspector function of Molegro [13].

N-acetylneuraminic acid (Neu5Ac) was generated with the Hyperchem database. 9-O-acetyl-N-acetylneuraminic acid (9-O-SIA) was retrieved from pdb file 6Q06 [18]. CLQ is N-(7-chloroquinolin-4-yl)-N,N-diethyl-pentane-1,4-diamine. Its three-dimensional structure was retrieved from pdb file # 4V20 [19]. CLQ-OH is (RS)-2-[[4-[(7-chloroquinolin-4-yl)amino]pentyl](ethyl)amino]ethanol. CLQ-OH was generated by hydroxylation of CLQ with Hyperchem. Both CLQ and CLQ-OH were energy-minimized and merged with water molecules as described below.

3. Results

3.1. Structural and conformational analysis of CLQ and CLQ-OH in water

The chemical structures of CLQ and CLQ-OH are shown in Fig. 1(a,b). The only difference between the two molecules is the presence of a terminal hydroxyl group in CLQ-OH. This OH group has a marked influence on the conformation and water-solubilization properties of the drug. CLQ-OH may adopt a wide range of conformations, the most stable being the extended one shown in Fig. 1(c). When immersed in a periodic box of 31.5 Å² with 1042 water molecules, the system reached, at equilibrium, an estimated energy of interaction of -92 kJ.mol⁻¹, accounting for 56 water molecules solvating CLQ-OH Fig. 1(d). In contrast, due to an intramolecular hydrophobic effect, CLQ appeared to be more condensed than CLQ-OH Fig. 1(e). At equilibrium, CLQ was surrounded by 58 water molecules with an energy of interaction of -79 kJ.mol⁻¹ Fig. 1(f).

These water-compatible conformations of CLQ and CLQ-OH were used as initial conditions for studying the interaction of these drugs with sialic acids and gangliosides.

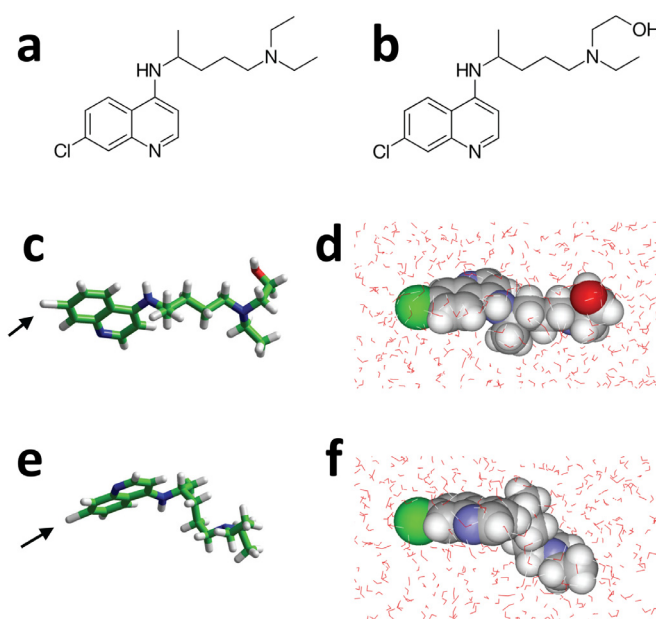


Fig. 1. Chemical structure of chloroquine (CLQ) and hydroxychloroquine (CLQ-OH). (a) CLQ. (b) CLQ-OH. (c) CLQ-OH extended conformer. (d) CLQ-OH in water. (e) Typical condensed conformer of CLQ. (f) CLQ in water. The molecules in (c–f) are shown in either tube or sphere rendering (carbon, green; nitrogen, blue; oxygen, red; hydrogen, white). In (c) and (e), the chlorine atom of CLQ and CLQ-OH is indicated by an arrow.

3.2. Sialic acids as molecular targets of CLQ and CLQ-OH

Neu5Ac is the predominant sialic acid found in human glycoproteins and gangliosides. When CLQ was merged with Neu5Ac, a quasi-instantaneous fit occurred between the two molecules, whose global shapes in water are geometrically complementary Fig. 2(a). This is particularly obvious in the views of the CLQ–Neu5Ac complex in mixed surface/balls and sticks rendition Fig. 2(a,b). The interaction was driven by the positioning of the negative charge of the carboxylate group of Neu5Ac and one of the two cationic charges of CLQ (pKa 10.2) Fig. 2(c). The energy of interaction of this complex was estimated to be -47 kJ.mol⁻¹. As coronaviruses preferentially interact with 9-O-acetyl-N-acetylneuraminic acid (9-O-SIA) [10], this study used a similar molecular modelling approach to assess whether CLQ could also interact with this specific sialic acid. A good fit between CLQ and 9-O-SIA was obtained Fig. 2(d–f), with an energy of interaction of -45 kJ.mol⁻¹. In this case, the carboxylate group of the sialic acid interacted with the cationic group of the nitrogen-containing ring of CLQ (pKa 8.1) Fig. 2(d). The complex was further stabilized by OH- π and van der Waals interactions.

Next, CLQ-OH was tested to assess whether it could, as CLQ, bind to 9-O-SIA (Fig. 3). The complex obtained with CLQ-OH was very similar to that obtained with CLQ [compare Fig. 3(a,b) with Fig. 2(e,f), although several conformational adjustments occurred during the simulations. Interestingly, the OH group of CLQ-OH reinforced the binding of CLQ to sialic acid through establishment of a hydrogen bond Fig. 3(c,d). Overall, this hydrogen bond compensated for the slight loss of energy caused by the conformational rearrangement, and the energy of interaction of the complex was estimated to be -46 kJ.mol⁻¹, which is very close to the value obtained for CLQ (-45 kJ.mol⁻¹).

