Binding mode of SARS-CoV-2 fusion peptide to human cellular membrane

Defne Gorgun,1 Muyun Lihan,1 Karan Kapoor,1 and Emad Tajkhorshid1,*
1Theoretical and Computational Biophysics Group, NIH Center for Macromolecular Modeling and Bioinformatics, Beckman Institute for Advanced Science and Technology, Department of Biochemistry, and Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois

ABSTRACT Infection of human cells by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV2) relies on its binding to a specific receptor and subsequent fusion of the viral and host cell membranes. The fusion peptide (FP), a short peptide segment in the spike protein, plays a central role in the initial penetration of the virus into the host cell membrane, followed by the fusion of the two membranes. Here, we use an array of molecular dynamics simulations that take advantage of the highly mobile membrane mimetic model to investigate the interaction of the SARS-CoV2 FP with a lipid bilayer representing mammalian cellular membranes at an atomic level and to characterize the membrane-bound form of the peptide. Six independent systems were generated by changing the initial positioning and orientation of the FP with respect to the membrane, and each system was simulated in five independent replicas, each for 300 ns. In 73% of the simulations, the FP reaches a stable, membrane-bound configuration, in which the peptide deeply penetrated into the membrane. Clustering of the results reveals three major membrane-binding modes (binding modes 1–3), in which binding mode 1 populates over half of the data points. Taking into account the sequence conservation among the viral FPs and the results of mutagenesis studies establishing the role of specific residues in the helical portion of the FP in membrane association, the significant depth of penetration of the whole peptide, and the dense population of the respective cluster, we propose that the most deeply inserted membrane-bound form (binding mode 1) represents more closely the biologically relevant form. Analysis of FP-lipid interactions shows the involvement of specific residues, previously described as the “fusion-active core residues,” in membrane binding. Taken together, the results shed light on a key step involved in SARS-CoV2 infection, with potential implications in designing novel inhibitors.

INTRODUCTION Coronavirus disease 2019 (COVID-19) emerged in late 2019 as a significant threat to human health. It became a global pandemic by March 2020 (1,2), and it continues to claim lives and to significantly impact all aspects of people’s lives around the globe. COVID-19 is caused by severe acute respiratory syndrome (SARS) coronavirus 2 (SARS-CoV2), a positive-strand RNA virus that causes severe respiratory complications, among other symptoms, in humans (3). SARS-CoV2 recognizes and infects human cells that express a cell surface receptor termed angiotensin-converting enzyme 2 (ACE2) (4), which is specifically recognized by the viral spike glycoprotein (S-protein). Binding of the two proteins is a prerequisite for the fusion of the viral and cellular membranes (5), one of the first and required steps in viral infection that facilitates the release of the viral genome into the infected cell (6–9).

Binding of the virus to the ACE2 receptor on the host cell is mediated by the S1 domain in the viral S-protein. The next key step, namely, virus-host membrane fusion, is mediated by the S2 domain (10). S2 consists of multiple proteolytic cleavage sites, including one at the boundary of S1/S2 and...
one at the S2' site, which are cleaved as part of the fusion process (11–13). Cleavages at these sites, downstream to the two heptad repeat regions (HR1 and HR2), induce the dissociation of the S1 subunits from the S-protein, followed by a series of conformational changes that trigger membrane fusion between the host cell membrane and the viral envelope (14,15). The remaining S2 trimer, a postfusion structural motif, is shared among all the class I viral fusion proteins (16,17). A critical part of any viral fusion protein in the coronavirus family is the relatively hydrophobic fusion peptide (FP), a segment in the S2 domain that is responsible for directly interacting with and inserting into the host cell membrane, thereby initiating the fusion process (8,18,19).

Viral FPs share several characteristics that help locate their position within their parent protein (the S-protein in SARS-CoV2, for example). The sequences of the FPs are highly conserved within each family of viruses (but not between families), the frequency of glycines and alanines in the sequence is relatively high, and their cleavage sites are often flanked by bulky hydrophobic residues as well as hydrophilic residues (20,21). Some FPs have a central kink produced by a proline and a helix-turn-helix structure, as observed in influenza virus hemagglutinin. In such cases, proteolytic cleavage occurs directly at the N-terminus to the FP, and the peptides are thus called external or N-terminal FPs. In other cases, the proteolytic cleavage site resides upstream from the FP, which is relatively longer (25–30 amino acids) and contains a prolonged α-helix in the fusion-active state (e.g., in the cases of Ebola virus or avian leukosis sarcoma virus (8,20,21)). Such FPs are referred to as internal FPs, which also is the case for SARS-CoV2.

To this date, three main FP regions have been proposed for severe acute respiratory syndrome coronavirus (SARS-CoV) S-proteins (22,23), which are all located between the N-terminus and the HR1 region of the S2 domain: 1) at the N-terminus of HR1, 2) near the S1/S2 cleavage site, the N-terminus and the HR1 region of the S2 domain: 1) and 3) at the C-terminus of the cleavage site S2' (13). Based on the criteria stated above, as well as other experimental evidence, most recent data suggest that immediately downstream of the S2' cleavage site of the SARS-CoV2 is the leading segment involved in the fusion process (11,24–27). Mutagenesis experiments showed the significance of the FP in this part of the protein, specifically terminal α-subfamily. Out of these, S-protein structures are available for HKU1, Middle East respiratory syndrome, SARS, and SARS-CoV2 from the β-subfamily. Out of these, S-protein structures are available for HKU1, Middle East respiratory syndrome, SARS, and SARS-CoV2. Additionally, bat coronavirus RaTG13, a closely related homolog of SARS-CoV2, was also included in the sequence alignment. Multiple sequence alignment for the above eight sequences was carried out using the MAFFT program with the L-INS-i method (50) and visualized using Jalview (51) (Fig. S1). The sequence alignment highlighted the high degree of residue conservation among the different coronaviruses, especially in the FP region, and informed the FP structural modeling as described below.

