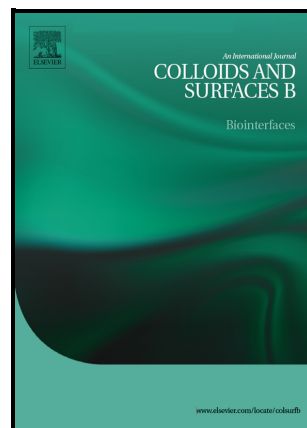


Role of cholesterol-recognition motifs in the infectivity of SARS-CoV-2 variants
Cholesterol consensus sequences in Coronaviruses

Carlos Javier Baier, Francisco J. Barrantes



PII: S0927-7765(22)00774-3

DOI: <https://doi.org/10.1016/j.colsurfb.2022.113090>

Reference: COLSUB113090

To appear in: *Colloids and Surfaces B: Biointerfaces*

Received date: 6 November 2022

Revised date: 2 December 2022

Accepted date: 10 December 2022

Please cite this article as: Carlos Javier Baier and Francisco J. Barrantes, Role of cholesterol-recognition motifs in the infectivity of SARS-CoV-2 variants
Cholesterol consensus sequences in Coronaviruses, *Colloids and Surfaces B: Biointerfaces*, (2022) doi:<https://doi.org/10.1016/j.colsurfb.2022.113090>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier.

Role of cholesterol-recognition motifs in the infectivity of SARS-CoV-2 variants

Carlos Javier Baier¹ and Francisco J. Barrantes²

¹Instituto de Ciencias Biológicas y Biomédicas del Sur (INBIOSUR), Universidad Nacional del Sur (UNS), Consejo de Investigaciones Científicas y Técnicas (CONICET), Departamento de Biología, Bioquímica y Farmacia, San Juan 670, B8000ICN, Bahía Blanca. ORCID: 0000-0001-6572-0664. E-mail: cjbaier@criba.edu.ar

²Laboratory of Molecular Neurobiology, BIOMED UCA-CONICET, 1600 Av. A. Moreau de Justo, C1107AAZ Buenos Aires, Argentina. ORCID: 0000-0002-4745-681X. E-mail: francisco_barrantes@uca.edu.ar

Short title: Cholesterol consensus sequences in Coronaviruses

Total number of words (including references): 5423

Total number of words (without references): 3837

Number of figures: 4

Abstract

The presence of linear amino acid motifs with the capacity to recognize the neutral lipid cholesterol, known as Cholesterol Recognition/interaction Amino acid Consensus sequence (CRAC), and its inverse or mirror image, CARC, has recently been reported in the primary sequence of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike S homotrimeric glycoprotein. These motifs also occur in the two other pathogenic coronaviruses, SARS-CoV, and Middle-East respiratory syndrome CoV (MERS-CoV), most conspicuously in the transmembrane domain, the fusion peptide,

the amino-terminal domain, and the receptor binding domain of SARS-CoV-2 S protein. Here we analyze the presence of cholesterol-recognition motifs in these key regions of the spike glycoprotein in the pathogenic CoVs. We disclose the inherent pathophysiological implications of the cholesterol motifs in the virus-host cell interactions and variant infectivity.

Keywords:

Cholesterol; cholesterol recognition motifs; viral infectivity; cell-surface phenomena; cholesterol-protein interactions; fusion proteins

Introduction

The appearance of ligand recognition in living organisms has occurred in various instances as a result of coevolution between interacting partners. In the case of pathogens like viruses, the interactions between their infective machinery and eukaryotic cell-surface receptors may not follow this coevolutionary mechanism [1]. This is because the virion receptors, transmembrane proteases in the case of coronaviruses (CoVs), fulfill preexisting physiological roles in the eukaryotic cells which the viruses exploit to serve their own pathophysiological cycle.

Soon after the irruption of the COVID-19 pandemic, structural work using cryo-electron microscopy (cryo-EM), and to a lesser extent X-ray diffraction and nuclear magnetic resonance (NMR) spectroscopy, produced spectacular advances in our knowledge of the causative agent, the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). These unprecedentedly vertiginous developments produced, in the course of a few months, atomistic depictions of the virus and its canonical cell-surface receptor, the metalloprotease angiotensin converting enzyme 2 (ACE2), and abundant hypotheses on the mechanistic interactions between the two (reviewed in [1]). A possible reason why ACE2 has an evolutionary advantage in SARS-CoV-2 as a target molecule is the ubiquitous cell-surface distribution of the enzyme in many host tissues, particularly mucosal epithelia [1, 2]. The spike (S) glycoprotein of SARS-CoV-2 and

other pathogenic CoVs is the one that interacts with ACE2. Binding to and fusion with the host-cell membrane appear to be sensitive to the lipid environment in which these two coupled processes occur [3].

Cholesterol plays important roles in orchestrating the biophysical properties of both proteins and lipids in the plasma membrane [4]. The observed cholesterol dependence of some viruses leads us to consider how they take advantage of cholesterol properties, in so doing optimizing viral infection or other steps of their life cycle. Several viral cholesterol-binding proteins relevant to virus entry processes have been described [5]. The amount of cholesterol in the virus envelope also appears to play a role: in members of the alphavirus genus (Chikungunya, Sindibis, Venezuelan equine encephalitis and Ross River virus), the cholesterol/phospholipid molar ratio is higher than in the host plasma membranes [6]. The higher cholesterol level in the viral particle is important in the organization of the viral envelope [7]. Cholesterol depletion in the influenza A virus envelope produces nicks and holes in the viral envelope, reducing virus infectivity [8]. Moreover, virions of influenza A virus and respiratory syncytial virus released from the cholesterol-depleted cells were less stable than those released from untreated cells [9]. The transmembrane gp41 subunit, an HIV-1 envelope protein, interacts directly with cholesterol in the viral membrane [10, 11].

Recent structural work is beginning to unravel the functional relationship between the S protein and cholesterol, present in both the host-cell membrane and the viral envelope bilayer [12]. ACE2 of Vero E6 and Caco-2 cells is recovered in the detergent-resistant phase in biochemical assays, an experimental finding that is conventionally, albeit not universally, taken as evidence that a protein resides in liquid-ordered (Lo) lipid domains, also termed “lipid rafts”. There is also substantial evidence in favor of the notion that SARS-CoV-2 entry requires the participation of cholesterol [13, 14]. In

this regard, Tang et al. [15] proposed that rather than simply organizing ACE2 and SARS-CoV-2 proteins into cholesterol-rich Lo lipid domains, cholesterol might be directly involved in membrane fusion dynamics, guaranteeing the formation of the fusion intermediate. Meher et al. [16] and Pattnaik et al. [17] demonstrated that membrane cholesterol is vital for the fusogenic activity of SARS-CoVs viruses. Nardacci et al. [18] reported that the accumulation of lipids in SARS-CoV-2 infected cells, both in vitro and in the lungs of patients, could be involved in SARS-CoV-2 pathogenesis. Furthermore, membrane cholesterol depletion of ACE2-expressing HEK293T cells with methyl- β -cyclodextrin reduced SARS-CoV-2 infection, suggesting that cholesterol-rich lipid domains, as well as endosomal acidification, are essential requirements for SARS-CoV-2 infection [13]. Sanders et al. [14] showed that pre-treatment of SARS-CoV-2 with methyl- β -cyclodextrin blocked virus infection, further reinforcing the notion that cholesterol content in the viral particle is critical for infectivity.