3.3. Molecular recognition of gangliosides by CLQ and CLQ-OH

In the respiratory tract, sialic acids are usually part of glycoproteins and gangliosides. Molecular modelling approaches were

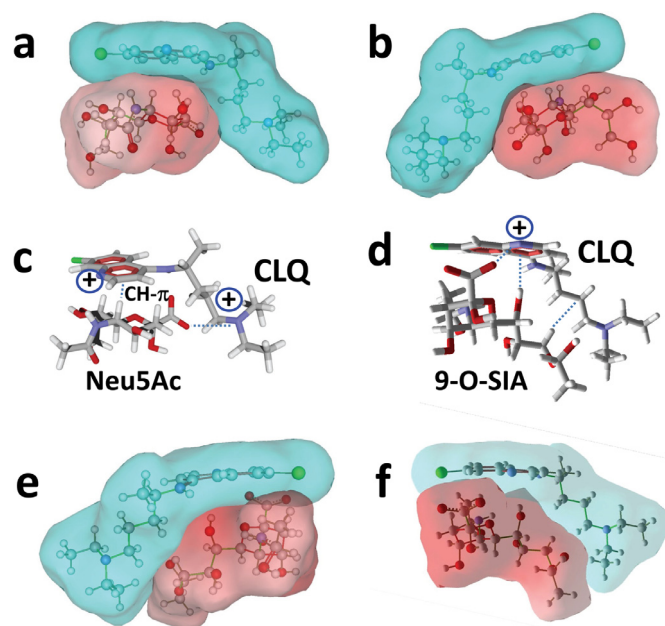


Fig. 2. Molecular modelling of chloroquine (CLQ) interaction with sialic acids. (a,b) Surface representation of the CLQ-sialic acid (Neu5Ac) complex. Two opposite views of the complex are shown. Note the geometric complementarity between the L-shape conformer of CLQ dissolved in water (in blue) and Neu5Ac (in red). (c) Neu5Ac bound to CLQ via a combination of CH-π and electrostatic interactions with one of the cationic groups of CLQ (+). (d) Molecular modelling of CLQ bound to N-acetyl-9-O-acetylneuraminic acid (9-O-SIA). From right to left, the dashed lines indicate a series of van der Waals, OH-π and electrostatic contacts with both cationic groups of CLQ (+). (e,f) Surface representations of the CLQ-9-O-SIA complex.

used to assess whether CLQ and CLQ-OH can recognize sialic acid units in their natural molecular environment. In these simulations, ganglioside GM1 was chosen as a representative example of human plasma membrane gangliosides. A first series of simulations was performed with CLQ. When merged with the ganglioside, CLQ had two distinct binding sites, both located in the polar saccharide part of GM1. The first site was located at the tip of the saccharide moiety of the ganglioside Fig. 4(a,b). The energy of interaction was estimated to be -47 kJ.mol^{-1} . CLQ retained the typical L-shape

structure of the water-soluble conformer bound to isolated sialic acids [compare Figs 2(c) and 4(a)]. From a mechanistic point of view, the carboxylate group of the sialic acid of GM1 was oriented towards the cationic groups of CLQ. The rings of CLQ faced the N-acetylgalactosamine (GalNAc) residue of GM1, establishing both OH-π interaction and hydrogen bonding Fig. 4(b). The second site was in a large area including both the ceramide-sugar junction and the saccharide moiety Fig. 4(c). The chlorine atom of CLQ was oriented towards the ceramide axis, allowing the nitrogen-containing ring of CLQ to stack on to the pyrane ring of the first sugar residue [i.e. glucose (Glc)]. The perfect geometric complementarity of the two partners Fig. 4(c,d) accounted for a particularly high energy of interaction in this case (-61 kJ.mol^{-1}). Interestingly, there was no overlap between the two CLQ-binding sites on GM1, so the ganglioside could accommodate two CLQ molecules Fig. 4(e), reaching a global energy of interaction of -108 kJ.mol^{-1} . A similar situation was observed with CLQ-OH, which occupies the same binding site as CLQ Fig. 4(f). In this case, the energy of interaction was further increased by stabilizing contacts established between the two CLQ-OH molecules, reaching -120 kJ.mol^{-1} . Overall, these data showed that CLQ and CLQ-OH have a good fit for sialic acids, either isolated or bound to gangliosides.

3.4. Structural analysis of the NTD of SARS-CoV-2 S protein

The next step of this study was to determine how SARS-CoV-2 could interact with plasma membrane gangliosides, and whether such interaction could be affected by CLQ and CLQ-OH. The global structure of the SARS-CoV-2 S protein [16] is shown in Fig. 5(a-d). It consists of a trimer of S proteins, each harbouring two distinct domains distant from the viral envelope: the receptor-binding region (RBD) and the NTD.

It was reasoned that if the RBD is engaged in functional interactions with the ACE-2 receptor, it would be interesting to search for potential ganglioside-binding sites on the other cell-accessible domain of the S glycoprotein (i.e. the NTD). The NTD contains approximately 290 amino acid residues. The tip of the NTD was of particular interest, as it displays a flat interface Fig. 5(f) ideally positioned for targeting a ganglioside-rich plasma membrane microdomain, such as a lipid raft.

