Received: 6 April 2022 | Revised: 21 June 2022 | Accepted: 22 June 2022

DOI: 10.1002/pro.4390

FULL-LENGTH PAPER

THE PROTEIN SOCIETY WILEY

# Exploring the pH dependence of the SARS-CoV-2 complete fusion domain and the role of its unique structural features

**Daniel Birtles** | **Anna E. Oh** | **Jinwoo Lee**

Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland, USA

**Correspondence**
Jinwoo Lee, Department of Chemistry and Biochemistry, University of Maryland, Biomolecular Sciences Building, College Park, MD 20742, USA.
Email: jinwoo@umd.edu

**Review Editor:** Hideo Akutsu

**Abstract**

SARS-CoV-2 may enter target cells through the process of membrane fusion at either the plasma (∼pH 7.4–7.0) or endosomal (∼pH 6.5–5.0) membrane in order to deliver its genetic information. The fusion domain (FD) of the spike glycoprotein is responsible for initiating fusion and is thus integral to the viral life cycle. The FD of SARS-CoV-2 is unique in that it consists of two structurally distinctive regions referred to as the fusion peptide (FP) and the fusion loop (FL); yet the molecular mechanisms behind how this FD perturbs the membrane to initiate fusion remains unclear. In this study via solution NMR, we witnessed only a slight conformational change in the FD between pH 7.4 and pH 5.0, resulting in a minor elongation of helix 1. However, we found that the FD's ability to mediate membrane fusion has a large and significant pH dependence, with fusion events being more readily induced at low pH. Interestingly, a biphasic relationship between the environmental pH and fusogenicity was discovered, suggesting a preference for the FD to initiate fusion at the late endosomal membrane. Furthermore, the conserved disulfide bond and hydrophobic motif "LLF" were found to be critical for the function of the complete FD, with minimal activity witnessed when either was perturbed. In conclusion, these findings indicate that the SARS-CoV-2 FD preferably initiates fusion at a pH similar to the late endosome through a mechanism that heavily relies on the internal disulfide bond of the FL and hydrophobic LLF motif within the FP.

**KEYWORDS**

CD, fusion domain, fusion loop, fusion peptide, membrane fusion, NMR, SARS-CoV-2, spike protein

## 1 | INTRODUCTION

The Coronavirus family contains several pathogenic viruses, the most infamous of which being severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2), the causative agent of coronavirus disease-2019 (COVID-19). Following the outbreak of SARS-CoV-2 in late 2019, COVID-19 quickly developed into a global pandemic,

which has infected more than 490 million people and led to over 6 million deaths.[1] Despite the arrival of several vaccines to combat the pandemic, it remains ongoing, mainly due to the rapid evolution of variants that contain increased rates of infectivity and an improved capacity to evade the immune response. Thus, a greater understanding of the molecular mechanisms behind SARS-CoV-2 infection, particularly concerning highly conserved regions, is necessity to help us better understand the unique characteristics of the coronavirus family and how they contribute to the virus's remarkable infectivity.

The success of a virus is determined by its ability to infect and propagate, the two primary goals of the viral lifecycle, the former of which is facilitated by the spike (S) glycoprotein in the coronavirus family. The spike protein can be subdivided into two functional units; S1, which is responsible for receptor binding and S2, which facilitates membrane fusion.[2,3] Following receptor binding, a cleavage event occurs within the S2 subunit at a site commonly referred to as S2', exposing the FD at the N-terminus of the cleaved S2 subunit.[4] From here, the FD is perfectly placed upon embedding into the target cells membrane to initiate the process of membrane fusion prior to the formation of the six-helix bundle; the most well accepted mechanism of fusion for Class I viruses.[5]As a result of this integral function, the FD is a well-conserved region throughout the coronavirus family (Figure 1) and is even found to be 100% conserved among all SARS-CoV-2 variants.

While FDs are generally well conserved within viral families, they can vary greatly across different viral families. For example, HIV and Influenza both contain N-terminal fusion peptides (FPs) but with different structures when embedded within the membrane.[6–8] Ebola virus contains an internal fusion loop (FL) that is held together by a disulfide bond and undergoes a large conformation change at low pH,[9] while the coronavirus FD contains both an N-terminal FP and an internal FL that has only ever been witnessed previously in the arenavirus family.[10,11] Nonetheless, all of the aforementioned FD's are responsible for the initiation of viral fusion through embedding within and perturbing their respective target membranes. It should be noted that different nomenclature has been used to describe the fusogenic domains found within the coronavirus family in the past, with the FP also referred to as FP1, and the FL referred to as the fusion peptide proximal region (FPPR) or FP2.[12,13] In this article, we will use the terminology FP and FL, this best captures the unique structural traits of these two defined regions that are thought to play important individual roles in the initiation of membrane fusion.

Previous work on the FD of SARS-CoV-1 displayed that both the FP and FL contained independent fusogenic activity but were most effective in unison, suggesting a synergistic relationship in the form of a bipartite fusion platform.[12] More recently, structural information gathered for the SARS-CoV-2 FD in dodecyl phosphocholine (DPC) micelles at pH 7.4[10] and 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine: 1,2-diheptanoyl-*sn*-glycero-3-phosphocholine ($DH_7PC$) bicelles at pH 5.0[14] has revealed further details. The FP was found to contain a helix-turn-helix motif that buries into the membrane, while the FL is more superficially associated with the lipid headgroups. The two structural models represent the FD embedded within the membrane following fusion at the plasma membrane or endosomal membrane respectively, as SARS-CoV-2 has been shown to be capable of entering target cells through both pathways.[15,16] Fusion at the endosomal membrane is often associated with low pH, as it is commonly found to be a major determinant for viral fusion.[17] The ability of SARS-CoV-2 to utilize either pathway in order to enter the cell via membrane fusion, is a potential contributing factor to the incredible infectivity of the virus. Hence, the molecular mechanisms of how SARS-CoV-2 initiates fusion via its FD at both plasma and endosomal membranes requires prioritized research to fully understand how these events transpire.

