Biochemistry

Article

Identifying Distinct Structural Features of the SARS-CoV-2 Spike Protein Fusion Domain Essential for Membrane Interaction

Daniel Birtles and Jinwoo Lee*

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: The SARS-CoV-2 spike protein is the primary antigenic determinant of the virus and has been studied extensively, yet the process of membrane fusion remains poorly understood. The fusion domain (FD) of viral glycoproteins is well established as facilitating the initiation of membrane fusion. An improved understanding of the structural plasticity associated with these highly conserved regions aids in our knowledge of the molecular mechanisms that drive viral fusion. Within the spike protein, the FD of SARS-CoV-2 exists immediately following S2′ cleavage at the N-terminus of the S2 domain. Here we have shown that following the introduction of a membrane at pH 7.4, the FD undergoes a transition from a random coil to a more structurally well-defined postfusion state. Furthermore, we have classified the domain into two distinct regions, a fusion peptide (FP, S^{816}−G^{838}) and a fusion loop (FL, D^{839}−F^{855}). The FP forms a helix−turn−helix motif upon association with a membrane, and the favorable entropy gained during this transition from a random coil is likely the driving force behind membrane insertion. Membrane depth experiments then revealed the FP is found inserted within the membrane below the lipid headgroups, while the interaction of the FL with the membrane is shallower in nature. Thus, we propose a structural model relevant to fusion at the plasma membrane in which the FP inserts itself just below the phospholipid headgroups and the FL lays upon the lipid membrane surface.
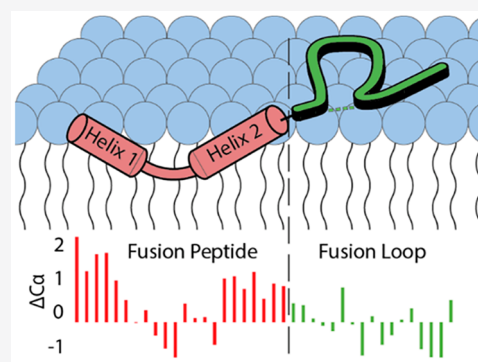
Following the previous outbreaks of severe acute respiratory syndrome coronavirus (SARS-CoV-1) in 2003 and Middle East respiratory syndrome coronavirus (MERS-CoV) in 2012, the threat of the coronavirus family was well established, yet the rapid emergence and remarkable infectivity of SARS-CoV-2, the causative agent of coronavirus disease 2019 (COVID-19), have resulted in a global pandemic unprecedented in modern times. With more than 200 million people infected and nearly 4.5 million deaths,[1] the importance of elucidating the underlying mechanisms behind pathogenic viruses has never been more prevalent than now.

Viral entry of the coronavirus family is mediated by the spike (S) glycoprotein, which is found on the exterior surface of the mature virion. The SARS-CoV-2 spike protein is a trimeric, class I viral glycoprotein that can be divided into two functional subunits, S1 and S2, responsible for host cell recognition and membrane fusion respectively.[2] The interaction between the receptor binding domain (RBD) in S1 and the host cell receptor, angiotensin converting enzyme 2 (ACE2), has been studied extensively.[3−5] However, research regarding the process of membrane fusion that immediately follows this binding event is much more limited. Membrane fusion is a natural phenomenon that relies on the organized structural rearrangement of specialized viral fusion machinery to overcome the energy barrier necessary for the virus and host cell membranes to coalesce. This fusing of opposing membranes allows the virus to deliver its genetic information

into the host cell, a fundamental component of the viral life cycle. The most well-accepted mechanism for viral membrane fusion is that of the six-helix bundle, a process that is initiated by the fusion domain (FD) after it embeds itself in the host cell membrane.[6] The coronavirus family has been found to fuse at both the plasma and endosomal membranes, highlighting the remarkable adaptability of these viruses.[7] Accordingly, the FD is found to be well conserved throughout, particularly in the most pathogenic members of the viral family (Figure 1). The amino acid sequence of the FD is even maintained through evolution, with no mutations witnessed in the most prevalent SARS-CoV-2 variants: alpha(B.1.1.7), beta(B.1.351), gamma(P.1), delta(B.1.617.2), kappa(B.1.617.1), and lambda-(C.37).[8]

Interestingly, the FD has two regions that both closely resemble fusogenic structures witnessed in other viruses (Figure 1). The first is an N-terminal fusion peptide (FP), which is considered to be a key characteristic of class I viral proteins due to the pioneering work carried out on the FDs of
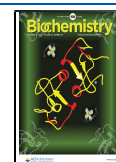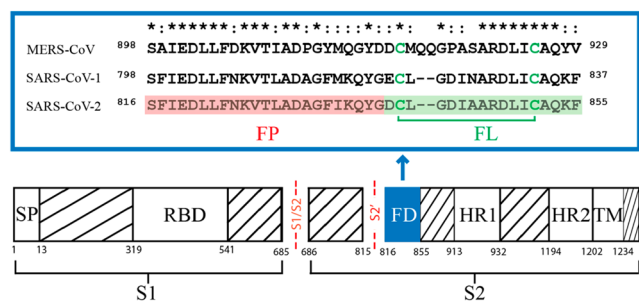
Figure 1. Functional regions of the SARS-CoV-2 spike protein. The fusion domain is found at the N-terminus of the S2 subunit following S2′ cleavage and shares a high degree of sequence conservation between the coronavirus family, with the three most pathogenic coronaviruses highlighted. Within the FD, there are two unique structural regions, the fusion peptide (FP, red) and the fusion loop (FL, green); the latter contains the highly conserved internal disulfide bond.

both HIV[9−11] and influenza.[12−14] This is followed by an internal disulfide-bonded region, which is akin to the fusogenic internal fusion loop (FL) seen within the Ebola virus (EBOV).[15] Due to this unique assembly, the region in SARS-CoV-2 spanning S^{816}−F^{855} will be known hereafter as the fusion domain (FD), which contains both a FP (S^{816}−G^{838}) and a FL (D^{839}−F^{855}).