Here we analyze the presence of cholesterol-recognition motifs in key regions of the spike glycoprotein in the pathogenic CoVs and disclose their possible pathophysiological implications in the virus-host cell interactions.

Material and Methods

Search for the presence and localization of the consensus sequence for the "CRAC" motif, [L/V]-[X](1-5)-[Y]-[X](1-5)-[R/K], or the "CARC" motif, [R/K]-[X](1-5)-[YFW]-[X](1-5)-[L/V], was carried out on the SARS-CoV-1 (Accession: AAU04646), SARS-CoV-2 (Accession: YP_009724390), MERS-CoV (Accession: K9N5Q8), and Omicron BA.1 (Accession: 7WP9_A) spike protein sequences, available at the National Center for Biotechnology Information (NCBI). To this end, we employed the Fuzzpro

application of the Jembooss software (European Molecular Biology Open Software Suite, EMBOSS) as in our original description of the consensus motifs [19]. Molecular representations were performed with Visual Molecular Dynamics (VMD) software, and the PDB files available in the protein data bank (PDB files 6XR8 and 7WP9, for SARS-CoV-2 and Omicron BA.1 spike proteins, respectively).

Results and Discussion

Consensus cholesterol-recognition sequences in the pathogenic CoVs

Cell-surface receptors for neurotransmitters have been instrumental in the discovery of linear amino acid sequences with the capacity to recognize cholesterol. Studies of the benzodiazepine receptor led to the identification of the sequence (L/V)-X-5-(Y)-X-5-(K/R), known as **Cholesterol Recognition/interaction Amino acid Consensus sequence (CRAC)** [20] (Fig. 1A). Work on the nicotinic acetylcholine receptor (nAChR) identified the inverse or mirror image of CRAC, which we coined “**CARC**”: (K/R)-X1-5-(Y/F/W)-X1-5-(L/V) [19] (Fig. 1A). In subsequent work, we found that these two cholesterol consensus motifs are present in a great variety of membrane proteins [21], including the superfamily of ligand-gated ion-channel (LGIC) proteins [21], other channels like the transient receptor potential (TRP) channel [22], and in the large superfamily of membrane-bound GPCR proteins [21]. CRAC/CARC sequences share a central aromatic residue, like tyrosine for CRAC, and tyrosine, phenylalanine, or tryptophan for CARC, flanked on both sides by one to five amino acid residues ending in a basic (arginine or lysine) and an apolar terminus (valine or leucine) (Fig. 1A). Both CARC and CRAC are vectorial motifs, with CARC preferentially, albeit not

exclusively, located in the exofacial membrane leaflet, whereas CRAC most often occurs in the cytoplasmic-facing leaflet of the plasma membrane [21].

The three pathogenic CoVs, namely the severe acute respiratory syndrome (SARS-CoV), the virus that gave rise to the epidemic at the brink of the 21st century, followed by the Middle-East respiratory syndrome CoV (MERS-CoV), and the contemporary SARS-CoV-2 associated with the pandemic, possess cholesterol-recognition motifs (Fig. 1A) in various regions of the virion, and most prominently in the spike S glycoprotein. Several cholesterol-recognition motifs occur in the transmembrane (TM) domain (Fig 1), in the fusion peptide (FP, Fig. 2), in the amino-terminal domain (NTD, Fig. 3), and the receptor binding domain (RBD, Fig. 4) of SARS-CoV-2 S protein. Fig. 1B depicts the SARS-CoV-2 spike (S) glycoprotein sequence, where the different functionally relevant regions of the virus can be identified, starting from the signal sequence (SS) in the S1 subunit, which also harbors the NTD and the important RBD. The key role of the S1 subunit is to prevent the praecox fusion of the S2 subunit to the target plasma membrane in the host cell and to bind to the mammalian cell-surface receptor, the metalloprotease angiotensin-converting enzyme-2 (ACE2). This step involves the proteolytic cleavage that activates the S protein by the host cell plasma membrane-resident transmembrane protease TMPRSS2 [23], a requisite for the conformational change of the S2 subunit [24] from a structurally disordered state into a wedge-shaped conformer [25] apt to undergo the fusion of the viral and plasma membranes via its FP.

CARC/CRAC cholesterol recognition motifs in the S protein transmembrane (TM) domain

The TM domain of the S protein anchors the glycoprotein in the virus envelope (Fig. 1B-C). It consists of three portions: a juxtamembrane aromatic region, a central hydrophobic region, and a cysteine-rich region [26]. As shown in Fig. 1C, the membrane-embedded regions of the S protein of the three pathogenic human CoVs share essentially the same cholesterol-recognition motifs, with a high degree of amino acid homology, especially in the two SARS virus. The CARC motif is deeper in the aromatic amino acid-rich part of the TM, with a CRAC laying immediately adjacent to it, partly embedded in the juxtamembrane portion, in a tail-to-tail disposition. We originally described this CARC-CRAC “back-to-back (or tail-to-tail)” mirror-image configuration in members of the GPCR superfamily [21]. The sequence homology observed in the TMs shown in Fig. 1C (and fusion peptide, see below), attests to the possible functional relevance of the cholesterol-recognition motifs in CoVs [26]. The interaction between cholesterol and SARS-CoV-2 TM regions could be impaired after pre-treatment of SARS-CoV-2 with the cholesterol-depleting compound methyl- β -cyclodextrin. This cholesterol-sequestering organic compound induces a conformational change in the viral particle that could be responsible, at least in part, for the reduced virus infectivity reported by Sanders et al. [14]. Wei and coworkers recently reported the occurrence of six such tail-to-tail cholesterol mirror motifs in the SARS-CoV-2 S protein and experimentally demonstrated, using a microscale thermophoresis assay, that the S protein binds cholesterol with a half-maximum inhibitory concentration IC_{50} of ~195 nM, but does not bind the cholesterol analogs campesterol or epicholesterol [27]. Contrary to expectations, the TM segment (residues 1203-1218) used in the article of Wei et al. [27] did not show interactions with

cholesterol. A preliminary study using atomistic molecular dynamics simulations found that cholesterol preferred the C-terminal half of the TM segment instead of binding to the CRAC/CARC motifs in the TM region [28]. Concerning SARS-CoV 2 S protein TM segment-cholesterol interactions, the differences between live cells experiments and in silico analyses or synthetic TM-segments-cholesterol interactions may be due to the role played by the rest of the S protein in live cells, acting as a final translator that defines virus infectivity.

