JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

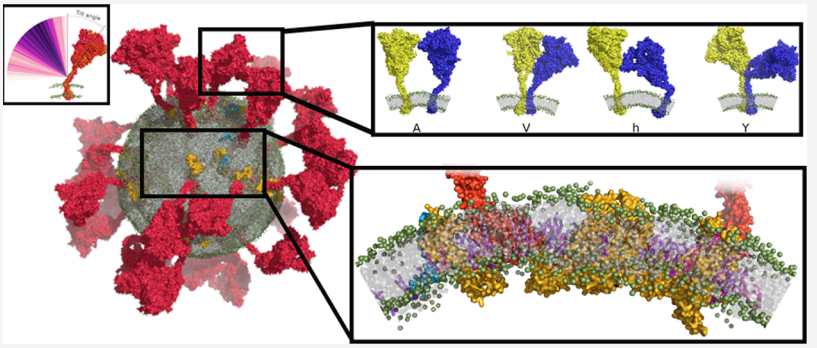
pubs.acs.org/jcim Article

Supramolecular Organization of SARS-CoV and SARS-CoV-2 Virions Revealed by Coarse-Grained Models of Intact Virus Envelopes

Beibei Wang,* Changqing Zhong, and D. Peter Tieleman

Cite This: https://doi.org/10.1021/acs.jcim.1c01240 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: The coronavirus disease 19 (COVID-19) pandemic is causing a global health crisis and has already caused a devastating societal and economic burden. The pathogen, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has a high sequence and architecture identity with SARS-CoV, but far more people have been infected by SARS-CoV-2. Here, combining the structural data from cryo-electron microscopy and structure prediction, we constructed bottom-up Martini coarse-grained models of intact SARS-CoV and SARS-CoV-2 envelopes. Microsecond molecular dynamics simulations were performed, allowing us to explore their dynamics and supramolecular organization. Both SARS-CoV and SARS-CoV-2 envelopes present a spherical morphology, with structural proteins forming multiple string-like islands in the membrane and clusters between the heads of spike proteins. Critical differences between the SARS-CoV and SARS-CoV-2 envelopes are the interaction pattern between the spike proteins and the flexibility of the spike proteins. Our models provide structural and dynamic insights into the SARS virus envelopes and could be used for further investigation, such as drug design and membrane fusion and fission processes.

INTRODUCTION

Coronaviruses, including severe acute respiratory syndrome coronavirus (SARS-CoV) and currently, in particular, SARS-CoV-2, are a major threat to public health.[1] They are enveloped positive-sense RNA viruses that can be transmitted from animals to humans and cause a variety of diseases ranging from common cold to severe diseases.[2]

The SARS-CoV and SARS-CoV-2 virions contain four main structural proteins: the nucleocapsid (N), spike (S), envelope (E), and membrane (M) proteins.[3] Ns are tightly packed with RNA in the viral lumen, while the S, M, and E proteins are located on the lipid bilayer of the viral envelope. The E (about 75 amino acids) is a small hydrophobic integral membrane protein and a multifunctional protein, supposed to be involved in virus assembly and release and pathogenesis.[4−6] The M (about 220 amino acids) is the most abundant structural protein in CoV virions and is composed of three parts: a short N-terminal domain (NTD) at the virion exterior region, three transmembrane (TM) helices, and a carboxy-terminal domain at the virion interior region.[2,7,8] The M is the primary driver of the virus budding process and directs the virion assembly by interacting with other structural proteins.[7,9,10] The S mediates the fusion process between viral and host membranes.[11] It is a homotrimer. Each monomer consists of two subunits: S1 (at the N-terminus, responsible for receptor recognition and binding) and S2 (at the C-terminus, directing the subsequent fusion process).[11,12] The cryo-electron microscopy (cryo-EM) structures of the SARS-CoV-2 S revealed its shared architecture with the SARS-CoV S, while the sequence of the SARS-CoV-2 S shares about 77% identity with that of the SARS-CoV S.[13] Both Ss recognize and bind angiotensin-converting enzyme 2 (ACE2), a zinc metallopeptidase involved in cardiovascular and immune system regulation, to enter and infect human cells.[14,15]

Received: October 12, 2021

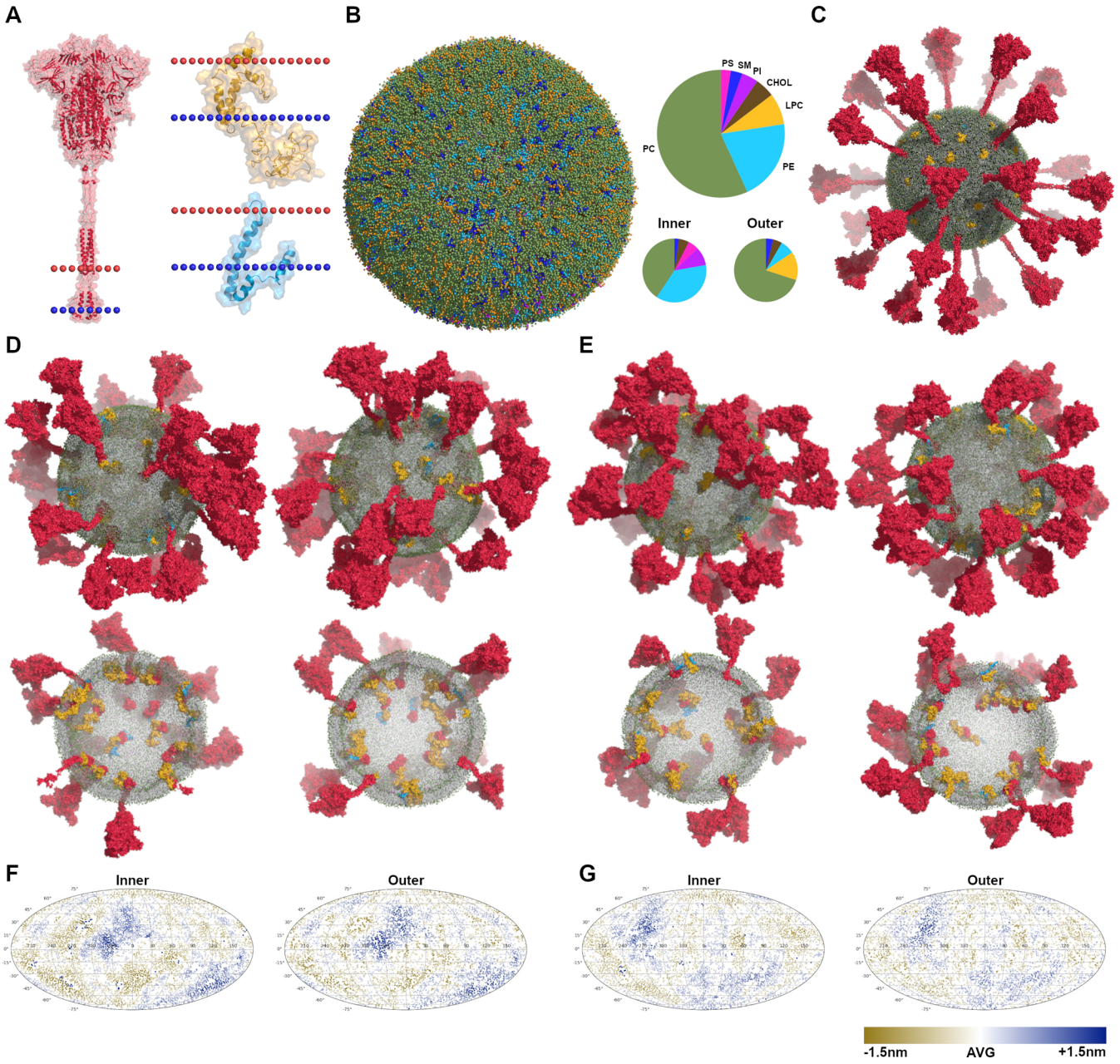


Figure 1. Martini CG models of virus envelopes. (A) Structures of the S (in red), M (in yellow), and E (in blue) with the dotted lines, indicating the position of lipid bilayers. (B) Equilibrated vesicle with the lipid composition according to that of the ER. The lipid bilayer is asymmetric, and the compositions of inner and outer leaflets are different. Lipid types considered here are phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), lyso-PC (LPC), sphingomyelin (SM), phosphatidylinositol (PI), and cholesterol (CHOL). (C) Multiple copies of proteins (E × 15 in blue, S × 30 in red, and M × 45 in yellow) were inserted into the equilibrated vesicle, making up the initial CG model of virus envelopes. (D,E) Equilibrated models after microsecond MD simulations of SARS-CoV-2-md1 and SARS-CoV-md3, illustrated by the side view and transverse section. (F,G) Mollweide projection maps of the envelope membranes of SARS-CoV-2-md1 and SARS-CoV-md3, illustrated using the distances between the lipid heads and the vesicle's center of mass.