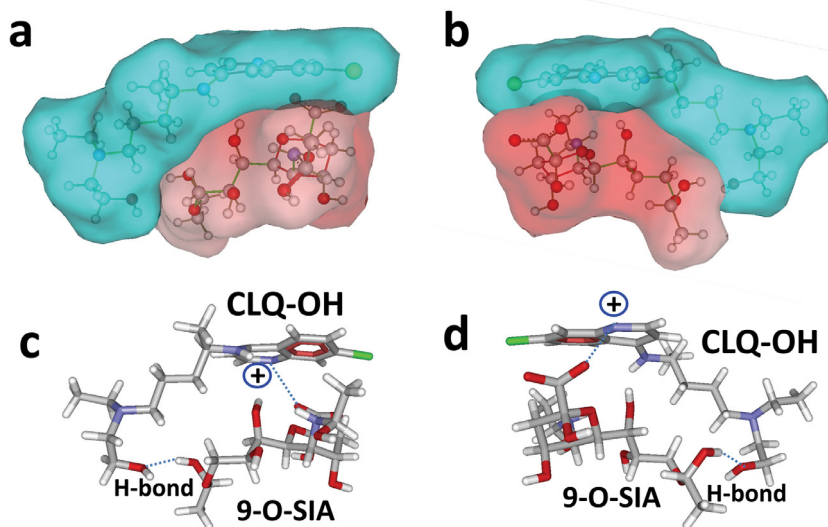


Fig. 3. Molecular modelling of hydroxychloroquine (CLQ-OH) interaction with sialic acids. (a,b) Surface representation of CLQ-OH bound to N-acetyl-9-O-acetylneuraminic acid (9-O-SIA). Two opposite views of the complex are shown. Note the geometric complementarity between CLQ-OH (in blue) and 9-O-SIA (in red). (c,d) Molecular mechanism of CLQ-OH binding to 9-O-SIA: combination of electrostatic interactions and hydrogen bonding.

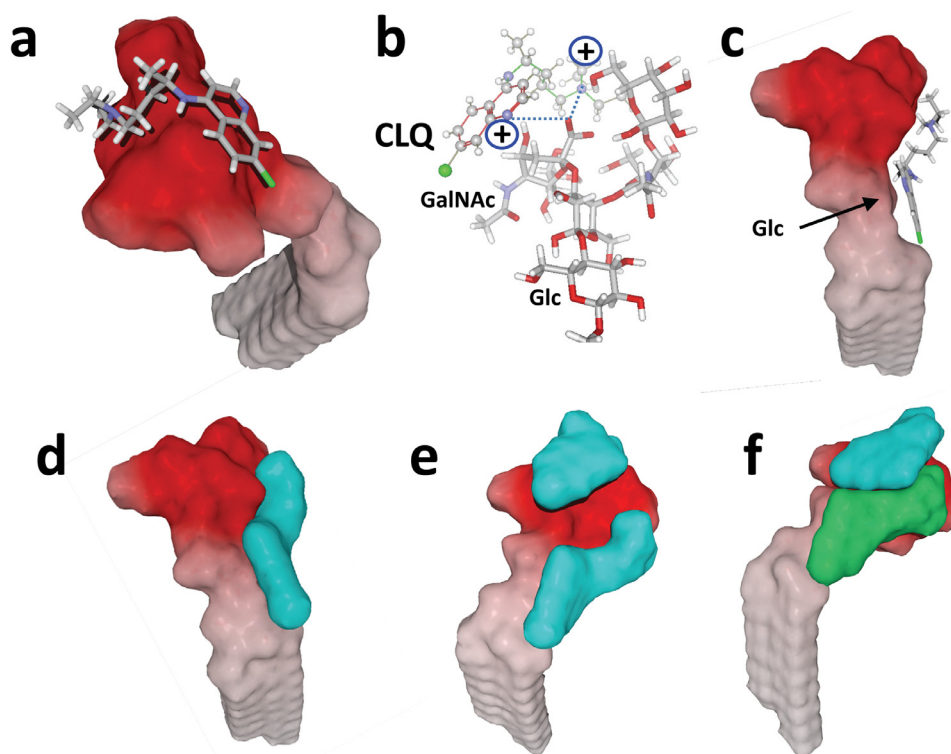


Fig. 4. Molecular modelling simulations of chloroquine (CLQ) and hydroxychloroquine (CLQ-OH) binding to ganglioside GM1. The surface electrostatic potential of GM1 indicates a non-polar, membrane-embedded part corresponding to ceramide (white areas), and an acidic part protruding in the extracellular space corresponding to the sialic-acid-containing saccharide part (red areas). (a) CLQ bound to the tip of the carbohydrate moiety of GM1. (b) Molecular mechanism of CLQ-ganglioside interactions. (c) Molecular dynamics simulations revealed a second site of interaction. In this case, the aromatic cycles of CLQ are positioned at the ceramide-sugar junction, whereas the nitrogen atoms interact with the acidic part of the ganglioside (not illustrated). (d,e) Surface views of GM1 complexed with one (d) or two (e) CLQ molecules (both in blue), illustrating the geometric complementarity of GM1 and CLQ molecules. (f) One GM1 molecule can also accommodate two distinct CLQ-OH molecules simultaneously, after slight rearrangement allowing increased fit due to CLQ-OH/CLQ-OH interactions. To improve clarity, CLQ-OH molecules bound to GM1 are represented in two distinct colours (blue and green).

