



OPEN

Interaction of Spike protein and lipid membrane of SARS-CoV-2 with Ursodeoxycholic acid, an in-silico analysis

Francisco Javier Rodal Canales¹, Laura Pérez-Campos Mayoral¹, María Teresa Hernández-Huerta², Luis Manuel Sánchez Navarro¹, Carlos Alberto Matias-Cervantes², Margarito Martínez Cruz³, Eli Cruz Parada³, Edgar Zenteno⁴, Edgar Gustavo Ramos-Martínez⁵, Eduardo Pérez-Campos Mayoral¹, Carlos Romero Díaz^{1,7}✉ & Eduardo Pérez-Campos^{3,6,7}✉

Numerous repositioned drugs have been sought to decrease the severity of SARS-CoV-2 infection. It is known that among its physicochemical properties, Ursodeoxycholic Acid (UDCA) has a reduction in surface tension and cholesterol solubilization, it has also been used to treat cholesterol gallstones and viral hepatitis. In this study, molecular docking was performed with the SARS-CoV-2 Spike protein and UDCA. In order to confirm this interaction, we used Molecular Dynamics (MD) in "SARS-CoV-2 Spike protein-UDCA". Using another system, we also simulated MD with six UDCA residues around the Spike protein at random, naming this "SARS-CoV-2 Spike protein-6UDCA". Finally, we evaluated the possible interaction between UDCA and different types of membranes, considering the possible membrane conformation of SARS-CoV-2, this was named "SARS-CoV-2 membrane-UDCA". In the "SARS-CoV-2 Spike protein-UDCA", we found that UDCA exhibits affinity towards the central region of the Spike protein structure of -386.35 kcal/mol, in a region with 3 alpha helices, which comprises residues from K986 to C1032 of each monomer. MD confirmed that UDCA remains attached and occasionally forms hydrogen bonds with residues R995 and T998. In the presence of UDCA, we observed that the distances between residues atoms OG1 and CG2 of T998 in the monomers A, B, and C in the prefusion state do not change and remain at 5.93 ± 0.62 and 7.78 ± 0.51 Å, respectively, compared to the post-fusion state. Next, in "SARS-CoV-2 Spike protein-6UDCA", the three UDCA showed affinity towards different regions of the Spike protein, but only one of them remained bound to the region between the region's heptad repeat 1 and heptad repeat 2 (HR1 and HR2) for 375 ps of the trajectory. The RMSD of monomer C was the smallest of the three monomers with a value of 2.89 ± 0.32 , likewise, the smallest RMSF was also of the monomer C (2.25 ± 0.56). In addition, in the simulation of "SARS-CoV-2 membrane-UDCA", UDCA had a higher affinity toward the virion-like membrane; where three of the four residues remained attached once they were close (5 Å, to the centre of mass) to the membrane by 30 ns. However, only one of them remained attached to the plasma-like membrane and this was in a cluster of cholesterol molecules. We have shown that UDCA interacts in two distinct regions of Spike protein sequences. In addition, UDCA tends to stay bound to the membrane, which could potentially reduce the internalization of SARS-CoV-2 in the host cell.

SARS-CoV-2 is a single-stranded RNA beta coronavirus, which causes COVID-19. The urgency of developing effective drugs for the COVID-19 treatment has been of the utmost importance¹. Great interest has been

¹Research Centre Faculty of Medicine UNAM-UABJO, Faculty of Medicine and Surgery, Autonomous University "Benito Juárez" of Oaxaca, 68020 Oaxaca, Mexico. ²CONACyT, Faculty of Medicine and Surgery, Autonomous University "Benito Juárez" of Oaxaca, 68020 Oaxaca, Mexico. ³National Technology of Mexico/IT Oaxaca, 68030 Oaxaca, Mexico. ⁴Faculty of Medicine, National Autonomous University of Mexico, 04360 Mexico City, Mexico. ⁵School of Sciences/Autonomous University "Benito Juárez" of Oaxaca, 68125 Oaxaca, Mexico. ⁶Clinical Pathology Laboratory, "Eduardo Pérez Ortega", 68000 Oaxaca, Mexico. ⁷These authors contributed equally: Carlos Romero Díaz and Eduardo Pérez-Campos. ✉email: carlos.rom.74he@gmail.com; perezcampos@prodigy.net.mx

generated to identify targets for drug repurposing by maps of proteins interactions with SARS-CoV-2², and many come from natural products of plant origin^{3–6}. According to the information released, as of November 5, 2021, SARS-CoV-2 has caused more than 248 million infections worldwide and more than 5 million deaths^{7,8}. COVID-19 manifests itself through heterogeneous clinical expression, with respiratory, cardiovascular, nervous system, liver and skin symptoms⁹.

SARS-CoV-2 is surrounded on average by 40 to 48 Spike (S) trimer proteins^{10,11}, these are heavily glycosylated and have three receptor-binding domains (RBDs), covered on top of the S trimer by the N-terminal domains (NTD)^{12,13}. These are mainly found in a closed prefusion conformation⁸. RBDs are exposed or move up the viral membrane in their closed conformation. The Spike protein, being connected to the membrane through a stalk region that has three flexible hinges, allows protein S to be dynamic⁵. Between the trimeric region and the virion membrane, there is a triple-stranded coiled-coil heptad repeat 2 (HR2) with three flexible hinges in the stem region, which can allow lateral displacement of the stalk^{8,12}. The S trimer of SARS-CoV-2 binds to its receptor ACE2 for subsequent membrane fusion and virus entry. Cleavage of protein S by furin and/or TMPRSS2 proteases is key to this binding¹⁴. The “fusion proteins” have an α -helical, coiled-coil structure, which requires proteolytic processing to generate a structure of fusion¹⁵.

Recently, it has been reported that SARS-CoV-2 budding is found in regions close to the endoplasmic reticulum and Golgi-like membrane arrangements¹⁰, due to the composition of the endoplasmic reticulum Golgi intermediate compartment cisternae (ERGIC), it has been proposed as the main composition of the virus membranes¹⁶. ERGIC is composed of approximately 58% phosphatidylcholine, 22% phosphatidylethanolamine, 3% phosphatidylserine, 10% phosphatidylinositol, 3% sphingomyelin, and 0.08% cholesterol¹⁷.

In the case of COVID-19, Ursodeoxycholic acid (UDCA), also called ursodiol, is a repurposed drug that has been suggested for clinical trials due to its properties in modulating apoptosis and the hyper-immune inflammatory response^{18,19}.

UDCA is a bile salt, it is amphipathic, thus, it has a hydrophobic component that is soluble in lipids and a hydrophilic component that is soluble in water. UDCA is also known as 3 α , 7 β -dihydroxy β -cholan-24-oic acid. It is not synthesized by the liver, it originates in the colon from bacterial epimerization 7 β (from *Ruminococcus gnavus* N53 in human gut microbiota) from primary bile to be absorbed later in the colon^{20,21}, this represents 4% of bile acids²².

Although UDCA is recognized as a non-antiviral hepatoprotective agent, it is used in combination with antivirals for the treatment of some viral diseases, such as viral hepatitis⁷. Also, UDCA alone or in combination with antivirals or immunomodulators, improve patients with hepatitis A virus (HAV), hepatitis B (HBV), hepatitis C (HCV), hepatitis D (HDV), hepatitis E (HEV), and Non-A-E acute viral hepatitis^{23,24}. In addition, potential beneficial effects in acquired immunodeficiency virus have been shown in pilot studies with UDCA²⁵. Previously, Yadav et al.²⁶ have explored the effect of ursodeoxycholate and chenodeoxycholate against SARS-CoV-2 envelope protein. They found that these molecules increase permeability just like surfactants/detergents.

The repositioning of old drugs arises as an advantage due to their previous development and application over the years²⁷. Among the methods used in drug repurposing studies are molecular docking and dynamics simulations²⁸. Molecular Dynamics is one of the computational strategies used for predicting the complementarity of the binding site between the ligand and the target^{29,30}.

In order to show the interaction of the SARS-CoV-2 virus with UDCA, we took into account the UDCA characteristics of hydrophobicity and its action on cholesterol. Here, we explored the interaction between both UDCA and the Spike protein through in silico studies, as well as UDCA in the presence of a model of the membrane of SARS-CoV-2. Also, according to the characteristics of the lipid constituents, we constructed different types of membrane and recorded possible UDCA/membrane interactions.