However, the toxicity and transmission capacities of SARS-CoV and SARS-CoV-2 are significantly different. In light of the ongoing global health emergency, there is an urgent need to clarify how the envelope of the CoVs fulfills its function and explore why the infection capacity is different. New structures of SARS-CoV-2 proteins have been obtained by cryo-EM nearly weekly.[16−21] Computational approaches have also been used for the structural prediction of unresolved sections,[22] molecular dockings of different drug molecules on virus proteins,[23−26] free-energy calculations of the S-ACE2 binding process,[27−32] and the down−up transition of the receptor-binding domain (RBD).[33,34] The larger-scale spatiotemporal processes, such as virion assembly, virus architecture, fusion, and budding, are still poorly understood and remain a challenge for experiments and all-atom molecular dynamics (MD) simulations.[35]

Coarse-grained (CG) models have been proven to be powerful to probe the spatiotemporal large-scale process of complicated biomolecular systems.[36] Martini CG models have been widely used to investigate protein−lipid/protein interactions.[37,38] In this study, we constructed Martini CG models of SARS-CoV and SARS-CoV-2 envelopes containing

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

multiple copies of E, M, and S proteins and thousands of different lipids; then, microsecond (μs) MD simulations were performed to equilibrate the CG models. Overall, our simulations revealed structural and dynamic details of the virus morphology, conformations of the three structural proteins (E, M, and S), and their protein−lipid/protein interactions for both SARS-CoV and SARS-CoV-2. Our results provide insight into structural and dynamic details of the critical difference between SARS-CoV and SARS-CoV-2 envelopes.

METHODS

Simulation System Setup. A CG vesicle with an inner diameter of approximately 45 nm, that is, an outer diameter of about 53 nm, was modeled by CHARMM-GUI Martini Maker.[39] According to the coronavirus biogenesis, the virus envelope buds from the endoplasmic-reticulum (ER)−Golgi intermediate compartment,[40] whose membrane shares properties with the ER−nuclear envelope (NE)−cis-Golgi lipid territory. Therefore, the lipid composition of the virus envelope is based on the lipid composition of the membranes of the ER−NE−cis-Golgi lipid territory[41] (Figure 1 and Table S1). The vesicle was equilibrated by a 2 μs MD simulation with its radius of gyration stabilizing at ∼22.3 nm.

The genes of the E and M proteins are conserved in SARS-CoV and SARS-CoV-2 (Figure S1), so the same structures of the M and E were used in both models. The oligomerization of the E differs from monomer[42,43] to pentamer[6,43] under different experimental conditions without the M and S proteins but is still not demonstrated in the CoV membrane. We chose to involve monomers of the E in our models. The monomer structure of the E was extracted from the SARS-CoV E protein pentameric structure (PDB: 5X29),[43] while the M structure downloaded from the website of Zhang Lab[22] was used with the adjustment of the carboxy-terminal domain to the virus lumen (Figure S3). The adjusted conformation is consistent with the newly released conformations by Zhang Lab and Feig Lab (Figure S4).

The main difference between SARS-CoV and SARS-CoV-2 in the gene sequence is centered on the S protein. We used the closed-state prefusion structure of SARS-CoV S (PDB: 5X58)[44] and SARS-CoV-2 S (PDB: 6VXX).[19] For the SARS-CoV S, the missing TM region of the S was taken from the predicted structure from I-TASSER website[22] and other missing fragments such as HR2 from other structures (PDBs: 6NB6, 6B3O, 2FXP). Also, the missing fragments of the SARS-CoV-2 S were modeled based on this complete SARS-CoV S structure. The models are generally consistent with those models reported previously[45] (Figure S5). The protonation states of all residues and lipids use their states at pH = 7. There is no clear consensus on the stoichiometric composition of the structural proteins of the envelopes so far. Here, a total of 90 proteins (15 Es, 30 Ss, and 45 Ms) were inserted into the vesicles randomly, and then the lipid molecules within 2 Å of proteins were removed, resulting in 16,000 lipid molecules being left. Then, the envelope models were solvated in a dodecahedron-shaped box with Martini waters and counterions to neutralize the overall charge and 0.15 M NaCl. The counterions are evenly distributed in the systems, and there is no charge imbalance between the interior and exterior regions of the virions (Figure S6). The final simulation systems contain about 7.3 million Martini beads.

MD Simulations. All CG MD simulations were performed using GROMACS version 2019.2[46] and with Martini 2.2 force field parameters[47,48] and the ELNEDYN elastic network model on the monomers using a force constant of 500 kJ/(mol·nm^2) and a distance cutoff of 0.9 nm. The temperature was maintained at 310 K using a velocity-rescaling thermostat[49] with a time constant for the coupling of 1 ps. Proteins and lipids and the solvent were coupled separately to the temperature bath. An isotropic pressure of 1 bar was maintained with the Parrinello-Rahman[50] barostat, with a compressibility of 4.5×10^{-5} bar^{-1} and a relaxation time constant of 5 ps. A cutoff of 12 Å was used for van der Waals and electrostatic interactions with a switching function from 10 Å for van der Waals. The systems were minimized for 15 000 steps with the steepest descent method and then equilibrated by short 3.6 ns *NVT* simulations with incremental time steps of 2, 5 10, 15, and 20 fs. Finally, the production simulations were performed with the *NPT* ensemble and a time step of 20 fs. Multiple replicates were simulated for each system, with a total simulation time of 60 μs (Table S2). All the analyses were performed over the last 1 μs trajectories. Visual MD (VMD)[51] was used for data analyses, visualization, and figure rendering. Details of data analyses can be found in the Supporting Information.

RESULTS

We constructed Martini CG models of SARS-CoV and SARS-CoV-2 envelopes with Es, Ss, and Ms inserted in a lipid vesicle (Figure 1A). The lipidomics of the vesicle is according to the composition of the human ER-related membrane[41] and is asymmetric between the outer and inner leaflets (Table S1). Microsecond MD simulations were carried out to equilibrate the models. Longer simulations (up to 12 μs) did not result in observable changes in the virions, so some simulations were performed for 3 μs to limit the computational cost (Table S2).

Physical CG Models of the Virus Envelopes. The virion size varies in different reports[17,21,52] in the range of 50−120 nm for the outer diameter. In order to minimize the required computing resources, the Martini CG models constructed here, with an average outer diameter of ∼53 nm, have a total of ∼7.3 million CG beads, containing 90 TM proteins embedded in a membrane of 16,000 lipid molecules with 68% PC, 7% PE, 16% LPC, 5% CHOL, and 4% SM in the outer leaflet and 41% PC, 37% PE, 9% PI, 5% PS, 5% CHOL, and 2% SM in the inner leaflet (Figure 1A−C and Table S1).

There is no clear consensus on the stoichiometric composition of structural proteins of the virus envelopes so far. The reported number of Ss per vision varies in the range of 20−40,[17,21] so 30 Ss were first randomly inserted into the equilibrated vesicle. The prefusion conformation was used for all Ss because the cryo-EM results suggested that more than 95% of Ss on the virions are in prefusion conformation. Then, 15 Es, the minimum abundance, were located randomly and finally limited by the size of our models; as much as 45 Ms were placed, which may be less than those previously inferred.[35,53] A total of 90 structural proteins were included, occupying about 15% of the virion surface and resulting in a 1:180 protein-to-lipid ratio. These structural models enable us to probe the nanoscale organization and dynamics of the SARS-CoV and SARS-CoV-2 envelopes.