Cryo-EM has revealed the structure of the spike protein in both pre- and postfusion states for SARS-CoV-2,[16] yet the postfusion state is missing a large disordered region (I^{770}−T^{912}) in the proximity of the membrane, which included the FD (S^{816}−F^{855}), with molecular dynamics providing some insight into this interface.[17,18] Previous work carried out on the FD within SARS-CoV-1, with the FP and FL investigated separately and together, revealed that both regions contain independent fusogenic activity but are most effective in synergy.[19,20] This led to the development of a bipartite fusion platform model for how the FD interacts with the membrane.[21] Moreover, the SARS-CoV-2 FD has been suggested to exhibit stronger membrane perturbing ability than the FD of SARS-CoV-1 through the use of electron paramagnetic resonance (EPR) and membrane ordering experiments.[22] The same study also displayed highly specific interactions between the FD and Ca^{2+}, which is thought to be an important cofactor in membrane fusion.[22]

In this work, we have expressed and purified a 40-amino acid construct that represents the complete FD of SARS-CoV-2 (S^{816}−F^{855}). The structure of the FD in both pre- and postfusion states has been investigated through solution NMR and CD. A large conformational change following association with a membrane mimic has been uncovered, which replicates the interaction with the host cell during the initial stages of membrane fusion. The FP (S^{816}−G^{838}) develops a helix−turn−helix motif, while the secondary structure of the FL (D^{839}−F^{855}) remains unchanged. Furthermore, the entire FD appears to interact with the membrane in a shallow manner and does not embed itself deeply within the membrane. Together, the data provided here supports a new model for the interaction between the SARS-CoV-2 FD and host cell membrane, one with unique structural characteristics compared to those of other well-established viral fusion domains.

MATERIALS AND METHODS

Expression. The SARS-CoV-2 fusion domain construct ^{816}(SFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI-CAQKF)^{855} was designed with an N-terminal six-His tag, followed by a small ubiquitin-like modifier (SUMO) tag from *Saccharomyces cerevisiae* (*SMT3* gene) to aid with solubility and expression.[23] This construct was made and inserted into the pET41 vector containing a T7 promoter by Genscript. The amino acid sequence was taken from the SARS-CoV-2 spike protein with UniProtKB accession number P0DTC2. The plasmid was transformed into *Escherichia coli* BL21 (DE3) cells, and a single colony was used to inoculate 5 mL of LB broth with the appropriate antibiotic. The starter culture was grown overnight at 37 °C and 225 rpm and then added to 1 L of LB medium. The large culture was grown to an OD_{600} between 0.6 and 0.8 and induced with a final isopropyl β-D-1-thiogalactopyranoside (IPTG) concentration of 1 mM for 4 h at 37 °C. For isotope labeling (^{13}C and ^{15}N), M9 minimal medium was used with 1 g of [^{15}N]ammonium chloride and 2 g of [^{13}C]glucose. For triply labeled samples (^{2}H, ^{13}C, and ^{15}N), 1 L of D_2O was used to make the minimal medium instead of H_2O. A 25 mL starter culture in the appropriate minimal medium was grown overnight at 37 °C and 225 rpm and then added to a 1 L volume. Following induction with a final IPTG concentration of 0.1 mM at an OD_{600} of 0.6−0.8, cells were grown overnight at 18 °C and 225 rpm. Cells were then pelleted at 4 °C for 30 min at 4500g and stored at −80 °C or purified immediately.

Purification. The cell pellet was resuspended in lysis buffer [8 M urea, 20 mM Tris, 300 mM NaCl, and 10 mM BME (pH 7.5)] and sonicated for 10 min on ice (40% power, 1 s on, 1 s off). The clear solution was then centrifuged for 30 min at 40000g and 4 °C. The supernatant was added to Ni-NTA resin and incubated for at least 2 h at 4 °C. Following incubation, the resin was washed [8 M urea, 20 mM Tris, 300 mM NaCl, and 10 mM imidazole (pH 7.5)], and then the bound construct was eluted [8 M urea, 20 mM Tris, 300 mM NaCl, and 250 mM imidazole (pH 7.5)]. SUMO protease ULP1 was added to the elution at a ratio of ∼1:10000 and then dialyzed overnight against 4 L of cleavage buffer [20 mM Tris, 300 mM NaCl, and 5 mM BME (pH 7.5)] at 4 °C. Protease vector pFGET19_Ulp1 was a gift from H. Iwai (Addgene plasmid 64697; http://n2t.net/addgene:64697; RRID: Addgene_64697).[24] The following day the sample was rebound to the Ni-NTA resin with an incubation of at least 2 h at 4 °C. The flow-through was collected, and the resin washed [20 mM Tris, 300 mM NaCl, and 10 mM imidazole (pH 7.5)] to obtain all of the cleaved construct. The rebind flow-through and wash were then pooled and dialyzed against 4 L portions of folding buffer [20 mM Tris and 300 mM NaCl (pH 7.5)] at 4 °C. Two buffer exchanges were made, both after 3 h intervals, before the sample was left overnight to allow for correct disulfide bond formation. The dialyzed sample was then concentrated with an Amicon Ultra Centrifugal Filter (3 kDa molecular weight cutoff) prior to further purification via size exclusion chromatography (SEC) using a SuperdexPeptide 10/30 HR column with 25 mM Na_2HPO_4 and 150 mM NaCl (pH 7.4) as the mobile phase. Samples were taken following SEC, and the identity of the FD was confirmed through matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry. Measurements were performed on a

Bruker Autoflex Speed instrument using a 2,5-dihydroxybenzoic acid (DHB) matrix.

NMR Experiments. NMR spectra were recorded using a Shigemi NMR tube with a sample volume of ~300 μL in 25 mM Na_2HPO_4, 150 mM NaCl buffer (pH 7.4) with 90% H_2O/10% D_2O. Experiments were carried out on a Bruker Ultrashield 600 MHz magnet with a CPTXI 600S3 H-C/N-D-05 Z cryoprobe at a temperature of 23 °C. Backbone assignment was performed following HNCA, HN(CO)CA, HNCO, HN(CA)CO, and HN(CA)CB experiments using 20% or 25% non-uniform sampling (NUS). All data were processed using NMRPipe[25] and NMRFAM-SPARKY[26] via NMRBox.[27] Triply labeled (^{2}H, ^{13}C, and ^{15}N) samples were produced solely for the HN(CA)CB experiments, to improve the signal-to-noise ratio. Dodecylphosphocholine (DPC) was applied as a membrane mimic at 100 mM for the backbone assignment. Accurate DPC concentrations were gathered via phosphorus NMR using an Ascend 800 MHz magnet with a CPQCI 1H-31P/13C/15N/D Z-GRD cryoprobe. The integral of the DPC peak was compared to that of the buffer (25 mM Na_2HPO_4) to ascertain the DPC concentration. T_1, T_2, and heteronuclear NOE experiments were carried out to determine the dynamic properties of both pre- and postfusion states of the FP. T_1 and T_2 data were processed using the Bruker Topspin Dynamics Center, while heteronuclear NOEs were processed via NMRFAM-SPARKY.[26] The error shown is propagated from the signal-to-noise ratio or standard error of the mean (SEM).