Materials and methods

Design and data set. For the “SARS-CoV-2 Spike protein-UDCA” analysis, docking and MD were applied to confirm the interaction between the SARS-CoV-2 Spike protein and UDCA. For the “SARS-CoV-2 Spike protein-6UDCA”, we used six UDCA residues around the Spike protein and performed MD. In addition, for the “SARS-CoV-2 membrane-UDCA”, we evaluated the interaction of UDCA with different types of membranes, including one with similar characteristics to the SARS-CoV-2 membrane.

The structure of UDCA was obtained from the Protein Data Bank (PDB)³¹, and its force fields were obtained with “reduce” and “antechamber”, utilities found within the AMBER14 package³². The trimeric structure of the Spike protein was taken from code 6VSB.pdb³³. The membrane model was designed and built as indicated below. Notably, we carried out MD with 7DF3.pdb³⁴ in a prefusion state, and we obtained the same results. To visualize MD’s trajectories, we used UCSF CHIMERA V.1.15³⁵, and VMD 1.9.3³⁶.

Docking study. We analysed the interaction between UDCA³⁷ and the trimeric structure of the Spike protein using HEX 6.3 software [Hex Protein Docking, loria.fr]³⁸ through geometric correlation, with an energy minimization post-processing. The dimension of the grid was 0.75 and steric scan 18, a random seed, receptor and ligand range of 180°, and a range of rotation of 360°.

Molecular dynamics simulation with SARS-CoV-2 Spike protein and UDCA. The “SARS-CoV-2 Spike protein-UDCA” was obtained by docking a structure with lower energy, which corresponded to – 386.35 kcal/mol. Next, the complex was subjected to MD with AMBER14 software²⁷. We did a minimization of 25,000 steps, with bond length restrictions for hydrogen atoms (ntc = 2), a cut-off radius of 8.0 Å, and positional restrictions (restraint_wt = 2). The structure was heated from 0 to 300 K for 250,000 steps with a 0.002 ps time step and bond length and positional constraints, a nonbonded cut-off of 8.0 Å, and using a Langevin thermostat.

Phospholipid	Virus	ERGIC	PLASM
DOPA	4	0	0
DOPC	7	40	21
DOPE	35	18	11
DOPS	18	5	8
CHL	0	1	24
TOTAL	64	64	64

Table 1. The number of phospholipids in each membrane. *DOPA* dioleoyl-phosphatidic acid, *DOPC* dioleoyl-phosphatidyl choline, *DOPE* dioleoyl-phosphatidyl ethanolamine, *DOPS* dioleoyl-phosphatidyl choline, *CHL* Cholesterol, *ERGIC* endoplasmic reticulum golgi intermediate compartment, *PLASM* plasmatic membrane.

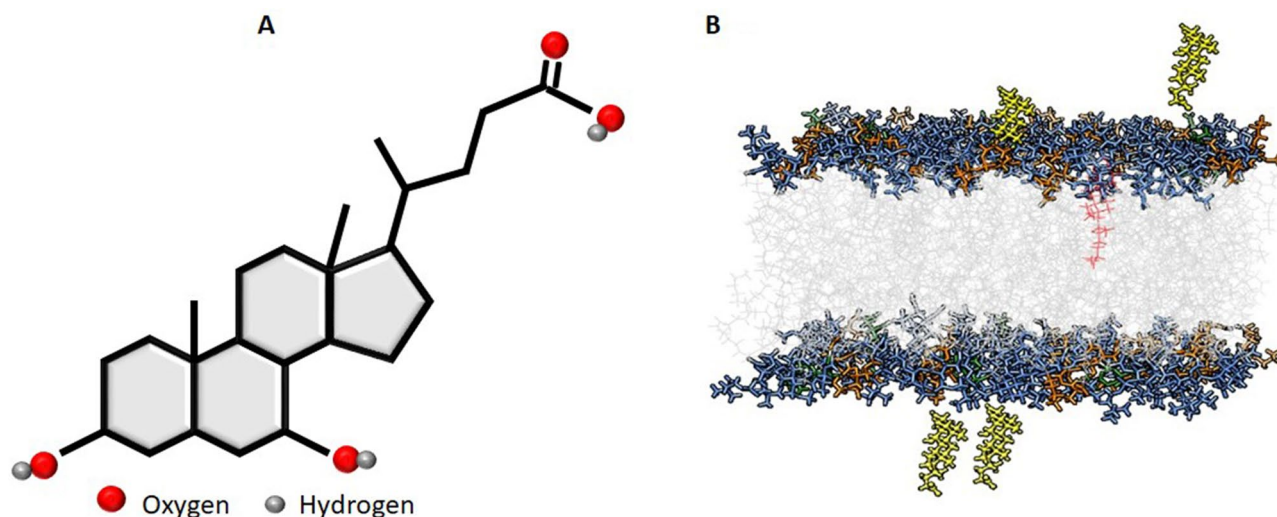


Figure 1. UDCA structure and identification and distribution of UDCA in the different membranes. (A) UDCA schematic representation, using the SMILE code (CC(CCC(=O)O)C1CCC2C1(CCC3C2C(CC4C3(CCC(C4)O)C)O)C). (B) Membrane ERGIC-like. All membranes (VIR, ERGIC, and PLASM) had a similar UDCA spatial distribution. This figure shows an ERGIC-like membrane built with Charmm-gui [CHARMM-GUI] server. Image edited with UCSF Chimera V 1.15. Grey lines: glycerol, orange lines: PE, blue lines: PC, green lines: PS, red lines: CHL and yellow lines: surrounding UDCA residues.

Then, density balancing was done over 750,000 steps with a time step of 0.002, 8.0 Å nonbonded cut-off, 1.0 pressure relaxation time, Langevin thermostat for temperature control, a collision frequency of 2.0/ps ($\gamma_{\text{maln}} = 2.0$) and positional constraints. Finally, a molecular dynamic of 4,000,000 steps (total 10 ns) was executed with a time step of 0.002 ps, a cut-off radius of 8.0 Å, Langevin thermostat for temperature control ($\text{ntt} = 3$) at 300 K, a collision frequency of 2.0 ($\gamma_{\text{ln}} = 2.0$) and positional constraints.

For the "SARS-CoV-2 Spike protein-6UDCA", we surround the trimeric structure of the Spike protein with six UDCA residues through MD. The same parameters were used for minimization, state solvation, heating, density balance and production of molecular dynamics. All sugar residues attached to Spike protein were removed. The trajectory analysis was visualized with VMD 1.9.3³⁷ and the distance and hydrogen bond calculations were carried out with CPPTRAJ³⁹.

Membrane design. Considering that SARS-CoV-2 buds in regions with a high vesicle density besides the endoplasmic reticulum (ER) and Golgi-like membrane¹⁰, we designed different membranes, including a SARS-CoV-2-like membrane (ERGIC-like). Virus (VIR)⁴⁰, ERGIC with cholesterol¹⁴, and eukaryotic plasmatic membranes (PLASM) with cholesterol²⁴ were constructed (Table 1). The same phospholipid distribution was used in all membranes and four UDCA residues were added, two in each leaf (Fig. 1). Thus, in "SARS-CoV-2 membrane-UDCA", we observed the response of four UDCA residues around the ERGIC-type, virion-type and plasma-type bilipid membranes. We constructed a model of six lipid bilayer membranes, three for reference, using dioleoyl-phosphatidic acid (DOPA), dioleoyl-phosphatidylcholine (DOPC), dioleoyl-phosphatidylethanolamine (DOPE), dioleoyl-phosphatidylserine (DOPS), and cholesterol (CHL) (128 each one, 64 in each leaf).