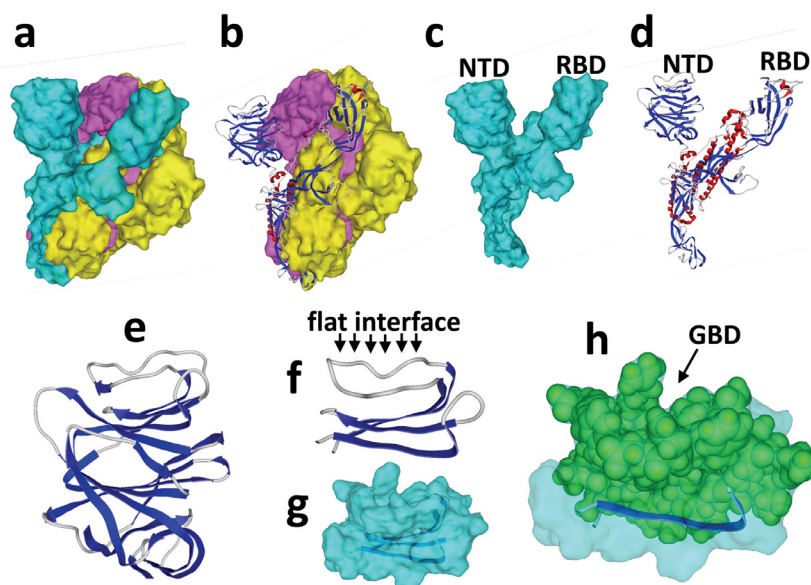


Fig. 5. Structural features of the SARS-CoV-2 spike (S) protein. (a) Trimeric structure (each S protein has a distinct surface colour, 'blue', 'yellow' and 'purple'). (b) Ribbon representation of 'blue' S protein in the trimer (α -helix, red; β -strand, blue; coil, grey). (c) Surface structure of the 'blue' S protein isolated from the trimer. (d) Ribbon structure of the 'blue' S protein. (e) Zoom on the N-terminal domain (NTD) of the 'blue' S protein. (f,g) Molecular model of a minimal NTD obtained with Hyperchem [ribbon in representation in (f), surface rendering in (g)]. (h) Highlighting of the amino acid residues of the NTD that could belong to a potential ganglioside-binding domain.

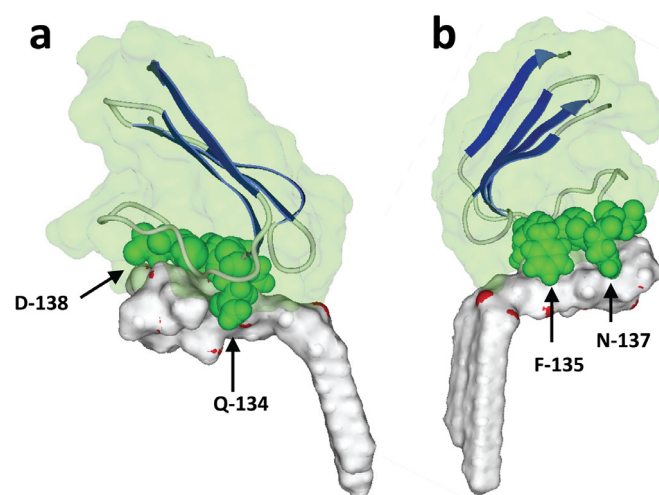


Fig. 6. Molecular complex between the N-terminal domain (NTD) of SARS-CoV-2 spike protein and a single GM1 ganglioside. The NTD is represented in ribbons superposed with a transparent surface rendering (light green). Two symmetric views of the complex are shown (a,b). The amino acid residues Q-134 to D-138 located in the centre of the ganglioside-binding domain are represented as green spheres. The saccharide part of the ganglioside forms a landing surface for the tip of the NTD.

The amino acid sequence of the planar interfacial surface located at the tip of the NTD was analysed for the presence of consensus ganglioside-binding domains [20]. These motifs are constituted by a triad of mandatory amino acid residues such as (K,R)-X_n-(F,Y,W)-X_n-(K,R). The X_n intercalating segments, usually four to five residues, may contain any amino acid, but often Gly, Pro and/or Ser residues. The strict application of this algorithm did not allow the detection of any potential ganglioside-binding domain in this region of the NTD. However, an intriguing over-representation of aromatic and basic residues was found in the 129–158 segment: 129-KVCEFOFCNDPFLGVYYHKNNKSWMESEFR-158. This 30-amino acid stretch also contains Gly, Pro and/or Ser residues that are often found in ganglioside-binding motifs. These observations supported the notion that the tip of the NTD could display a large ganglioside-attachment interface.

3.5. Molecular interactions between gangliosides and the NTD of SARS-CoV-2 S protein

Molecular dynamic simulations of a structural motif encompassing amino acid residues 100–175 of the NTD Fig. 5(f–h) merged with ganglioside GM1 further supported this concept. As shown in Fig. 6, the large flat area of this structural domain fitted very well with the protruding oligosaccharide part of the ganglioside. Several amino acid residues appear to be critical for this interaction, especially Phe-135, Asn-137 and Arg-158 (Table 1). Overall, the complex involved 10 amino acid residues for a total energy of interaction of -100 kJ.mol^{-1} . At this stage, it was observed that approximately 50% of the interface was involved in the complex, leaving the remaining 50% available for interaction with a second GM1 molecule. As expected, merging a second GM1 molecule with the preformed GM1–NTD complex led to a trimolecular complex consisting of two gangliosides in a typical symmetrical chalice-like structure into which the NTD could insert its interfacial ganglioside-binding domain (Fig. 7). The formation of this trimolecular complex was progressive, starting with a conformational rearrangement of the first ganglioside–NTD complex triggered by the second GM1 molecule. The energy of interaction of the new complex was consistently increased by 37%, reaching an estimated value of -137 kJ.mol^{-1} . At this stage, attachment of the NTD to the ganglioside-rich microdomain involved the whole interface (i.e.

Table 1

Energy of interaction of each amino acid residue of SARS-CoV-2 spike protein in contact with GM1 molecules.

Amino acid residues	Energy of interaction (kJ.mol ⁻¹)
First step: one GM1 molecule	
Asp-111	–5.6
Lys-113	–8.2
Gln-134	–8.6
Phe-135	–20.1
Cys-136	–7.0
Asn-137	–15.2
Asp-138	–6.4
Arg-158	–17.4
Ser-161	–9.7
Ser-162	–2.0
Total	–100.2
Second step: two GM1 molecules	
Asp-111	–15.8
Ser-112	–10.7
Lys-113	–9.2
Gln-134	–11.2
Phe-135	–10.5
Cys-136	–6.2
Asn-137	–4.7
Phe-140	–5.2
Gly-142	–5.6
Glu-156	–9.0
Phe-157	–13.8
Arg-158	–19.8
Tyr-160	–3.2
Ser-161	–9.7
Ser-162	–2.0
Total	–136.6

15 surface-accessible residues from Asp-111 to Ser-162). The critical residues were Asp-111, Gln-134, Phe-135, Arg-158 and Ser-161 (Table 1).