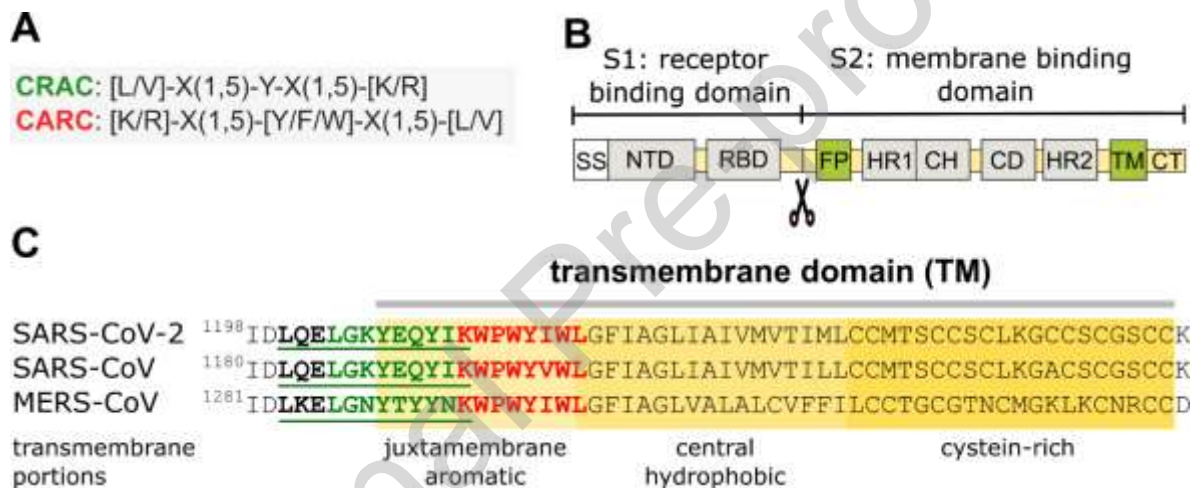


Figure 1. Cholesterol-recognition motifs in the TM region of spike S glycoprotein from SARS-CoV-2, SARS-CoV, and MERS-CoV. A) Sequences of the cholesterol-recognition motifs. B) Schematic diagram of SARS-CoV-2 primary structure. SS, signal sequence; NTD, N-terminal domain; RBD, receptor binding domain; FP, fusion peptide; HR1, heptad repeat 1; CH, central helix; CD, connector domain; HR2, heptad repeat 2; TM, transmembrane domain; CT, cytoplasmic tail. C) Cholesterol-recognition motifs along the linear sequences of the TMs. The TM region exhibits back-to-back mirror images of highly conserved CRAC (green) and CARC (red) motifs at the N-term juxtamembrane aromatic region.

Cholesterol recognition motifs in the fusion peptide (FP) of the S glycoprotein

For enveloped viruses, the release of their genome into the host cell requires the fusion of their membrane bilayer to either the plasma membrane or the endocytic vesicle membranes of the target cell. The FP is a short segment (~28 amino acid residues long)

of the S2 domain that constitutes the functional fusogenic element in SARS-CoV-2 needed for direct fusion to the host lipid membrane [15] (Figs. 2A). Dacon et al. [29] have recently highlighted the potential of the FP as a target epitope to design next-generation CoV vaccines. Cholesterol affects the extent of SARS-CoV-2 binding and fusion to cellular membranes [16, 17, 30, 31]. In Figs. 2A-B, the SARS-CoV-2 S protein FP is highlighted in red in the pre-fusion conformation of the S glycoprotein. The fusion peptide is precluded from fusing to the target cell, well hidden in the S protein core.

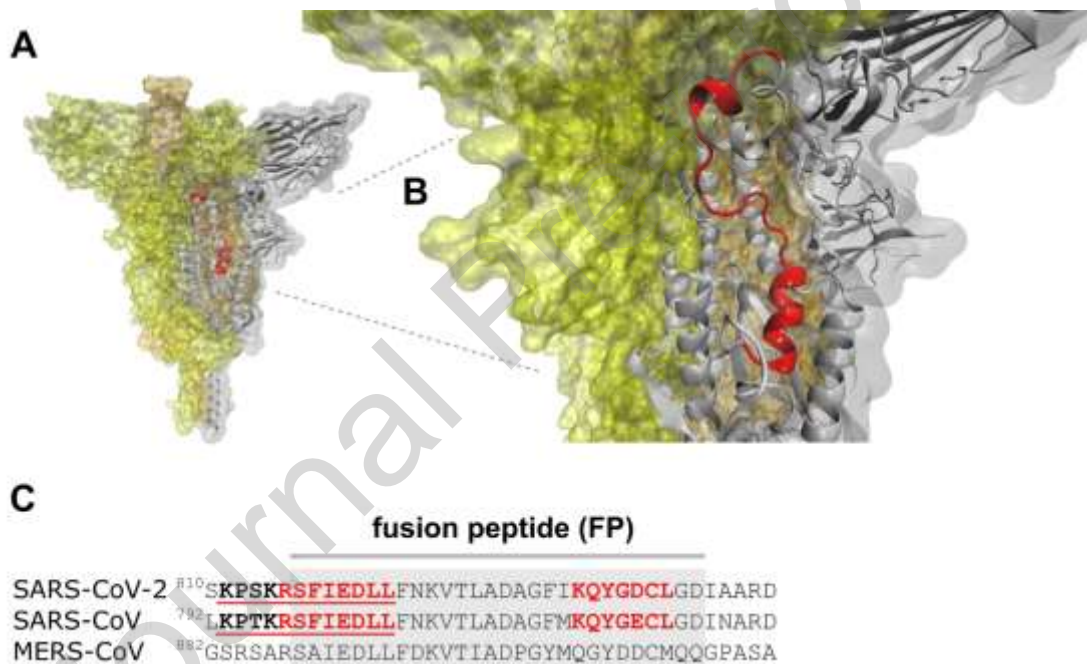


Figure 2. Cholesterol-recognition motifs in the fusion peptide of the spike S glycoprotein from the SARS-CoV-2, SARS-CoV, and MERS-CoV, the three highly pathogenic members of the heptad human CoVs. A) FP segment (red ribbon) in the viral envelope of the SARS-CoV-2 S glycoprotein B) Close-up view of the FP in the SARS-CoV-2 S protein. C) Cholesterol-recognition motifs along the linear sequences of the FPs.