Here, we provide evidence that the FD of SARS-CoV-2 initiates membrane fusion more readily at low pH,

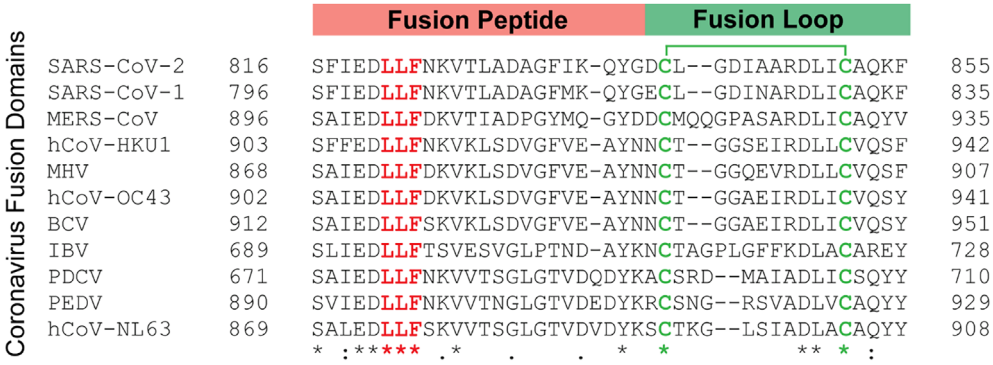


**FIGURE 1** Multiple sequence alignment of the FD in the coronavirus family. The structurally distinct regions that make up the FD, the FP and FL are shown. Highlighted in green is the conserved disulfide bond found within the FL and in red is the conserved hydrophobic motif LLF found within the FP

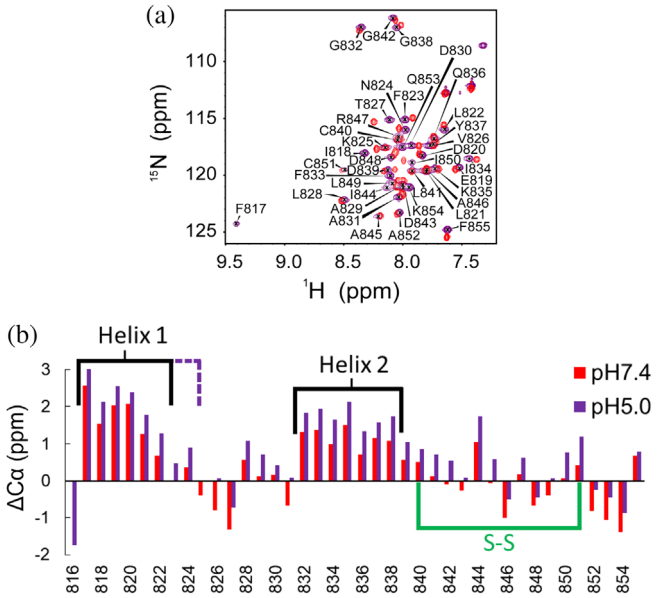


**FIGURE 2** A slight conformational change is witnessed for the SARS-CoV-2 FD embedded within DPC micelles between pH 7.4 (Red) and pH 5.0 (Purple). (a) A minor change in the tertiary conformation of the FD was detected via 1H-15 N HSQC chemical shift perturbations, with the backbone assignment shown at pH 5.0. (b) CSI with Cα data revealed the same secondary structure elements at both pH 7.4 and pH 5.0, with a slight elongation of helix 1 at pH 5.0. All experiments were carried out in 25 mM sodium phosphate, 150 mM NaCl and 100 mM DPC

suggesting a possible preference for the virus to enter the cell via the endocytic pathway. Through solution NMR a slight change in the conformation of the FD was observed between pH 7.4 and pH 5.0. However, utilizing a FRET-based lipid mixing assay the SARS-CoV-2 FD displayed significantly more efficient fusion at pH 5.0 than at pH 7.0, with a biphasic relationship found to exist between the pH and the fusogenicity of the FD. Furthermore, the perturbation of either the FP or the FL results in a drastic decrease in activity, highlighting the importance and synergistic relationship of both regions. We propose a model to explain the structure and function relationship involved in how the SARS-CoV-2 FD initiates membrane fusion.

## 2 | RESULTS

### 2.1 | A minor change in the conformation of the FD embedded within the membrane is present between pH 7.4 and pH 5.0

The majority of viruses may only enter the target cell via a single pathway, yet SARS-CoV-2 has been shown to fuse at both the plasma and endosomal membrane,[15,16] with one of the largest differences between the two membranes being the pH of the local environment. Previous work on other viral proteins demonstrates that the decreasing pH of the endocytic pathway can induce a structural transition to a more fusogenic conformation of the viral FD.[9,18] Lowering the local pH environment of the SARS-CoV-2 FD from pH 7.4 to pH 5.0 induces a slight conformational change within the peptide (Figure 2).

We previously showed that when associated with either a small unilamellar vesicle (SUV) or DPC micelle at pH 7.4, the conformation of the complete FD ($S^{816}$-$F^{855}$) undergoes a drastic conformational change compared to in solution.[10] When the FD is embedded within a DPC micelle, decreasing the pH to that present in the acidic environment of the late endosome (pH 5.0) appears to induce a minor structural rearrangement when compared to neutral pH (Figure 2). This lack of global structural change is reiterated by CD utilizing the more physiologically relevant bilayer environment of SUVs (Figure S1). Comparing the $^{1}$H-$^{15}$N HSQC spectra at pH 7.4 and pH 5.0, we witness minor chemical shift perturbations, suggestive of a change in tertiary conformation (Figure 2a). Then, utilizing Cα chemical shift indexing (CSI), we further examined the change in secondary structure. The defining structural features within the FP that allow the complete FD to embed within the DPC micelle at pH 7.4, the helix-turn-helix, also exists at pH 5.0 (Figure 2b). Within helix 1, a slight elongation was detectable from $L^{822}$ to $N^{824}$ in the CSI; this correlates well with the published structure of the FD embedded in a lipid bilayer in a wedge conformation at pH 5.0 (PDB: 7MY8).[14] A potential elongation of helix 2 from $G^{838}$ to $G^{842}$ was also detected in the pH 5.0 CSI (Figure 2b). However, we believe that this elongation is negated due to residues immediately downstream of $G^{838}$ serving as a flexible linker that connects the FP and the FL. Without this linker region, the wedge conformation seen in the published pH 5.0 structure (PDB: 7MY8) would be unable to form, offering an explanation as to why helix 2 ends at $G^{838}$ in all available experimental structural data of the complete FD.[10,14] We also compared the dynamic nature and depth of insertion within the DPC micelle of the complete FD at pH 7.4 and pH 5.0. Both $T_1$ and HN-NOEs revealed no significant changes in the dynamics of the peptide; yet, the $T_2$ data displayed more dynamic behavior when embedded in the membrane at pH 5.0 than at pH 7.4 (Figure S2). The 16-DSA paramagnetic probe also revealed little change based on changing pH alone, but the 5-DSA probe did suggest that the complete FD embeds deeper into the membrane at pH 5.0 (0.20 ± 0.02) than at pH 7.4 (0.32 ± 0.03) (Figure S3).