Paramagnetic Relaxation Enhancement (PRE). To probe membrane association and depth, gadolinium(III) diethylenetriaminepentaacetic acid [Gd(DTPA)] (Sigma), 16-doxyl stearic acid (16-DSA) (Sigma) and 5-DSA (Avanti Polar Lipids) were used in PRE experiments. Gd(DTPA) was solubilized in autoclaved H_2O at a concentration of 100 mM and titrated into the NMR sample to achieve concentrations of 0.1, 0.5, 1, 2, 3, and 5 mM. For the 5- and 16-DSA titrations, known concentrations of DSA were dried down via a nitrogen stream while vortexing to remove chloroform. The NMR sample was then added to the DSA film to resuspend the paramagnetically tagged stearic acid. The 5- and 16-DSA titrations were carried out for concentrations of 0.5, 1, 2, 3, and 5 mM. All data were collected through ^{1}H–^{15}N HSQC at 23 °C, and none of the paramagnetic agents altered the structure of the peptide, which is evident in the lack of chemical shift perturbations. For all non-overlapping signals at each titration point, the relative intensity was calculated as $(I_0 - I_n)/I_0$, where I_0 is the peak intensity prior to addition of paramagnetic agents and I_n is the peak intensity at titration point n. For the Gd(DTPA) data, I_0 was the first titration point due to increasing signal intensity as a result of the paramagnetic effect.[28] The error shown is propagated from the signal-to-noise ratio or SEM.

Circular Dichroism (CD) Spectroscopy. CD spectra were recorded on a Jasco J810 spectropolarimeter using a quartz cuvette with a 1 mm path length. All experiments were carried out at 23 °C in 2.5 mM Na_2HPO and , 15 mM NaCl (pH 7.4) with 20 μM protein. Concentrations of 1 mM SUVs and 100 mM DPC was used where applicable. Data were collected from 260 to 190 nm with a step size of 1 nm at 20 nm/min and averaged over three accumulations. The spectra were converted from millidegrees to mean residue molar ellipticity (θ) to calculate the percentage helix at 222 nm using the term $(\theta_{222} - \theta_C)/(\theta_H - \theta_C)$, where $\theta_H = (250T - 44000)(1 - 3/n)$

and $\theta_C = 2220 - 53T$,[29] where T is the temperature in degrees Celsius and n is the number of residues in the peptide.

Preparation of Small Unilamellar Vesicles. Small unilamellar vesicles (SUVs) were produced by mixing specified amounts of lipid stock solutions in glass test tubes. Chloroform was removed via gently vortexing while applying a continuous stream of nitrogen, and residual solvent was further evaporated under vacuum overnight. The lipid film was then resuspended in 2.5 mM Na_2HPO_4 and 15 mM NaCl (pH 7.4) and sonicated for 15 min at 10% duty cycle (1 s on and 1 s off) with the sample saturated in ice−water, using a Branson ultrasonicator microtip.

■ RESULTS

Structural Investigation of the Pre- and Postfusion States of the Fusion Domain. The process of viral membrane fusion begins with the insertion of a fusogenic domain into the host cell membrane. Due to the hydration layer of water molecules and lipid packing, this is an energetically unfavorable process.[30] The role of the FD is to undergo a structural change to embed within and disrupt the host cell lipid bilayer, initiating the process of membrane fusion. Here we show the presence of a prefusion and postfusion state of the SARS-CoV-2 FD (Figure 2).

In solution, the FD displays an almost completely random coil secondary structure, which is evident by the large negative dip at 200 nm when assessed through CD (Figure 2A). Following the introduction of a membrane mimic, whether that be DPC or small unilamellar vesicles (SUV), a change in secondary structure is witnessed (Figure 2A). This can be characterized by the distinctive double dip at 208 and 222 nm in the spectra, indicative of α-helical content. In the presence of POPC/POPG (1:1) SUVs, the FD displayed 19% α-helical content that agrees well with the witnessed value of 20% with 100 mM DPC.[29] The presence of a similar structural transition with both DPC micelles and the bilayer-forming SUVs validates the use of DPC micelles for further structural investigation.

To further examine this conformational change, solution NMR was utilized. Through ^{1}H−^{15}N HSQC, a stark contrast between the pre- and postfusion states can be witnessed (Figure 2B). These large chemical shift perturbations following the addition of the membrane mimic DPC suggest that the FD is interacting with the membrane and undergoing significant structural changes, similar to that witnessed via CD. Following assignment of the backbone for both prefusion (Figure S2) and postfusion states (Figure 2B), a more detailed investigation of the change in secondary structure was pursued. Utilizing the chemical shift of C_α for all residues present within the FD, a chemical shift index (CSI) was generated from which the secondary structure was ascertained.[31] The chemical shifts are compared to that of a random coil, where a positive value indicates an α-helical configuration and a negative value a β-strand. The CSI reiterates the random coil secondary structure for the prefusion state as all values are <1 (Figure 2C, blue). However, for the postfusion state, it suggests that two α-helices separated by a short break (Figure 2C, red) form within the FP region. Specifically, we see helix 1 from residues F^{817}−L^{822} and helix 2 from residues G^{832}−G^{838}, with the intervening residues (F^{823}−A^{831}) serving as a linker region. This correlates well with the data acquired from CD, as both display an increase in α-helical character in the postfusion conformation. Interestingly, large secondary structure changes are evident toward the FP
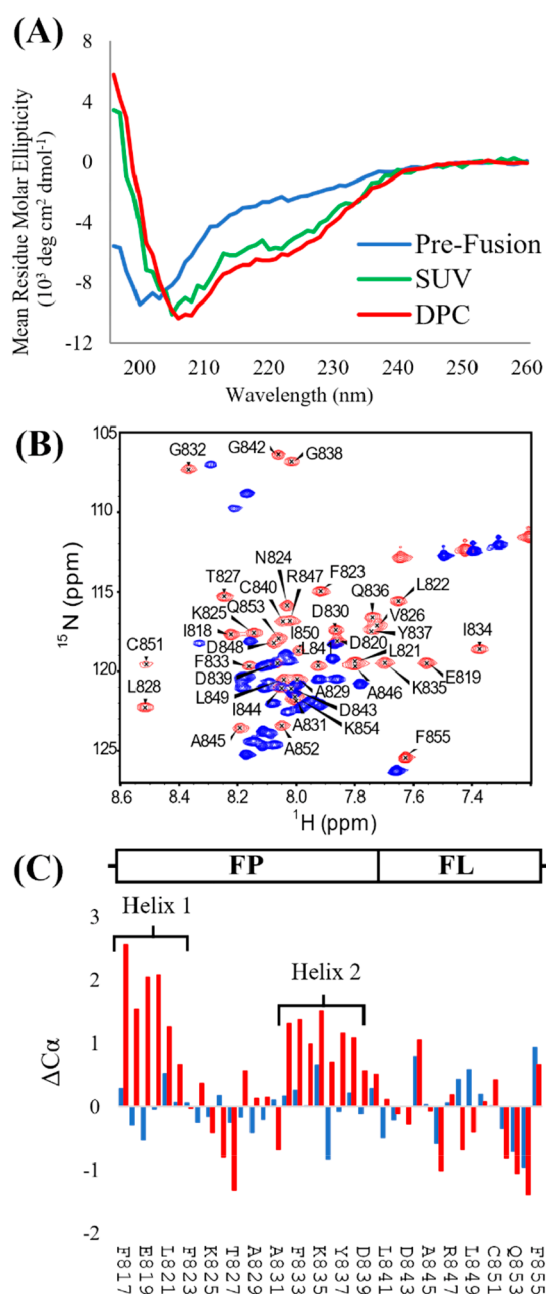
Figure 2. Prefusion (blue) and postfusion (red) states of the SARS-CoV-2 FD. (A) CD highlights the secondary structure change between the prefusion and postfusion states, caused by the introduction of 100 mM DPC micelles (red) or 1 mM small unilamellar vesicles (SUVs) (green). (B) Overlay of the ^1H–^{15}N HSQC spectra of the prefusion and postfusion states with the backbone assignment shown for the postfusion state. (C) Chemical shift index using $C\alpha$ values confirming the presence of two α-helices in the postfusion conformation of the FD.