The cholesterol-recognition motifs CARC and its mirror image CRAC of the three pathogenic viruses are shown in Fig. 2C. Interestingly, the S glycoprotein FP region contains a conserved CARC motif in the FP N-term (Fig. 2C), and an almost identical

CARC motif (KQYG[D/E]CL) in SARS-CoV and SARS-CoV-2 in their C-term region. Conversely, CARC/CRAC motifs are absent in the FP C-term of another pathogenic CoV, MERS-CoV (Fig. 2C). Madu et al. [32] reported that the SARS-CoV S2 fusion peptide ⁷⁹⁸*SFIEDLLFNKVTLADAGFMKQY*⁸¹⁹*GCGKKKK* (*linker*), which includes ~90% of the N-term, and ~40% of the C-term CARC motifs, respectively, promoted a greater extent of lipid mixing in POPC-POPS-cholesterol (1:3:1) liposomes. Mahajan et al. [31], demonstrated that a 64-residue long fusion peptide (LFP) [31, 33], ⁷⁵⁸*RNTREVFAQVKQMYKTPTLKYFGGFNFSQILPSPLKPTKRSFIEDLLFNKVTLADAGFMKQYGE*⁸²¹, which includes 7 CARC-motifs and a CRAC motif, exhibits a dose-dependent lipid mixing activity in DMPC liposomes, highlighting its fusogenic potential. An NMR spectroscopy study has recently shown that the FP transforms from an intrinsically disordered structure in solution into a wedge-shaped structure in lipid bicelles, with the hydrophobic, narrow end inserted in the lipid moiety [34]. The FP region that interacts more strongly with bicelles contains the CARC regions shown in Fig. 2C. Santamarina et al. [35] reported that an FP fragment containing the first CARC motif (Fig. 2C) penetrates into the hydrophobic acyl region of the host cell plasma membrane, increasing the dynamics of the fatty acid acyl tails and weakening the membrane structural integrity, possibly a requisite towards viral penetration. Importantly, the increase in membrane flexibility is more pronounced in the more rigid, Lo plasma membrane regions rich in cholesterol [35]. Shen et al. [36] have recently compared the binding modes of SARS-CoV and SARS-CoV-2 FPs using in silico molecular dynamics. They showed that SARS-CoV-2 FP binds to a synthetic POPC/POPE/cholesterol bilayer membrane more effectively than the SARS-CoV FP [36]. Although the amino acid sequences of both FPs are quite similar (Fig. 2C), the

corresponding cryo-electron microscopy structures show that the helix length of SARS-CoV FP is longer than that of the SARS-CoV-2 FP [36]. The higher hydrophobicity of the SARS-CoV-2 FP may also contribute to disrupting the target cell membrane [37]. All in all, these structural differences could be responsible for the differential affinities between SARS-CoV and SARS-CoV-2 FPs [36]. The amino acids F817, I818, and L821 present in the short α -helix of the SARS-CoV-2 FP exhibited a stronger interaction with the cell membrane than those of the long α -helix from SARS-CoV [36]. The three hydrophobic amino acid residues involved are part of the N-term CARC motif present in the SARS-CoV-2 FP (Fig. 2C) and are an indispensable requirement for cholesterol binding in the CARC motif [19, 38]. In addition, SARS-CoV FP has two favorable membrane-binding modes, which overlap with the presence of CARC motifs in the N-term and C-term portions of the PF [36] (Fig. 2C). The presence of CARC motifs in the SARS-CoV and SARS-CoV-2 FP could be indicative of an absolute requirement for cholesterol-protein interactions during the early stages of the fusion of the viral particle with the host-cell plasma membrane. The occurrence of a cholesterol-recognition motif in a key region like the FP suggests a relevant biosensor role, guiding the S glycoprotein to cholesterol-rich domains in the host plasma membrane, and improving the efficiency of the virus infectivity.

Cholesterol recognition motifs in the NTD of the S glycoprotein

Scavenger receptor class B type 1 (SR-B1) is a multifunctional membrane-bound protein mainly expressed in liver and one of the heavy-density lipoprotein (HDL) receptors [3]. Wei and coworkers [27] demonstrated that expression of SR-B1 confers susceptibility to SARS-CoV-2 infection, facilitating the attachment and entry of the virus. Since SR-B1 only enhanced viral uptake in the presence of ACE2, the authors

interpreted these last results as an indication that SR-B1 is an entry cofactor of SARS-CoV-2 (reviewed in [3]). Inhibitors of SR-B1 or silencing of its expression abrogates SARS-CoV-2 infection, and the presence of HDL significantly increases viral infection. Blocking the cholesterol/HDL binding site of SARS-CoV-2 with mAb 1D2 strongly reduced HDL-enhanced SARS-CoV-2 infection [27]. The mAb 1D2 overlaps antigenic sites with mAb 4A8, a neutralizing human antibody that recognizes the NTD of the S protein [39] (Fig. 1A). Utilizing cryo-electron microscopy, Chi et al. [39] determined that mAb 4A8 mainly binds to the NTD through three complementarity-determining regions (CDRs). The amino acids Lys147, Lys150, and Tyr145, located in a CRAC motif of NTD-SARS-2-S (Fig. 3A), were identified as important antigenic sites for recognition by mAb 4A8 antibody [27, 39]. Through an in vitro binding assay, Wei et al. [27] found that three cholesterol-binding peptides encompassing the CRAC/CARC region are present in the NTD of SARS-CoV-2: amino acid residues 24-32, 129-150, and 267-277, respectively [27] (Figs. 3A-B). Interestingly, the peptide region that displays the highest interaction with cholesterol is the segment 129-150 [27]. This segment presents two cholesterol recognition motifs along its sequence, i.e., a CARC motif $^{129}\text{KVCEFQFCNDPFL}^{141}$, and a CRAC motif $^{141}\text{LGVYYHKNNK}^{150}$ (Fig. 3 C-D; CARC¹²⁹⁻¹⁴¹ in yellow, CRAC¹⁴¹⁻¹⁵⁰ in orange). According to Wei et al. [27], SARS-CoV-2 first binds cholesterol and high-density lipoprotein (HDL) (or one of its components). This complex is then recognized by SR-B1 as an entry cofactor, helping to increase the complex concentration in host-cell membrane regions where ACE2 is expressed and facilitating successful encounters with the receptor [3, 27].