We designed the following membranes: VIR, VIR-IU5, ERGIC, ERGIC-IU5, PLASM, PLASM-IU5, where IU5 denotes UDCA. Additionally, for comparison, we built two membrane systems with the same amount of phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), and cholesterol, however, instead of using DOPC, DOPE and DOPS, we used POPC (1-palmitoyl-2-oleylphosphatidylcholine), POPE (1-palmitoyl-2-oleylphosphatidylethanolamine), and POPS (1-palmitoyl-2-oleylphosphatidylserine). The percentage of PC,

Frame	Acceptor	Donor		RMSF	Distance (nm)—angle (°)
41	T998@OG1 (A)	IU5_2906@HOA1	IU5_2906@O1A	0.77	2.8286 153.1908
3	IU5_2906@O4	T998@HG1 (C)	T998@OG1 (C)	0.79	2.8305 160.6405
3	IU5_2906@O4A	G999@H (B)	G999@N (B)	0.79	2.9078 140.9356
2	IU5_2906@O1B	R995@HH12 (C)	R995@NH1 (C)	0.94	2.8931 151.5614
1	D994@OD2 (A)	IU5_2906@HOA1	IU5_2906@O1A	0.74	2.5852 154.5614
	D994@OD2 (A)	IU5_2906@HOB1	IU5_2906@O1B	0.74	2.8389 162.9852
	D994@O(A)	IU5_2906@HOA1	IU5_2906@O1A	0.74	2.8648 147.6706
	IU5_2906@O1A	R995@HH12 (A)	R995@NH1	0.94	2.8746 136.8333
	R995@NH1 (C)	IU5_2906@HOA1	IU5_2906@O1A	0.94	2.9528 167.5659

Table 2. Interactions between UDCA and T998 of central helices of S2 of Spike protein of SARS-CoV-2. These distances and angles correspond to H bonds formed during the trajectory of molecular dynamics. As observed in Fig. 2B, once these atoms are at a close distance, they remain together along the trajectory. In molecular dynamics, most of the H-bonds are at the beginning of the trajectory. H-bonds are formed between UDCA (denoted as IU5_2906 according to AMBER files) and T998, R995, D994 and G999 from three monomers (H-bonds obtained with CPPTRAJ).

PE, PS, and CHL is shown in Table 1. Sphingomyelin was not used in the systems because AMBER lacks force fields for these phospholipids.

The membrane systems were built with CHARMM-GUI^{41,42} using an average of 35 water molecules per lipid. To equalize the charges to zero, we employed symmetric lipidic bilayers with potassium ions with the force fields lipid17, TIP3, GAFF, and NTP assembly, in AMBER format.

Minimizations of 50,000 steps were performed at constant volume, no pressure scale, no restriction on hydrogen atoms and a nonbonded cut-off of 10.0 Å. After this, a preheating from 0 to 100 °K, 250,000 steps, with a 0.002-time step, hydrogen bond length restrictions, no evaluation for hydrogen bonds, relative geometric tolerance for coordinates of 0.0000001, Langevin thermostat and a collision rate of 1.0/ps. Then, we simulated a warm-up from 100 to 300 °K with the same parameters as in the previous preheat but for 250,000 steps, periodic boundary conditions at constant pressure, anisotropic pressure coupling, and a pressure relaxation time of 2.0 ps, all according to the AMBER14 instructions⁴³. For the density balance, periodic limit conditions were used at constant pressure, Langevin's thermostat, a collision frequency of 1.0/ps, relative geometric tolerance for coordinates of 0.0000001, a cut-off radius of 10.0 Å during 1, 750,000 steps with a time step of 0.002 (for a total of 3.5 ps). The molecular dynamics was done during 60,000,000 steps with a time step of 0.002 (total 120 ps), with restrictions for hydrogen bonds, relative geometric tolerance for coordinates of 0.0000001, Langevin thermostat, with a collision frequency of 1.0/ps at temperature 300 °K, 10 Å cut-off, anisotropic pressure coupling, and periodic boundary conditions at constant pressure.

We used VMD 1.9.3³⁶ for the visualization of the trajectories and CPPTRAJ³⁹ for the calculation of distances, RMSD, and the calculation of hydrogen bonds.

Results

Interaction between the trimeric structure of the Spike protein and UDCA by docking. In the "SARS-CoV-2 Spike protein-UDCA" analysis, we employed the trimeric structure of the Spike protein and UDCA using the HEX 6.3 program. UDCA showed an affinity energy of -386.35 kcal/mol with the residues conforming the three alpha central helices region. This set of three alpha helices is formed by the contribution of one alpha helix of every monomer (residues K986 to C1032). We think that this set of three alpha helices hold together the three monomers of the Spike protein before a post-fusion conformation.

When applying MD to this structure (spike protein-UDCA), we observed that the UDCA remained in the same position during the whole trajectory, and there were even moments when hydrogen bonds were formed with residues close to it (Table 2). UDCA had affinity mainly with residues R995 and T998 of monomers A, B, and C. Residues R995 and T998 were found in the central helices from subunit S2 (S2) of the protein (Fig. 2).

The oxygen atom OG1 of T998 of the monomer A interacts with the hydrogen HOA1 of UDCA. The hydrogen atom HGB of T998 of monomer B interacts with the oxygen atom O4 of the UDCA, and the oxygen atom OG1 of the residue T998 of monomer C interacts with the hydrogen H222 of UDCA with average distances of 2.67 ± 0.41 Å. On the other hand, the distances between monomers A, B, and C in the pre-fusion state were 5.6 and 8.6 Å for OG1 and CG2 of every T998, respectively, and in the post-fusion state, it is 9.58 and 9.23 Å for the same atoms (Fig. 3A,B). In presence of UDCA, these distances are not modified and remained at 5.93 ± 0.62 and 7.78 ± 0.51 Å, respectively (Fig. 2A). This leads us to consider that UDCA could act as a connecting bridge that would not allow an adequate conformational shape for the post-fusion state. The RMSD, RMSF as well as Rg of monomers A, B, and C were maintained stable as shown in Fig. 2B–E, indicating an equilibrated system. The monomer C showed the smallest RMSD compared with monomer A and B (RMSD (A), 4.03 ± 0.6 Å; RMSD (B), 3.14 ± 0.42 Å and RMSD (C), 2.89 ± 0.32 Å). RMSF average of monomer C was 2.25 ± 0.56 and the average Rg of monomer C was 44.31 ± 0.28 .

Likewise, the distances between UDCA and residues R995, T998 and D994, R995 and T998 of the second and third monomers were preserved. In Table 2, we observed that ARG 995, THR 998 and GLY 999 atoms showed a