The available cryoEM structure of the SARS-CoV2 S-protein at the time (Protein Data Bank, PDB: 6VSB) (52) contained only a partial
structure (12 residues) of the FP, making it unsuitable for construction of the initial SARS-CoV2 FP model (Fig. 1 A). Because the S2 domain of the S-protein containing the FP is well-conserved among the SARS coronaviruses, we used the S-protein from SARS-CoV as a template for modeling the FP of the SARS-CoV2. Accordingly, the cryoEM structure of the S-protein from SARS-CoV (PDB: 5XLR) (53) containing the FP structure was used as a template for constructing the initial SARS-CoV2 FP model. Our SARS-CoV2 FP model also contained the loop connecting the FP to the neighboring proximal region, suggested to be important in the fusion process (24–26) but not modeled fully in this study because of a lack of a suitable structural template at the time. The only two different residues between the sequences of SARS-CoV and SARS-CoV2 FPs (I834/M816 and D839/E821 in Fig. S1) were mutated to the SARS-CoV2 residues (Fig. 1 B).

The initial FP model was then solvated and ionized with 0.15 M NaCl using SOLVATE and AUTOIONIZE plugins of VMD, respectively (54). Energy minimization was carried out using the steepest descent method for 100 steps followed by an equilibrium MD of the FP in solution for 20 ns in an NPT ensemble, to obtain a fully relaxed initial SARS-CoV2 FP in an aqueous environment. The resulting equilibrated FP was used in the subsequent membrane-binding simulations.

Membrane preparation

To obtain membrane-bound models of the SARS-CoV2 FP, we performed multiple independent simulations employing the HMMM model (Fig. 1 C) (33,34,36), in which the diffusion of lipids is enhanced, allowing for better sampling and convergence of FP-lipid interactions. As shown in previous studies, compared with conventional (full) membrane simulations, physical characteristics of membrane binding remain conserved for peripheral proteins, but the binding timescale is reduced by an order of magnitude with the use of HMMM (38,41,55). The HMMM model has been extensively used in a broad spectrum of systems providing mechanistic details of lipid-protein interactions (37–41,43–49,56–65) and can accurately reproduce the energetics associated with partitioning of amino acids within the headgroups and in the most peripheral section of the leaflets (35). We use this technique to facilitate binding and insertion of the FP into the membrane, which would allow us to focus our computational effort on running many independent membrane-binding simulations.

In this study, symmetric full membranes were first constructed using CHARMM-GUI (66), with a lipid composition (POPC/POPE/CHL: 60/6/34 mol%) that included major lipids in the outer leaflet of the human plasma membrane (28–31), except for sphingomyelin, which has not been sufficiently tested in a cholesterol-containing HMMM model. The initial, full membranes generated with CHARMM-GUI were converted to HMMM membranes by removing the atoms after the fifth carbon in the phospholipid acyl chains, resulting in short-tailed PC and PE lipids while keeping the cholesterol molecules intact. To mimic the membrane core, a previously developed in silico solvent, termed SCSE (including two carbon-like interaction centers) (36), was used to match the number of heavy atoms removed from the lipid tails in the previous step. The resulting HMMM membranes contained 2178 SCSE molecules (36) and 150 lipids in each leaflet.

To further expand the sampling of the phase space, we varied the initial placement and orientation of the FP (i.e., six rotated orientations (P, A, N, S, I, R) (Fig. 1 C)), to minimize the initial bias. The systems were solvated and ionized using the SOLVATE and AUTOIONIZE plugins in VMD (54) with 0.15 M NaCl, resulting in total system sizes of ~70,000–90,000 atoms and box sizes of 103 × 103 × 90–103 × 103 × 110 Å3. Multiple replicas were simulated for each FP orientation because the diffusion and mixing of lipids and the process of membrane binding and insertion of the FP can slow, even when using HMMM membranes, and more sampling would ensure the reliability of the obtained membrane-bound configurations. Five independent HMMM membranes with each specified FP orientation were generated using a Monte Carlo-based lipid-mixing protocol developed in our group to further enhance variation of initial membrane configurations for each FP orientation.