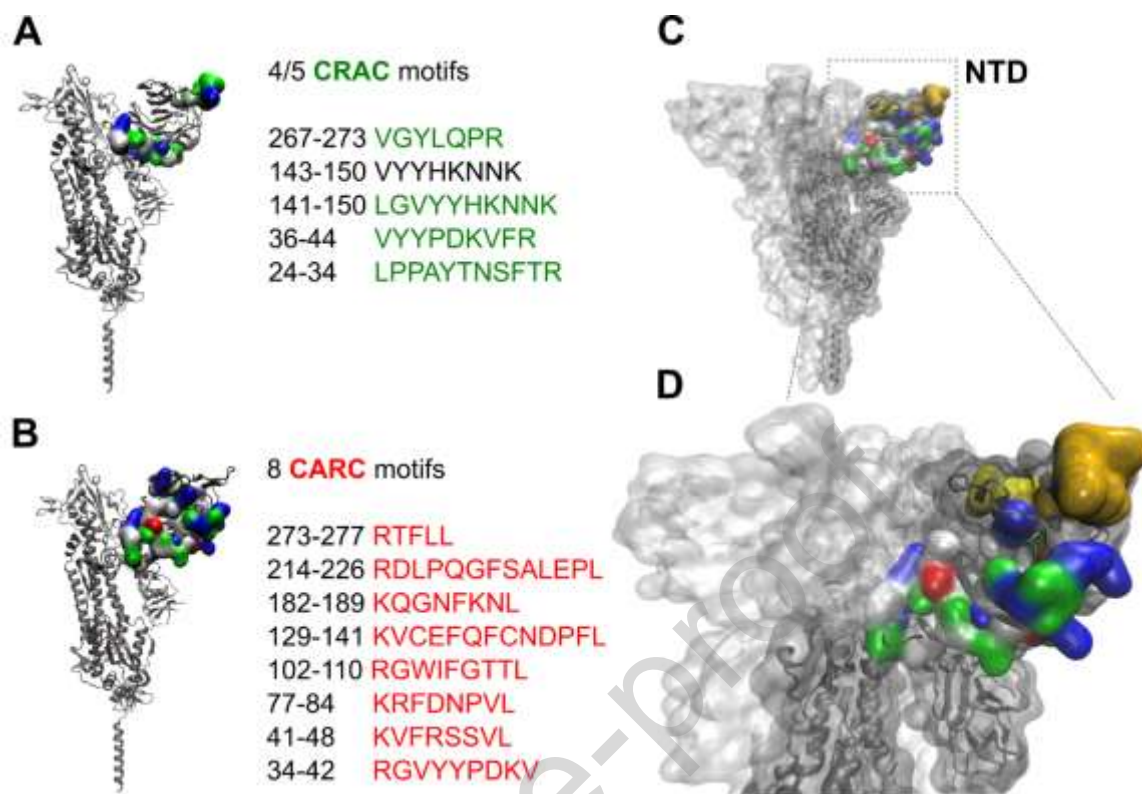


Figure 3. Cholesterol-recognition motifs in the N-terminal domain (NTD) region of the spike S glycoprotein from SARS-CoV-2. Sequences of the cholesterol-recognition motifs, A) CRAC and b) CARC, localized in the NTD region of SARS-CoV-2 spike protein. C) CARC and CARC motifs are colored in the NTD segment. The yellow and orange segments represent the CARC129-141 and CRAC141-150 motifs, respectively. D) Close-up view of the segments shown in C).

It is intriguing how a simple lipid molecule like cholesterol can constitute a molecular target and subsequently bind to a protein motif located outside the plasma membrane. There is, however, a precedent of this molecular interaction: in the Smoothed (SMO) protein, a G protein-coupled receptor, cholesterol binds to the extracellular cysteine-rich domain, located outside the extracellular leaflet of the plasma membrane [40]. Fantini et al. [41, 42] described a ganglioside-binding domain in the NTD, which could enable the binding of SARS-CoV-2 to the plasma membrane-lipid L_0 domains. Each spike protein can simultaneously bind three ganglioside molecules. This multivalent binding process could modify the membrane curvature,

facilitate lipid coalescence and ACE2 receptor recruitment, thus increasing the chances of finding functional ACE2 receptor molecules in the host membrane [41].

Several mutations in the spike protein contribute to the increased transmissibility of some SARS-CoV-2 variants. The most accepted explanation is that mutations conferring higher affinity of the spike protein for the ACE2 receptor result in more transmissible virus variants. This does not necessarily account for the success of the Omicron variant, suggesting that spike ectodomains directly interact with host plasma membrane bilayers [43]. Overduin et al. [43] recently undertook a detailed *in silico* study of the membrane-binding surfaces of the spike protein variants, showing that membrane binding propensities increase through different versions, over time, impacting the S protein's affinity for cell membranes. Spike protein trimers are shown to shift from initial perpendicular stances to increasingly tilted positions that draw viral particles alongside host cell membranes before engaging ACE2 receptors. This culminates in the assembly of the fusion apparatus, enhancing membrane interactions of the more infective variants [43]. The flexibility of the S glycoprotein has been documented experimentally by cryo-electron microscopy [44] and by coarse-grained molecular dynamics studies showing that the protein can lay almost parallel to the membrane surface [45]. In an end-on view of the spike protein ectodomain, as seen from the host cell (Figure 4), the area of CRAC and CARC exposed motifs in the NTD and RBD increase from the wild-type Wuhan SARS-CoV-2 to the highly infective Omicron BA.1 variant. The RBD segment 445-461 from SARS-CoV-2 [27], which contains in its sequence the CRAC⁴⁴²⁻⁴⁵⁴ /CARC⁴⁴¹⁻⁴⁵² motifs exposed in the spike protein surface, has been experimentally shown to interact with cholesterol [27] (Figure 4). Even though the segments shown in Fig. 4 are conserved in both SARS-CoV-2 444-461 and Omicron BA.1 441-458 virus variants, differences in the protein conformation

may result in a better CRAC/CARC surface exposure in the Omicron variant. Higher exposure of CRAC/CARC motifs on the virus S protein binding surface could contribute to the enhanced ability of the spike protein to interact with the host plasma membrane, and consequently, augment virus infectivity.