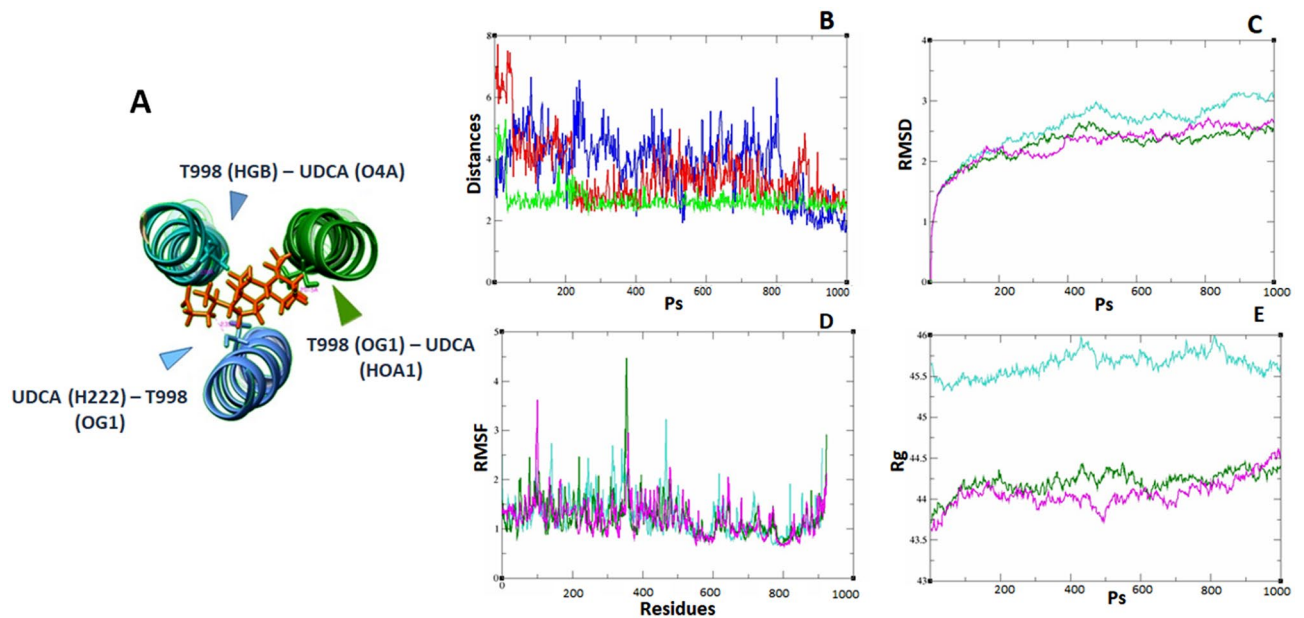


Figure 2. Schematic representation of UDCA in the central region of the alpha-helices. The UDCA—Spike protein interaction is maintained along the molecular dynamics' trajectory due to the interaction of the bonds between the THR 998 atoms of monomers A, B, and C with UDCA. (A) UDCA interaction is shown with monomers A, B and C. UDCA in orange. Image edited with UCSF Chimera V 1.15. (B) The distances of the atoms that maintain this interaction are observed, the intervening atoms are shown in Table 2; blue line UDCA-T998 (chain B), red line UDCA-T998 chain C and green line UDCA-T998. (C) RMSD calculated by monomer (A, turquoise; B, green and C, magenta line) which shows stability in the simulation. (D) RMSF calculation, fluctuation of each residue per monomer (A, turquoise; B, green and C, magenta line) in the region of alpha helices, from lysine 986 to cysteine 1032. (E) Calculation of the radius of gyration that shows the degree of compactness of the monomers: A (45.64 ± 0.14 ; turquoise line) > B (44.21 ± 0.13 ; green line) > C (44.07 ± 0.17 ; magenta line).

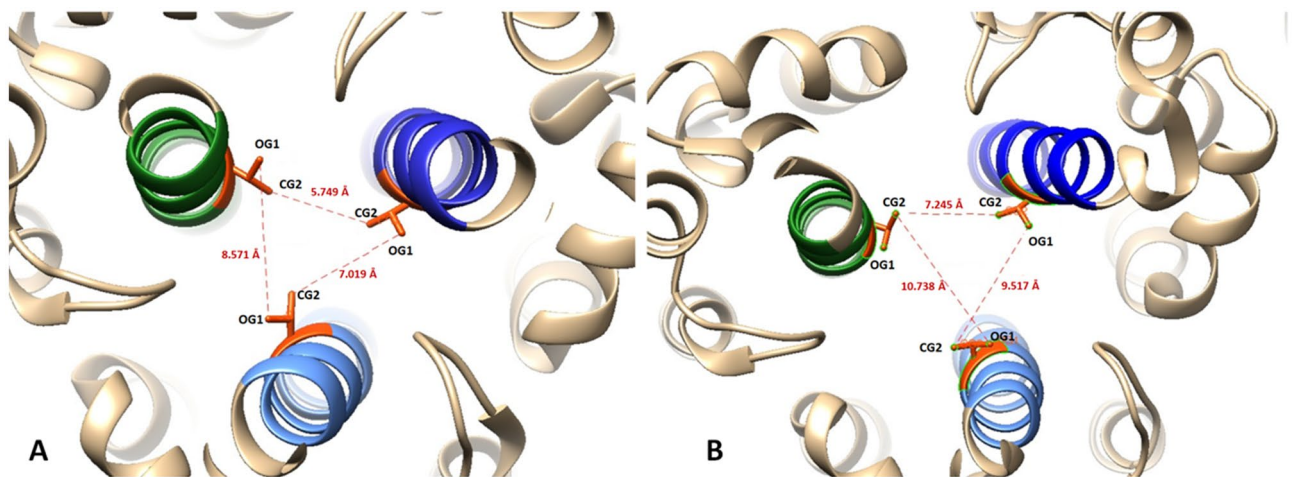


Figure 3. Distances between OG1 and CG2 of the monomers A (green), B (blue), and C (light blue) of THR 988 in Spike protein. We observed that there is a slight increase in distance suggesting that these helices could separate from each other once the membrane fusion process begins. In (A) the closed conformation is observed (smaller distances) and in (B) the open conformation. With the presence of UDCA, it remains in a conformation similar to (A). Taken from 7DF3.pdb and 7BNO.pdb and edited with chimera V. 1.15.

low RMSF (0.94, 0.77 and 0.79, respectively). The lowest RMSF of THR could explain the interactions established with UDCA since it had a lower value and major stability.

Furthermore, we compared the closed (pre-fusion) and open (post-fusion) conformation of the alpha-helices of the Spike protein^{34,43} using UCSF CHIMERA V.1.15³⁶. We observed that in the closed conformation (7DF3.pdb) (Fig. 3A), the OG1 and CG2 atoms of every T998 are separated for 5.6 and 8.6 Å from each other,

respectively. However, for the open conformation (7BNO.pdb) (Fig. 3B), those same atoms had a distance of 9.58 and 9.23 Å, respectively. This suggests that in post-fusion conformation, these alpha-helices are slightly separated, and this separation is important for correcting fusion between viral and host membranes; although, in the presence of UDCA, we observed that the conformation remains closed, as seen in Fig. 2A. In accordance with the observed distances, we think that if UDCA binds to this region and serves as a bridge between these residues, it could prevent the Spike protein from folding correctly to its open conformation, and thus not carry out membrane fusion, Fig. 2A.

UDCA showed affinity towards a region between HR1 and HR2. In the "SARS-CoV-2 Spike protein-6UDCA", we placed six UDCA around the trimeric structure of the Spike protein in its closed conformation (6VSB.pdb). We visually analyse the trajectory using VMD 1.9.3. We found that three of the six residues were attracted to three different regions of the Spike protein in different areas to the first analysis. One of them was attracted to an ASN536 residue of the C monomer for a very short period of 50 ps, while the rest of the trajectory was kept repelled. Another UDCA was attracted to a residue of THR of chain C with a very short period of 25 ps. The third one was attracted by CYS1126 of chain C with a longer period. The last interaction UDCA-CYS1126 is conserved in different periods of the trajectory. Also, these residues interact for long periods of time, for up to 375 ps, which suggests a strong attraction. The distances between UDCA and N536 was 10.94 ± 4.69 Å, between UDCA and T602 was 10.94 ± 4.69 Å, and between UDCA and C1126 was 6.99 ± 3.15 Å. In monomer C, we found 1.39 of RMSF with CYS1126, less than the average (1.81 ± 0.55) of the whole monomer C; monomer A with 4.03 ± 0.6 , monomer B with 3.14 ± 0.42 and monomer C with 2.89 ± 0.32 . The RMFS's and Rg of each monomer were similar (Fig. 4A–D). A graphical representation is shown in Fig. 4E, where we see the UDCA interacting with the region between HR1 and HR2.

Lipid membranes and UDCA. In "SARS-CoV-2 membrane-UDCA", we used dioleoyl- or palmitoyl- as controls and observed that the electron density remained stable on both sides of the membrane, in a range of ± 20 Å along all trajectories (Figs. 5A and 6), indicating balanced simulations.

Three UDCA were attracted to the virion-like membrane, which was retained for a long time (30 ns), as was observed in the distances of the centres of mass of the UDCA with respect to the PC and PE. UDCA bound to these residues from the beginning of the trajectory and remained together along the trajectory (Fig. 5B). In the Supplementary files, we show the simulation of a membrane virion-like with UDCA.

In the ERGIC-like membrane, only one leaf contained a cholesterol residue, the other side contained 62.5% PC, 28.2% PE and 8% PS. Possibly because of cholesterol, the UDCA that were attracted to the membrane only remained attached for a short time and then were detached, as shown in Fig. 5C,D.

In a eukaryotic PLASM-like membrane, there was 37,5% CHL, 33% PC, 17% PE, and 12,5% PS. We found that two UDCA residues maintained a close distance to the phospholipids, Fig. 5E,F. However, they did not stay together along the membrane. The behaviour of PC and PE is the same, they tend to attract UDCA residues. It is notorious that in the PLASM-like membrane, a cluster of three cholesterol is formed (charmm-gui.org server aleatory constructed that way), and only this way, cholesterol attracted UDCA, since if cholesterol is found alone, it does not attract UDCA.