(S^{816}–G^{838}), but much less so toward the FL (D^{839}–F^{855}) (Figure 2C). The internal disulfide bond within this region (C^{840}–C^{851}) most likely plays a large role in confining the FL to a similar structure within the pre- and postfusion states. Correct folding of the FL to ensure internal disulfide bond formation was confirmed by size exclusion chromatography, with a single peak representing a monomeric FD (Figure S1B) and C_β chemical shifts of the oxidized cysteines, all of which are >35 ppm (Figure S3).[32] Furthermore, large chemical shift

perturbations are witnessed via ^1H–^{15}N HMQC after the disulfide bond is severed through addition of 5 mM TCEP to the NMR sample (Figure S4).

Elucidation of the Interactions between the Fusion Domain and Membrane. To better understand how the postfusion state of the FD interacts with the membrane, the dynamic behavior of the protein was investigated. It is to be expected that the prefusion state would be significantly more dynamic than the postfusion state, with the latter restricted by its association with the membrane. Through spin–lattice T_1, spin–spin T_2, and ^1H–^{15}N NOE experiments, clear differences in the dynamics of the pre- and postfusion states can be witnessed (Figure 3).

The overall relaxation times of the FD in its prefusion state [^1H–^{15}N NOEs = 0.207 ± 0.040 (Figure 3A), $T_1 = 0.563 \pm 0.010$ s (Figure 3B), and $T_2 = 0.272 \pm 0.008$ s (Figure 3C)] agree very well with the dynamics of a random coil.[33] The postfusion state on the contrary [^1H–^{15}N NOEs = 0.625 ± 0.026 (Figure 3A), $T_1 = 0.808 \pm 0.007$ s (Figure 3B), and $T_2 = 0.090 \pm 0.005$ s (Figure 3C)] represents that of a much more restricted structural conformation as expected. These values correlate excellently with previous work carried out on the influenza FP that undergoes a transition to a more structured form when bound to DPC micelles.[34] Moreover, PRE experiments were conducted to explore the extent to which the FD inserts itself within the membrane. The water-soluble Gd(DTPA) broadens the signals of those residues that are solvent-exposed and further emphasizes the difference between the pre- and postfusion states (Figure 4A). In the prefusion state, we see much more severe quenching by Gd(DTPA) due to the entire peptide being exposed to solvent in the absence of DPC. Interestingly, the postfusion conformation also displays quenched signals throughout (Figure 4A). Although to a lesser extent than the prefusion state, this still suggests that no residues within the FD are found to be sufficiently deep within the membrane to remain unaffected by Gd(DTPA). To further investigate this hypothesis, the paramagnetic probes 5-DSA and 16-DSA, which can relay specific membrane depth information, were incorporated into the DPC micelles. Intriguingly, similar trends were witnessed with both 5-DSA (Figure 4B) and 16-DSA (Figure 4C). However, if we look more closely at the FP and FL regions independently, we witness something interesting. The relative intensities from the 16-DSA probe (Figure 4C) are very similar, 0.569 ± 0.010 and 0.582 ± 0.138 for the FP and FL, respectively. However, for the 5-DSA probe, we see a noticeable decrease in the relative intensity for the FP (0.507 ± 0.009) than we do for the FL (0.685 ± 0.012). This can be attributed to the FP region existing in the proximity of the 5-DSA probe compared to the FL. Furthermore, residues F^{823}–D^{830}, a linker region found between the α-helices of the FP (Figure 2B), are the most severely quenched by both DSA probes. This suggests that F^{823}–D^{830} is the deepest point of insertion of the FD within the membrane. Taking all of the PRE data into consideration, we believe that the FD associates with the membrane in a largely shallow manner. Specifically, the FP embeds into the membrane to approximately the 5-DSA probe, with the stretch of amino acids between the two α-helices (F^{823}–D^{830}) being the deepest point of insertion. As the FL is not as severely quenched by 5-DSA, we believe that it does not insert itself within the membrane but rather interacts with the lipid headgroups in a more superficial manner.
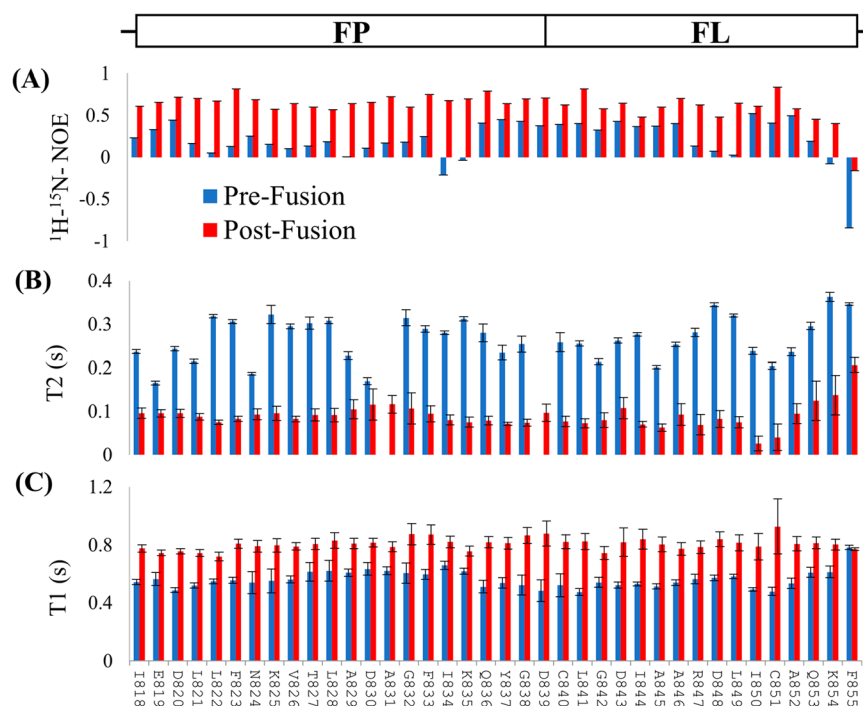
Figure 3. Solution NMR relaxation experiments measuring the (A) ^{1}H−^{15}N NOE, (B) T_2, and (C) T_1 relaxation rates for both the prefusion (blue) and postfusion (red) states. In all measurements, the prefusion state can be found to be significantly more dynamic than the postfusion state due to the latter being restricted by its association with the membrane. Amino acids A831 and D839 are missing from the prefusion T_2 measurements due to signal overlap. Error bars are from the signal-to-noise ratio.