In all simulations, both the SARS-CoV and SARS-CoV-2 envelopes are stable and present very similar morphology globally (Figures 1D,E and S7−S14). The shapes of both virus

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

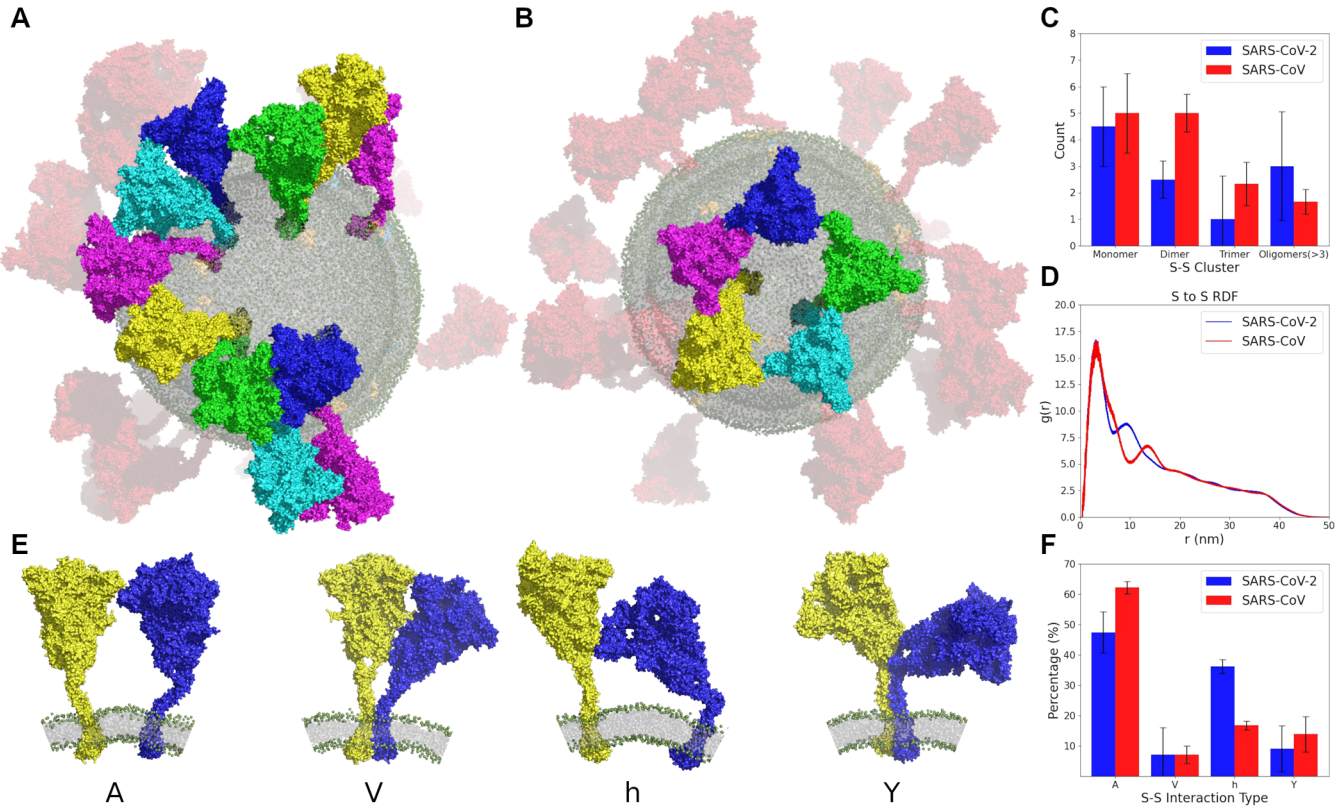


Figure 2. Spike−spike (S−S) interactions. (A,B) Examples of the curvilinear string and five-membered rings spatially organized by S heads. (C) Number of S−S clusters in curvilinear strings with different lengths. (D) RDF, $g(r)$ between Ss. (E,F) Four types of S−S interactions and their percentages.

envelopes maintain a relatively regular sphere. A spherical morphology has been observed in several cryo-EM studies.[17,20,21] The distances between individual lipid heads and the center of mass of the entire virion present slight fluctuations less than 1.5 nm. The landscapes of distance fluctuations of the inner leaflet show almost the same pattern as that of the outer leaflet (Figures 1 and S7−S14), indicating no significant variation of the bilayer thickness and curvature. The local curvature induced by TM proteins was observed in 500 ns atomistic MD simulations and was conjectured to be the driving force of the deviation from an ideal sphere.[54] In our virus models, however, no curvature changes in the membrane were observed in all simulations, which may be the reason that the virions maintain a spherical shape.

The average curvature radius (R) of the equilibrated vesicle, calculated with MemCurv software[55] before inserting the proteins is about 22.2 nm. After inserting the proteins, the vesicle shrinks slightly to about 21.9 nm in radius. During all simulations, except for the simulation of SARS-CoV-2-md1, R and the radius of gyration (R_g) of envelopes without the ectodomains of the S are almost constant with only slight fluctuations, but R_g of intact envelopes decreases from 32.8 to ∼29.1 nm, resulting from the swaying motions of the Ss at all angles (Figure S15). In the simulation SARS-CoV-2-md1, a remarkable decrease of 0.3 nm appears in the evolution of R at around 1.5 μs, induced by the stalk of one S falling on the membrane artificially (Figures S15 and S16).

Spike−Spike Interactions on the Envelopes. Overall, the homotrimer Ss organize into curvilinear strings with varying length (up to 11 Ss) and a small number of five-membered rings spatially through extensive protein−protein interactions between the huge heads (S1 domains) but without obvious aggregation (Figures 1, 2A,B and S7−S14). The morphology of spatial organization is consistent with recent cryo-EM results.[17,18,21] The SARS-CoV Ss prefer the conformations of a dimer and trimer, while SARS-CoV-2 Ss tend to connect into longer oligomers (Figure 2C). A global picture of S−S interactions may be obtained from radial distribution functions (RDFs) of Ss around Ss. The difference between SARS-CoV and SARS-CoV-2 Ss is also reflected in the average RDFs (Figure 2D). RDFs show that both first peaks appear at about 3.3 nm, but the intensity and position of the second peaks are quite different. The average second peak of SARS-CoV Ss at about 13.5 nm is weaker than that of SARS-CoV-2 Ss at about 8.7 nm. The positions of the second peak of each Ss are statistically significantly different between SARS-CoV and SARS-CoV-2 with a p-value of 0.02. These results manifest that SARS-CoV-2 Ss have a higher tendency of aggregation than SARS-CoV Ss.

The stalks of Ss in both SARS-CoV and SARS-CoV-2 systems show high flexibility, resulting in orientations in all directions and a variety of interaction patterns. According to interaction sites, the S−S interactions can be classified into four main types: A-shaped, V-shaped, h-shaped, and Y-shaped (Figure 2E,F). The A-shaped interaction is dominant (about 50% for SARS-CoV-2 and 60% for SARS-CoV), mainly through the interactions of adjacent NTDs, while the V-shaped structure is minor (about 7%) and has extensive interactions. The proportion of the h-shaped interaction (NTD-CD) has the greatest difference between SARS-CoV-2 (36%) and SARS-CoV (17%). The Y-shaped interaction contributes about 10% and has intertwining stalks and TMs.