## 2.2 | The environmental pH is critical for the FD to initiate membrane fusion

Functionally, environmental pH can be associated with the fusogenicity of the FD at the plasma membrane (∼pH 7.4–7.0), as well as different stages of the endosomal membrane (∼pH 6.5–5.0). The endocytic pathway sees a large decrease in pH as the lysosomal compartment forms, ranging from pH 6.5–6.0 in the early endosome to pH 5.5–5.0 in the late endosome.[17] Here, we provide evidence that the environmental pH plays a significant role in the fusogenicity of the SARS-CoV-2 FD, with the virus preferentially fusing at a pH corresponding to that of the late endosomal membrane (Figure 3).

At pH 5.0, the FD can initiate membrane fusion at peptide/lipid (*P*/*L*) ratios as low as 0.01, with the relationship between the *P*/*L* ratio and the amount of membrane fusion found to increase in a linear fashion (Figure 3a). A *P*/*L* ratio of 0.01 is close to that of physiological conditions based on electron tomography analysis of the viral surface.[19,20] However, at pH 7.0 a dramatic decrease in the fusogenicity of the FD is observed, where even at much higher *P*:*L* ratios the amount of fusion witnessed is significantly lower than at pH 5.0 (Figure 3a). A linear correlation up to a *P*/*L* ratio of 0.2 is still observed at pH 7.0 signifying a positive relationship between the concentration of peptide and the amount of lipid mixing witnessed (Figure S4). This relationship eventually tails off at higher *P*/*L* ratios (>0.2), most likely due to oversaturation of the peptide on the liposome surface.

To probe the relationship between pH and fusogenicity further, we carried out the same functional assay over a range of several different environmental pH's that are relevant within the endocytic pathway (Figure 3b). A biphasic relationship was revealed where we see an initial increase in fusogenicity at a pH corresponding to that of the early endosome (pH 6.5–6.0) followed by a brief plateau. The jump in fusion from pH 7.0 to pH 6.0 is approximately sixfold, and thus signifies a dramatic increase in the fusogenic capacity of the FD shortly after being endocytosed. A second sharp increase in activity is then found at a pH resembling that of the late endosomal compartment (pH 5.0), and when we compare the ability of the FD to initiate fusion at pH 5.0 to that at pH 7.0 with the same *P*/*L* ratio, we see almost a ninefold increase in activity. Our results suggest that the FD can initiate fusion in the pH environment of the plasma membrane (∼pH 7.0) and early endosome (pH 6.5–6.0), yet as the pH decreases in the later stages of the endocytic pathway (pH 5.5–5.0) the probability of a successful membrane fusion event increases.

Our findings show that the SARS-CoV-2 FD preferentially initiates fusion in an environment where the pH is lower than that of the plasma membrane (<pH 7.0). Moreover, the two-stage fusion profile witnessed (Figure 3b) demonstrates that the FD can initiate fusion in the early endosome (∼pH 6.0), yet the potency of the FD increases by ∼50% at a pH corresponding to that of the late endosome (∼pH 5.0). This suggests that the SARS-CoV-2 not only favorably fuses via the endocytic pathway, but its likelihood of initiating a successful fusion event increases as the pathway progresses.

## 2.3 | The hydrophobic motif "LLF" of the FP and the internal disulfide bond within the FL are both integral for full fusogenicity of the complete FD

We earlier found that the complete FD consists of two structurally defined regions, the FP and the FL, with the former proving to be the primary site of interaction with

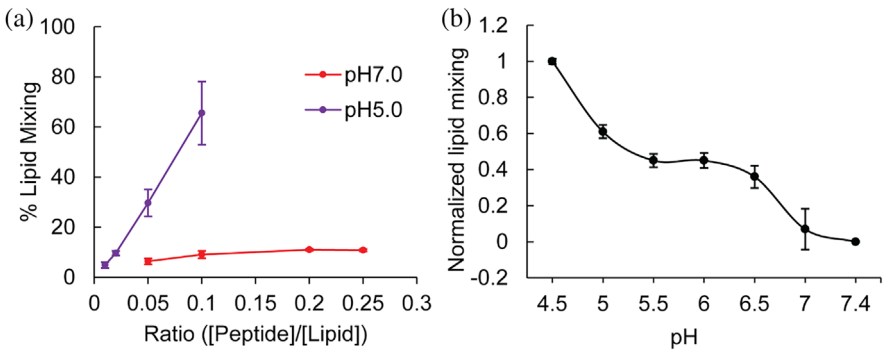


**FIGURE 3** The SARS-CoV-2 FD displays a large pH dependence to fusion in simple lipid compositions of 3:1 POPC:POPG. (a) A linear relationship can be seen with the P:L ratio at pH 5.0 (Purple), but drastically reduced fusion is present at pH 7.0 (Red) even at much higher peptide concentrations. (b) A positive, biphasic relationship is witnessed between decreasing environmental pH and lipid mixing with a *P*/*L* ratio of 0.05. All lipid mixing experiments were all carried out in 10 mM HMA, 100 mM NaCl with $n \geq 4$

the DPC micelle.[10] Previous work on the FP of SARS-CoV-1 found that a "hydrophobic motif" consisting of $L^{821}$, $L^{822}$, and $F^{823}$ played an important role in perturbing the membrane.[12,21] On the other hand, although the FL is known to be important for the activity of the complete FD, how it contributes to the initiation of membrane fusion at different pH's is still not completely understood.[21,22] Several previous studies have assessed the FP and FL in isolation, thus removing any potential synergistic interplay between the two with regards to the mechanism of initiating membrane fusion. Here, we have assessed the complete FD comprising both the FP and FL, where we alternately disrupted the LLF motif within the FP and the disulfide bond within the FL.

First, we mutated each individual amino acid in the LLF motif to alanine creating three single mutation constructs as well as a triple mutant; all of which displayed a remarkable loss of function (Figure 4a). We then explored whether or not the aromatic functional group of $F^{823}$ plays a key role in fusion by replacing this phenylalanine with a tyrosine ($F^{823}Y$) and then a tryptophan ($F^{823}W$) (Figure 4b). The $F^{823}Y$ mutant recovered very little functionality and was equivalent to $F^{823}A$. Opposingly, $F^{823}W$ recovered significantly more fusogenic activity than both $F^{823}A$ and $F^{823}Y$, although this was still approximately fivefold lower compared to that of the WT (Figure 4b). These results echo that witnessed in SARS-CoV-1, which also found each residue in the LLF motif to be critical toward membrane perturbation via viral infectivity assays.[21] All of the LLF mutants displayed the same global secondary structure as the WT (Figure 4d), ruling out global secondary structure perturbations as a reason for the large loss in function witnessed. Therefore, it is clear from the data that every residue within the LLF motif is integral toward the function of the complete FD, with a potential key role for the aromatic amino acid of $F^{823}$.