■ DISCUSSION

There have been a plethora of structural studies, both physical and computational, carried out on the SARS-CoV-2 spike protein since its discovery in December 2019.[4,5,16,35−41] While this has allowed the binding of the spike protein to the host cell receptor, ACE2, to be investigated extensively, the molecular mechanisms behind membrane fusion remain elusive. Imperative in this pursuit is the elucidation of the spike protein while associated with a lipid membrane; however, this poses its own challenges. The structure of the membrane-associated (postfusion) SARS-CoV-2 spike protein currently available (Protein Data Bank entry 6XRA[16]) is missing a large region in the S2 subunit (I^{770}−T^{912}), which contains the complete FD (S^{816}−F^{855}). Insight into how the FD of the spike protein associates with and perturbs the host cell membrane to initiate fusion is integral in understanding how the virus enters the cell. This is evident in previous research for influenza,[12−14,34,42,43] HIV,[9,10,44] and EBOV,[15,45] where structural elucidation of the FD provided much needed molecular insight into the process of membrane fusion, something that has been severely lacking in the coronavirus family, with a clear gap in the knowledge regarding the structure of the FD in a membrane. Research into the coronavirus FD has been hindered by debate regarding its location, with several proposed candidates.[46−50] However, the region S^{816}−F^{855} in SARS-CoV-2 has the strongest case as the FD due to a high degree of sequence conservation, N-terminal location, and high fusogenic activity.[19−21] These characteristics are all key requisites for the FDs of viral proteins.[30] The FD of SARS-CoV-2 (S^{816}−F^{855}) consists of two structurally distinct regions, the fusion peptide (FP, S^{816}−G^{838}) and the fusion loop (FL, D^{839}−F^{855}). Both the FP and the FL have been shown to display independent fusogenic activity but are most effective in

synergy.[21] Hence, we believe that the FP and FL work in concert to create a highly fusogenic structure, one that resembles two different classes of fusion machinery. The FP is akin to the N-terminal namesakes witnessed in HIV[10,44] and influenza,[12,42] while the internal FL resembles the structure seen in EBOV.[15] This novel arrangement of fusogenic machinery within a single virus glycoprotein is highly unique and could potentially explain the incredible fusogenicity of SARS-CoV-2.

Structural plasticity is a key feature of N-terminal FPs that allows them to insert themselves within and perturb host cell membranes.[51] Our data show that the N-terminal FP undergoes structural rearrangement in the presence of a membrane (Figure 2). A transition from an unstructured (prefusion) to structured (postfusion) state as the FP interacts with the membrane results in the reorganization of lipids, decreasing the activation energy necessary for membrane fusion.[30] Additionally, CSI revealed that the FP forms two α-helices upon interaction with a lipid membrane (Figure 2C). A general propensity to form α-helical secondary structure has been seen previously for the FD of SARS-CoV-2, as well as for the FP of SARS-CoV-1 when separated from the FL.[19,22] Between helix 1 (F^{817}−L^{822}) and helix 2 (G^{832}−G^{838}) exists a linker region (F^{823}−A^{831}). The linker not only connects the two α-helices but also through PRE investigation appears to be the deepest point of membrane insertion and thus likely plays an integral role in perturbing the host cell membrane, in agreement with models created through computational studies.[17,52] We are proposing that the FP undergoes a transition into an obtuse helix−turn−helix motif upon interacting with a lipid membrane. This conformation is often witnessed in membrane proteins,[53] and a similar structural transition is witnessed in the boomerang model of
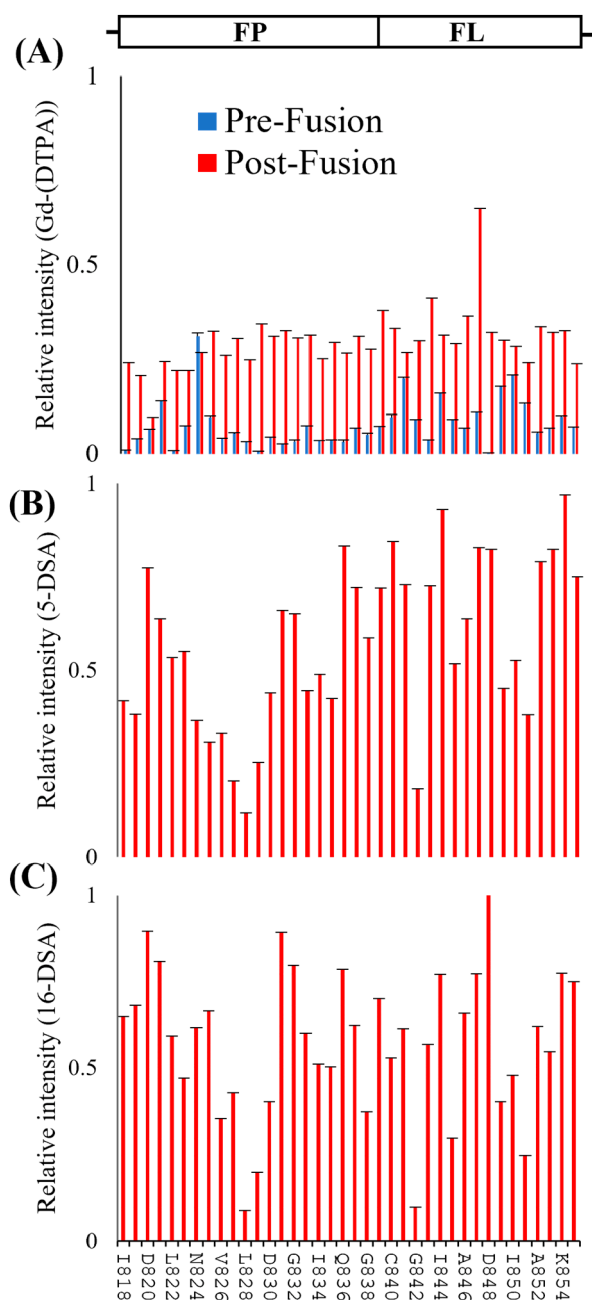
Figure 4. Paramagnetic relaxation enhancement (PRE) experiments with (A) Gd(DTPA), (B) 5-DSA, and (C) 16-DSA. (A) With Gd(DTPA), significantly more quenching is witnessed for the prefusion state (blue) than for the postfusion (red), indicating it is more exposed to solvent. Relative intensities witnessed in the presence of (B) 1 mM 5-DSA are more significantly quenched than with (C) 1 mM 16-DSA, although both show similar trends. Specifically, residues F^{823}–D^{830} display the largest reduction with both DSA probes. Data are capped at 1, although this is applied to only D848 in panel C. Error bars shown are propagated from the signal-to-noise ratio.