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

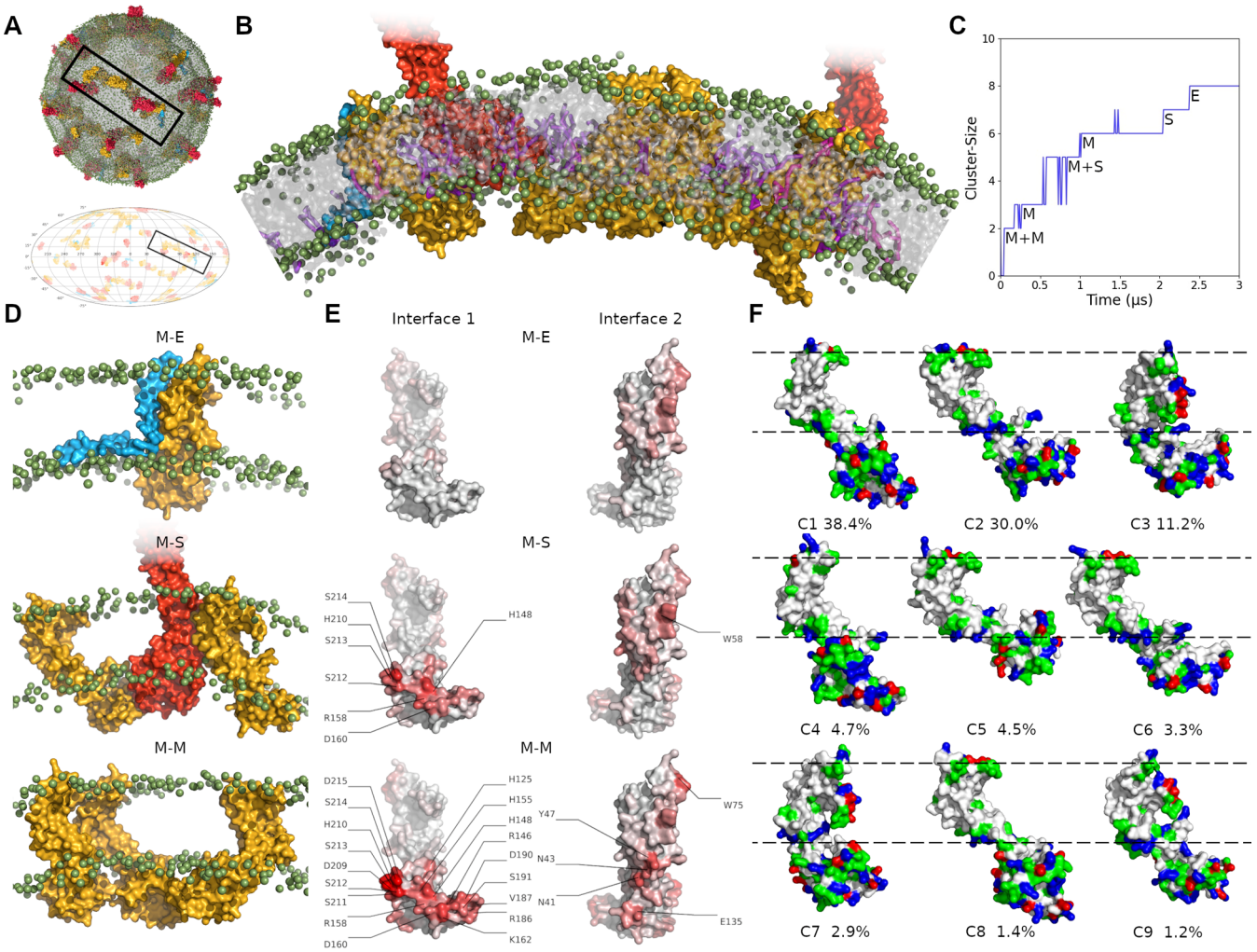


Figure 3. Protein−protein interactions in the envelope membranes. (A) Example of the equilibrated CG model of virus envelopes with the S showing only the TM domains and its Mollweide projection map of protein distributions. The islands in the black frame are enlarged in (B). The color scheme is the same as Figure 1. The lipid heads are shown in green spheres and PI lipids in purple sticks. (C) Formation process of this island over the simulation time. (D) Examples of M−E, M−S, and M−M interactions. (E) Residues involving in protein−protein interactions on the two M interfaces (interface-1 and interface-2). The shade of red indicates the percentage of the residue involved in the M−E, M−S, and M−M interaction in the simulations, averaged over all copies of Ms and all replicas of simulations. (F) Representative M structure of nine clusters accounting for more than 1%, resulting from a RMSD-based clustering analysis. The structure was colored according to the physicochemical properties of amino acids: polar residues in green, basic residues in blue, acidic residues in red, and nonpolar residues in white.

Both A-shaped[52] and Y-shaped[21] dimers of S trimers have been reported in recent cryo-EM studies.

M Mediating Protein−Protein Interactions on Virus Envelopes. Visualizing the envelopes without the S ectodomains, it is clear that the TMs of different structural proteins form string-like islands with variable size (up to 18 different proteins) in the envelopes of both SARS-CoV and SARS-CoV-2 (Figures 3A and S17−S24). The string-like islands were also observed in previous simulations of several different complex membrane models.[38,56] Most of the islands are made up of a mixture of the E, M, and S proteins. The promiscuous interactions stabilize the protein islands and will slow their diffusion. The average diffusion coefficients of the proteins, calculated from mean-square displacements of the last 1 μs (1.02 ± 0.51 for the E, 0.58 ± 0.28 for the S, and 0.80 ± 0.42 for the M), reduced to about 60% of these calculated from the first 1 μs (1.50 ± 0.42 for the E, 1.01 ± 0.58 for the S, and 1.33 ± 0.34 for the M) (Figure S25). The degree of slowing down is the same as that caused by the crowding of G-protein-coupled receptors.[57]

The Es, in the monomers in our models, have extensive contacts with the Ms and Ss, but no tendency to form pentamers was observed. It is noteworthy that more SARS-CoV Es distribute around the M and S than SARS-CoV-2 Es do (Figure S26).

The recruitment of islands is mainly driven by extensive interactions between the E/M/S and M proteins, the most abundant protein in the virus envelopes. Taking this randomly selected cluster (the one shown in Figure 3B) as an example, two Ms form a dimer in the first 0.1 μs, and then the third M, a M−S cluster, the fourth M, a S and an E were recruited one by one (Figure 3C). Two preferred interfaces of Ms mediate both homologous and heterologous protein−protein interactions (Figure 3D,E). Interface-1 involves the residues located at the ectodomain, while interface-2 involves the residues located at the TM domain (Figure 3E). Promiscuous protein−protein interactions are largely, but not exclusively, mediated by

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Table 1. D−E Index Matrix, with Average Depletion/Enrichment for Different Structural Proteins[a]

SARS-CoV	PC	PE	Chol	SM	PS	PI	LPC
S	0.56±0.02	1.67±0.07	1.03±0.04	0.21±0.03	1.72±0.06	3.43±0.09	1.26±0.03
M	0.55±0.01	1.56±0.04	0.88±0.04	0.25±0.03	2.03±0.07	3.92±0.21	1.31±0.05
E	0.57±0.01	1.21±0.03	2.22±0.34	0.51±0.04	2.18±0.32	4.67±0.21	0.81±0.12
SARS-CoV-2	PC	PE	Chol	SM	PS	PI	LPC
S	0.56±0.01	1.72±0.05	1.02±0.10	0.22±0.01	1.65±0.25	3.27±0.22	1.19±0.04
M	0.53±0.01	1.61±0.02	0.83±0.01	0.21±0.02	1.98±0.12	4.03±0.04	1.28±0.03
E	0.59±0.02	1.17±0.04	2.25±0.09	0.46±0.07	1.95±0.30	4.66±0.23	0.77±0.01

[a]The D−E index is computed by dividing the lipid composition of the first 0.7 nm shell by the bulk membrane composition.

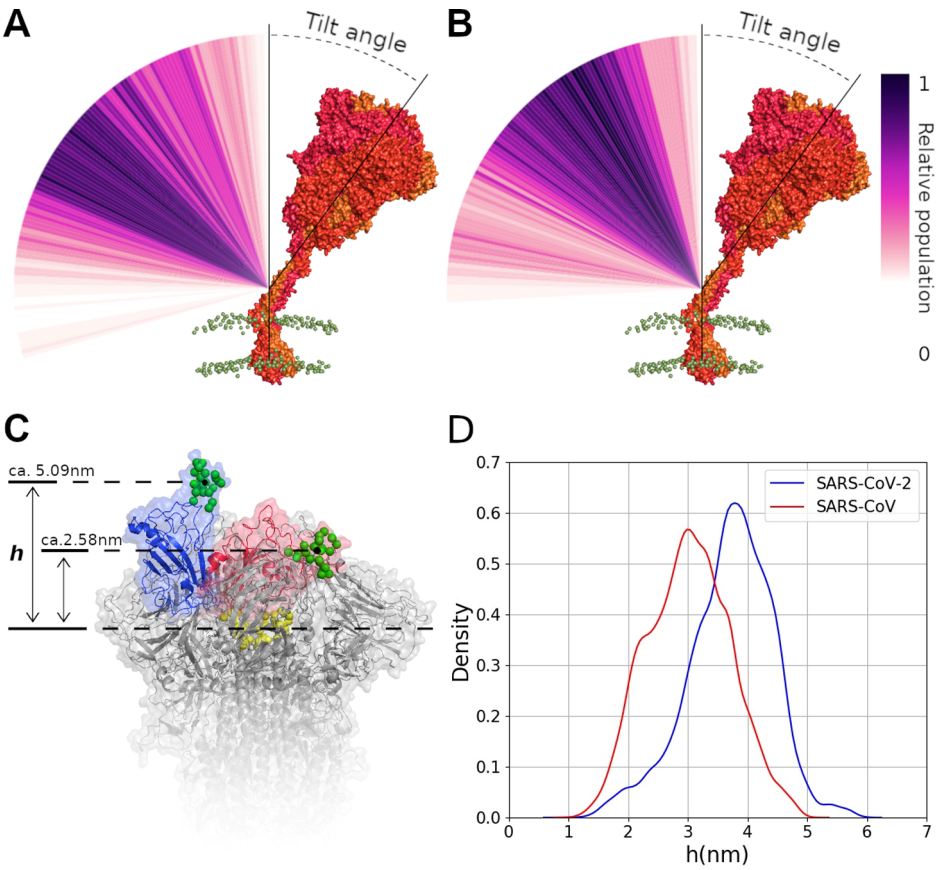


Figure 4. Flexibility of Ss. Distribution of the S tilt angle of SARS-CoV (A) and SARS-CoV-2 (B). (C) Distance (h) between the RBD tip (in green spheres) and the top of the central helix (in yellow spheres) of the closed (in red) and up-states (in blue). (D) h distributions of SARS-CoV-2 and SARS-CoV.