Membrane-binding simulations

The systems were energy minimized using steepest descent method for 10,000 steps and simulated for 10 ns with the Ca atoms of the peptide harmonically restrained (k = 0.5 kcal mol⁻¹ Å⁻²), followed by a production run of 300 ns after removing the Ca restraints. A harmonic restraint along the z axis, with a force constant of k = 0.05 kcal mol⁻¹ Å⁻² was applied to the C2, C26, and C36 atoms of the phospholipids and to the O3 atom of cholesterol to reproduce the atomic distributions of these lipids in a full
lipid bilayer more closely and to prevent the occasional escape of short-tailed lipids into the aqueous phase, which is expected for these surfactant-like molecules. To prevent SCSE molecules from diffusing out of the core of the membrane, we subjected them to a grid-based restraining potential, applied using the gridForce (67) feature of NAMD (68,69). Five replicas, each with an independently generated HMMM membrane (different lipid mixings) and a starting orientation of the FP, were simulated, resulting in a total of 30 independent membrane-binding simulations for a cumulative time of 9 μs. The initial and final structures for all the simulations, as well as representative membrane-bound structures, topology/parameter files, and configuration files for the runs are all made available in Open Science Framework, https://osf.io/gqfdis/?view_only=8101f10dcd3140e690bb0d4fa50e0af6. Trajectory files were prohibitively large to be deposited, but will be made available upon request.

**Simulation protocol**

The systems were simulated using an NPT ensemble with a constant x/y ratio, using a target pressure and temperature of 1.0 atm and 310 K, respectively, and a time step of 2 fs. Although scaling up the area combined with a constant-area protocol is commonly used for HMMM simulations to allow faster membrane binding (33,34,36), given that the FP is considerably small, we did not expect that our NPT setup would significantly hamper the process of membrane binding. Simulations were performed using NAMD (68,69) with the CHARMM36 (70) force field parameters for lipids and proteins. The TIP3P model was used for water (71). The Nosé-Hoover Langevin piston method (72,73) was utilized to maintain the constant pressure, and constant temperature was maintained by Langevin dynamics with a damping coefficient $\gamma$ of 0.5 ps$^{-1}$ applied to all atoms. Nonbonded interactions were cut off after 12 Å, starting at a switching distance of 10 Å. The particle mesh Ewald method (74) was used for long-range electrostatic calculations with a grid density greater than 1 Å$^{-3}$.

**Analysis**

We define stable membrane binding with the criteria described below (see Fig. 3). First, a contact between the FP and the membrane lipids is defined for any heavy atom in the FP that is within 3.5 Å of any lipid heavy atom. Any contiguous segment of the simulation trajectory with a length of at least 30 ns, during which at least one contact between the lipids and the FP existed, was considered stable binding (marked with a blue background in Fig. 3).

To characterize the binding orientation of the FP with respect to the membrane, the first principal axis (PA) of the FP helical segment (residues 816–823) was calculated (Fig. S2). The angle between the first PA and the membrane normal, $\theta_{PA}$, was used to represent the tilting of the FP. We then identify two additional angles to describe the orientation of two phenylalanine residues on the FP helical segment with respect to the membrane, $\theta_{PA17}$ and $\theta_{PA23}$, which are defined by the angle between the membrane normal and the perpendicular component (to the PA) of F817 and F823 Cα-Cα vector, respectively. Combining $\theta_{PA}$ with these two additional angles, which take into account the rotational degrees of freedom of the FP, we obtain a full description of the PF-binding orientation to the membrane.

For all the frames identified as membrane bound (see above for definition), a vector composed of the z distances between the centers of masses (COMs) of individual FP side chains and the lipid phosphate plane was calculated. A dissimilarity matrix was then constructed using the Euclidean distances between the COMs of individual residues and the phosphate plane was computed and diagonalized. The resulting eigenvectors (i.e., the principal components (PCs)) represent the coordinates that maximize the variance of the projected data. The first two PCs were selected to project the original z-distance data of membrane-bound frames along with the clustering result onto the reduced dimension. The PCA was performed using the scikit-learn package (76).

**RESULTS AND DISCUSSION**

The SARS-CoV2 attachment to and its penetration into the human plasma membrane are key steps in its viral infection. These steps rely on the interaction of the viral FP with the host cellular membrane. This aspect is the central phenomenon that we study here, using atomistic simulations to characterize the binding pose and the conformation of the SARS-CoV2 FP when bound to the membrane. In the following sections, we first describe the results of our membrane-binding simulations in terms of the overall binding of the FP using some coarse parameters. We then use the depth of insertion of individual FP residues (COMs) in the membrane in a clustering analysis. We then examine the clustering results in a reduced dimension, offered by PCA, to better classify the different microstates that arise during the membrane-binding simulations and discuss the biological relevance of the resulting bound states.