It is noteworthy that UDCA does not show attraction to membranes containing palmitoyl (phospholipids with 20 carbons), whatever its constitution.

Discussion

UDCA interacts in the central helix region of S2 and HR1 and HR2. We observed that UDCA interacts in two regions of Spike protein. One of the regions which shown attraction is in the area with three alpha-helices of three monomers of the Spike protein (residues R995, T998, V991, D994, T998, T998, and R995), and the other, with a CYS1126 residue in a zone between the HR1 and HR2 regions.

Also, we compared the pre-fusion and post-fusion conformation of the alpha-helices of the Spike protein^{34,36,43} and observed that in the pre-fusion conformation (7DF3.pdb, Fig. 3A) and the post-fusion conformation (7BNO.pdb, Fig. 3B), those same atoms have a distance of 9.58 and 9.23 Å, respectively. As was measured before⁴⁴ in post-fusion conformation, these alpha-helices are slightly separated; this separation is important for a correct fusion between viral and host membranes, but in the presence of UDCA, we observed that the conformation remains closed.

This leads us to consider that UDCA could join in two regions of the Spike protein, in the S2 region, at residue T998 found in the central helices formed by residues K986–C1032. Precisely in this region, there are two mutations of interest, the K986 and V987, these are already being induced with proline substitutions in several vaccines that are in development⁴⁵. In the joint of the central helices of S2, UDCA could prevent conformational changes in the Spike protein once it binds to ACE2. UDCA could act as a link between the three monomers, since as we observed in Fig. 2A,B, UDCA forms hydrogen bonds with the T998 residues of each of the monomers.

UDCA interacts in the central helix region of S2 and HR1 and HR2. A key step in the coronavirus infection process is membrane fusion⁴⁶. In this stage, the virus protein fusion core is formed by the union of three copies of HR1 and three of HR2. This stage had been verified in a certain way by the inhibition of fusion through the introduction of peptides, analogous to HR1 and HR2. According to the current model, HR1 and HR2 are joined to form the coiled-coil bundle necessary for fusion^{46,47}.

The post-fusion stage takes place after the receptor-binding domain (RBD) has contact with its target protein (ACE2). After this contact, the Spike protein has a series of conformational changes, among these, the union of the three HR1 and three HR2⁴⁸. Another region where the UDCA has a high probability of binding is an

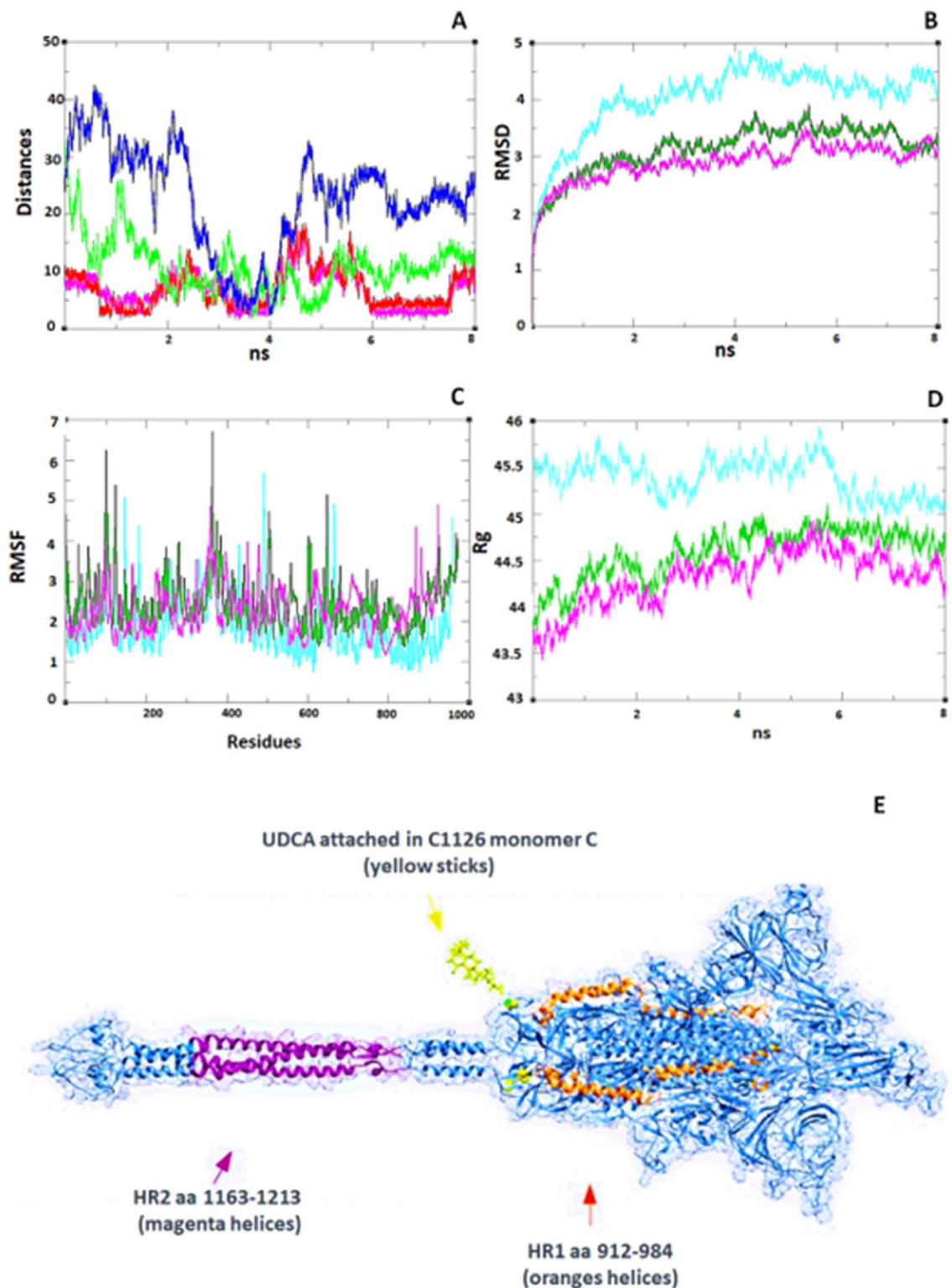


Figure 4. Interaction of UDCA with cysteine residue C1126. (A) Distances between UDCA 2908 and C1126 of monomer C (atom O4-UDCA vs HB2-C1126. Atom O4A-UDCA vs HB3-C1126 in magenta and red, respectively). UDCA-2916 interaction with N536 (atom HC11-UDCA vs OD1-N536 in green line). UDCA 2912 interaction with T602 (O4-UDCA vs HB-T602 in blue line). Residues N536 and T602 bind to UDCA momentarily, instead C1126 binds to UDCA and tends to retain it, thus we think that if UDCA interferes in this region near to HR1-HR2, it could prevent the coiled-coil heptad repeat binding in the area between the HR1 and HR2 regions, necessary for membrane fusion. (B–D) (monomer A in turquoise line, B in green line and C in magenta line), respectively. (E) General Overview of the Spike protein (taken from charm-gui.org, 6VSB.1.1.1). The HR1 regions of each of the A, B, and C monomers are shown in orange tones. The HR2s are shown in purple tones and the regions to which the UDCA binds in the simulations are shown in yellow (stick representation). Values of RMSD, RMSF and Rg indicates a balanced simulation, although there is a slightly higher value for monomer A, B, and C are very similar. The three UDCA in the MD were differentiated using the Amber nomenclature as UDCA-2908, UDCA-2916 and UDCA-2912.