To assess the importance of the FL within the fusion mechanism of the complete FD, we used the reducing agent tris(2-carboxyethyl)phosphine (TCEP) to break the

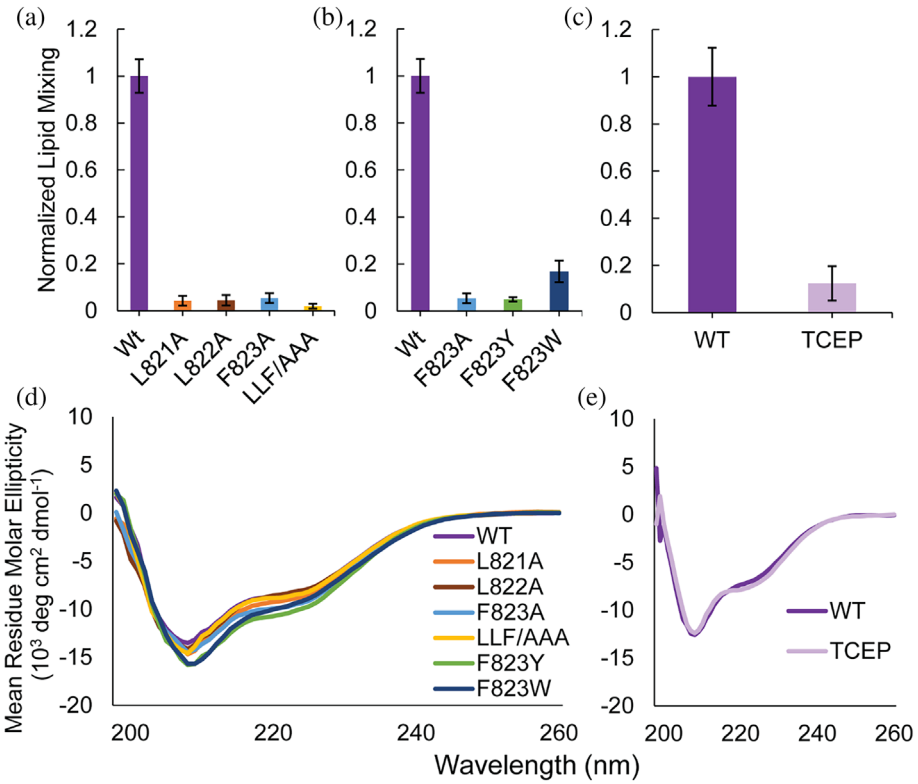


**FIGURE 4** The hydrophobic motif "LLF" and the FL play a key role in the fusogenicity of the complete FD at pH 5.0. (a) Mutating the LLF motif residues to alanine individually, or all three simultaneously, leads to an almost complete loss of function ($n \geq 12$). (b) Conserving the aromatic functional group at position 823 replenished some of the peptide's activity only when tryptophan was present, but not with tyrosine ($n = 16$). (c) The presence of 1 mM TCEP in the environment decreases the amount of lipid mixing witnessed substantially ($n = 16$). CD then confirmed that the global secondary structure of all (d) LLF mutants and the (e) WT in the presence of 500 μM TCEP was closely aligned with that of the WT in DPC micelles indicating that no significant changes in the global secondary structure had occurred. Liposomes consisting of POPC:POPG 3:1 were utilized for all lipid mixing experiments in 10 mM HMA 100 mM NaCl. All CD experiments were carried out in 1 mM HMA 10 mM NaCl 50 mM DPC

conserved internal disulfide bond. In the presence of 1 mM TCEP a significant decrease in fusogenic activity is witnessed (Figure 4c), with close to a 10-fold loss in fusogenicity observed if the internal disulfide bond is severed. Furthermore, the addition of TCEP to the WT brought about no change in the global secondary structure of the complete FD as witnessed through CD (Figure 4e). This suggests that the helix-turn-helix motif of the FP remains intact even after the disulfide bond of the FL is lost. Interestingly, we previously found chemical shift perturbations in the FL when comparing the complete WT FD to that present under reducing conditions via $^{1}H$-$^{15}N$ HSQC, indicating large tertiary structure changes.[10] Put together, this indicates that the helix-turn-helix motif of the FP still forms within the membrane even without the disulfide bonded FL, and the decrease in fusogenicity witnessed is likely due to a change in the tertiary orientation of the FL and its association/interaction with the membrane and/or with the FP.

## 3 | DISCUSSION

Much of the research involving the activity of the coronavirus FD has been carried out in SARS-CoV-1, which with a sequence conservation of ~95% would appear to be pertinent when it comes to the FD of SARS-CoV-2. To date, research surrounding the FD has either concentrated on the peptides activity at pH 5.0 or assessed the FP and FL independently.[14,21,23,24] Initially, both domains were found to contain some fusogenic properties in isolation, with the FP even displaying a preference for fusion at low pH.[21] Since then, it has been proven that the FP and FL act in a synergistic manner to exert a membrane perturbing effect that is greater than that witnessed when the two components were assessed independently.[12] Thus, in order to fully understand the molecular mechanisms behind how the FD of SARS-CoV-2 interacts with the membrane and initiates membrane fusion, it is integral to assess both the FP and FL together as a single complete FD at different functionally relevant pH's that represent the local environments of the plasma and endosomal membranes.

When investigating the mechanistic details of viral fusion, it is important that the contributions of other environmental factors toward membrane fusion such as $Ca^{2+}$ and lipids are also considered. In previous work across several different viruses, both $Ca^{2+}$ and the lipid composition of the target cell membrane have proven to be significant toward membrane fusion.[8,25–32] For the FD of SARS-CoV-2, limited research has been conducted regarding the impact of lipid compositions on membrane perturbation; yet, $Ca^{2+}$ has been shown to be significant at low pH.[32] Interestingly when investigating the role of $Ca^{2+}$ at neutral pH, we found that no significant effect was witnessed either structurally or functionally (Figure S5). In this study in an attempt to understand the importance of pH and the individual domains of the SARS-CoV-2 FD without the added layer of protein: $Ca^{2+}$, protein:lipid interactions, and/or changing membrane properties, we opted to use no $Ca^{2+}$ and a simple lipid composition of POPC:POPG 3:1, which is akin to that used previously in the study of several different viral FDs.[9,33–35]