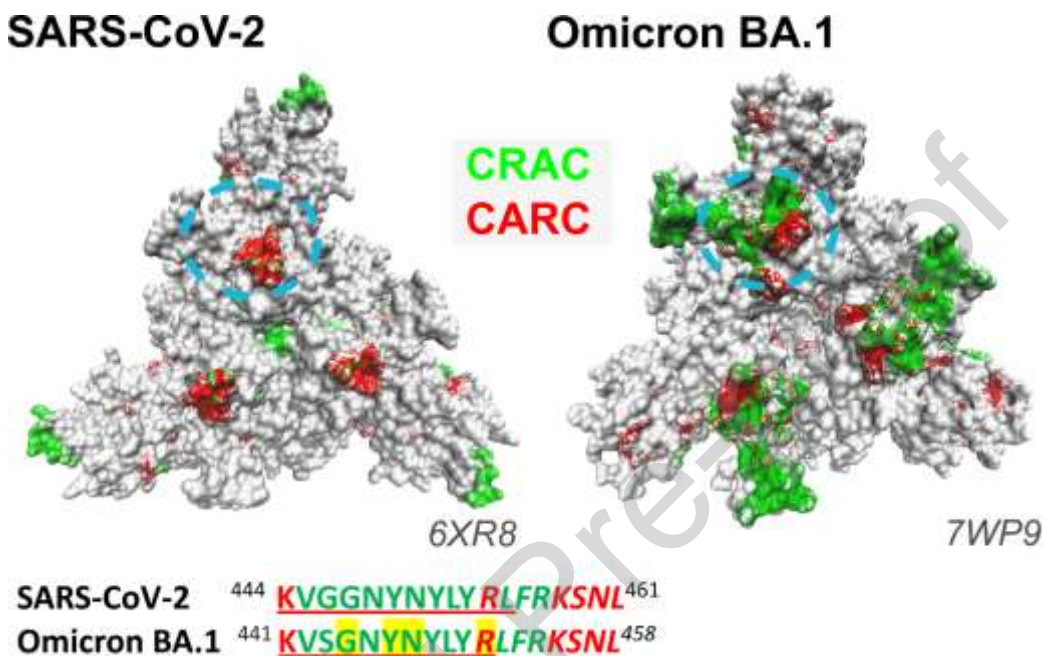


Figure 4. Top view of SARS-CoV-2, and Omicron BA.1 spike protein. CRAC and CARC motifs are colored green and red, respectively. The approximate position of the sequences SARS-CoV-2 444-461, and Omicron BA.1 441-458 are highlighted by a dotted light-blue circle. Letters highlighted in yellow correspond to plasma membrane-interacting residues proposed by Overduin et al. [43].

Conclusions

From the data presented here, we propose two functional attributes of the cholesterol-recognition motifs in the S glycoprotein of the SARS-CoV-2: 1) the presence of CARC and CRAC motifs in the surface of the S trimer (NTD and RBD) that docks onto the host membrane surface is indicative of actual S protein-cholesterol interactions occurring at the earliest stage of virus infection, that is, during the landing of the virion on the host cell surface; 2) the presence of cholesterol-recognition motifs

in the FP segment strongly suggests their involvement in the subsequent fusion step. Thus, the presence of the cholesterol consensus regions provides the structural basis for the sequential i) reduction of dimensionality of the virion binding step, i.e., from the random walk in the 3D space to the 2D of the membrane surface, ii) the accelerated diffusion of the S trimer in the target plasma membrane to iii) dock in cholesterol-rich domains containing the ACE2 receptor. The combination of these cholesterol-dependent steps appears to bear direct relevance to the higher infectivity of the Omicron SARS-CoV-2 variant.

Funding:

Project PIP 222021-2023 GI to FJB.

Author contributions:

Conceptualization: FJB, CJB; Methodology: FJB, CJB; Investigation: FJB, CJB;

Visualization: CJB; Writing: original draft: FJB; review & editing: FJB, CJB.

Competing interests:

Authors declare that they have no competing interests.

Data and materials availability:

All data are available in the main text.

References

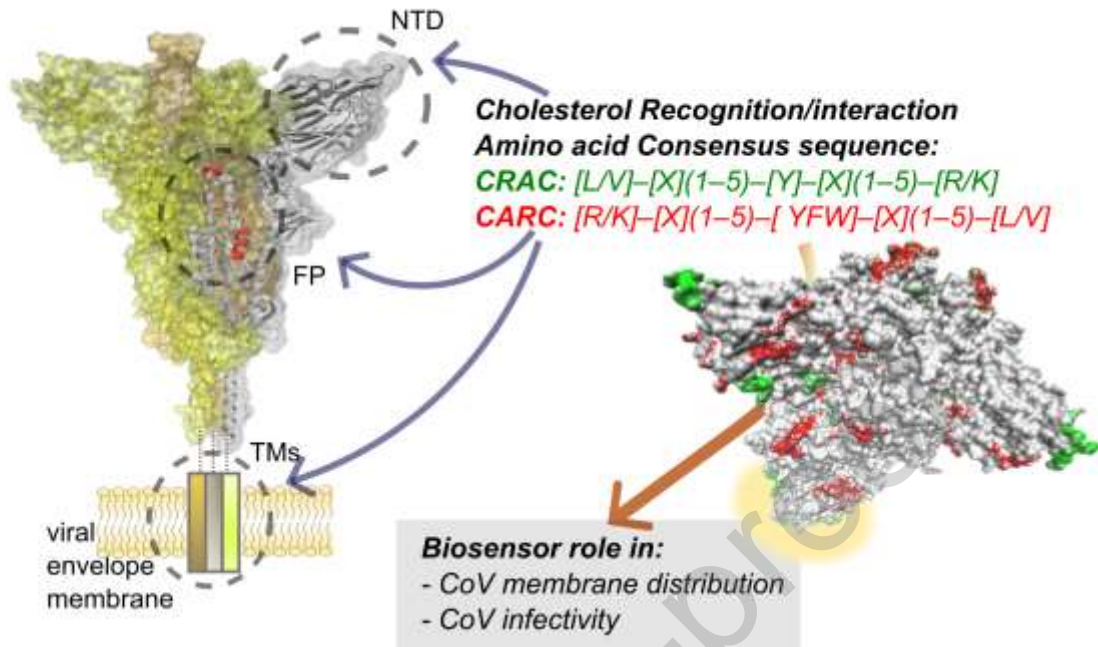
- [1] F.J. Barrantes, The Contribution of Biophysics and Structural Biology to Current Advances in COVID-19, *Annu Rev Biophys*, 50 (2021) 493-523.
- [2] F.J. Barrantes, The unfolding palette of COVID-19 multisystemic syndrome and its neurological manifestations, *Brain Behav Immun Health*, 14 (2021) 100251.
- [3] F.J. Barrantes, The constellation of cholesterol-dependent processes associated with SARS-CoV-2 infection, *Prog Lipid Res*, 87 (2022) 101166.
- [4] F.R. Maxfield, G. van Meer, Cholesterol, the central lipid of mammalian cells, *Curr Opin Cell Biol*, 22 (2010) 422-429.
- [5] C. Schroeder, Cholesterol-binding viral proteins in virus entry and morphogenesis, *Subcell Biochem*, 51 (2010) 77-108.
- [6] I.P. Sousa, Jr., C.A.M. Carvalho, A.M.O. Gomes, Current Understanding of the Role of Cholesterol in the Life Cycle of Alphaviruses, *Viruses*, 13 (2020).
- [7] I.P. Sousa, Jr., C.A. Carvalho, D.F. Ferreira, G. Weissmuller, G.M. Rocha, J.L. Silva, A.M. Gomes, Envelope lipid-packing as a critical factor for the biological activity and stability of alphavirus particles isolated from mammalian and mosquito cells, *The Journal of biological chemistry*, 286 (2011) 1730-1736.
- [8] S. Barman, D.P. Nayak, Lipid raft disruption by cholesterol depletion enhances influenza A virus budding from MDCK cells, *Journal of virology*, 81 (2007) 12169-12178.
- [9] S. Bajimaya, T. Frankl, T. Hayashi, T. Takimoto, Cholesterol is required for stability and infectivity of influenza A and respiratory syncytial viruses, *Virology*, 510 (2017) 234-241.