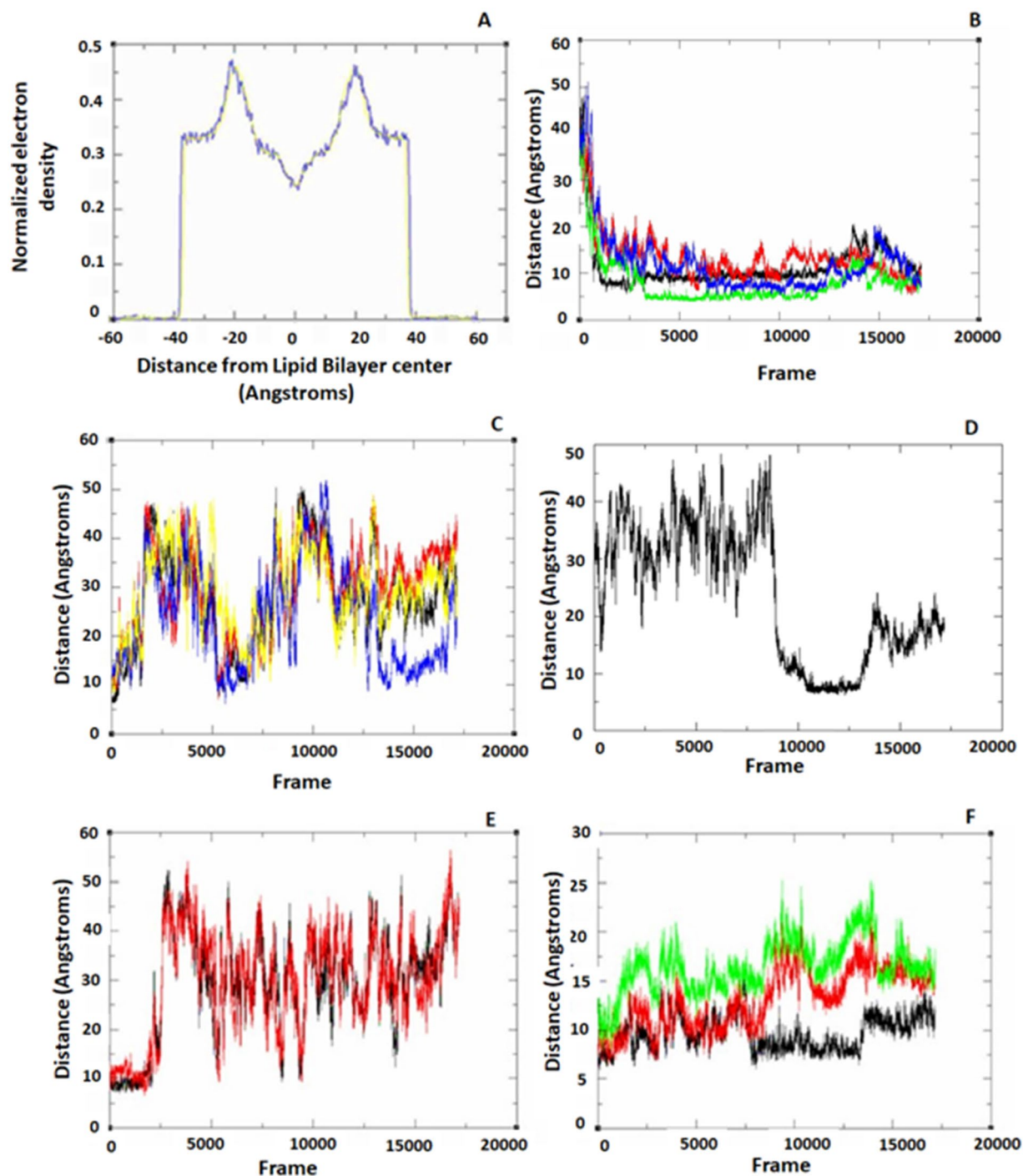


Figure 5. Electron density profile bilayer membranes and distances between the centre of mass of UDCA and PE. (A) Electron density profile bilayer membranes was very similar in all membranes along the trajectory, showing stable Molecular Dynamics. The virion membrane has a major percentage of DOPE and DOPS. The electron density remains stable on both leaves of the membrane (± 20 Angstroms). (B) Distances from the centre of mass of UDCA in relation to four centres of mass of PE, PE272 (black), PE218 (red), PE251 (green) and PC221 (blue), to which the UDCA molecules were attracted. We observed that in the virus membrane, once the UDCA residue is attracted, it tends to remain attached to the membrane. Distances calculated with CPPTRAJ. Also, in the Golgi intermediate compartment membrane (ERGIC), we have shown the distances from the centre of mass of UDCA residues to the phospholipid residues, for which it exhibits affinity. (C) It is observed that residue 401 of UDCA is close to the phospholipids PE17 (black), PE20 (red), PC59 (blue) and PC47 (yellow). (D) The distance of residue UDCA 403 towards PE293 (black line) is observed. Both figures show affinity, however, they do not remain together over time. In addition, we evaluate the distances between the UDCA residues and PC, PE and CHL phospholipids. (E) We observed the distances between UDCA and PE. It is noteworthy that the only way that UDCA was attached to cholesterol was because 3 cholesterol residues were found together, if they were separate, UDCA and cholesterol tended to separate. (F) The distances between residues of UDCA and cholesterol are shown, we observed that from the beginning the UDCA and PE/CHL were united but over time they tended to separate.

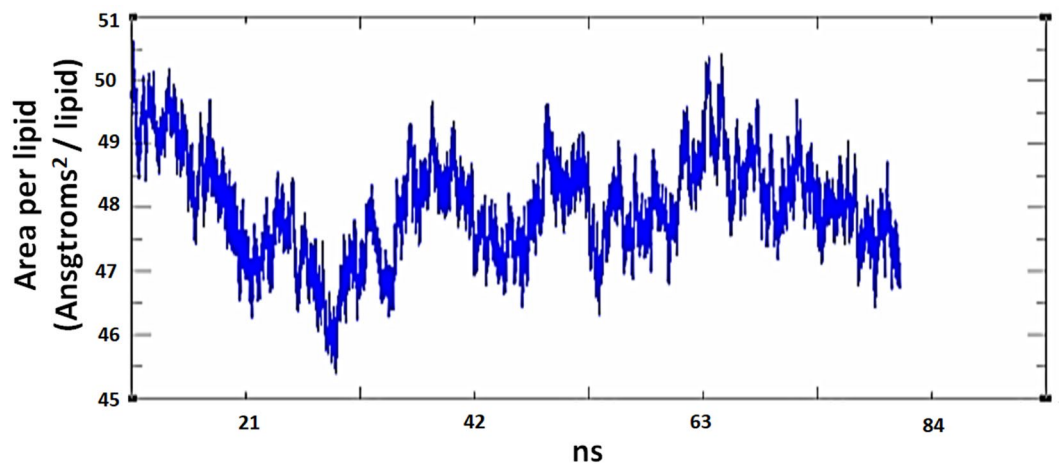


Figure 6. Area per lipid for PLASM-IU5-like membrane. As shown in the graph, it tends to be stable along the trajectory. This parameter was similar for all of the rest of the membranes simulated. Obtained with CCPTRAJ.

intermediate region located between the monomers HR1 and HR2, since it is more exposed, as seen in Fig. 4C. In this region, UDCA interacts with the residue C1126. This residue is part of the C1082-C1126 disulphide bridge, which could modulate the flexibility of the protein^{49,50}.

The residue C1126 is in the region between this HR1 and HR2 and in our results of "SARS-CoV-2 protein-6UDCA", the oxygen atoms O4 and O4A of UDCA stays bound to the hydrogen atoms HB2 and HB3 of C1126, respectively, up to 375 ps. UDCA has a β -hydroxylation that provides it with greater hydrophilicity and thus could interact with hydrogens^{51,52}. This causes that UDCA and cysteine C1126 to suddenly separate, however, nanoseconds later, both molecules tend to bound, as shown in Fig. 4A (red and magenta line).

UDCAs interact in a stable time interval with the virion-like membrane and PLASM-like membrane. In membranes, UDCA interacts with PE and PC of virus-like and ERGIC-like membranes. Although with the virion-like membranes the interaction show more affinity, due to three of four residues are attracted to it, whereas in ERGIC-like membrane, only show that one UDCA is attracted.

The finding that three cholesterol molecules are present and UDCA is easily attracted, is related to published studies stating that eukaryotic membranes are rich in cholesterol, which is used by SARS-CoV-2 to enter cells, as it depends on a cholesterol-rich lipid raft⁵³. Previously, another study found that cholesterol increases interactions with viral membranes due to its electrostatic and solvation properties⁵⁴, moreover, cholesterol is also known to promote oligomerization of the SARS-CoV-2 FP⁵⁵. Therefore, we think that if UDCA binds to cholesterol lipid rafts in eukaryotic cells, UDCA could hinder the interaction between cholesterol and FP and avoid oligomerization. According to Fig. 5A,B, UDCA may form a cluster with cholesterol-rich lipid rafts in eukaryotic membranes, and thus prevent that the fusion peptide of the Spike protein of SARS-CoV-2 from oligomerizes.