**Spontaneous membrane binding and insertion of SARS-CoV2 FP**

To characterize the membrane-binding mode of the SARS-CoV2 FP to human cellular membranes, we performed 30 independent membrane-binding simulations using HMMM lipid bilayers. Spontaneous diffusion and membrane binding of the FP in each simulation replica can be monitored by tracking the position of the COM of the peptide with respect to the membrane as a coarse metric. The z component of the COM was tracked with respect to the phosphate layer (the average z position of all the phosphate groups) of each leaflet (blue and red lines, respectively, in Fig. 2). Because of the applied periodic boundary conditions and the free diffusion of the FP in the solution, the FP was able to diffuse toward either the upper or the lower leaflet of the membrane (Fig. 2), both containing lipid compositions representing the outer leaflet of human plasma membrane (28–31,77). Among the 30 performed membrane-binding simulations, instances of both stable (majority) and transient binding events, as well as cases with no membrane binding, were observed. Because the lipids in our simulated membranes are all neutral, we do not expect any major electrostatic forces driving the binding. Rather, the major driving forces between the two stable states are hydrophobic effects, and the FP diffusion in the solution can make the peptide take a longer time to make the initial encounter with the membrane.
For further analysis of the membrane interactions, we selected only the portions of the trajectories where “stable membrane binding” or a “membrane-bound state” was defined (see Materials and methods for details). Contacts between the FP and the lipid bilayer are shown in Fig. 3, in which the blue segments of the graphs are considered to be stable binding (lipid contact for at least 30 ns). In 22 of the 30 independent membrane-binding simulations (73%) membrane-bound configurations are observed in 22 of the 30 independent membrane-binding simulations (73%) membrane-bound configurations are observed (Fig. 3). For example, we observe stable membrane binding in parallel replicas P2, P3, P4, and P5, and in antiparallel replicas A1, A2, A3, and A5 (Fig. 3). In some simulations, nearly the entire length of the FP was observed to be engaged with membrane lipids (e.g., replicas P2, P5, A1, N1, I4, I5, and R1), whereas in other cases, only a specific part of the peptide makes contact with the membrane. Randomizing the initial placement of the peptides prevented biasing and allowed for better sampling of the membrane interaction of the FP. In some simulations, the FP bound to the lipid bilayer as early as 10 ns, whereas, in the others, it diffused and tumbled longer before interacting with the membrane, naturally allowing the peptide to further decouple itself from the effect of initial placement. For example, the initially reclined peptide in simulation replica R5 results in a membrane-bound configuration in which the FP interacts with the lipids through its loop part, or initially standing-placed peptides in S1 and S2 simulations end up being buried in the bilayer via their helical segment (Fig. 3).

The overall structure of the FP plays a major role in its mode and depth of membrane binding. Therefore, we monitored the structural evolution of the FP throughout the simulations. The N-terminal \(\alpha\)-helical segment (residues 816–823) remains largely unperturbed during the simulations, whereas the rest of the FP structure undergoes conformational changes. Within the membrane-bound configurations, the most common FP structure is still a hybrid \(\alpha\)-helix/loop structure in which either the initial structure is mostly preserved or the \(\alpha\)-helical segment is extended (Fig. S4B, panels 2 and 5), which has been experimentally observed in many viral FPs when inserted into the membrane, for example for the case of herpesvirus and Hendra virus (78–80). In some replicas, we observed FPs evolving into hairpin-like structures (helix-hinge-helix) (as shown, for example, in panels 1 and 2 of Fig. S4B), another structural arrangement also commonly reported for viral FPs (81,82). In a few cases, we also observed partial unfolding of the \(\alpha\)-helical segment (Fig. S4B, panel 4). From an energetic perspective, insertion of helical segments into the membrane might be less costly because backbone amide groups

![FIGURE 2](https://example.com/figure2.png)
in a helix can satisfy their hydrogen bonds internally (83), which explains the preservation and even extension/formation of helical segments in most of our membrane-binding simulations. Because the secondary structure evolves in time, hereon for simplicity, the terms α-helical segment and loop segment will refer to the initial α-helix (residues 817–827) and loop (residues 828–841) parts of the FP.

Using the stable membrane-bound states defined above, we further examined the average position and lipid interaction of individual membrane-bound FP residues (Fig. 4). The average depth of insertion of individual residues reaches as deep as 10 Å below the phosphate layer, indicating interaction with the hydrophobic core of the membrane. From the pattern of residue-lipid interactions, we observe three major classes of membrane-bound states in which either the whole peptide, the α-helical segment, or the loop part of the FP interacts with the lipid bilayer, respectively. For example, in replicas P3, P5, and A5, the FP is almost completely buried into the membrane; in replicas P2, N1, and N2, the FP is bound to the membrane primarily through its α-helical segment; and in replicas P4, A1, A2, and A3, the loop is mostly engaged with the lipids. To better classify these binding modes in a reduced space, we next performed clustering of the obtained membrane-bound configurations.

Clustering and characterization of FP binding modes

To analyze the ensemble of FP membrane-bound configurations captured in the 22 HMMM simulations in which stable membrane binding was observed, clustering was performed using the insertion depth of individual side chains with respect to the membrane (see Materials and methods for details). Despite the large variance in membrane-bound poses among MD snapshots from different simulation replicas, three major clusters can be identified, representing the three distinct binding patterns mentioned briefly above. The three binding modes (termed 1, 2, and 3) make up 50.4, 11.2, and 38.4%, respectively, of the whole population of membrane-bound configurations observed in the simulations.

For better visualization, we performed PCA and reduced the dimension to the first two PCs, which together cover 82.4% of the variance (Fig. 5 A). The nature of these two PCs, PC1 and PC2, was examined by their eigenvectors, which quantitatively evaluate the contribution of membrane insertion by each residue (Fig. S3). Based on the eigenvectors, PC1 and PC2 can be viewed as a measure of the membrane insertion of the helical segment (PC1) and that of the C-terminal