- [10] J.A. Nieto-Garai, A. Arboleya, S. Otaegi, J. Chojnacki, J. Casas, G. Fabrias, F.X. Contreras, H.G. Krausslich, M. Lorizate, Cholesterol in the Viral Membrane is a Molecular Switch Governing HIV-1 Env Clustering, *Adv Sci (Weinh)*, 8 (2021) 2003468.
- [11] R.F. Epand, A. Thomas, R. Brasseur, S.A. Vishwanathan, E. Hunter, R.M. Epand, Juxtamembrane protein segments that contribute to recruitment of cholesterol into domains, *Biochemistry*, 45 (2006) 6105-6114.
- [12] J. Glende, C. Schwegmann-Wessels, M. Al-Falah, S. Pfefferle, X. Qu, H. Deng, C. Drosten, H.Y. Naim, G. Herrler, Importance of cholesterol-rich membrane microdomains in the interaction of the S protein of SARS-coronavirus with the cellular receptor angiotensin-converting enzyme 2, *Virology*, 381 (2008) 215-221.
- [13] X. Li, W. Zhu, M. Fan, J. Zhang, Y. Peng, F. Huang, N. Wang, L. He, L. Zhang, R. Holmdahl, L. Meng, S. Lu, Dependence of SARS-CoV-2 infection on cholesterol-rich lipid raft and endosomal acidification, *Comput Struct Biotechnol J*, 19 (2021) 1933-1943.
- [14] D.W. Sanders, C.C. Jumper, P.J. Ackerman, D. Bracha, A. Donlic, H. Kim, D. Kenney, I. Castello-Serrano, S. Suzuki, T. Tamura, A.H. Tavares, M. Saeed, A.S. Holehouse, A. Ploss, I. Levental, F. Douam, R.F. Padera, B.D. Levy, C.P. Brangwynne, SARS-CoV-2 requires cholesterol for viral entry and pathological syncytia formation, *Elife*, 10:e65962 (2021).
- [15] T. Tang, M. Bidon, J.A. Jaimes, G.R. Whittaker, S. Daniel, Coronavirus membrane fusion mechanism offers a potential target for antiviral development, *Antiviral Res*, 178 (2020) 104792.
- [16] G. Meher, S. Bhattacharjya, H. Chakraborty, Membrane Cholesterol Modulates Oligomeric Status and Peptide-Membrane Interaction of Severe Acute Respiratory Syndrome Coronavirus Fusion Peptide, *J Phys Chem B*, 123 (2019) 10654-10662.
- [17] G.P. Pattnaik, S. Bhattacharjya, H. Chakraborty, Enhanced Cholesterol-Dependent Hemifusion by Internal Fusion Peptide 1 of SARS Coronavirus-2 Compared to Its N-Terminal Counterpart, *Biochemistry*, 60 (2021) 559-562.
- [18] R. Nardacci, F. Colavita, C. Castilletti, D. Lapa, G. Matusali, S. Meschi, F. Del Nonno, D. Colombo, M.R. Capobianchi, A. Zumla, G. Ippolito, M. Piacentini, L. Falasca, Evidences for lipid involvement in SARS-CoV-2 cytopathogenesis, *Cell Death Dis*, 12 (2021) 263.
- [19] C.J. Baier, J. Fantini, F.J. Barrantes, Disclosure of cholesterol recognition motifs in transmembrane domains of the human nicotinic acetylcholine receptor, *Scientific reports*, 1 (2011) 69.
- [20] H. Li, V. Papadopoulos, Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern, *Endocrinology*, 139 (1998) 4991-4997.
- [21] J. Fantini, R.M. Epand, F.J. Barrantes, Cholesterol-Recognition Motifs in Membrane Proteins, *Adv Exp Med Biol*, 1135 (2019) 3-25.
- [22] A.G. Lee, Interfacial Binding Sites for Cholesterol on TRP Ion Channels, *Biophysical Journal*, 117 (2019) 2020-2033.
- [23] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Müller, C. Drosten, S. Pöhlmann, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor, *Cell*, 181 (2020) 271-280.e278.
- [24] Y. Cai, J. Zhang, T. Xiao, H. Peng, S.M. Sterling, R.M. Walsh, S. Rawson, S. Rits-Volloch, B. Chen, Distinct conformational states of SARS-CoV-2 spike protein, *Science (New York, N.Y.)*, 369 (2020) 1586-1592.
- [25] R.K. Koppiseti, Y.G. Fulcher, S.R. Van Doren, Fusion Peptide of SARS-CoV-2 Spike Rearranges into a Wedge Inserted in Bilayered Micelles, *Journal of the American Chemical Society*, (2021).
- [26] X. Xia, Domains and Functions of Spike Protein in SARS-Cov-2 in the Context of Vaccine Design, 13 (2021) 109.
- [27] C. Wei, L. Wan, Q. Yan, X. Wang, J. Zhang, X. Yang, Y. Zhang, C. Fan, D. Li, Y. Deng, J. Sun, J. Gong, X. Yang, Y. Wang, X. Wang, J. Li, H. Yang, H. Li, Z. Zhang, R. Wang, P. Du, Y. Zong, F. Yin, W. Zhang, N. Wang, Y. Peng, H. Lin, J. Feng, C. Qin, W. Chen, Q. Gao, R. Zhang, Y. Cao, H. Zhong, HDL-scavenger receptor B type 1 facilitates SARS-CoV-2 entry, *Nature metabolism*, 2 (2020) 1391-1400.
- [28] S. Lall, P. Balaram, S. Gosavi, M.K. Mathew, Dynamics and self-assembly of the SARS-CoV-2 spike transmembrane domain, (2021) 2021.2006.2007.447334.
- [29] C. Dacon, C. Tucker, L. Peng, C.D. Lee, T.H. Lin, M. Yuan, Y. Cong, L. Wang, L. Purser, J.K. Williams, C.W. Pyo, I. Kosik, Z. Hu, M. Zhao, D. Mohan, A.J.R. Cooper, M. Peterson, J. Skinner, S. Dixit, E. Kollins, L. Huzella, D. Perry, R. Byrum, S. Lembirik, D. Drawbaugh, B. Eaton, Y. Zhang, E.S. Yang, M. Chen, K. Leung, R.S. Weinberg, A. Pegu, D.E. Geraghty, E. Davidson, I. Douagi, S. Moir, J.W. Yewdell, C. Schmaljohn, P.D. Crompton, M.R. Holbrook, D. Nemazee, J.R. Mascola, I.A. Wilson, J. Tan, Broadly neutralizing antibodies target the coronavirus fusion peptide, *Science*, (2022) eabq3773.
- [30] Y. Correa, S. Waldie, M. Thepaut, S. Micciula, M. Moulin, F. Fieschi, H. Pichler, V. Trevor Forsyth, M. Haertlein, M. Cardenas, SARS-CoV-2 spike protein removes lipids from model membranes and interferes with the capacity of high density lipoprotein to exchange lipids, *J Colloid Interface Sci*, 602 (2021) 732-739.