SARS-CoV-2 is well established as being able to fuse at both the plasma membrane[15] and endosomal membrane[16] with the route taken thought to be due to the cell type being infected and the presence of particular proteases.[36] For the majority of endocytosed enveloped viruses, pH plays a major role in determining the site of viral fusion; yet, the direct impact of pH on the fusion machinery, specifically the function of the FD, is yet to be established. The complete FD of SARS-CoV-2 has been shown here to contain a clear pH dependency in order to initiate membrane fusion (Figure 3). The biphasic nature of this dependency points to a complex functional landscape of the FD in the endocytic pathway, where fusion can be initiated in the early endosome (pH 6.5–6.0) but may more readily occur in the late endosome as the pH continues to decrease (~pH 5.0) (Figure 3b). A similar profile has been witnessed previously for the internal fusion peptide (IFP) of the avian sarcoma/leukosisvirus subtype A virus.[37] Here, the two-step process of fusion was attributed to the more complete transition from hemi-fusion to full fusion at the lower pH, which is consistent with other studies on the pH requirements for fusion.[38] It is therefore reasonable to assume that this also occurs for the FD of SARS-CoV-2 where an increasing number of complete fusion events are witnessed as the environmental pH decreases, although further experimentation is necessary for this to be proven. Furthermore, a preference for SARS-CoV-1 and SARS-CoV-2 to fuse in the late endosome (~pH 5.0) has also been shown in vivo.[39,40] While one hypothesis is that this may be due to the buildup of cathepsin L and the subsequent cleavage of S2'; our data suggests that this ability to fuse later in the endocytic pathway could also be due to the increased fusogenicity of the FD at decreasing pH.

Structurally, we witnessed very little change in secondary structure in the FD between pH 7.4 and pH 5.0, with only a small elongation of helix 1 ($L^{822}$-$N^{824}$) apparent. The FP ($S^{816}$-$G^{838}$) remains the primary region that interacts with and perturbs the DPC micelle through a helix-turn-helix motif, with the turn in the middle of this motif being the deepest point embedded upon insertion. Meanwhile, the FL ($D^{839}$-$F^{855}$) appears to interact

superficially with the DPC micelle surface at both pH's. Based on our structural data in combination with the solved structure at pH 5.0 (PDB: 7MY8)[14] and MD simulations,[41,42] we are proposing that slight changes to the orientation of the FP alongside dynamic motions of the FL on the surface of the lipid headgroups are responsible for the drastic increase in functional activity of the complete FD witnessed at low pH (Figure 5). In this model at pH 5.0, the slight elongation of helix 1 sees a minor change in the orientation of the helix-turn-helix motif where the entire FP subsequently embeds deeper within the membrane. Simultaneously, we believe the FL becomes more dynamic and samples several orientations atop the membrane at lower pH, with the wedge conformation favored for membrane perturbation. Within this wedge conformation, the FL has several interactions with the FP[14] and thus as the FP embeds deeper, the FL is also drawn further into the membrane. This results in a much greater local membrane perturbing effect, as it would greatly impact the lipid packing in the immediate membrane environment. This increased depth of insertion throughout the complete FD within the lipid membrane at lower pH is also corroborated by our PRE data with 5-DSA (Figure S3). MD simulations assessing the binding of the complete SARS-CoV-2 FD to the endosomal membrane identified the disulfide bonded FL as a potential mechanical stabilizer, which served to further stabilize the FP inside the membrane and increase the force needed to remove it.[41] This gives credence to our hypothesis that the FL and FP work synergistically at low pH, increasing their collective membrane perturbing effects and associating more tightly with the membrane. Furthermore, this idea that the FL could be dynamic is reinforced by cryo-EM data of the whole spike protein carried out at a series of low pH's.[16] Zhou et al., identified the FL as a "pH switch," which underwent large structural changes prior to receptor binding in the endocytic pathway due solely to the decreasing pH. We believe the FL may undergo similar dynamic motions while it is associated with the endosomal membrane, with more conformations sampled at increasingly lower pH's until an appropriate orientation is found that allows for the initiation of fusion.

Within SARS-CoV-1 the hydrophobic motif LLF in the FP was shown to be integral to the regions capacity to initiate fusion.[12] This motif also appears to be critical for fusion in SARS-CoV-2, at both pH 5.0 (Figure 4a) and pH 7.0 (Figure S6), although the mechanistic details are still unclear. LLF is found immediately before the deepest embedded component of the complete FD, the turn region. Hence, this hydrophobic motif may be integral toward the correct orientation and subsequent insertion of this region within the membrane. The FD of influenza also contains a helix-turn-helix motif, although it is inversed in comparison to that of SARS-CoV-2 FD, with the turn the most shallow portion inserted in the membrane.[6,43] Crucial to the turn region within influenza are two aromatic sidechains that flank either side.[44] Within the SARS-CoV-2 FD, $F^{823}$ could be one such residue that plays a similar integral role and alongside $F^{833}$ at the beginning of helix 2, could be responsible for ensuring the turn region is correctly formed and embedded within the membrane. The small gain of function witnessed in $F^{823}W$ compared to that of $F^{823}A$ further strengthens this argument (Figure 2d), while simultaneously pointing to a necessity for phenylalanine at that position. A possible explanation for the lack of function witnessed with the $F^{823}Y$ mutant is that the hydroxyl group could form non-

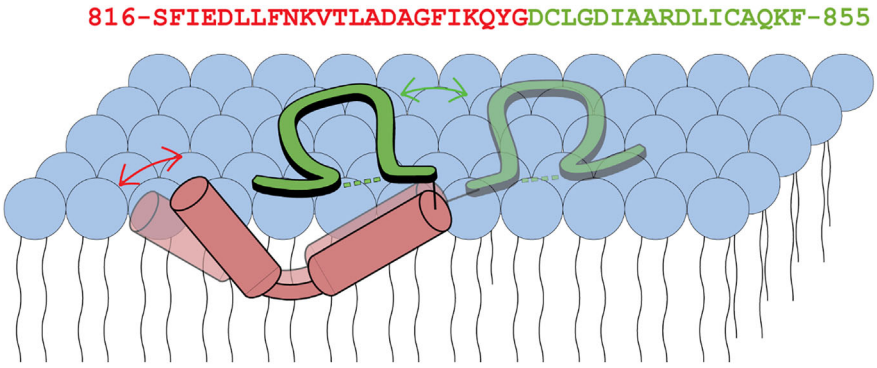


**FIGURE 5** Structural model of the FD at pH 7.4 (Opaque) and pH 5.0 (Full). The FP (Red) embeds in the membrane with a helix-turn-helix conformation at both pH's; yet at pH 5.0, we see a slight elongation of helix 1 resulting in the turn region burying deeper into the membrane. The FL (Green) interacts more superficially with the surface of the membrane and may display more dynamic motions at different pH's. A wedge conformation, with the FL between the two helices of the FP, has been shown previously to exist at pH 5.0. As the FP embeds deeper at low pH, it simultaneously draws the FL further into the lipid headgroups, leading to a greater membrane perturbing effect

native hydrogen bonds either within the FD or with the lipid headgroups, which may then prevent the FD from successfully initiating fusion. A similar proposal for the importance of $F^{823}$ and $F^{833}$ was put forward based on MD simulations prior to the full structural elucidation of the complete FD.[42] Within our model (Figure 5) the LLF motif plays an integral role in the pH dependent rearrangement of helix 1, as it becomes encapsulated in the elongated helix, which may be key to the increased fusogenicity of the FD witnessed at low pH.