the influenza FP.[12,42] Previous work had identified the hydrophobic motif L^{821}–F^{823} as being essential for the FD's membrane interaction and perturbing effects.[19,21] Our structural data place this motif at the end of helix 1 and the start of the turn, which suggests it is one of the deepest points of the FD within the membrane. Interestingly, if F^{823} was in the

helix, then it would be placed on the opposite side to that of L^{821} and L^{822} (Figure S5), whereas a turn may allow all three residues to face into the membrane core. Deeper insertion within the membrane often suggests stronger perturbing effects and fusion activity, due to stronger destabilization of lipid packing and subsequent lowering of the activation energy necessary for membrane fusion.[54] We believe that the bulky aromatic ring of F^{823} may play a key role in destabilizing the outer leaflet of the membrane and should be investigated further.

On the contrary, in the disulfide-bonded FL we see that the secondary structure remains undefined in both pre- and postfusion states (Figure 2C). This disulfide bond is highly conserved throughout the entire coronavirus family,[20] implying that it is in some way important for the virus. Intriguingly, a large portion of the FD encompassing the entire FL (N^{824}–F^{855}) has been implicated as a pH-dependent refolding region.[5] This region found at the spike–interdomain interface within the spike protein trimer on the viral surface was observed to undergo a dramatic conformational change at endosomal pH (pH 5.5), resulting in coordinated movements of the receptor binding domain (RBD) and potential involvement in the evasion of the immune response.[5] While this suggests that the FL may play an important role in the conformation of the glycoprotein trimer, it also points toward functional implications within the endocytic pathway, similar to those witnessed within the internal FL of EBOV.[15] An interesting idea about the FL's function surrounds cleavage by endosomal cathepsins. An auxiliary cleavage site has been identified between A^{829} and D^{830}, which is in the middle of helix 2 of the FP and before the FL.[55] If a cleavage event occurred here, then that would leave only the internal FL available to initiate membrane fusion for the spike protein. Although this theory requires further investigation, this could provide the virus with a potentially different mechanism to initiate fusion, one that may be preferred within the endocytic pathway.

On the basis of the data, we have created a structural model for the FD embedded within the membrane (Figure 5). Upon
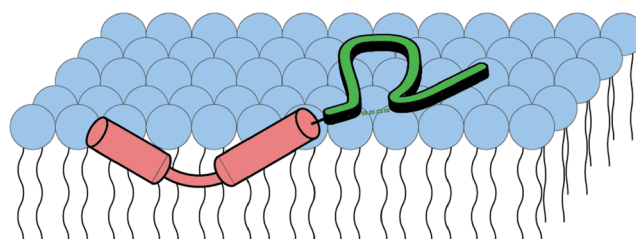

Figure 5. Structural model of the fusion domain of SARS-CoV-2 embedded in the host cell membrane. The fusion peptide (S^{816}–G^{838}) is colored red, with a helix–turn–helix motif, and embeds itself within the membrane in a shallow manner, just below the lipid headgroups. Alternatively, the fusion loop (D^{839}–F^{855}) is colored green and lays upon the surface of the membrane.

examining how deeply the FD inserts itself into the membrane, we witnessed that the FP (0.507 ± 0.009) was quenched by the 5-DSA probe more so than the FL (0.685 ± 0.012). Interestingly, both the FP and the FL displayed almost identical quenching data from the 16-DSA probe. Our hypothesis is that the FP is found within the membrane around the 5-DSA region, just below the lipid headgroups. As the FL experiences less quenching by 5-DSA, we believe it to

be more superficially associated with the surface of the membrane. The similar quenching data witnessed with 16-DSA may be due to both regions lying at the limit of the probe's paramagnetic effect, making it difficult to interpret the data gathered from the experiment. Therefore, within our model, the FP is embedded within the membrane just below the lipid headgroups, while the FL lays upon the lipid surface in an orientation perpendicular to the membrane (Figure 5). The FL may play a key role interacting with environmental factors such as divalent ions, and unique lipid headgroups, which are easily accessible for the FL within our model. Ca^{2+} has already been implicated as a cofactor in the membrane perturbing activity of the FD with binding sites found within both the FP and the FL, while specific lipids have been shown previously to have dramatic effects on the structure and activity of FDs in several viruses.[43,44,56]

In conclusion, we have provided structural evidence of a working model for the FD of SARS-CoV-2 when associated with a plasma membrane (Figure 5). Our model consists of both an N-terminal FP and an internal FL. The FP forms a helix−turn−helix motif upon membrane association and is embedded just below the lipid headgroups, while the FL interacts with the membrane in a more superficial manner, which we believe is in order to remain available to interact with environmental factors that may increase fusogenicity. The structural insights provided here can be used to better understand the mechanistic process of SARS-CoV-2 membrane fusion, and further work to determine the tertiary structure of the complete FD will build upon this work to provide a more in-depth understanding of how the FD interacts with the host cell membrane. Structural elucidation of the FD has the potential to act as a platform for the development of novel antiviral therapeutics to target the FD and prevent membrane fusion by SARS-CoV-2.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.1c00543.

Expression and purification of the fusion domain from the SARS-CoV-2 spike protein (Figure S1), backbone assignment for the prefusion state of the SARS-CoV-2 fusion domain (Figure S2), $C\beta$ strips confirming the presence of an internal disulfide bond via verification of oxidized Cys residues (Figure S3), conformational change induced following breaking of the disulfide bond (Figure S4), and amphipathic nature of helices 1 and 2 in the fusion peptide (Figure S5) (PDF)

Accession Codes

UniProtKB P0DTC2.