3.6. Potential coordinated interactions between SARS-CoV-2 and the plasma membrane of a host cell: key role of gangliosides in lipid rafts

Taken together, these data strongly support the concept of a dual receptor/attachment model for SARS-CoV-2, with the RBD domain being involved in ACE-2 receptor recognition, and the NTD interface responsible for finding a ganglioside-rich landing area (lipid raft) at the cell surface.

Such a dual receptor model, consistent with the topology of the SARS-CoV-2 S protein, is proposed in Fig. 8. With this model in mind, the potential effects of CLQ and CLQ-OH were studied, both of which, according to the molecular modelling data, have a high affinity for sialic acids and gangliosides.

3.7. Molecular mechanism of CLQ and CLQ-OH antiviral effect: preventing SARS-CoV-2 S protein access to cell surface gangliosides

With the aim of establishing whether CLQ and CLQ-OH could prevent the attachment of SARS-CoV-2 to plasma membrane gangliosides, the initial NTD–GM1 complex was superposed with a drug–GM1 complex (Fig. 9). To improve clarity, the ganglioside is not presented in Fig. 9. This superposition shows that the NTD and the drug (CLQ-OH in this case) share the same spatial position when bound to GM1, so GM1 cannot bind the viral protein and the drug simultaneously. This is due to the fact that the NTD and the drugs (CLQ and CLQ-OH) bind to GM1 with a similar mechanism controlled by a dyad of functional interactions: a hydrogen bond and a geometrically perfect CH- π stacking interaction. In the case of the NTD, the hydrogen bond involves Asn-167, whereas CH- π stacking is mediated by the aromatic ring of Phe-135 Fig. 6(b). On

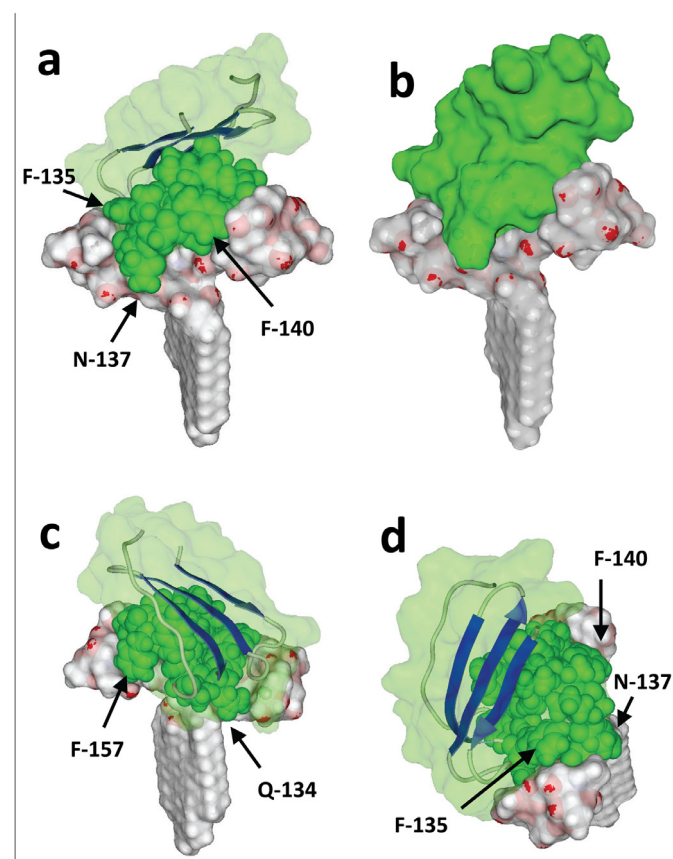


Fig. 7. Molecular complex between the N-terminal domain (NTD) of SARS-CoV-2 spike protein and a dimer of GM1. In (a), (c) and (d), the NTD is represented as in Fig. 4. In (b), the surface of the NTD is shown without any transparency. The amino acid residues Q-134 to S-162 belonging to the ganglioside-binding domain (GBD) are represented as green spheres. Compared with a single GM1 molecule, the dimer of gangliosides forms a larger attractive surface for the NTD. In the above view of (d), the anchorage of the NTD to the gangliosides is particularly obvious. As chloroquine also interacts with the saccharide part of GM1, its presence would clearly mask most of the landing surface available for the NTD, preventing attachment of the virus to the plasma membrane of host cells.

one hand, Asn-167 establishes a network of hydrogen bonds with the GalNAc residue of GM1. On the other hand, the flat aromatic ring of Phe-135 stacks on to the cycle of the Glc residue of GM1. In the case of CLQ and CLQ-OH, it is the nitrogen-containing ring of the drug that stacks on to the Glc ring Fig. 4(c). Note that both the Phe-135 (in red) and CLQ-OH (in green) rings are located in the same position (Fig. 9). The other CLQ-OH molecule, which covers the tip of the sugar part of the ganglioside, interacts with the GalNAc ganglioside Fig. 4(b). When the NTD is bound to the ganglioside, the side chain of Asn-137 is found in this exact position (Fig. 9). Thus, once two CLQ-OH (or two CLQ) molecules are bound to a ganglioside Fig. 4(e,f), any binding of a SARS-Cov-2 S protein to the same ganglioside is totally prevented. The energy required to overcome this steric incompatibility is estimated to be several hundred $\text{kJ}\cdot\text{mol}^{-1}$, which is far too high to occur.