An interesting idea regarding the molecular mechanism behind the initiation of membrane fusion is that SARS-CoV-2 undergoes a variation on the six-helix bundle that involves not just the FD but several domains in perturbing the membrane.[45] This would involve another region downstream of the FD in the S2 subunit that has also been identified as being able to initiate fusion and perturb the membrane in isolation, known as the IFP.[23,35,46–48] The mechanism requires the FP to serve as the primary anchor embedded within the target cell membrane, which is then reinforced by more superficial interactions between both the FL and IFL with the lipid headgroup region; all of which have been proven to occur experimentally.[10,14,23] In a similar fashion, HR1 would also perturb the lipid headgroup region of the target cell membrane via insertion of its amphipathic helix, with HR2 performing a similar role on the viral membrane. As both HR regions perturb the membrane alongside the complete FD and internal FL, the large energy barrier associated with membrane fusion is significantly lowered. Membrane fusion is then driven by the formation of the 6HB which physically forces the opposing membranes to come into close proximity and ultimately coalesce.[45] While this hypothesis is an enticing one and goes some way to explaining the reason for several fusogenic regions to exist within a single spike glycoprotein, more evidence is needed to fully establish this as the mechanism for SARS-CoV-2 membrane fusion.

In conclusion, the SARS-CoV-2 FD displays a remarkable pH dependence to its activity suggesting a preference to fuse at the late endosomal membrane. Structural information at different pH's indicates that slight changes in the orientation of the FP and dynamic motions of the FL could have large functional significance on this pH dependent mechanism for the initiation of membrane fusion. Additionally, the LLF motif of the FP and the internal disulfide bond of the FL have been shown to be critical toward the fusogenecity of the complete FD at both neutral and low pH. Here, we have provided further insight into the molecular mechanism behind how the SARS-CoV-2 virus initiates membrane fusion via its structurally distinct FD.

# 4 | MATERIALS AND METHODS

## 4.1 | Expression and purification

The SARS-CoV-2 FD construct $^{816}$SFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKF$^{855}$ was designed with an N-terminal 6x His-tag, followed by a small ubiquitin like-modifier tag from *Saccharomyces cerevisiae* (*SMT3* gene) to aid with solubility and expression.[49] This sequence was gathered from the SARS-CoV-2 spike protein with the UniProtKB accession number P0DTC2. Ni-NTA affinity chromatography was employed and following cleavage of both tags, the isolated FD was subjected to a total of 12 L of dialysis buffer to ensure the correct formation of the internal disulfide bond, with the presence of a pure monomer confirmed via gel filtration chromatography. The expression and purification of the SARS-CoV-2 FD has been described in detail previously.[10] All mutants in this study were purified in the same manner as the wild type.

## 4.2 | NMR experiments

NMR spectra were acquired using a shigemi NMR tube with a sample volume of ~300 µl in 25 mM Sodium Phosphate, 150 mM NaCl, pH 7.4 or pH 5.0 buffer with 90%$H_2O$/10%$D_2O$. All experiments were carried out on the 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% nonuniform sampling. All data was processed using the software NMRPipe and NMRFAM-SPARKY via NMRBox.[50–52] DPC was applied as a membrane mimic at 100 mM, with accurate DPC concentrations gathered via phosphorous NMR using the Bruker Ascend 800 MHz magnet with a CPQCI 1H-31P/13C/15 N/D Z-GRD Cryoprobe. Spin–lattice ($T_1$), spin–spin ($T_2$) and heteronuclear NOE experiments were carried out to determine the dynamic properties of the FD at both pH 7.4 and pH 5.0. $T_1$ and $T_2$ data was processed using the Bruker Topspin Dynamics Center, while heteronuclear NOEs were processed via NMRFAM-SPARKY. Paramagnetic relaxation enhancement (PRE) experiments were performed for both 5-doxyl stearic acid (5-DSA) (Avanti polar lipids, Alabaster, AL) and 16-DSA (Sigma Aldrich, St Louis, MO), with the procedure and data analysis described in detail previously.[10] Data shown here is using final concentrations of 2 mM for both PRE probes and the error shown is propagated from the signal to noise or standard error of the mean (SEM).

### 4.3 | Preparation of large Unilamellar vesicles (LUVs)

LUVs were produced by mixing specified amounts of lipid stock solutions in glass test tubes. Chloroform was removed via gentle vortexing while applying a continuous stream of nitrogen and residual solvent was further evaporated under vacuum overnight. The lipid film was resuspended in 10 mM HEPES/MES/Sodium acetate (HMA), 100 mM NaCl, pH 7.4 buffer and subjected to 10 freeze–thaw cycles between liquid nitrogen and a 42°C water bath. Finally, the liposomes were extruded using a liposofast extrusion kit (Avestin™), a total of 21 times through two polycarbonate membranes with a 100 nm pore size. LUVs were either used immediately or stored for a maximum of 72 hr at 4°C prior to use.