FIGURE 3 Identifying stable membrane binding using FP-lipid contacts. The plots show contacts between individual FP residues (y axis) and the lipids (heavy atom distance of less than 3.5 Å) over time (x axis) in a sampling step size of 1 ns. All contacts are labeled with black dots. The segments of the trajectories in which, for a minimum of 30 ns, at least one contact existed between the FP and lipids are marked with a blue background. We observe stable binding in 22 out of 30 replicas. To see this figure in color, go online.
loop (PC2), with more negative values indicating deeper insertion. The clustering centers (i.e., the medoids) of the three binding modes are selected as representative binding poses (Fig. 5, B–D). The three selected poses here represent the most sampled forms (clusters) of membrane-bound FP and thus the energetically most favored states. The convergence of these three binding states can be observed from MD snapshots that are closest to the clustering centers while at the same time belonging to different replicas (Fig. S5). Binding mode 1, which accounts for over half of the data, corresponds to a pattern in which the helical segment or the C-terminal loop segment being the primary membrane-interacting region, respectively. We note that there are snapshots of membrane-bound FP poses located at the boundaries between the clusters or at the far edge of each cluster, which are much less sampled compared with the most representative region of the cluster. We believe these snapshots correspond to intermediate membrane-bound poses that are less biologically relevant.

Orientation of the FP in its membrane-bound configuration

To further analyze the membrane-binding configurations for SARS-CoV2 FP in our simulations, we calculated the orientations of the FP throughout the simulations (Fig. S2). We defined three angles, $\theta_{PA}$, $\theta_{PH17}$, and $\theta_{PH23}$, describing the FP helix orientation (see Materials and methods for the definition of the angles) over the simulation trajectories (Fig. 6 A) in replicas showing stably bound configurations. Although occasionally the FP can switch between binding modes 1, 2, and 3 within the same simulation replica, the analysis of these angles in different replicas will provide a better qualitative description of the major binding modes observed in our simulations. For the replicas where binding mode 1 is the dominant binding configuration (e.g., P3, P5, A5, I3, R1), the average $\theta_{PA}$ angle is $90.4^\circ$ (Fig. 6 B), which corresponds to a parallel orientation with respect to the membrane. In replicas in which binding mode 2 is dominant (N1, N2, N4, R4), $\theta_{PA}$ averages at $104.7^\circ$ (Fig. 6 B), indicating an oblique configuration with the N-terminus facing the membrane core. Because the C-terminal segment is more flexible and the orientation angles are defined for the
α-helical segment, it is more difficult to provide an equally clear description for binding mode 3 (e.g., P4, A1, A2, A3, S3, S4, I2, I4, I5). For this binding mode, in most of the replicas toward the end of the trajectory, $\theta_{PA}$ is less than 90°, averaging at 51°. This means that the N-terminal of the helix is facing up, and as the angle decreases, the helix becomes more orthogonal to the membrane. The $\theta_{F817}$ angle in binding modes 1 and 2 stabilized at average values of 134.6 and 137.2°, respectively, with a peak around 160°, where F817 is facing the membrane core almost directly (Fig. 6B). F823 is located on the opposite side of the helix with respect to the F817; therefore, in binding modes 1 and 2, this residue is mostly facing up and away from the membrane with average angles of 53.6 and 59.6°, respectively.

**Physiologically relevant membrane-bound configuration of SARS-CoV2 FP**

To discuss our results within a biology/physiology context, we compare our findings with experimental observations. We also assess the mechanistic relevance of the observed binding modes within the context of the whole S-protein by evaluating the way specific FP residues interact with the membrane.

In our simulations, we observe three major membrane-binding modes for SARS-CoV2 FP: binding mode 1, in which the peptide is nearly parallel to deeply buried in the membrane; binding mode 2, in which the helical segment of the FP is the primary site engaging with the membrane; and binding mode 3, in which the C-terminal of the peptide is primarily interacting with the membrane (Fig. 5). Mutagenesis experiments have shown that highly conserved residues L821, L822, and F823, termed the fusion-active core, play a major role in viral fusion of SARS-CoV (21,84). Given the high sequence similarity between SARS-CoV and SARS-CoV2 FPs, including L821/L822/F823 (Fig. S1), the fusion-active core is also likely to be crucial for the binding and insertion of SARS-CoV2 FP. Consistent with these results, in binding modes 1 and 2, residues L821, L822, and F823 are deeply inserted into the hydrophobic core of the membrane, closely interacting with the membrane lipids (Figs. 5C and 7A). Additionally, electron spin resonance spectroscopy experiments have shown that the highly conserved FP in SARS-CoV forms a “fusion platform.” In this predicted structure, the majority of the FP residues are deeply inserted into the membrane, whereas C-terminal residues are free in the solution (24,25), in a
similar fashion to what we observe in binding mode 1 (Fig. 7A).

Evaluating the binding positions of the FP in the context of the postfusion S-protein is another important component in identifying the most physiologically relevant binding form of the FP. The cleavage taking place before and aiding the conformational change of the S-protein at the S2' cleavage site makes S816 the N-terminus of the whole postfusion S-protein. Therefore, FP in binding modes 1 and 2, which both enter the membrane with their N-terminus facing the membrane core at an oblique angle, are mechanistically viable and compatible with the nature of the predicted conformational changes in the S-protein. This is in sharp contrast to binding mode 3, in which the C-terminal segment is buried in the membrane.

Furthermore, binding mode 1 is the most highly sampled binding mode in our simulations, accounting for more than half of the membrane-bound forms we observe in our entire data set. It is also the most deeply inserted binding form among the three major bound configurations. Given the established role of several hydrophobic residues in the helical segment of the FP in membrane interaction, the resemblance to the fusion platform, and its depth of penetration in the membrane, we strongly believe and propose that the binding mode 1 represents the physiologically relevant membrane-bound form of SARS-CoV2 FP, which we term here as the deeply inserted membrane-bound form.