AUTHOR INFORMATION

Corresponding Author

Jinwoo Lee − *Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, United States;* orcid.org/0000-0002-1433-2810; Email: jinwoo@umd.edu

Author

Daniel Birtles − *Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742, United States*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.biochem.1c00543

ABBREVIATIONS

SARS-CoV, severe acute respiratory syndrome coronavirus; FD, fusion domain; FP, fusion peptide; FL, fusion loop; NMR, nuclear magnetic spectroscopy; CD, circular dichroism; PRE, paramagnetic relaxation enhancement; DPC, dodecylphosphocholine; SUVs, small unilamellar vesicles; HSQC, heteronuclear single-quantum coherence; HMQC, heteronuclear multiple-quantum coherence; EPR, electron paramagnetic resonance.

REFERENCES

(1) New Cases of COVID-19 in World Countries. Johns Hopkins Coronavirus Resource Center, Johns Hopkins University: Baltimore, MD, 2021.

(2) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.-Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* 2020, 181, 894−904.e9.

(3) 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.

(4) 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.

(5) Zhou, T.; Tsybovsky, Y.; Gorman, J.; Rapp, M.; Cerutti, G.; Chuang, G. Y.; Katsamba, P. S.; Sampson, J. M.; Schön, A.; Bimela, J.; Boyington, J. C.; Nazzari, A.; Olia, A. S.; Shi, W.; Sastry, M.; Stephens, T.; Stuckey, J.; Teng, I. T.; Wang, P.; Wang, S.; Zhang, B.; Friesner, R. A.; Ho, D. D.; Mascola, J. R.; Shapiro, L.; Kwong, P. D. Cryo-EM Structures of SARS-CoV-2 Spike without and with ACE2 Reveal a pH-Dependent Switch to Mediate Endosomal Positioning of Receptor-Binding Domains. *Cell Host Microbe* 2020, 28, 867−879.

(6) Skehel, J. J.; Wiley, D. C. Coiled Coils in Both Intracellular Vesicle and Viral Membrane Fusion. *Cell* 1998, 95, 871−874.

(7) 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.

(8) Elbe, S.; Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Chall* 2017, 1, 33−46.

(9) Li, Y.; Tamm, L. K. Structure and plasticity of the human immunodeficiency virus gp41 fusion domain in lipid micelles and bilayers. *Biophys. J.* 2007, 93, 876−885.

(10) Dimitrov, A. S.; Rawat, S. S.; Jiang, S.; Blumenthal, R. Role of the Fusion Peptide and Membrane-Proximal Domain in HIV-1 Envelope Glycoprotein-Mediated Membrane Fusion. *Biochemistry* **2003**, *42*, 14150−14158.

(11) Peisajovich, S. G.; Epand, R. F.; Pritsker, M.; Shai, Y.; Epand, R. M. The polar region consecutive to the HIV fusion peptide participates in membrane fusion. *Biochemistry* **2000**, *39*, 1826−1833.

(12) Han, X.; Bushweller, J. H.; Cafiso, D. S.; Tamm, L. K. Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. *Nat. Struct. Biol.* **2001**, *8*, 715−720.

(13) Gray, C.; Tatulian, S. A.; Wharton, S. A.; Tamm, L. K. Effect of the N-terminal glycine on the secondary structure, orientation, and interaction of the influenza hemagglutinin fusion peptide with lipid bilayers. *Biophys. J.* **1996**, *70*, 2275−2286.

(14) Gething, M. J.; Doms, R. W.; York, D.; White, J. Studies on the mechanism of membrane fusion: site-specific mutagenesis of the hemagglutinin of influenza virus. *J. Cell Biol.* **1986**, *102*, 11−23.

(15) Gregory, S. M.; Harada, E.; Liang, B.; Delos, S. E.; White, J. M.; Tamm, L. K. Structure and function of the complete internal fusion loop from Ebolavirus glycoprotein 2. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11211−11216.

(16) Cai, Y.; Zhang, J.; Xiao, T.; Peng, H.; Sterling, S. M.; Walsh, R. M.; Rawson, S.; Rits-Volloch, S.; Chen, B. Distinct conformational states of SARS-CoV-2 spike protein. *Science* **2020**, *369*, 1586−1592.

(17) Gorgun, D.; Lihan, M.; Kapoor, K.; Tajkhorshid, E. Binding mode of SARS-CoV-2 fusion peptide to human cellular membrane. *Biophys. J.* **2021**, *120*, 2914−2926.

(18) Schaefer, S. L.; Jung, H.; Hummer, G. Binding of SARS-CoV-2 Fusion Peptide to Host Endosome and Plasma Membrane. *J. Phys. Chem. B* **2021**, *125*, 7732−7741.

(19) Madu, I. G.; Roth, S. L.; Belouzard, S.; Whittaker, G. R. 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.* **2009**, *83*, 7411−7421.

(20) Madu, I. G.; Belouzard, S.; Whittaker, G. R. SARS-coronavirus spike S2 domain flanked by cysteine residues C822 and C833 is important for activation of membrane fusion. *Virology* **2009**, *393*, 265−271.

(21) Lai, A. L.; Millet, J. K.; Daniel, S.; Freed, J. H.; Whittaker, G. R. The SARS-CoV Fusion Peptide Forms an Extended Bipartite Fusion Platform That Perturbs Membrane Order in a Calcium-Dependent Manner. *J. Mol. Biol.* **2017**, *429*, 3875−3892.

(22) Lai, A. L.; Freed, J. H. SARS-CoV-2 Fusion Peptide has a Greater Membrane Perturbating Effect than SARS-CoV with Highly Specific Dependence on Ca2+. *J. Mol. Biol.* **2021**, *433*, 166946.

(23) Malakhov, M. P.; Mattern, M. R.; Malakhova, O. A.; Drinker, M.; Weeks, S. D.; Butt, T. R. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J. Struct. Funct. Genomics* **2004**, *5*, 75−86.

(24) Guerrero, F.; Ciragan, A.; Iwaï, H. Tandem SUMO fusion vectors for improving soluble protein expression and purification. *Protein Expression Purif.* **2015**, *116*, 42−49.

(25) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, *6*, 277−293.

(26) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015**, *31*, 1325−1327.

(27) Maciejewski, M. W.; Schuyler, A. D.; Gryk, M. R.; Moraru, I. I.; Romero, P. R.; Ulrich, E. L.; Eghbalnia, H. R.; Livny, M.; Delaglio, F.; Hoch, J. C. NMRbox: A Resource for Biomolecular NMR Computation. *Biophys. J.* **2017**, *112*, 1529−1534.

(28) Kuriashkin, I. V.; Losonsky, J. M. Contrast enhancement in magnetic resonance imaging using intravenous paramagnetic contrast media: A review. *Veterinary Radiology* **2000**, *41*, 4−7.