### 4.4 | Lipid mixing assay

The FRET based lipid mixing assay utilized here is well established in the field to assess the fusogenicity of viral FD's.[8,9,35,53] All LUVs used in this study were comprised of 75 mol% 16:0–18:1 POPC (1-palmitoyl-2-oleoyl-glycero-3-phosphocholine) and 25 mol% 16:0–18:1 POPG (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-[1′-rac-glycerol]). LUVs composed of these specified lipids were mixed with vesicles of a similar composition, but with 1 mol% of the fluorescently labeled 18:1 Liss Rhod PE (1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine-*N*-[lissamine rhodamine B sulfonyl]) and 18:1 NBD PE (1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine-N-[7-nitro-2-1,3-benzoxadiazol-4-yl]) at a ratio of 9:1 unlabeled: labeled. All lipids used in this study were supplied by Avanti polar lipids (Alabaster, AL). Experiments were conducted in a Corning Costar black walled, clear bottom 96 well plates with excitation and emission wavelengths at 460 nm and 538 nm, respectively. To prevent spectral overlap from the unwanted fluorophore a cut-off at 530 nm was present. Volumes of 150 µl per well were employed, with 5 µM FD and 100 µM LUVs (protein/lipid ratio of 0.05) unless stated otherwise, with all conditions containing at least four replicas ($n \geq 4$). Both liposomes and FD were suspended in 10 mM HMA, 100 mM NaCl, pH 7.4 prior to the experiment, with acidification to the described pH controlled by addition of a specific amount of 1 M HCl. Fluorescence was recorded using a SpectraMax M5 microplate with all readings taken at room temperature (~22°C). Percent lipid mixing was calculated as $\frac{(I_F - I_B)}{(I_{100} - I_B)} \times 100$, where $I_F$ is fluorescence intensity measured, $I_B$ is the background fluorescence, and $I_{100}$ is the 100% fluorescence intensity value gathered after complete vesicle rupture following the addition of 1% Triton X-100. No fusion was witnessed when the FD was incubated at pH 7.4 and thus this was used as the baseline measurement for all experiments prior to decreasing the pH. Controls containing no protein were ran alongside all experimental conditions and subtracted from the final values. Error was propagated from either the standard deviation of the sample and the control or SEM.

### 4.5 | Circular dichroism (CD) spectroscopy

CD spectra were collected on a Jasco J810 Spectro-Polarimeter using a quartz cuvette with a 2 mm path length. All experiments were carried out at room temperature (~22°C) in 1 mM HMA, 10 mM NaCl, 50 mM DPC, pH 7.4 or pH 5.0 with a protein concentration of ~20 µM. Data was collected from 260 nm to 190 nm with a step size of 1 nm at 50 nm/min and averaged over three accumulations, with all smoothing carried out using the program CDToolX.[54]

## AUTHOR CONTRIBUTIONS

**Daniel Birtles:** Conceptualization (equal); formal analysis (equal); investigation (lead); writing – original draft (equal); writing – review and editing (equal). **Anna E. Oh:** Formal analysis (supporting); investigation (supporting); writing – original draft (supporting). **Jinwoo Lee:** Conceptualization (lead); formal analysis (equal); supervision (lead); writing – review and editing (equal).

## ACKNOWLEDGMENTS

The authors would like to thank the University of Maryland, College Park for startup funds and are grateful to D. Zhang for assistance with solution NMR experimentation. They are also appreciative of the entire Lee lab for their help in the construction of this manuscript.

## CONFLICT OF INTEREST

The authors report no conflict of interest.

## ORCID

*Jinwoo Lee* https://orcid.org/0000-0002-1433-2810

## REFERENCES

1. Johns Hopkins Coronavirus Resource Center. New cases of covid-19 in world countries. MD, Maryland: Johns Hopkins, 2022. https://coronavirus.jhu.edu/map.html.
2. Hoffmann M, Kleine-Weber H, Pöhlmann S. A multibasic cleavage site in the spike protein of sars-cov-2 is essential for infection of human lung cells. Mol Cell. 2020;78(4):779–784.e775.

3. Jaimes JA, Millet JK, Whittaker GR. Proteolytic cleavage of the sars-cov-2 spike protein and the role of the novel s1/s2 site. iScience. 2020;23(6):101212.
4. Belouzard S, Chu VC, Whittaker GR. Activation of the sars coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proc Natl Acad Sci U S A. 2009;106(14):5871–5876.
5. Skehel JJ, Wiley DC. Coiled coils in both intracellular vesicle and viral membrane fusion. Cell. 1998;95(7):871–874.
6. Smrt ST, Draney AW, Lorieau JL. The influenza hemagglutinin fusion domain is an amphipathic helical hairpin that functions by inducing membrane curvature. J Biol Chem. 2015;290(1):228–238.
7. Li Y, Tamm LK. Structure and plasticity of the human immunodeficiency virus gp41 fusion domain in lipid micelles and bilayers. Biophys J. 2007;93(3):876–885.
8. Lai AL, Moorthy AE, Li Y, Tamm LK. 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(1–2):3–15.
9. Gregory SM, Harada E, Liang B, Delos SE, White JM, Tamm LK. Structure and function of the complete internal fusion loop from ebolavirus glycoprotein 2. Proc Natl Acad Sci U S A. 2011;108(27):11211–11216.
10. Birtles D, Lee J. Identifying distinct structural features of the sars-cov-2 spike protein fusion domain essential for membrane interaction. Biochemistry. 2021;60:2978–2986. https://doi.org/10.1021/acs.biochem.1c00543.
11. Pennington HN, Lee J. Lassa virus glycoprotein complex review: Insights into its unique fusion machinery. Biosci Rep. 2022;42(2):BSR20211930.
12. Lai AL, Millet JK, Daniel S, Freed JH, Whittaker GR. 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(24):3875–3892.
13. Cai Y, Zhang J, Xiao T, et al. Distinct conformational states of sars-cov-2 spike protein. Science. 2020;369(6511):1586–1592.
14. Koppisetti RK, Fulcher YG, Van Doren SR. Fusion peptide of sars-cov-2 spike rearranges into a wedge inserted in bilayered micelles. J Am Chem Soc. 2021;143(33):13205–13211.
15. Hoffmann M, Kleine-Weber H, Schroeder S, et al. Sars-cov-2 cell entry depends on ace2 and tmprss2 and is blocked by a clinically proven protease inhibitor. Cell. 2020;181(2):271–280.e278.
16. Zhou T, Tsybovsky Y, Gorman J, et al. 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(6):867–879.e865.
17. White JM, Whittaker GR. Fusion of enveloped viruses in endosomes. Traffic. 2016;17(6):593–614.
18. Han X, Bushweller JH, Cafiso DS, Tamm LK. Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. Nat Struct Biol. 2001;8(8):715–720.
19. Klein S, Cortese M, Winter SL, et al. Sars-cov-2 structure and replication characterized by in situ cryo-electron tomography. Nat Commun. 2020;11(1):5885.
20. Ke Z, Oton J, Qu K, et al. Structures and distributions of sars-cov-2 spike proteins on intact virions. Nature. 2020;588(7838):498–502.
21. Madu IG, Roth SL, Belouzard S, Whittaker GR. 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(15):7411–7421.
22. Madu IG, Belouzard S, Whittaker GR. Sars-coronavirus spike s2 domain flanked by cysteine residues c822 and c833 is important for activation of membrane fusion. Virology. 2009;393(2):265–271.
23. Santamaria A, Batchu KC, Matsarskaia O, et al. Strikingly different roles of sars-cov-2 fusion peptides uncovered by neutron scattering. J Am Chem Soc. 2022;144(7):2968–2979.
24. Lai AL, Freed JH. Negatively charged residues in the membrane ordering activity of sars-cov-1 and -2 fusion peptides. Biophys J. 2022;121(2):207–227.
25. Yang S-T, Kiessling V, Simmons JA, White JM, Tamm LK. Hiv gp41–mediated membrane fusion occurs at edges of cholesterol-rich lipid domains. Nat Chem Biol. 2015;11(6):424–431.
26. Takeda M, Leser GP, Russell CJ, Lamb RA. Influenza virus hemagglutinin concentrates in lipid raft microdomains for efficient viral fusion. Proc Natl Acad Sci. 2003;100(25):14610–14617.
27. Biswas S, Yin S-R, Blank PS, Zimmerberg J. Cholesterol promotes hemifusion and pore widening in membrane fusion induced by influenza hemagglutinin. J Gen Physiol. 2008;131(5):503–513.
28. Barz B, Wong TC, Kosztin I. Membrane curvature and surface area per lipid affect the conformation and oligomeric state of hiv-1 fusion peptide: A combined ftir and md simulation study. Biochim Biophys Acta. 2008;1778(4):945–953.
29. Roth SL, Whittaker GR. Promotion of vesicular stomatitis virus fusion by the endosome-specific phospholipid bis(monoacylglycero)phosphate (bmp). FEBS Lett. 2011;585(6):865–869.
30. Nathan L, Lai AL, Millet JK, et al. Calcium ions directly interact with the ebola virus fusion peptide to promote structure–function changes that enhance infection. ACS Infect Dis. 2020;6(2):250–260.
31. Dubé M, Rey FA, Kielian M. Rubella virus: First calcium-requiring viral fusion protein. PLoS Pathog. 2014;10(12):e1004530.
32. Lai AL, Freed JH. 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(10):166946.
33. Cheng SF, Chien MP, Lin CH, Chang CC, Liu YT, Chang DK. The fusion peptide domain is the primary membrane-inserted region and enhances membrane interaction of the ectodomain of hiv-1 gp41. Mol Membr Biol. 2010;27(1):31–44.
34. Gregory SM, Larsson P, Nelson EA, Kasson PM, White JM, Tamm LK. Ebolavirus entry requires a compact hydrophobic fist at the tip of the fusion loop. J Virol. 2014;88(12):6636–6649.
35. Basso LGM, Zeraik AE, Felizatti AP, Costa-Filho AJ. Membranotropic and biological activities of the membrane fusion peptides from sars-cov spike glycoprotein: The importance of the complete internal fusion peptide domain. Biochim Biophys Acta Biomembr. 2021;1863(11):183697.
36. Jackson CB, Farzan M, Chen B, Choe H. Mechanisms of sars-cov-2 entry into cells. Nat Rev Mol Cell Biol. 2022;23(1):3–20.