The deeply inserted membrane-bound form is anchored in the membrane with majority of its residues interacting with the lipids. In addition to the fusion-active core discussed above, we also observe direct lipid interactions for other specific residues that might be targeted in future mutagenesis studies. Most of the hydrophobic residues (e.g., F817/I818, V826/T827/L828/A829, and F833/I834) are involved in binding and insertion into the membrane along with a few charged residues (Fig. 7B). Insertion of some of these residues is key to reducing hydrophobic exposure of the FP to the water environment, and hence, is likely to increase the binding affinity of the FP. In addition, deeply inserted F817 and F833 are distinctive to and conserved among the SARS coronaviruses, which are currently the most infectious coronavirus species S1 (85). Hence, we can speculate that F817 and F833 may function as an important factor in enhancing the infection rate of SARS-CoV2 by increasing the binding affinity of FP to the membrane. Future experimental studies will be necessary to further validate this hypothesis.

A recent complementary study has investigated the SARS-CoV2 FP membrane binding in the presence of Ca²⁺ ions (86). They observe different binding modes, including one in which the N-terminal helical segment binds...
the membrane along with L822/F823 penetrating the membrane, as well as a binding mode with insertion of F833/I834 into the lipid bilayer, which is in agreement with our findings (86). Another interesting aspect is FP’s interactions with cholesterol. Although this study does not cover cholesterol specificity in the binding of FP, binding affinity measurements for the FP in the presence of cholesterol supports that the FP favors interaction with cholesterol (32). Therefore, as exemplified in Fig. 7A, we can speculate that the penetrating aromatic residues (F817, F823, F833, and Y837) in the membrane core are likely to interact with membrane cholesterol and increase the binding affinity.

In addition to the simulated peptide in this study (which we have referred to as the FP), there are other regions (named FP2 in SARS-CoV and fusion peptide proximal region (FPPR) in SARS-CoV2) (11,24–27), which are implicated in membrane fusion. The FPPR of SARS-CoV2 downstream of the FP was later resolved and claimed to be involved in the structural rearrangement of the S-protein before membrane fusion (27). An internal disulfide bond within the FPPR, between C840 and C851, was observed and suggested to increase membrane-ordering activity (25,27). The membrane-ordering activity of the FP, due to the fusion-active core, is significantly higher than the FPPR, and the activity of FP/FPPR together is only slightly increased compared with their activity separately (25). Although our deeply inserted membrane-bound form supports the concept in which both the FP and FPPR can interact with the membrane simultaneously as two subdomains (25), to characterize such a platform interacting with the membrane a longer peptide, including the FPPR, should be studied in the future.

CONCLUSION
COVID-19, which has emerged as a severe pandemic worldwide, calls for a need to accelerate the development of novel therapeutic intervention strategies. The S-protein of the SARS-CoV2 contains the key machinery necessary for the infection of human cells, including the FP, a highly conserved segment that inserts into the human cellular membrane and initiates the fusion of the virus. Yet, there are no postfusion S-protein structures illustrating the binding of FP to human cellular membrane. In this study, using an extensive set of simulations, we describe how the SARS-CoV2 FP binds lipid bilayers representing mammalian cellular membranes and characterize, at an atomic level, lipid-protein interactions important for the stability of its bound state.

We capture different membrane-binding configurations from these simulations, which are classified using a detailed clustering analysis and based on geometrical evaluation of the peptide with respect to the membrane, resulting in three major membrane-bound configurations. Further analysis of these configurations in a mechanistic context, taking into account the structural requirement of the entire S-protein, comparison of the results to previous experimental characterization of specific residues in the FP in the coronavirus family, and the degree of membrane engagement of the FP all support the first binding mode (binding mode 1) to be the most likely and the most physiologically relevant form for the membrane-bound SARS-CoV2 FP.

Characterizing the mechanism of the fusion-driving FP-host membrane interactions is key to our understanding of the critical steps involved in viral infection, paving the way for potential development of novel therapeutics against SARS-CoV2 in addition to targeting the binding and interaction of the S-protein with the ACE2 host receptor. These include modulation of the FP membrane-binding interface through small molecules with high affinity and specificity for this region of the S-protein or inhibiting the key lipid-protein interactions observed.

Based on the suggested binding mode elucidated in our study, mutagenesis experiments can be designed to further confirm the role of the important residues implicated in

Biophysical Journal 120, 2914–2926, July 20, 2021 2923
membrane binding. Given the close similarity of the fusion peptides in coronaviruses in general, these results can also be applicable to infections caused by other members of this life-threatening family of pathogens.

SUPPORTING MATERIAL

Supporting material can be found online at https://doi.org/10.1016/j.bpj.2021.02.041.

AUTHOR CONTRIBUTIONS

D.G., K.K., and E.T. designed the research. D.G. carried out all simulations. D.G. and M.L. analyzed the data. All authors wrote the article.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (grants P41-GM104601 and R01-GM123455 to E.T.). Molecular dynamics simulations were performed using Blue Waters and computational resources provided by Microsoft Azure.