charged residues displayed at interface-1 and aromatic residues (W58, W75, and Y47) at interface-2. Primarily based on these two interfaces, the TM domains form M−S (interface-1—S), M−M (interface-1−Interface-2), and M−E (interface-2—E) interactions and assemble into long curvilinear islands. The residues involved in M−S and M−M interactions are almost the same. Previous morphological studies have shown that the S protein does not appear in the envelope region in the absence of the M proteins,[53] and the extensive interactions between the S and M proteins are consistent with the observation in our simulations. This indicates that the Ms may prevent excessive aggregation of the Ss.

Previous morphological studies[53,58] suggested that a part of the Ms form dimers, which also was confirmed in our simulations. The M dimer is tightly bound by the strong and extensive interactions between the two interfaces (Figure 3D,E). In the process of dimer formation, the endodomains contact first, and then the Ms gradually move closer until the TM region has contacts forming the dimer (Figure S27). The

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

endodomain of the M mediates the interaction between the dimers.

None of the previous studies provided a clear picture of the structure of the M. In the simulations, in spite of the restriction of elastic networks imposed on proteins, the M presents a variety of structures. A root-mean-square-difference-based clustering with a cutoff of 0.5 nm resulted in 31 clusters. The representative structures of nine clusters accounting for more than 1% are shown in Figure 3F. The cryo-EM study has classified the M structures into elongated and compact conformations according to the length of the endodomain.[53] Our simulations clearly demonstrated these two conformations. The endodomain has interactions with the lipid heads, resulting in the compact conformation (∼55%, clusters C2, C3, C5, C6, C7, and C9), while the elongated conformation (∼45%, clusters C1, C4, and C8) appears as an extended endodomain domain, which may interact with the RNA package.[59]

Lipid Microenvironments of Structural Proteins. A protein modulates its local lipid environment in a unique way, and lipid−protein interactions are regarded as unique fingerprints for membrane proteins.[60] We calculated the lipid depletion−enrichment (D−E) index of the three structural proteins separately (Table 1). A D−E index greater than 1 means enrichment and less than 1 means depletion. Consistent with the protein−protein interactions discussed above, the lipid environments of SARS-CoV and SARS-CoV-2 models are consistent due to the highly similar TMs.

Compared with the M and S, the E presents a significantly different lipid microenvironment. The depletion of PC and SM and the enrichment of PE, PS, and PI are consistent for all the three proteins. However, it is noteworthy that the Es have more SM and CHOL in the first layer of lipids. In the simulations, CHOLs bind nonspecifically and dissociate on a sub-microsecond timescale. The protein tends to be located in its unique lipid environment. The difference in the lipid microenvironments between the M/S and E is also reflected in our simulations with most of the Es being located at the end or inflection point of the string-like protein islands (Figures 3A,B and S17−S24). In other words, the lipid microenvironment of the E monomer may signal to terminate the further extension of the protein clusters and may contribute to the pattern of protein distribution.

PIs, with four negative charges, are highly enriched around all proteins, mediating the protein clusters (Figure 3B). Positively charged residues are abundant at the layers of lipid heads and have strong electrostatic interactions with the PI heads. These interactions ensure the stable embedding of proteins in the lipid bilayers, especially for the S, which has a large and highly flexible head but only three TM helices.

Flexibility of the Ss on the Envelopes. The Ss show high flexibility with the heads orienting in all different directions (Figures 1, S7−S14, 4A,B). The tilt angle α is defined as the angle between the orientation of the central helix and the normal axis of the envelope bilayer. The α distribution of the SARS-CoV-2 Ss ranges from 1.1 to 116.2° with the densest population at about 30°, while the α distribution of the SARS-CoV Ss shows an almost equally wide range (0.3−100°) and the densest population at about 55°. The tilt angles of the SARS-CoV-2 Ss, in the recent cryo-EM results[18,21] and all-atom MD simulations,[45] are distributed mostly at 40−50°, which is larger than our results. This inconsistency may be due to the different calculation methods of the tilt angle, different degrees of the crowding of the Ss, and the usage of the elastic network in the Martini models, which restricts the protein motions.

The distribution of α indicates that the SARS-CoV-2 Ss prefer a more standing conformation compared with the SARS-CoV Ss. The sequences of the stalks are conserved between SARS-CoV and SARS-CoV-2, so the difference in distribution may result from the different interaction patterns. The flexibility of the S makes it easier to search for receptor proteins, but excessive flexibility is not conducive to the formation of stable interactions. Therefore, compared with that of SARS-CoV, the lower flexibility of the S heads of SARS-CoV-2 may be one of the reasons why SARS-CoV-2 has a stronger infection ability.

Two conformations of the Ss, "RBD down" and "one RBD up", were observed in our simulations, same as the observation of the virions by cryo-EM.[17,18,21] We calculated the distance (h) between residues of the RBD tip (residues 470−490 of SARS-CoV-2 and residues 457−477 of SARS-CoV) and residues of the central helix (residues 986−996 of SARS-CoV-2 and residues 968−978 of SARS-CoV) (Figure 4C). The populations of h in SARS-CoV systems are between 1.2 and 5.0 nm with a peak at 3.0 nm, while in SARS-CoV-2 systems, they are between 1.3 and 6.0 nm with a peak at 3.8 nm (Figure 4D). The distributions clearly reveal that SARS-CoV-2 has more RBDs in the up-state than SARS-CoV. Only the Ss with the RBD in the up-state can bind to ACE2 and infect cells.[16]

DISCUSSION

Microsecond MD simulations with Martini CG models were carried out to explore the dynamics and supramolecular organization of SARS-CoV and SARS-CoV-2 envelopes at the experimental length scale. The presence of multiple copies of different membrane proteins allows us to get statistically significant information on protein−protein/lipid interactions. The system complexity is comparable to those accessible by experiments such as cryo-EM. These CG models therefore provide a valuable complement to investigate the architectures of SARS-CoV and SARS-CoV-2 envelopes.

The virus envelopes maintain their spherical morphology in all simulations. Cryo-EM experiments showed both spherical and ellipsoidal viruses.[17,21] This inconsistency may be because our simulations are limited to several microseconds, and we did not consider the crowding between viruses. The change in the envelope morphology corresponds to the change of the bilayer curvature. The local curvature introduced by proteins is related to their structural symmetry,[60] while the TM domains of the E, M, and S are small and not symmetric. The oligomerization of E in the virus membrane is still uncertain. The monomer E was adopted in our models, but some works suggested that the E may form a homopentamer,[6] which may induce a local membrane curvature. However, the content of the E is the least, so its influence on the membrane curvature should be limited. Therefore, the ellipsoidal morphology may be mainly caused by the crowding between viruses.

In the simulations, we are able to observe the formation of S clusters and supramolecular islands by the TM domains of different proteins and lipids in the envelope membrane. The interaction patterns and flexibility of the S stalks show agreements with the experimental observations[17,21,52] and all-atom simulations,[18,61] indicating the reliability of the structural and dynamic details obtained from our simulations. The

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

formation of S clusters can restrict the orientation of the S heads, while the supramolecular islands slow down the S diffusion in the membrane. Both these points favor the virus infection by stabilizing the interactions between S and ACE2. Previous experiments have confirmed that a nanocluster is required for efficient pathogen binding and internalization.[62]

Our models, of course, are just approximate to the virus envelopes. In particular, the lipidomics and stoichiometric composition of structural proteins of the virus envelopes have not been accurately determined. The numbers of Ss and Ms per virion, reported in the experimental results, were speculated by assuming uniform distribution over the virus surface and also depend on experimental conditions.[53] Our current models described the virus envelopes under only one stoichiometric composition. Compared with the stoichiometry adopted by Yu and his collaborators in their SARS-CoV-2 model,[35] our models have more Ss and less Ms, which may influence the sizes of protein clusters. However, the size of the string-like protein island does not affect the mode of protein interactions.