In Fig. 5A,B, we observed that in virion-like membranes, the distance of the centres of mass between UDCA and PE, at the beginning of the simulation are far apart, although they tend to attract each other and once they are joined, they no longer separate. Whereas ERGIC-like membrane, the distance from centres of mass of UDCA towards PE, shows that they are not attracted, except when three cholesterol molecules are together, the UDCA tends to be attracted (PLASM-like membrane, Fig. 5A–F).

Likewise, bile acids are endogenous inhibitors of the NLRP3 inflammasome via NLRP3 ubiquitination, through the TGR5-cAMP-PKA axis. Bile acids could act through this pathway and regulate the activation of the NLRP3 inflammasome⁵⁶ and attenuate COVID-19⁵⁷. Furthermore, it has been reported that UDCA presents anti-apoptotic activity by modulating mitochondrial membrane permeability transition⁵⁸.

However, within the limitations of this work, we were unable to build a more realistic membrane due to the lack of force fields of various lipids. In addition, the sugars that accompany the Spike protein could serve as protection against various compounds including UDCA. Therefore, removing them for simulation reduces reliability.

In conclusion, the interaction between UDCA and various residues of the Spike protein and its membrane indicates that UDCA has an affinity and remains attached to the central helix of S2 of the Spike protein monomers. UDCA interacts in a zone between HR1 and HR2 regions, it also tends to remain membrane-bound in the presence of cholesterol molecules. The data showed that UDCA could interact with the SARS-CoV-2 Spike protein and a membrane model like SARS-CoV-2, destabilizing the interaction of the virus its target cells.

Received: 4 July 2021; Accepted: 1 November 2021

Published online: 15 November 2021

References

- Lu, H. Drug treatment options for the 2019–new coronavirus (2019–nCoV). *Biosci. Trends* **14**(1), 69–71. <https://doi.org/10.5582/bst.2020.01020> (2020).
- Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**(7816), 459–468. <https://doi.org/10.1038/s41586-020-2286-9> (2020).
- Bhardwaj, V. K. et al. Bioactive molecules of tea as potential inhibitors for RNA-dependent RNA polymerase of SARS-CoV-2. *Front. Med. (Lausanne)* **8**, 684020. <https://doi.org/10.3389/fmed.2021.684020> (2021).
- Bhardwaj, V. K. et al. Identification of bioactive molecules from tea plant as SARS-CoV-2 main protease inhibitors. *J. Biomol. Struct. Dyn.* **39**(10), 3449–3458. <https://doi.org/10.1080/07391102.2020.1766572> (2021).
- Singh, R., Bhardwaj, V. K., Sharma, J., Purohit, R. & Kumar, S. In-silico evaluation of bioactive compounds from tea as potential SARS-CoV-2 nonstructural protein 16 inhibitors. *J. Tradit. Complement. Med.* <https://doi.org/10.1016/j.jtcme.2021.05.005> (2021) (Epub ahead of print).
- Zaki, A. A., Ashour, A., Elhady, S. S., Darwish, K. M. & Al-Karmalawy, A. A. Calendulaglycoside A showing potential activity against SARS-CoV-2 main protease: Molecular docking, molecular dynamics, and SAR studies. *J. Tradit. Complement. Med.* <https://doi.org/10.1016/j.jtcme.2021.05.001> (2021).
- Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet. Infect. Dis* **20**(5), 533–534 (2020).
- Johns Hopkins University & Medicine (2020) “COVID-19 Dashboard”. <https://coronavirus.jhu.edu/map.html> Accessed 11 November 2021.
- Esakandari, H. et al. A comprehensive review of COVID-19 characteristics. *Biol. Proced. Online* **22**, 19. <https://doi.org/10.1186/s12575-020-00128-2> (2020).
- Klein, S. et al. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. *Nat. Commun.* **11**(1), 5885. <https://doi.org/10.1038/s41467-020-19619-7> (2020).
- Turoňová, B. et al. In situ structural analysis of SARS-CoV-2 Spike reveals flexibility mediated by three hinges. *Science* **370**(6513), 203–208. <https://doi.org/10.1126/science.abd5223> (2020).
- Ke, Z. et al. Structures and distributions of SARS-CoV-2 Spike proteins on intact virions. *Nature* **588**(7838), 498–502. <https://doi.org/10.1038/s41586-020-2665-2> (2020).
- Bhardwaj, V. K., Singh, R., Das, P. & Purohit, R. Evaluation of acridinedione analogs as potential SARS-CoV-2 main protease inhibitors and their comparison with repurposed anti-viral drugs. *Comput. Biol. Med.* **128**, 104117. <https://doi.org/10.1016/j.combiomed.2020.104117> (2021).
- Gioia, M. et al. Role of proteolytic enzymes in the COVID-19 infection and promising therapeutic approaches. *Biochem. Pharmacol.* **182**, 114225 (2020).
- White, J. M., Delos, S. E., Brecher, M. & Schornberg, K. Structures and mechanisms of viral membrane fusion proteins: Multiple variations on a common theme. *Crit. Rev. Biochem. Mol. Biol.* **43**(3), 189–219. <https://doi.org/10.1080/10409230802058320> (2008).
- O'Donnell, V. B. et al. Potential role of oral rinses targeting the viral lipid envelope in SARS-CoV-2 infection. *Function (Oxford, England)* **1**(1), zqaa002. <https://doi.org/10.1093/function/zqaa002> (2020).
- van Meer, G. Lipids of the Golgi membrane. *Trends Cell Biol.* **8**(1), 29–33. [https://doi.org/10.1016/s0962-8924\(97\)01196-3](https://doi.org/10.1016/s0962-8924(97)01196-3) (1998).
- Subramanian, S., Iles, T., Ikramuddin, S. & Steer, C. J. Merit of an ursodeoxycholic acid clinical trial in COVID-19 patients. *Vaccines* **8**(2), 320. <https://doi.org/10.3390/vaccines8020320> (2020).
- Abdulrab, S., Al-Maweri, S. & Halboub, E. Ursodeoxycholic acid as a candidate therapeutic to alleviate and/or prevent COVID-19-associated cytokine storm. *Med. Hypotheses* **143**, 109897. <https://doi.org/10.1016/j.mehy.2020.109897> (2020).
- Angulo, P. Use of ursodeoxycholic acid in patients with liver disease. *Curr. Gastroenterol. Rep.* **4**(1), 37–44. <https://doi.org/10.1007/s11894-002-0036-9> (2002).
- Lee, J. Y. et al. Contribution of the 7 β -hydroxysteroid dehydrogenase from *Ruminococcus gnavus* N53 to ursodeoxycholic acid formation in the human colon. *J. Lipid Res.* **54**(11), 3062–3069. <https://doi.org/10.1194/jlr.M039834> (2013).
- Lazaridis, K. N., Gores, G. J. & Lindor, K. D. Ursodeoxycholic acid “mechanisms of action and clinical use in hepatobiliary disorders”. *J. Hepatol.* **35**(1), 134–146. [https://doi.org/10.1016/s0168-8278\(01\)00092-7](https://doi.org/10.1016/s0168-8278(01)00092-7) (2001).
- Fabris, P. et al. Effect of ursodeoxycholic acid administration in patients with acute viral hepatitis: A pilot study. *Aliment. Pharmacol. Ther.* **13**(9), 1187–1193. <https://doi.org/10.1046/j.1365-2036.1999.00592.x> (1999).
- Lirussi, F. et al. Long-term treatment of chronic hepatitis C with ursodeoxycholic acid: Influence of HCV genotypes and severity of liver disease. *Liver* **19**(5), 381–388. <https://doi.org/10.1111/j.1478-3231.1999.tb00066.x> (1999).
- Kurktschiev, D., Temelkova-Kurktschiev, T., Horn, K. & Schentke, K. Successful immunomodulating in AIDS patients with ursodeoxycholic acid—A pilot study. *Clin. Exp. Immunol.* **115**(1), 144–146. <https://doi.org/10.1046/j.1365-2249.1999.00760.x> (1999).
- Yadav, R., Choudhury, C., Kumar, Y. & Bhatia, A. Virtual repurposing of ursodeoxycholate and chenodeoxycholate as lead candidates against SARS-Cov2-Envelope protein: A molecular dynamics investigation. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2020.1868339> (2020) (Epub ahead of print).
- Pushpakom, S. et al. Drug repurposing: progress, challenges and recommendations. *Nat. Rev. Drug Discov.* **18**(1), 41–58. <https://doi.org/10.1038/nrd.2018.168> (2019).
- Brogi, S. Computational approaches for drug discovery. *Molecules* **24**(17), 3061. <https://doi.org/10.3390/molecules24173061> (2019).
- Borkotoky, S., Dey, D. & Banerjee, M. Computational insight into the mechanism of SARS-CoV-2 membrane fusion. *J. Chem. Inf. Model.* **61**(1), 423–431. <https://doi.org/10.1021/acs.jcim.0c01231> (2021).
- Gorgun, D., Lihan, M., Kapoor, K. & Tajkhorshid, E. Binding mode of SARS-CoV-2 fusion peptide to human cellular membrane. *Biophys. J.* <https://doi.org/10.1016/j.bpj.2021.02.041> (2021).
- Berman, H. M. et al. The protein data bank. *Nucleic Acids Res.* **28**(1), 235–242. <https://doi.org/10.1093/nar/28.1.235> (2000).
- Case, D. A. et al. *AMBER 14* (University of California, 2014).
- Wrapp, D. et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *BioRxiv Preprint Serv. Biol.* <https://doi.org/10.1101/2020.02.11.944462> (2020).
- Xu, C. et al. Conformational dynamics of SARS-CoV-2 trimeric spike glycoprotein in complex with receptor ACE2 revealed by cryo-EM. *Sci. Adv.* **7**(1), eabe5575. <https://doi.org/10.1126/sciadv.abe5575> (2021).
- Petersen, E. F. et al. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**(13), 1605–1612. <https://doi.org/10.1002/jcc.20084> (2004).
- Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**(1), 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5) (1996).
- PubChem [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004-. PubChem Compound Summary for CID 31401, Ursodeoxycholic acid. <https://pubchem.ncbi.nlm.nih.gov/compound/Ursodeoxycholic-acid> Accessed 14 June 2021.
- Ritchie, D. W. & Kemp, G. J. Protein docking using spherical polar Fourier correlations. *Proteins* **39**(2), 178–194 (2000).
- Roe, D. R. & Cheatham, T. E. 3rd. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**(7), 3084–3095. <https://doi.org/10.1021/ct400341p> (2013).