REFERENCES

1. Huang, C., Y. Wang, …, B. Cao. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 395:497–506.
2. Chan, J. F.-W., S. Yuan, …, K.-Y. Yuen. 2020. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet. 395:514–523.
3. Singhal, T. 2020. A review of coronavirus disease-2019 (COVID-19). Indian J. Pediatr. 87:281–286.
4. Wan, Y., J. Shang, …, F. Li. 2020. Receptor recognition by the novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS coronavirus. J. Virol. 94:e00127-20.
5. Tang, T., M. Bidon, …, S. Daniel. 2020. Coronavirus membrane fusion mechanism offers a potential target for antiviral development. Antiviral Res. 178:104792.
6. Skehel, J. J., and D. C. Wiley. 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu. Rev. Biochem. 69:531–569.
7. Eckert, D. M., and P. S. Kim. 2001. Mechanisms of viral membrane fusion and its inhibition. Annu. Rev. Biochem. 70:777–810.
8. White, J. M., S. E. Delos, …, K. Schornberg. 2008. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. Crit. Rev. Biochem. Mol. Biol. 43:189–219.
9. Kielian, M., and F. A. Rey. 2005. Antigenic and functional diversity of the membrane-active regions of the severe acute respiratory syndrome coronavirus spike protein. J. Virol. 79:1474–1475.
10. Lai, A. L., Y. Li, and L. K. Tamm. 2005. Interplay of proteins and lipids in virus entry by membrane fusion. In Protein-Lipid Interactions. L. K. Tamm, ed. Wiley-VCH, pp. 279–305.
11. Madu, I. G., S. L. Roth, …, G. R. Whittaker. 2009. Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide. J. Virol. 83:7411–7421.
12. Guillén, J., P. K. Kinnunen, and J. Villalain. 2008. Membrane insertion of the three main membranotropic sequences from SARS-CoV S2 glycoprotein. Biochim. Biophys. Acta. 1778:2765–2774.
13. Guillén, J., A. J. Pérez-Berná, …, J. Villalain. 2005. Identification of the membrane-active regions of the severe acute respiratory syndrome coronavirus spike glycoprotein using a 16/18-mer peptide scan: implications for the viral fusion mechanism. J. Virol. 79:1743–1752.
14. Millet, J. K., and G. R. Whittaker. 2018. Physiological and molecular triggers for SARS-CoV membrane fusion and entry into host cells. Virology. 517:3–8.
15. Lai, A. L., J. K. Millet, …, G. R. Whittaker. 2017. The SARS-CoV fusion peptide forms an extended bipartite fusion platform that perturbs membrane order in a calcium-dependent manner. J. Mol. Biol. 429:3875–3892.
16. Kirchdoerfer, R. N., N. Wang, …, A. B. Ward. 2018. Stabilized coronavirus spikes are resistant to conformational changes induced by receptor recognition or proteolysis. Sci. Rep. 8:15701.
17. Cai, Y., J. Zhang, …, B. Chen. 2020. Distinct conformational states of SARS-CoV-2 spike protein. Science. 369:1586–1592.
18. van Meer, G., D. R. Voelcker, and G. W. Feigenson. 2008. Membrane lipids: where they are and how they behave. Nat. Rev. Mol. Cell Biol. 9:112–124.
19. Casares, D., P. V. Escritá, and C. A. Rosselló. 2019. Membrane lipid composition: effect on membrane and organelle structure, function and compartmentalization and therapeutic avenues. Int. J. Mol. Sci. 20:2167.
20. Mouritsen, H., U. Janssen-Bienhold, …, R. Weiler. 2004. Cryptochromes and neuronal-activity markers colocalize in the retina of migratory birds during magnetic orientation. Proc. Natl. Acad. Sci. USA. 101:14294–14299.
21. Dodge, J. T., and G. B. Phillips. 1967. Composition of phospholipids and of phospholipid fatty acids and aldehydes in human red cells. J. Lipid Res. 8:667–675.
22. Moshfegh, G., S. Bhattacharjya, and H. Chakraborty. 2019. Membrane cholesterol modulates oligomeric status and peptide-membrane interaction of severe acute respiratory syndrome coronavirus fusion peptide. J. Phys. Chem. B. 123:10654–10662.
23. Arcario, M. J., Y. Z. Ohkubo, and E. Tajkhorshid. 2011. Capturing spontaneous partitioning of peripheral proteins using a biphasic membrane-mimetic model. J. Phys. Chem. B. 115:7029–7037.
Membrane binding of viral fusion peptide

34. Ohkubo, Y. Z., T. V. Pogorelov, ..., E. Tajkhorshid. 2012. Accelerating membrane insertion of peripheral proteins with a novel membrane mimetic model. *Biophys. J.* 102:2130–2139.

35. Pogorelov, T. V., J. V. Vermaas, ..., E. Tajkhorshid. 2014. Partitioning of amino acids into a model membrane: capturing the interface. *J. Phys. Chem. B.* 118:1481–1492.

36. Vermaas, J. V., T. V. Pogorelov, and E. Tajkhorshid. 2017. Extension of the highly mobile membrane mimetic to transmembrane systems through customized in silico solvents. *J. Phys. Chem. B.* 121:3764–3776.

37. Baylon, J. L., I. L. Lenov, ..., E. Tajkhorshid. 2013. Characterizing the membrane-bound state of cytochrome P450 3A4: structure, depth of insertion, and orientation. *J. Am. Chem. Soc.* 135:8542–8551.