The lipid composition of the ER-related membrane features less CHOLs and anionic lipids than the plasma membrane-related membrane.[41,63] The enrichment of CHOLs was only observed around Es, mainly due to nonspecific interactions. The anionic lipids, PS and PI, make up about 15% of the inner leaflet and are enriched around all E, M, and S proteins as most proteins do.[37] Therefore, the content of CHOL and anionic lipids appears to have little effect on protein−protein interactions in the SARS-CoV and SARS-CoV-2 envelopes.

The S protein has a dense coating of glycans to evade the host immune system. Atomistic simulations also demonstrated that the glycans play an essential role in modulating the conformational transitions of the S protein.[18,33,34,61,64] Our models in this study did not include glycans. Apparently, glycans are not involved in the protein−protein interactions of the TM domains, which form the heterogeneous protein islands in the virus membrane, while the glycosylation may affect protein−protein interactions between the S heads and the size of the S clusters. However, the patterns of protein−protein interactions in our models are consistent with these observed in the cryo-EM, as well as these from all-atom simulations, which contain four copies of glycosylated S proteins. It indicates that glycosylation may have little effect on the interaction modes between S proteins. Anyway, to get a deeper understanding of the virion structures, further studies are underway: careful parameterization of the glycosylation and modeling and simulation of virus envelopes with glycosylation, different protein and lipid stoichiometric compositions, and expanded size dimensions.

Another possible limitation of our simulations is the use of Martini CG force field 2.2, which may limit the protein conformational changes due to the elastic network,[65] such as the RBD opening in our simulations, and may tend to excessive protein aggregation because of excessive interprotein interactions. Future simulations will try the latest Martini force fields 3.0.[66]

CONCLUSIONS

Our Martini CG models illustrate SARS-CoV and SARS-CoV-2 envelopes at the atomistic level and the experimental complexity and scale. The structural proteins are not uniformly distributed over the envelopes. Most of the Ss form oligomers with extensive interactions between their huge heads, while the TM domains of structural proteins clusters into heterogeneous string-like islands mediated by negatively charged lipids. Our simulations also revealed that the critical difference between SARS-CoV and SARS-CoV-2 envelopes lies in the S−S interaction patterns and the intrinsic flexibilities of the Ss. The SARS-CoV-2 Ss have more inclination to interact with each other and higher intrinsic flexibility to recognize and bind the receptors than the SARS-CoV Ss, which may be two of the reasons that SARS-CoV-2 caused more infections than SARS-CoV. The structural and dynamic details of our models provide an improved understanding of the virus envelopes and could be used for further studies, such as drug design and fusion and fission processes.

DATA AND SOFTWARE AVAILABILITY

The initial vesicle was generated with the webserver CHARMMGUI (https://charmm-gui.org/), which is free for academic use. GROMACS 5.1.2 and VMD 1.9.2 are also free software for performing MD simulations and for displaying and analyzing MD trajectories. Mollweide projection maps were plotted by the free Python module, Matplotlib (https://matplotlib.org/). Coordinate files of the CG models of the last frames from all simulations, without solvent molecules, are available at GitHub (https://github.com/ChangqingZhong/Martini-MD-of-SARS-CoV-and-SARS-CoV-2).

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c01240.

Methods providing details of data analysis, lipid composition of our membrane models, simulation information, sequence alignment of the three TM proteins, details of initial structural models of three TM proteins, distribution of the counterions, details of the equilibrated CG models of SARS-CoV and SARS-CoV-2 of all simulations, average curvature radius and radius of gyration over the simulation time, structural details of the simulation of SARS-CoV-2-md1, protein distributions in all simulations, lateral diffusion coefficient of three TM proteins, interprotein RDFs, and formation of M dimers (PDF)

One of the SARS-CoV-2 simulations (MP4)

AUTHOR INFORMATION

Corresponding Author

Beibei Wang − *Centre for Advanced Materials Research, Advanced Institute of Natural Sciences, Beijing Normal University at Zhuhai, Zhuhai 519087, China;* orcid.org/0000-0003-2193-3777; Email: bbwang@bnu.edu.cn

Authors

Changqing Zhong − *Centre for Informational Biology, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 611731, China*

D. Peter Tieleman − *Department of Biological Sciences and Centre for Molecular Simulation, University of Calgary, Calgary T2N 1N4, Canada;* orcid.org/0000-0001-5507-0688

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c01240

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council (Canada), the Canada Research Chairs Program, the National Natural Science Foundation of China (nos. 31971176 and 31800616), and the Fundamental Research Funds for the Central Universities (no. A03018023601045). Calculations were run on Compute Canada facilities, funded by the Canada Foundation for Innovation and partners.

■ REFERENCES

(1) Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; Chen, H.-D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R.-D.; Liu, M.-Q.; Chen, Y.; Shen, X.-R.; Wang, X.; Zheng, X.-S.; Zhao, K.; Chen, Q.-J.; Deng, F.; Liu, L.-L.; Yan, B.; Zhan, F.-X.; Wang, Y.-Y.; Xiao, G.-F.; Shi, Z.-L. A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin. *Nature* **2020**, *579*, 270−273.

(2) Masters, P. S. The Molecular Biology of Coronaviruses. *Adv. Virus Res.* **2006**, *66*, 193−292.

(3) Alsaadi, E. A. J.; Jones, I. M. Membrane Binding Proteins of Coronaviruses. *Future Virol.* **2019**, *14*, 275−286.

(4) Wilson, L.; Gage, P.; Ewart, G. Hexamethylene Amiloride Blocks E Protein Ion Channels and Inhibits Coronavirus Replication. *Virology* **2006**, *353*, 294−306.

(5) Ruch, T. R.; Machamer, C. E. The Coronavirus E Protein: Assembly and Beyond. *Viruses* **2012**, *4*, 363−382.

(6) Mandala, V. S.; McKay, M. J.; Shcherbakov, A. A.; Dregni, A. J.; Kolocouris, A.; Hong, M. Structure and Drug Binding of the SARS-CoV-2 Envelope Protein Transmembrane Domain in Lipid Bilayers. *Nat. Struct. Mol. Biol.* **2020**, *27*, 1202−1208.

(7) Arndt, A. L.; Larson, B. J.; Hogue, B. G. A Conserved Domain in the Coronavirus Membrane Protein Tail Is Important for Virus Assembly. *J. Virol.* **2010**, *84*, 11418−11428.

(8) Ujike, M.; Taguchi, F. Incorporation of Spike and Membrane Glycoproteins into Coronavirus Virions. *Viruses* **2015**, *7*, 1700−1725.

(9) Boscarino, J. A.; Logan, H. L.; Lacny, J. J.; Gallagher, T. M. Envelope Protein Palmitoylations Are Crucial for Murine Coronavirus Assembly. *J. Virol.* **2008**, *82*, 2989−2999.

(10) Kuo, L.; Masters, P. S. Genetic Evidence for a Structural Interaction between the Carboxy Termini of the Membrane and Nucleocapsid Proteins of Mouse Hepatitis Virus. *J. Virol.* **2002**, *76*, 4987−4999.

(11) Belouzard, S.; Millet, J. K.; Licitra, B. N.; Whittaker, G. R. Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. *Viruses* **2012**, *4*, 1011−1033.

(12) Cai, Y.; Zhang, J.; Xiao, T.; Peng, H.; Sterling, S. M.; Walsh, R. M., Jr.; Rawson, S.; Rits-Volloch, S.; Chen, B. Distinct Conformational States of SARS-CoV-2 Spike Protein. *Science* **2020**, *369*, 1586−1592.

(13) Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; Chen, H.-D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R.-D.; Liu, M.-Q.; Chen, Y.; Shen, X.-R.; Wang, X.; Zheng, X.-S.; Zhao, K.; Chen, Q.-J.; Deng, F.; Liu, L.-L.; Yan, B.; Zhan, F.-X.; Wang, Y.-Y.; Xiao, G.-F.; Shi, Z.-L. A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin. *Nature* **2020**, *579*, 270−273.