- [31] M. Mahajan, D. Chatterjee, K. Bhuvaneshwari, S. Pillay, S. Bhattacharjya, NMR structure and localization of a large fragment of the SARS-CoV fusion protein: Implications in viral cell fusion, *Biochim Biophys Acta Biomembr*, 1860 (2018) 407-415.
- [32] I.G. Madu, S.L. Roth, S. Belouzard, G.R. Whittaker, Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide, *Journal of virology*, 83 (2009) 7411-7421.
- [33] M. Mahajan, S. Bhattacharjya, NMR structures and localization of the potential fusion peptides and the pre-transmembrane region of SARS-CoV: Implications in membrane fusion, *Biochim Biophys Acta*, 1848 (2015) 721-730.
- [34] R.K. Koppiseti, Y.G. Fulcher, S.R. Van Doren, Fusion Peptide of SARS-CoV-2 Spike Rearranges into a Wedge Inserted in Bilayered Micelles, *J Am Chem Soc*, 143 (2021) 13205-13211.
- [35] A. Santamaria, K.C. Batchu, O. Matsarskaia, S.F. Prevost, D. Russo, F. Natali, T. Seydel, I. Hoffmann, V. Laux, M. Haertlein, T.A. Darwish, R.A. Russell, G. Corucci, G. Fragneto, A. Maestro, N.R. Zaccari, Strikingly Different Roles of SARS-CoV-2 Fusion Peptides Uncovered by Neutron Scattering, *J Am Chem Soc*, 144 (2022) 2968-2979.
- [36] H. Shen, Z. Wu, L. Chen, Different Binding Modes of SARS-CoV-1 and SARS-CoV-2 Fusion Peptides to Cell Membranes: The Influence of Peptide Helix Length, *J Phys Chem B*, (2022).
- [37] A.L. Lai, J.H. Freed, SARS-CoV-2 Fusion Peptide has a Greater Membrane Perturbing Effect than SARS-CoV with Highly Specific Dependence on Ca²⁺, *Journal of Molecular Biology*, 433 (2021) 166946.
- [38] C. Di Scala, C.J. Baier, L.S. Evans, P.T.F. Williamson, J. Fantini, F.J. Barrantes, Relevance of CARC and CRAC Cholesterol-Recognition Motifs in the Nicotinic Acetylcholine Receptor and Other Membrane-Bound Receptors, *Current topics in membranes*, 80 (2017) 3-23.
- [39] X. Chi, R. Yan, J. Zhang, G. Zhang, Y. Zhang, M. Hao, Z. Zhang, P. Fan, Y. Dong, Y. Yang, Z. Chen, Y. Guo, J. Zhang, Y. Li, X. Song, Y. Chen, L. Xia, L. Fu, L. Hou, J. Xu, C. Yu, J. Li, Q. Zhou, W. Chen, A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2, *Science*, 369 (2020) 650-655.
- [40] M. Kinnebrew, R.E. Woolley, T.B. Ansell, E.F.X. Byrne, S. Frigui, G. Luchetti, R. Sircar, S. Nachtergaele, L. Mydock-McGrane, K. Krishnan, S. Newstead, M.S.P. Sansom, D.F. Covey, C. Siebold, R. Rohatgi, Patched 1 regulates Smoothed by controlling sterol binding to its extracellular cysteine-rich domain, *Sci Adv*, 8 (2022) eabm5563.
- [41] J. Fantini, H. Chahinian, N. Yahi, Leveraging coronavirus binding to gangliosides for innovative vaccine and therapeutic strategies against COVID-19, *Biochemical and biophysical research communications*, 538 (2021) 132-136.
- [42] J. Fantini, H. Chahinian, N. Yahi, Synergistic antiviral effect of hydroxychloroquine and azithromycin in combination against SARS-CoV-2: What molecular dynamics studies of virus-host interactions reveal, *Int J Antimicrob Agents*, 56 (2020) 106020.
- [43] M. Overduin, T.A. Kervin, A. Tran, Progressive membrane-binding mechanism of SARS-CoV-2 variant spike proteins, *iScience*, 25 (2022) 104722.
- [44] B. Turoňová, M. Sikora, C. Schürmann, W.J.H. Hagen, S. Welsch, F.E.C. Blanc, S. von Bülow, M. Gecht, K. Bagola, C. Hörner, G. van Zandbergen, J. Landry, N.T.D. de Azevedo, S. Mosalaganti, A. Schwarz, R. Covino, M.D. Mühlebach, G. Hummer, J. Krijnse Locker, M. Beck, In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges, (2020) eabd5223.
- [45] B. Wang, C. Zhong, D.P. Tieleman, Supramolecular Organization of SARS-CoV and SARS-CoV-2 Virions Revealed by Coarse-Grained Models of Intact Virus Envelopes, *Journal of Chemical Information and Modeling*, 62 (2022) 176-186.

Graphical abstract

SARS-CoV-2 S glycoprotein



Journal Pre-proof

Credit author statement

CJB and FJB contributed equally in data curation; FJB conceptualization and writing the initial draft; both authors wrote/edited subsequent multiple versions of the manuscript. CJB conducted analysis of the data and produced all illustrations.

Journal Pre-proof

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Journal Pre-proof

Highlights

- Cholesterol-recognition sequences are present in pathogenic coronaviruses (CoVs).
- These motifs are localized in key regions of the spike glycoprotein in CoVs.
- Cholesterol motifs may be implicated in the varying infectivity of CoV variants.

Journal Pre-proof