40. Ivanova, P. T. *et al.* Lipid composition of viral envelope of three strains of influenza virus—Not all viruses are created equal. *ACS Infect. Dis.* **1**(9), 399–452. <https://doi.org/10.1021/acscinfecdis.5b00040> (2015).
41. Jo, S. *et al.* CHARMM-GUI 10 years for biomolecular modeling and simulation. *J. Comput. Chem.* **38**(15), 1114–1124. <https://doi.org/10.1002/jcc.24660> (2017).
42. Lee, J. *et al.* CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.* **153**(3), 035103. <https://doi.org/10.1063/5.0012280> (2020).
43. Benton, D. J. *et al.* The effect of the D614G substitution on the structure of the Spike glycoprotein of SARS-CoV-2. *Proc. Natl. Acad. Sci. U.S.A.* **118**(9), e2022586118. <https://doi.org/10.1073/pnas.2022586118> (2021).
44. Martí, D. *et al.* Temperature effect on the SARS-CoV-2: A molecular dynamics study of the spike homotrimeric glycoprotein. *Comput. Struct. Biotechnol. J.* **19**, 1848–1862. <https://doi.org/10.1016/j.csbj.2021.03.037> (2021).
45. Dai, L. & Gao, G. F. Viral targets for vaccines against COVID-19. *Nat. Rev. Immunol.* **21**(2), 73–82. <https://doi.org/10.1038/s41577-020-00480-0> (2021).
46. Siddell, S., Wege, H. & Ter Meulen, V. The biology of coronaviruses. *J. Gen. Virol.* **64**(Pt 4), 761–776. <https://doi.org/10.1099/0022-1317-64-4-761> (1983).
47. Eckert, D. M. & Kim, P. S. Mechanisms of viral membrane fusion and its inhibition. *Annu. Rev. Biochem.* **70**, 777–810. <https://doi.org/10.1146/annurev.biochem.70.1.777> (2001).
48. Xu, Y. *et al.* Characterization of the heptad repeat regions, HR1 and HR2, and design of a fusion core structure model of the Spike protein from severe acute respiratory syndrome (SARS) coronavirus. *Biochemistry* **43**(44), 14064–14071. <https://doi.org/10.1021/bi049101q> (2004).
49. Ghorbani, M. *et al.* Exploring dynamics and network analysis of spike glycoprotein of SARS-COV-2. *Biophys. J.* **120**(14), 2902–2913. <https://doi.org/10.1016/j.bpj.2021.02.047> (2021).
50. Al-Zyoud, W. & Hazem, H. Dynamics prediction of emerging notable spike protein mutations in SARS-CoV-2 implies a need for updated vaccines. *Biochimie* **191**, 91–103. <https://doi.org/10.1016/j.biochi.2021.08.011> (2021).
51. Alvarez, M. *et al.* Crystal structure of chenodeoxycholic acid, ursodeoxycholic acid and their two 3beta,7alpha- and 3beta,7beta-dihydroxy epimers. *Steroids* **72**, 535–544. <https://doi.org/10.1016/j.steroids.2007.03.007> (2007).
52. Poša, M. Heuman indices of hydrophobicity of bile acids and their comparison with a newly developed and conventional molecular descriptors. *Biochimie* **97**, 28–38. <https://doi.org/10.1016/j.biochi.2013.09.010> (2014).
53. Li, X. *et al.* Dependence of SARS-CoV-2 infection on cholesterol-rich lipid raft and endosomal acidification. *Comput. Struct. Biotechnol. J.* **19**, 1933–1943. <https://doi.org/10.1016/j.csbj.2021.04.001> (2021).
54. Doktorova, M. *et al.* Cholesterol promotes protein binding by affecting membrane electrostatics and solvation properties. *Biophys. J.* **113**(9), 2004–2015. <https://doi.org/10.1016/j.bpj.2017.08.055> (2017).
55. Meher, G. *et al.* Membrane cholesterol modulates oligomeric status and peptide-membrane interaction of severe acute respiratory syndrome coronavirus fusion peptide. *J. Phys. Chem.* **123**(50), 10654–10662. <https://doi.org/10.1021/acs.jpcc.9b08455> (2019).
56. Guo, C. *et al.* Bile acids control inflammation and metabolic disorder through inhibition of NLRP3 inflammasome. *Immunity* **45**(4), 944. <https://doi.org/10.1016/j.immuni.2016.10.009> (2016).
57. Zhao, N. *et al.* The NLRP3 inflammasome and COVID-19: Activation, pathogenesis and therapeutic strategies. *Cytokine Growth Factor Rev.* **61**, 2–15. <https://doi.org/10.1016/j.cytogfr.2021.06.002> (2021).
58. Rodrigues, C. M. *et al.* Ursodeoxycholic acid may inhibit deoxycholic acid-induced apoptosis by modulating mitochondrial transmembrane potential and reactive oxygen species production. *Mol. Med. (Cambridge, Mass.)* **4**(3), 165–178 (1998).

Acknowledgements

This work was supported by the Faculty of Medicine of the Benito Juárez University of Oaxaca, Mexico. The authors would like to thank Charlotte Grundy for her support throughout the work, the National Technology of Mexico (TecNM)/IT Oaxaca and also CONACYT- BP-PA-2021050723-4900732-959110 for financial support.

Author contributions

F.J.R.C., C.R.D. and E.P.C. conceptualized the study and wrote the main manuscript text; L.P.C.M., M.T.H.H. and C.R.D. contributed to discussions and/or data and prepared figures; L.M.S.N., C.A.M.C., M.M.C., E.C.P. and E.Z. provided pathological expertise; E.G.R.M. and E.P.C.M. performed the research. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-01705-5>.

Correspondence and requests for materials should be addressed to C.R.D. or E.P.-C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021