(14) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; Müller, M. A.; Drosten, C.; Pöhlmann, S. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **2020**, *181*, 271−280.

(15) Turner, A. J.; Hiscox, J. A.; Hooper, N. M. ACE2: from Vasopeptidase to SARS Virus Receptor. *Trends Pharmacol. Sci.* **2004**, *25*, 291−294.

(16) Benton, D. J.; Wrobel, A. G.; Xu, P.; Roustan, C.; Martin, S. R.; Rosenthal, P. B.; Skehel, J. J.; Gamblin, S. J. Receptor Binding and Priming of the Spike Protein of SARS-CoV-2 for Membrane Fusion. *Nature* **2020**, *588*, 327−330.

(17) Ke, Z.; Oton, J.; Qu, K.; Cortese, M.; Zila, V.; McKeane, L.; Nakane, T.; Zivanov, J.; Neufeldt, C. J.; Cerikan, B.; Lu, J. M.; Peukes, J.; Xiong, X.; Kräusslich, H.-G.; Scheres, S. H. W.; Bartenschlager, R.; Briggs, J. A. G. Structures and Distributions of SARS-CoV-2 Spike Proteins on Intact Virions. *Nature* **2020**, *588*, 498−502.

(18) Turoňová, B.; Sikora, M.; Schuermann, C.; Hagen, W. J. H.; Welsch, S.; Blanc, F. E. C.; von Buelow, S.; Gecht, M.; Bagola, K.; Hoerner, C.; van Zandbergen, G.; Landry, J.; de Azevedo, N. T. D.; Mosalaganti, S.; Schwarz, A.; Covino, R.; Muehlebach, M. D.; Hummer, G.; Locker, J. K.; Beck, M. In Situ Structural Analysis of SARS-CoV-2 Spike Reveals Flexibility Mediated by Three Hinges. *Science* **2020**, *370*, 203−208.

(19) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**, *181*, 281−292.

(20) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural Basis for the Recognition of SARS-CoV-2 by Full-length Human ACE2. *Science* **2020**, *367*, 1444−1448.

(21) Yao, H.; Song, Y.; Chen, Y.; Wu, N.; Xu, J.; Sun, C.; Zhang, J.; Weng, T.; Zhang, Z.; Wu, Z.; Cheng, L.; Shi, D.; Lu, X.; Lei, J.; Crispin, M.; Shi, Y.; Li, L.; Li, S. Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **2020**, *183*, 730−738.

(22) Zhang, C.; Zheng, W.; Huang, X.; Bell, E. W.; Zhou, X.; Zhang, Y. Protein Structure and Sequence Reanalysis of 2019-nCoV Genome Refutes Snakes as Its Intermediate Host and the Unique Similarity between Its Spike Protein Insertions and HIV-1. *J. Proteome Res.* **2020**, *19*, 1351−1360.

(23) Alfaro, M.; Alfaro, I.; Angel, C. Identification of Potential Inhibitors of SARS-CoV-2 Papain-like Protease from Tropane Alkaloids from Schizanthus Porrigens: A Molecular Docking Study. *Chem. Phys. Lett.* **2020**, *761*, 138068.

(24) Elzupir, A. O. Inhibition of SARS-CoV-2 Main Protease 3CLpro by Means of Alpha-ketoamide and Pyridone-containing Pharmaceuticals Using in Silico Molecular Docking. *J. Mol. Struct.* **2020**, *1222*, 128878.

(25) Pandey, P.; Rane, J. S.; Chatterjee, A.; Kumar, A.; Khan, R.; Prakash, A.; Ray, S. Targeting SARS-CoV-2 Spike Protein of COVID-19 with Naturally Occurring Phytochemicals: an in Silico Study for Drug Development. *J. Biomol. Struct. Dyn.* **2021**, *39*, 6306−6316.

(26) Sahu, S. N.; Mishra, B.; Sahu, R.; Pattanayak, S. K. Molecular Dynamics Simulation Perception Study of the Binding Affinity Performance for Main Protease of SARS-CoV-2. *J. Biomol. Struct. Dyn.* **2020**, DOI: 10.1080/07391102.2020.1850362.

(27) Bai, C.; Warshel, A. Critical Differences between the Binding Features of the Spike Proteins of SARS-CoV-2 and SARS-CoV. *J. Phys. Chem. B* **2020**, *124*, 5907−5912.

(28) García-Iriepa, C.; Hognon, C.; Francés-Monerris, A.; Iriepa, I.; Miclot, T.; Barone, G.; Monari, A.; Marazzi, M. Thermodynamics of the Interaction between the Spike Protein of Severe Acute Respiratory Syndrome Coronavirus-2 and the Receptor of Human Angiotensin-Converting Enzyme 2. Effects of Possible Ligands. *J. Phys. Chem. Lett.* **2020**, *11*, 9272−9281.

(29) Laurini, E.; Marson, D.; Aulic, S.; Fermeglia, M.; Pricl, S. Computational Alanine Scanning and Structural Analysis of the SARS-CoV-2 Spike Protein/Angiotensin-Converting Enzyme 2 Complex. *ACS Nano* **2020**, *14*, 11821−11830.

(30) Spinello, A.; Saltalamacchia, A.; Magistrato, A. Is the Rigidity of SARS-CoV-2 Spike Receptor-Binding Motif the Hallmark for Its Enhanced Infectivity? Insights from All-Atom Simulations. *J. Phys. Chem. Lett.* **2020**, *11*, 4785−4790.