38. Vermaas, J. V., and E. Tajkhorshid. 2014. A microscopic view of phospholipid insertion into biological membranes. *J. Phys. Chem. B.* 118:1754–1764.

39. Arcario, M. J., and E. Tajkhorshid. 2014. Membrane-induced structural rearrangement and identification of a novel membrane anchor in talin F2F3. *Biophys. J.* 107:2059–2069.

40. Blanchard, A. E., M. J. Arcario, ..., E. Tajkhorshid. 2014. A highly tilted membrane configuration for the pre-fusion state of synaptobrevin. *Biophys. J.* 107:2112–2121.

41. Vermaas, J. V., and E. Tajkhorshid. 2014. Conformational heterogeneity of α-synuclein in membrane. *Biochim. Biophys. Acta.* 1838:3107–3117.

42. Wu, Z., and K. Schulten. 2014. Synaptotagmin’s role in neurotransmitter release likely involves Ca(2+)-induced conformational transition. *Biophys. J.* 107:1156–1166.

43. Rihauté, J.-F., E. Gagné, ..., P. Lagüé. 2015. Molecular model of hemoglobin N from mycobacterium tuberculosis bound to lipid bilayers: a combined spectroscopic and computational study. *Biochemistry.* 54:2073–2084.

44. Madsen, J. J., Y. Z. Ohkubo, ..., O. H. Olsen. 2015. Membrane interaction of the factor VIIIa discoidin domains in atomistic detail. *Biochemistry.* 54:6123–6131.

45. Vermaas, J. V., J. L. Baylon, ..., E. Tajkhorshid. 2015. Efficient exploration of membrane-associated phenomena at atomic resolution. *J. Membr Biol.* 248:563–582.

46. Baylon, J. L., J. V. Vermaas, ..., E. Tajkhorshid. 2016. Atomic-level description of protein-lipid interactions using an accelerated membrane model. *Biochim. Biophys. Acta.* 1858:1573–1583.

47. Muller, M. P., Y. Wang, ..., E. Tajkhorshid. 2017. Lipid specificity of the membrane binding domain of coagulation factor X. *J. Thromb. Haemost.* 15:2005–2016.

48. Vermaas, J. V., and E. Tajkhorshid. 2017. Differential membrane binding mechanics of synaptotagmin isoforms observed at atomic detail. *Biochemistry.* 56:281–293.

49. Baylon, J. L., and E. Tajkhorshid. 2015. Capturing spontaneous membrane insertion of the influenza virus hemagglutinin fusion peptide. *J. Phys. Chem. B.* 119:7882–7893.

50. Katoh, K., K. Misawa, ..., T. Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30:3059–3066.

51. Clamp, M., J. Cuff, ..., G. J. Barton. 2004. The Jalview Java alignment editor. *Bioinformatics.* 20:426–427.

52. Wrapp, D., N. Wang, ..., J. S. McLellan. 2020. Cryo-EM structure of the 2019-nCoV spike in the pre-fusion conformation. *Science.* 367:1260–1263.

53. Gui, M., W. Song, ..., X. Wang. 2017. Cryo-electron microscopy structures of the SARS-CoV spike glycoprotein reveal a prerequisite conformational state for receptor binding. *Cell Res.* 27:119–129.

54. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27-28.

55. Wildermuth, K. D., V. Monje-Galvan, ..., J. B. Klauda. 2019. Effect of membrane lipid packing on stable binding of the ALPS peptide. *J. Chem. Theory Comput.* 15:1418–1429.
78. Epand, R. M. 2003. Fusion peptides and the mechanism of viral fusion. Biochim. Biophys. Acta. 1614:116–121.

79. Smith, E. C., S. M. Gregory, . . ., R. E. Dutch. 2012. Role of sequence and structure of the Hendra fusion protein fusion peptide in membrane fusion. J. Biol. Chem. 287:30035–30048.

80. Gianni, T., L. Menotti, and G. Campadelli-Fiume. 2007. A heptad repeat in herpes simplex virus 1 gH, located downstream of the alpha-helix with attributes of a fusion peptide, is critical for virus entry and fusion. J. Virol. 81:2540.

81. Han, X., J. H. Bushweller, . . ., L. K. Tamm. 2001. Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. Nat. Struct. Biol. 8:715–720.

82. Hsu, C.-H., S.-H. Wu, . . ., C. Chen. 2002. Structural characterizations of fusion peptide analogs of influenza virus hemagglutinin. Implication of the necessity of a helix-hinge-helix motif in fusion activity. J. Biol. Chem. 277:22725–22733.

83. Wimley, W. C., and S. H. White. 1996. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. Nat. Struct. Biol. 3:842–848.

84. Madu, I. G., S. Belouzard, and G. R. Whittaker. 2009. SARS-coronavirus spike S2 domain flanked by cysteine residues C822 and C833 is important for activation of membrane fusion. Virology. 393:265–271.

85. Hu, T., Y. Liu, . . ., Q. He. 2020. A comparison of COVID-19, SARS and MERS. PeerJ. 8:e9725.

86. Khelashvili, G., A. Plante, . . ., H. Weinstein. 2021. Ca2+-dependent mechanism of membrane insertion and destabilization by the SARS-CoV-2 fusion peptide. Biophys. J. 120:1105–1119.