(29) Han, X.; Tamm, L. K. A host-guest system to study structure-function relationships of membrane fusion peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 13097−13102.

(30) Epand, R. M. Fusion peptides and the mechanism of viral fusion. *Biochim. Biophys. Acta, Biomembr.* **2003**, *1614*, 116−121.

(31) Wishart, D. S.; Sykes, B. D.; Richards, F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **1992**, *31*, 1647−1651.

(32) Sharma, D.; Rajarathnam, K. 13C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* **2000**, *18*, 165−171.

(33) Wirmer, J.; Peti, W.; Schwalbe, H. Motional properties of unfolded ubiquitin: a model for a random coil protein. *J. Biomol. NMR* **2006**, *35*, 175−186.

(34) Smrt, S. T.; Draney, A. W.; Lorieau, J. L. The Influenza Hemagglutinin Fusion Domain Is an Amphipathic Helical Hairpin That Functions by Inducing Membrane Curvature. *J. Biol. Chem.* **2015**, *290*, 228−238.

(35) 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.

(36) Barrett, C. T.; Neal, H. E.; Edmonds, K.; Moncman, C. L.; Thompson, R.; Branttie, J. M.; Boggs, K. B.; Wu, C.-Y.; Leung, D. W.; Dutch, R. E. Effect of clinical isolate or cleavage site mutations in the SARS-CoV-2 spike protein on protein stability, cleavage, and cell-cell fusion. *J. Biol. Chem.* **2021**, *297*, 100902.

(37) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **2020**, *369*, 330−333.

(38) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci. Rep.* **2020**, *10*, 14991.

(39) 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.

(40) Woo, H.; Park, S.-J.; Choi, Y. K.; Park, T.; Tanveer, M.; Cao, Y.; Kern, N. R.; Lee, J.; Yeom, M. S.; Croll, T. I.; Seok, C.; Im, W. Developing a Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein Model in a Viral Membrane. *J. Phys. Chem. B* **2020**, *124*, 7128−7137.

(41) Zhang, X.-Y.; Guo, J.; Wan, X.; Zhou, J.-G.; Jin, W.-P.; Lu, J.; Wang, W.-H.; Yang, A.-N.; Liu, D. X.; Shi, Z.-L.; Yuan, Z.-M.; Li, X.-G.; Meng, S.-L.; Duan, K.; Wang, Z.-J.; Yang, X.-M.; Shen, S. Biochemical and antigenic characterization of the structural proteins and their post-translational modifications in purified SARS-CoV-2 virions of an inactivated vaccine candidate. *Emerging Microbes Infect.* **2020**, *9*, 2653−2662.

(42) Lai, A. L.; Park, H.; White, J. M.; Tamm, L. K. Fusion Peptide of Influenza Hemagglutinin Requires a Fixed Angle Boomerang Structure for Activity. *J. Biol. Chem.* **2006**, *281*, 5760−5770.

(43) Takeda, M.; Leser, G. P.; Russell, C. J.; Lamb, R. A. Influenza virus hemagglutinin concentrates in lipid raft microdomains for efficient viral fusion. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 14610−14617.

(44) Lai, A. L.; Moorthy, A. E.; Li, Y.; Tamm, L. K. Fusion activity of HIV gp41 fusion domain is related to its secondary structure and depth of membrane insertion in a cholesterol-dependent fashion. *J. Mol. Biol.* **2012**, *418*, 3−15.

(45) Nathan, L.; Lai, A. L.; Millet, J. K.; Straus, M. R.; Freed, J. H.; Whittaker, G. R.; Daniel, S. Calcium Ions Directly Interact with the Ebola Virus Fusion Peptide To Promote Structure-Function Changes That Enhance Infection. *ACS Infect. Dis.* **2020**, *6*, 250−260.

(46) Mahajan, M.; Bhattacharjya, S. NMR structures and localization of the potential fusion peptides and the pre-transmembrane region of SARS-CoV: Implications in membrane fusion. *Biochim. Biophys. Acta, Biomembr.* **2015**, *1848*, 721−730.

(47) Mahajan, M.; Chatterjee, D.; Bhuvaneswari, K.; Pillay, S.; Bhattacharjya, S. NMR structure and localization of a large fragment of the SARS-CoV fusion protein: Implications in viral cell fusion. *Biochim. Biophys. Acta, Biomembr.* **2018**, *1860*, 407−415.

(48) Sainz, B.; Rausch, J. M.; Gallaher, W. R.; Garry, R. F.; Wimley, W. C. Identification and Characterization of the Putative Fusion Peptide of the Severe Acute Respiratory Syndrome-Associated Coronavirus Spike Protein. *J. Virol.* **2005**, *79*, 7195−7206.

(49) Guillén, J.; Kinnunen, P. K. J.; Villalaín, J. Membrane insertion of the three main membranotropic sequences from SARS-CoV S2 glycoprotein. *Biochim. Biophys. Acta, Biomembr.* **2008**, *1778*, 2765−2774.

(50) Guillén, J.; Pérez-Berná, A. J.; Moreno, M. R.; Villalaín, J. A Second SARS-CoV S2 Glycoprotein Internal Membrane-Active Peptide. Biophysical Characterization and Membrane Interaction. *Biochemistry* **2008**, *47*, 8214−8224.

(51) White, J. M.; Delos, S. E.; Brecher, M.; Schornberg, K. Structures and Mechanisms of Viral Membrane Fusion Proteins: Multiple Variations on a Common Theme. *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 189.

(52) Khelashvili, G.; Plante, A.; Doktorova, M.; Weinstein, H. Ca2+-dependent mechanism of membrane insertion and destabilization by the SARS-CoV-2 fusion peptide. *Biophys. J.* **2021**, *120*, 1105−1119.

(53) Oh, K. I.; Smith-Dupont, K. B.; Markiewicz, B. N.; Gai, F. Kinetics of peptide folding in lipid membranes. *Biopolymers* **2015**, *104*, 281−290.

(54) Qiang, W.; Sun, Y.; Weliky, D. P. A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 15314−15319.

(55) Bollavaram, K.; Leeman, T. H.; Lee, M. W.; Kulkarni, A.; Upshaw, S. G.; Yang, J.; Song, H.; Platt, M. O. Multiple sites on SARS-CoV-2 spike protein are susceptible to proteolysis by cathepsins B, K, L, S, and V. *Protein Sci.* **2021**, *30*, 1131−1143.

(56) Roth, S. L.; Whittaker, G. R. Promotion of vesicular stomatitis virus fusion by the endosome-specific phospholipid bis-(monoacylglycero)phosphate (BMP). *FEBS Lett.* **2011**, *585*, 865−869.