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

(31) Wang, Y.; Liu, M.; Gao, J. Enhanced Receptor Binding of SARS-CoV-2 through Networks of Hydrogen-bonding and Interactions. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 13967−13974.
(32) Andrade, J.; Gonçalves, P. F. B.; Netz, P. A. Why Does the Novel Coronavirus Spike Protein Interact so Strongly with the Human ACE2? A Thermodynamic Answer. *ChemBioChem* **2021**, *22*, 865−875.
(33) Fallon, L.; Belfon, K. A. A.; Raguette, L.; Wang, Y.; Stepanenko, D.; Cuomo, A.; Guerra, J.; Budhan, S.; Varghese, S.; Corbo, C. P.; Rizzo, R. C.; Simmerling, C. Free Energy Landscapes from SARS-CoV-2 Spike Glycoprotein Simulations Suggest that RBD Opening Can Be Modulated via Interactions in an Allosteric Pocket. *J. Am. Chem. Soc.* **2021**, *143*, 11349−11360.
(34) Sztain, T.; Ahn, S.-H.; Bogetti, A. T.; Casalino, L.; Goldsmith, J. A.; Seitz, E.; McCool, R. S.; Kearns, F. L.; Acosta-Reyes, F.; Maji, S.; Mashayekhi, G.; McCammon, J. A.; Ourmazd, A.; Frank, J.; McLellan, J. S.; Chong, L. T.; Amaro, R. E. A Glycan Gate Controls Opening of the SARS-CoV-2 Spike Protein. *Nat. Chem.* **2021**, *13*, 963−968.
(35) Yu, A.; Pak, A. J.; He, P.; Monje-Galvan, V.; Casalino, L.; Gaieb, Z.; Dommer, A. C.; Amaro, R. E.; Voth, G. A. A Multiscale Coarse-grained Model of the SARS-CoV-2 Virion. *Biophys. J.* **2021**, *120*, 1097−1104.
(36) Ingólfsson, H. I.; Lopez, C. A.; Uusitalo, J. J.; de Jong, D. H.; Gopal, S. M.; Periole, X.; Marrink, S. J. The Power of Coarse Graining in Biomolecular Simulations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 225−248.
(37) Corradi, V.; Sejdiu, B. I.; Mesa-Galloso, H.; Abdizadeh, H.; Noskov, S. Y.; Marrink, S. J.; Tieleman, D. P. Emerging Diversity in Lipid-Protein Interactions. *Chem. Rev.* **2019**, *119*, 5775−5848.
(38) Marrink, S. J.; Corradi, V.; Souza, P. C. T.; Ingólfsson, H. I.; Tieleman, D. P.; Sansom, M. S. P. Computational Modeling of Realistic Cell Membranes. *Chem. Rev.* **2019**, *119*, 6184−6226.
(39) Qi, Y.; Ingólfsson, H. I.; Cheng, X.; Lee, J.; Marrink, S. J.; Im, W. CHARMM-GUI Martini Maker for Coarse-Grained Simulations with the Martini Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 4486−4494.
(40) Klein, S.; Cortese, M.; Winter, S. L.; Wachsmuth-Melm, M.; Neufeldt, C. J.; Cerikan, B.; Stanifer, M. L.; Boulant, S.; Bartenschlager, R.; Chlanda, P. SARS-CoV-2 Structure and Replication Characterized by in Situ Cryo-Electron Tomography. *Nat. Commun.* **2020**, *11*, 5885.
(41) Jackson, C. L.; Walch, L.; Verbavatz, J.-M. Lipids and Their Trafficking: An Integral Part of Cellular Organization. *Dev. Cell* **2016**, *39*, 139−153.
(42) Li, Y.; Surya, W.; Claudine, S.; Torres, J. Structure of a Conserved Golgi Complex-targeting Signal in Coronavirus Envelope Proteins. *J. Biol. Chem.* **2014**, *289*, 12535−12549.
(43) Surya, W.; Li, Y.; Torres, J. Structural model of the SARS coronavirus E channel in LMPG micelles. *Biochim. Biophys. Acta, Biomembr.ranes* **2018**, *1860*, 1309−1317.
(44) Yuan, Y.; Cao, D.; Zhang, Y.; Ma, J.; Qi, J.; Wang, Q.; Lu, G.; Wu, Y.; Yan, J.; Shi, Y.; Zhang, X.; Gao, G. F. Cryo-EM Structures of MERS-CoV and SARS-CoV Spike Glycoproteins Reveal the Dynamic Receptor Binding Domains. *Nat. Commun.* **2017**, *8*, 15092.
(45) Choi, Y. K.; Cao, Y.; Frank, M.; Woo, H.; Park, S.-J.; Yeom, M. S.; Croll, T. I.; Seok, C.; Im, W. Structure, Dynamics, Receptor Binding, and Antibody Binding of the Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein in a Viral Membrane. *J. Chem. Theory Comput.* **2021**, *17*, 2479−2487.
(46) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(47) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 7812−7824.
(48) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.
(49) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
(50) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(51) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.
(52) Bangaru, S.; Ozorowski, G.; Turner, H. L.; Antanasijevic, A.; Huang, D.; Wang, X.; Torres, J. L.; Diedrich, J. K.; Tian, J.-H.; Portnoff, A. D.; Patel, N.; Massare, M. J.; Yates, J. R., III; Nemazee, D.; Paulson, J. C.; Glenn, G.; Smith, G.; Ward, A. B. Structural Analysis of Full-length SARS-CoV-2 Spike Protein from an Advanced Vaccine Candidate. *Science* **2020**, *370*, 1089−1094.
(53) Neuman, B. W.; Kiss, G.; Kunding, A. H.; Bhella, D.; Baksh, M. F.; Connelly, S.; Droese, B.; Klaus, J. P.; Makino, S.; Sawicki, S. G.; Siddell, S. G.; Stamou, D. G.; Wilson, I. A.; Kuhn, P.; Buchmeier, M. J. A Structural Analysis of M Protein in Coronavirus Assembly and Morphology. *J. Struct. Biol.* **2011**, *174*, 11−22.
(54) Singharoy, A.; Maffeo, C.; Delgado-Magnero, K. H.; Swainsbury, D. J. K.; Sener, M.; Kleinekathöfer, U.; Vant, J. W.; Nguyen, J.; Hitchcock, A.; Isralewitz, B.; Teo, I.; Chandler, D. E.; Stone, J. E.; Phillips, J. C.; Pogorelov, T. V.; Mallus, M. I.; Chipot, C.; Luthey-Schulten, Z.; Tieleman, D. P.; Hunter, C. N.; Tajkhorshid, E.; Aksimentiev, A.; Schulten, K. Atoms to Phenotypes: Molecular Design Principles of Cellular Energy Metabolism. *Cell* **2019**, *179*, 1098−1111.
(55) Bhaskara, R. M.; Grumati, P.; Garcia-Pardo, J.; Kalayil, S.; Covarrubias-Pinto, A.; Chen, W.; Kudryashev, M.; Dikic, I.; Hummer, G. Curvature Induction and Membrane Remodeling by FAM134B Reticulon Homology Domain Assist Selective ER-phagy. *Nat. Commun.* **2019**, *10*, 2320.
(56) Rassam, P.; Copeland, N. A.; Birkholz, O.; Tóth, C.; Chavent, M.; Duncan, A. L.; Cross, S. J.; Housden, N. G.; Kaminska, R.; Seger, U.; Quinn, D. M.; Garrod, T. J.; Sansom, M. S. P.; Piehler, J.; Baumann, C. G.; Kleanthous, C. Supramolecular Assemblies Underpin Turnover of Outer Membrane Proteins in Bacteria. *Nature* **2015**, *523*, 333−336.
(57) Koldsø, H.; Sansom, MS Organization and Dynamics of Receptor Proteins in a Plasma Membrane. *J. Am. Chem. Soc.* **2015**, *137*, 14694−14704.
(58) de Haan, C. A. M.; Vennema, H.; Rottier, P. J. M. Assembly of the Coronavirus Envelope: Homotypic Interactions between the M Proteins. *J. Virol.* **2000**, *74*, 4967−4978.
(59) Narayanan, K.; Chen, C.-J.; Maeda, J.; Makino, S. Nucleocapsid-independent Specific Viral RNA Packaging via Viral Envelope Protein and Viral RNA Signal. *J. Virol.* **2003**, *77*, 2922−2927.
(60) Corradi, V.; Mendez-Villuendas, E.; Ingólfsson, H. I.; Gu, R.-X.; Siuda, I.; Melo, M. N.; Moussatova, A.; DeGagné, L. J.; Sejdiu, B. I.; Singh, G.; Wassenaar, T. A.; Delgado Magnero, K.; Marrink, S. J.; Tieleman, D. P. Lipid-Protein Interactions Are Unique Fingerprints for Membrane Proteins. *ACS Cent. Sci.* **2018**, *4*, 709−717.
(61) Sikora, M.; von Bülow, S.; Blanc, F. E. C.; Gecht, M.; Covino, R.; Hummer, G. Computational Epitope Map of SARS-CoV-2 Spike Protein. *PLoS Comput. Biol.* **2021**, *17*, No. e1008790.
(62) Manzo, C.; Torreno-Pina, J. A.; Joosten, B.; Reinieren-Beeren, I.; Gualda, E. J.; Loza-Alvarez, P.; Figdor, C. G.; Garcia-Parajo, M. F.; Cambi, A. The Neck Region of the C-type Lectin DC-SIGN Regulates Its Surface Spatiotemporal Organization and Virus-binding Capacity on Antigen-presenting Cells. *J. Biol. Chem.* **2012**, *287*, 38946−38955.
(63) Casares, D.; Escribá, P. V.; Rosselló, C. A. Membrane Lipid Composition: Effect on Membrane and Organelle Structure, Function and Compartmentalization and Therapeutic Avenues. *Int. J. Mol. Sci.* **2019**, *20*, 2167.
(64) Casalino, L.; Gaieb, Z.; Goldsmith, J. A.; Hjorth, C. K.; Dommer, A. C.; Harbison, A. M.; Fogarty, C. A.; Barros, E. P.; Taylor, B. C.; McLellan, J. S.; Fadda, E.; Amaro, R. E. Beyond Shielding: The

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX

Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent. Sci.* **2020**, *6*, 1722−1734.
(65) Periole, X.; Cavalli, M.; Marrink, S.-J.; Ceruso, M. A. Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition. *J. Chem. Theory Comput.* **2009**, *5*, 2531−2543.
(66) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H.; Wassenaar, T. A.; Kroon, P. C.; Melcr, J.; Nieto, V.; Corradi, V.; Khan, H. M.; Domański, J.; Javanainen, M.; Martinez-Seara, H.; Reuter, N.; Best, R. B.; Vattulainen, I.; Monticelli, L.; Periole, X.; Tieleman, D. P.; de Vries, A. H.; Marrink, S. J. Martini 3: a General Purpose Force Field for Coarse-grained Molecular Dynamics. *Nat. Methods* **2021**, *18*, 382−388.

https://doi.org/10.1021/acs.jcim.1c01240
J. Chem. Inf. Model. XXXX, XXX, XXX−XXX