3.8. Sequence alignment analysis of SARS-CoV-2 and related coronavirus: evolution of the ganglioside-binding domain at critical amino acid residues

As CLQ and CLQ-OH are potential therapies for SARS-CoV-2 infection, it is important to check whether the amino acid residues identified as critical for ganglioside binding are conserved among clinical isolates. The alignment of the 111–162 domain of 11 clinical

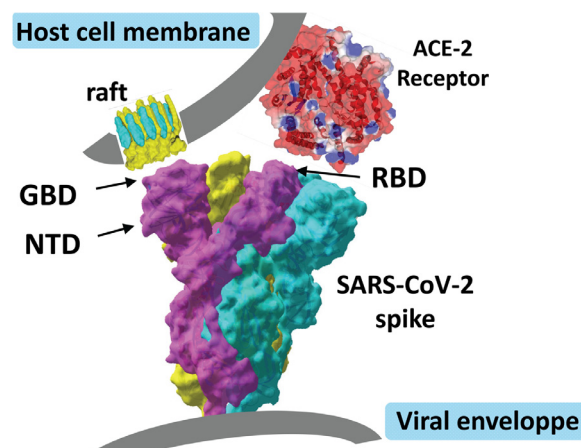


Fig. 8. Dual recognition of gangliosides and angiotensin-converting enzyme-2 (ACE-2) by SARS-CoV-2 spike (S) protein. The viral protein displays two distinct domains, the tips of which are available for distinct types of interactions. The receptor-binding domain binds to the ACE-2 receptor, and the N-terminal domain (NTD) binds to the ganglioside-rich domain of the plasma membrane. Lipid rafts, which are membrane domains enriched in gangliosides (in yellow) and cholesterol (in blue), provide a perfect attractive interface for adequately positioning the viral S protein at the first step of the infection process. These structural and molecular modelling studies suggest that amino acid residues 111–162 of the NTD form a functional ganglioside-binding domain, the interaction of which with lipid rafts can be efficiently prevented by chloroquine and hydroxychloroquine.

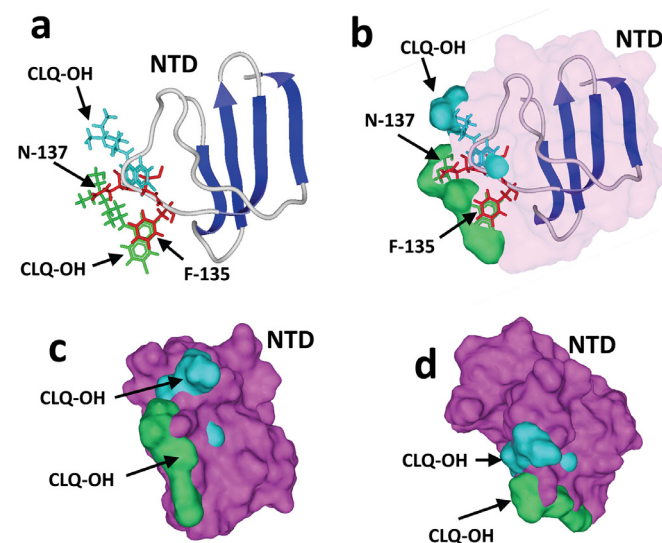


Fig. 9. Mechanism of action of hydroxychloroquine (CLQ-OH). The N-terminal domain (NTD) bound to GM1 was superposed onto GM1 in interaction with two CLQ-OH molecules. The models only show the NTD and both CLQ-OH molecules (not GM1, to improve clarity). (a,b) The aromatic ring of F-135 (in red), which stacks onto the glucose cycle of GM1, overlaps the aromatic CLQ-OH rings (in green) which also bind to GM1 via a CH- π stacking mechanism. The N-137 residue of the NTD interacts with the N-acetylgalactosamine residue of GM1, but this interaction cannot occur in the presence of CLQ-OH as this part of GM1 is totally masked by the drug. (c,d) Superposition of the NTD surface (in purple) with the two CLQ-OH molecules bound to GM1, illustrating the steric impossibility that prevents NTD binding to GM1 when both CLQ-OH molecules are already interacting with the ganglioside.

cal isolates of SARS-CoV-2 from various geographic origins (including Asia and USA) is shown in Fig. 10. In this region, which contains the ganglioside-binding domain identified in the present report, all amino acids are fully conserved. Interestingly, the motif is built like a giant consensus ganglioside-binding domain: a central region displaying the critical aromatic residue (Phe-135) and a basic residue at each end (Lys-113 and Arg-158). In the middle of

Virus	a.a.	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	a.a.
6VSB_A	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 Reference
QIA98596	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 Taiwan
QH200379	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 South Korea
QH200389	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 USA
QIA98606	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 Taiwan
QIA20044	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 China
QH287592	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 USA
QH287582	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 USA
Q82464	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 USA
QHU36844	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 China
QH200358	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 China
QH200399	111	DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	162 USA
SARS-CoV-2		DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	6VSB_A
Bat RaTG13		DSKTQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	QHR63300.2
Bat SARS-like		DNTSQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	AVP78031.1
Bat SARS-like		DNTSQSL	LLIVN	NATNV	IKVCE	FQFCND	PFLGV	YYHKN	KSWM	SEFRV	YSS	AVP78042.1
Human SARS-related		NNKSQSV	IIINN	STNV	VIRAC	NFEL	CDN	FFAV	SKPM	GTQ	THTM	IFDN 5X4S_A
Human SARS BJ01		NNKSQSV	IIINN	STNV	VIRAC	NFEL	CDN	FFAV	SKPM	GTQ	THTM	IFDN AAS48453.1
Bat SARS-like		NNKSQSV	IIINN	STNV	VIRAC	NFEL	CDN	FFAV	SKPM	GTQ	THTM	IFDN ATO98157.1
Bat SARS-like		NNKSQSV	IIINN	STNV	VIRAC	NFEL	CDN	FFAV	SKPM	GTQ	THTM	IFDN ATO98205.1
Bat SARS-like		NNKSQSV	IIINN	STNV	VIRAC	NFEL	CDN	FFAV	SKPM	GTQ	THTM	IFDN ALK02457.1
PalM civet SARS-like		DNKSQSV	IIINN	STNV	VIRAC	NFEL	CDN	FFV	SKPM	GTQ	THTM	IFDN AAV98002.1

Fig. 10. Amino acid sequence alignments of the ganglioside-binding domain (GBD) of the SARS-CoV-2 spike protein. (a) Clinical SARS-CoV-2 isolates aligned with the reference sequence (6VSB_A, fragment 111–162). The amino acid residues involved in GM1 binding are indicated in red. Two asparagine residues acting as glycosylation sites are highlighted in yellow. (b) Alignments of human and animal viruses compared with SARS-CoV-2 (6VSB_A, fragment 111–162). Deletions are highlighted in green, amino acid changes in residues involved in ganglioside binding are highlighted in blue, conserved residues of the GBD are highlighted in red, and asparagine residues acting as glycosylation sites are highlighted in yellow.

each stretch separating this typical triad, there is a N-glycosylation site (Asn-122 and Asn-149). These last regions are not directly involved in ganglioside binding, so the oligosaccharide linked to these asparagine residues could be perfectly intercalated between the sugar head group of gangliosides.