37. Delos SE, Brecher MB, Chen Z, Melder DC, Federspiel MJ, White JM. Cysteines flanking the internal fusion peptide are required for the avian sarcoma/leukosis virus glycoprotein to mediate the lipid mixing stage of fusion with high efficiency. J Virol. 2008;82(6):3131–3134.
38. Melikyan GB, Barnard RJ, Markosyan RM, Young JA, Cohen FS. Low ph is required for avian sarcoma and leukosis virus env-induced hemifusion and fusion pore formation but not for pore growth. J Virol. 2004;78(7):3753–3762.
39. Mingo RM, Simmons JA, Shoemaker CJ, et al. Ebola virus and severe acute respiratory syndrome coronavirus display late cell entry kinetics: Evidence that transport to npc1+ endolysosomes is a rate-defining step. J Virol. 2015;89(5): 2931–2943.
40. Zang R, Case JB, Yutuc E, et al. Cholesterol 25-hydroxylase suppresses sars-cov-2 replication by blocking membrane fusion. Proc Natl Acad Sci U S A. 2020;117(50):32105–32113.
41. Schaefer SL, Jung H, Hummer G. Binding of sars-cov-2 fusion peptide to host endosome and plasma membrane. J Phys Chem B. 2021;125(28):7732–7741.
42. Gorgun D, Lihan M, Kapoor K, Tajkhorshid E. Binding mode of sars-cov-2 fusion peptide to human cellular membrane. Biophys J. 2021;120(14):2914–2926.
43. Lai AL, Park H, White JM, Tamm LK. Fusion peptide of influenza hemagglutinin requires a fixed angle boomerang structure for activity. J Biol Chem. 2006;281(9):5760–5770.
44. Lai AL, Tamm LK. Locking the kink in the influenza hemagglutinin fusion domain structure. J Biol Chem. 2007;282(33): 23946–23956.
45. Chiliveri SC, Louis JM, Ghirlando R, Bax A. Transient lipid-bound states of spike protein heptad repeats provide insights into sars-cov-2 membrane fusion. Sci Adv. 2021;7((41)): eabk2226.
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. 2015;1848(2):721–730.
47. Sainz B, Rausch JM, Gallaher WR, Garry RF, Wimley WC. Identification and characterization of the putative fusion peptide of the severe acute respiratory syndrome-associated coronavirus spike protein. J Virol. 2005;79(11):7195–7206.
48. Pattnaik GP, Bhattacharjya S, Chakraborty H. Enhanced cholesterol-dependent hemifusion by internal fusion peptide 1 of sars coronavirus-2 compared to its n-terminal counterpart. Biochemistry. 2021;60(8):559–562.
49. Guerrero F, Ciragan A, Iwaï H. Tandem sumo fusion vectors for improving soluble protein expression and purification. Protein Expr Purif. 2015;116:42–49.
50. Maciejewski MW, Schuyler AD, Gryk MR, et al. Nmrbox: A resource for biomolecular nmr computation. Biophys J. 2017; 112(8):1529–1534.
51. Lee W, Tonelli M, Markley JL. Nmrfam-sparky: Enhanced software for biomolecular nmr spectroscopy. Bioinformatics. 2015; 31(8):1325–1327.
52. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. Nmrpipe: A multidimensional spectral processing system based on unix pipes. J Biomol NMR. 1995;6(3):277–293.
53. Struck DK, Hoekstra D, Pagano RE. Use of resonance energy transfer to monitor membrane fusion. Biochemistry. 1981; 20(14):4093–4099.
54. Miles AJ, Wallace BA. Cdtoolx, a downloadable software package for processing and analyses of circular dichroism spectroscopic data. Protein Sci. 2018;27(9):1717–1722.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Birtles D, Oh AE, Lee J. Exploring the pH dependence of the SARS-CoV-2 complete fusion domain and the role of its unique structural features. Protein Science. 2022;31(9): e4390. https://doi.org/10.1002/pro.4390