It was also noted that the ganglioside-binding domain of the NTD is fully conserved in bat RaTG13, which indicates a close relationship between the bat coronavirus and the human isolates that are currently circulating around the world. However, the motif is slightly different in other bat- and human-related coronaviruses (Fig. 10), suggesting a recent evolution which could explain, at least in part, why SARS-CoV-2 is more contagious than previously characterized human coronaviruses.

4. Discussion

Sialic acids linked to glycoproteins and gangliosides are used by a broad range of viruses as receptors and/or attachment factors for cell entry [10]. These viruses include major human pathogens affecting the respiratory tract, such as influenza [21] and coronaviruses [22,23]. The attachment to sialic-acid-containing cell surface structures is mediated by receptor-binding proteins that belong to the viral spike. In the case of coronaviruses, this function is fulfilled by the S glycoprotein [9,24]. SARS-CoV and SARS-CoV-2 interact with the ACE-2 protein, which has been identified as a key determinant of the contagiousness of viruses [8]. However, considering the increased transmissibility of SARS-CoV-2 compared with SARS-CoV, binding to ACE-2 alone might not be enough to ensure robust infection of the upper respiratory tract. Thus, it is likely that SARS-CoV-2 might also bind to other cell surface attachment factors, such as sialic-acid-containing glycoproteins and gangliosides. Consistent with this notion, it has been shown that depletion of cell surface sialic acids by neuraminidase treatment inhibited MERS-CoV entry of human airway cells [25]. These data, which provide direct evidence that sialic acids play a critical role in human coronavirus attachment, broaden the therapeutic options to block the replication of SARS-CoV-2 that is responsible for the COVID-19 pandemic.

Few drugs have shown consistent antiviral efficiency *in vitro* together with reported efficiency in patients infected with SARS-CoV-2 [3,12]. Of these, CLQ is of interest as its chemical structure is based on a combination of cationic nitrogen atoms and aromatic rings. Both features have been shown to be key determi-

nants of the recognition of sialic acids and gangliosides by proteins [20,26]. Modelling approaches have been used successfully to decipher various molecular mechanisms of protein–sugar interactions accounting for the interaction of virus [27], bacteria [28], membrane [13] and amyloid proteins [20] with cell surface glycolipids. This *in-silico* strategy was applied to unravel the molecular mechanisms underlying the antiviral mechanisms of CLQ and CLQ-OH against SARS-CoV-2 infection. First, it was shown that CLQ and CLQ-OH bind readily to sialic acids with high affinity, including the typical 9-O-SIA subtype recognized by coronaviruses [23]. Next, it was shown that CLQ and CLQ-OH also bind to sialic-acid-containing gangliosides. Based on these data, the drugs may also recognize the sialic acid residues of glycoproteins. Further studies will help clarify this point.

This molecular modelling study has identified a new type of ganglioside-binding domain in the NTD of the SARS-CoV-2 S protein. This ganglioside-binding domain consists of a large flat interface enriched in aromatic and basic amino acid residues. It covers a stretch of 52 amino acid residues (111–162), which is longer than all linear ganglioside-binding domains characterized to date [29]. However, the new SARS-CoV-2 motif is organized into three distinct regions, including a central aromatic domain and two terminal basic domains (Fig. 10). Thus, this motif displays the typical features that determine optimal binding to gangliosides (i.e. CH- π stacking and electrostatic interactions).

A major outcome of this study is the demonstration that CLQ and CLQ-OH display molecular groups that fully mimic the way in which the S protein binds to gangliosides. Two CLQ (or two CLQ-OH) molecules can bind simultaneously to the polar head group of ganglioside GM1. Interestingly, these simulations indicated that CLQ-OH is more potent than CLQ, in line with the reported increased antiviral activity of CLQ-OH against SARS-CoV-2 [30]. Once bound to GM1, the drugs prevent any access to the Glc and GalNAc units of the ganglioside, which are the critical binding residues for Phe-135 and Asn-137, respectively. This amino acid dyad, as well as all the other residues that mediate ganglioside binding by the SARS-CoV-2 spike, is fully conserved among clinical isolates worldwide. It is also conserved in the bat RaTG13 isolate, which reinforces the hypothesis of bat-to-human transmission. From an epidemiological point of view, it can be hypothesized that the evolution of this motif has conferred an enhanced attachment capacity of human coronaviruses to the respiratory tract through improved S–ganglioside interactions.

5. Conclusion

Given the global health emergency, drug repurposing is obviously the option of choice [2,3]. However, a considerable amount of time could be saved by in-silico testing to determine the capability of any potential anti-SARS-CoV-2 to disrupt the interaction of the S protein with the host cell membrane. Applied to both RBD-ACE-2 and NTD-ganglioside interactions, this molecular modelling method will help select those drugs that are likely to interfere with the initial attachment of virus particles to the respiratory tract surface epithelium. The study data support the use of CLQ, and preferentially CLQ-OH, as initial therapy for patients infected with SARS-CoV-2.

Funding: None.

Competing interests: None declared.

Ethical approval: Not required.

References

- [1] Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020;579:270–3.
- [2] Mullard A. Drug repurposing programmes get lift off. *Nat Rev Drug Discov* 2012;11:505–6.
- [3] Colson P, Rolain JM, Raoult D. Chloroquine for the 2019 novel coronavirus SARS-CoV-2. *Int J Antimicrob Agents* 2020;55:105923.
- [4] Gao J, Tian Z, Yang X. Breakthrough: chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci Trends* 2020 Feb 19 [Epub ahead of print]. doi:10.5582/bst.2020.01047.
- [5] Savarino A, Di Trani L, Donatelli I, Cauda R, Cassone A. New insights into the antiviral effects of chloroquine. *Lancet Infect Dis* 2006;6:67–9.
- [6] Mauthe M, Orhon I, Rocchi C, Zhou X, Luhr M, Hijlkema KJ, et al. Chloroquine inhibits autophagic flux by decreasing autophagosome-lysosome fusion. *Autophagy* 2018;14:1435–55.
- [7] Vincent MJ, Bergeron E, Benjannet S, Erickson BR, Rollin PE, Ksiazek TG, et al. Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virol J* 2005;2:69.
- [8] Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 2003;426:450–4.
- [9] Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of the SARS-CoV-2 by full-length human ACE2. *Science* 2020 Mar 4 pii: eabb2762 [Epub ahead of print]. doi:10.1126/science.abb2762.
- [10] Matrosovich M, Herrler G, Klenk HD. Sialic acid receptors of viruses. *Top Curr Chem* 2015;367:1–28.
- [11] Di Scala C, Fantini J. Hybrid in silico/in vitro approaches for the identification of functional cholesterol-binding domains in membrane proteins. *Methods Mol Biol* 2017;1583:7–19.
- [12] Colson P, Rolain JM, Lagier JC, Brouqui P, Raoult D. Chloroquine and hydroxychloroquine as available weapons to fight COVID-19. *Int J Antimicrob Agents* 2020 Mar 4:105932 [Epub ahead of print]. doi:10.1016/j.ijantimicag.2020.105932.
- [13] Flores A, Ramirez-Franco J, Desplantes R, Debreux K, Ferracci G, Wernert F, et al. Gangliosides interact with synaptotagmin to form the high-affinity receptor complex for botulinum neurotoxin B. *Proc Natl Acad Sci USA* 2019;116:18098–108.
- [14] Huang Y, Huang S, Di Scala C, Wang Q, Wandall HH, Fantini J, et al. The glycosphingolipid MacCer promotes synaptic bouton formation in *Drosophila* by interacting with Wnt. *Elife* 2018;7 pii:e38183.
- [15] Lee J, Patel DS, Stähle J, Park SJ, Kern NR, Kim S, et al. CHARMM-GUI membrane builder for complex biological membrane simulations with glycolipids and lipoglycans. *J Chem Theory Comput* 2019;15:775–86.
- [16] Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 2020;367:1260–3.
- [17] Guvench O, Greene SN, Kamath G, Brady JW, Venable RM, Pastor RW, et al. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* 2008;29:2543–64.
- [18] Park YJ, Walls AC, Wang Z, Sauer MM, Li W, Tortorici MA, et al. Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. *Nat Struct Mol Biol* 2019;26:1151–7.
- [19] Huta BP, Mehlenbacher MR, Nie Y, Lai X, Zubieta C, Bou-Abdallah F, et al. The lysosomal protein saposin B binds chloroquine. *Chem Med Chem* 2016;11:277–382.
- [20] Yahi N, Fantini J. Deciphering the glycolipid code of Alzheimer's and Parkinson's amyloid proteins allowed the creation of a universal ganglioside-binding peptide. *PLoS One* 2014;9:e104751.
- [21] Verma DK, Gupta D, Lal SK. Host lipid rafts play a major role in binding and endocytosis of influenza A virus. *Viruses* 2018;10 pii:E650.
- [22] Lu Y, Liu DX, Tam JP. Lipid rafts are involved in SARS-CoV entry into Vero E6 cells. *Biochem Biophys Res Commun* 2008;369:344–9.
- [23] Tortorici MA, Walls AC, Lang Y, Wang C, Li Z, Koerhuis D, et al. Structural basis for human coronavirus attachment to sialic acid receptors. *Nat Struct Mol Biol* 2019;26:481–9.
- [24] Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 2020 pii:S0092-8674(20)30229-4 [Epub ahead of print]. doi:10.1016/j.cell.2020.02.052.
- [25] Li W, Hulsmit RJG, Widjaja I, Raj VS, McBride R, Peng W, et al. Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. *Proc Natl Acad Sci USA* 2017;114:E8508–17.
- [26] Fantini J, Yahi N. Molecular insights into amyloid regulation by membrane cholesterol and sphingolipids: common mechanisms in neurodegenerative diseases. *Expert Rev Mol Med* 2010;12:e27.
- [27] Mahfoud R, Garmy N, Maresca M, Yahi N, Puigserver A, Fantini J. Identification of a common sphingolipid-binding domain in Alzheimer, prion, and HIV-1 proteins. *J Biol Chem* 2002;277:11292–6.
- [28] Fantini J, Garmy N, Yahi N. Prediction of glycolipid-binding domains from the amino acid sequence of lipid raft-associated proteins: application to HpaA, a protein involved in the adhesion of *Helicobacter pylori* to gastrointestinal cells. *Biochemistry* 2006;45:10957–62.
- [29] Fantini J, Yahi N. Brain lipids in synaptic function and neurological disease. Clues to innovative therapeutic strategies for brain disorders. San Francisco: Elsevier; 2015.
- [30] Yao X, Ye F, Zhang M, Cui C, Huang B, Niu P, et al. In vitro antiviral activity and projection of optimized dosing design of hydroxychloroquine for the treatment of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). *Clin Infect Dis* 2020 Mar 9 pii:ciaa237. doi:10.1093/cid/ciaa237.