The SARS-CoV-2 Fusion Domain Provides Clues Towards the Molecular Mechanism for Membrane Fusion

Daniel Birtles, Jinwoo Lee*

Department of Chemistry and Biochemistry, University of Maryland, College Park 20742, Maryland, USA

*Email: jinwoo@umd.edu

The viral lifecycle is a highly optimized series of intricate processes that allows a virus to take advantage of a host cell's own molecular machinery to produce additional viral particles. Every step of the lifecycle is integral to the virus and thus, in order to understand how to better protect ourselves against viral infections, we must first understand how these viruses can effectively survive on a molecular level.

Entry into a target cell is a critical component of the viral lifecycle, which in the coronavirus family is facilitated by the spike glycoprotein (Figure 1). This process can be split into receptor binding and membrane fusion, which are carried out by the two functional subunits of the spike glycoprotein, S1 and S2 respectively. The receptor binding domain (RBD) found in S1 interacts with the target cell receptor, angiotensin converting enzyme 2 (ACE2). Once bound, a cleavage event then occurs at the S2' site which releases the fusion domain (FD) at the N-terminus of the cleaved S2 subunit (S2'). The FD is then free to interact with the target cell membrane, initiating a cascade of structural

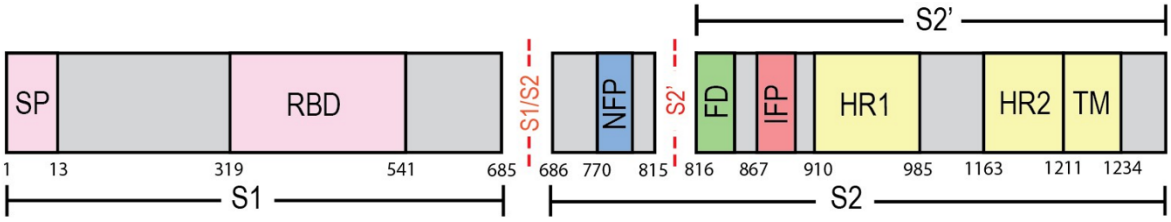

Figure 1: SARS-CoV-2 Spike protein domains. Three potential FDs were initially identified: N-terminal fusion peptide (NFP 770-788; Blue), the fusion domain (FD 816-855; Green) and the internal fusion peptide (IFP 867-909; Red). Other domains highlighted are the signaling peptide (SP) and receptor binding domain (RBD) in S1. The heptad repeat 1 (HR1) and 2 (HR2) as well as the transmembrane domain (TM) in S2'.

rearrangements in S2' that ultimately results in membrane fusion and delivery of the viral genome into the host cell.

Following the initial outbreak of SARS-CoV-1 in 2002, a significant increase in research surrounding the coronavirus family took place to identify the FD. Several membrane interacting regions were quickly identified, characterized and proposed, with most evidence pointing towards the N-terminus of S2' as the most likely candidate. Over the next 10 years, Lai and coworkers discovered that this FD consisted of two functional units that could interact with and perturb lipid membranes both when independently synthesized, and in synergy to even greater effect.[1] Whilst this research was carried out in SARS-CoV-1, strong sequence conservation suggested that this was applicable to all members of the coronavirus family. This led to the coronavirus FD being labelled as a bipartite system that forms an extended fusion platform, due to it containing two regions that can elicit fusion.

Further investigation into the SARS-CoV-2 FD led to the discovery of two structurally independent regions. At the N-terminus of S2', exists the FP which comprises the first 22 amino acids of the FD in a helix-turn-helix motif. This structural motif allows the

FP to bury within the membrane, with the turn being the deepest point of insertion. Immediately following the FP is the internal FL, formed by a conserved disulfide bond. The FL contains no discernible secondary structure and only superficially interacts with the membrane. It is our belief that the nomenclature FP and FL, most clearly describes the two regions, as it reinforces the fact that they are two distinct structures within a single domain that retain key similarities to previously described and well established fusogenic regions found in other viruses. Around the same time, the structure of the FD in a lipid bilayer was published and verified these structural findings.[2] Further work was then undertaken to understand the molecular details of how the FD initiates fusion, through the individual perturbation of the FP and FL. A significant decrease in fusogenic ability was witnessed for the FD when a key hydrophobic motif within the FP, 'LLF', was disrupted via mutagenesis and also when the disulfide bond within the FL was severed.[3] Those results indicated a complex, synergistic mechanism that involves both the FP and FL to such a capacity where neither region can efficiently initiate fusion without the other fully intact.

Atomic resolution structures of the full spike protein acquired via Cryo-EM, alongside solution NMR studies of individual fusogenic regions have provided key insights into the molecular mechanism for SARS-CoV-2 membrane fusion (Figure 2). Our current understanding suggests that following S2' cleavage, the FD interacts with the target cell membrane to perturb the local lipid environment. Heptad repeat 1 (HR1) also perturbs the lipids in the target membrane, whilst HR2 interacts with the viral membrane in what can be considered as a metastable intermediate state that serves as the initial interaction between the spike protein and target cells lipid membrane.[4] Thermodynamically favorable structural changes in S2' then drive HR1 and HR2 together into a six-helix bundle, pulling the opposing membranes into close proximity. At some point during this process, the FP portion of the FD leaves the membrane, making way for the internal fusion peptide (IFP; D^{867} - I^{909}), which following the completion of the fusion event, is found to interact with the transmembrane domain and the FL within the fused membrane.[5] It should be noted that whilst no atomic structure is available for HR2 bound to a membrane, there is structural evidence supporting its existence in a slow dynamic equilibrium between lipid bound and the six-helix bundle states.[4] Furthermore, the post-fusion structure remains uncleaved at

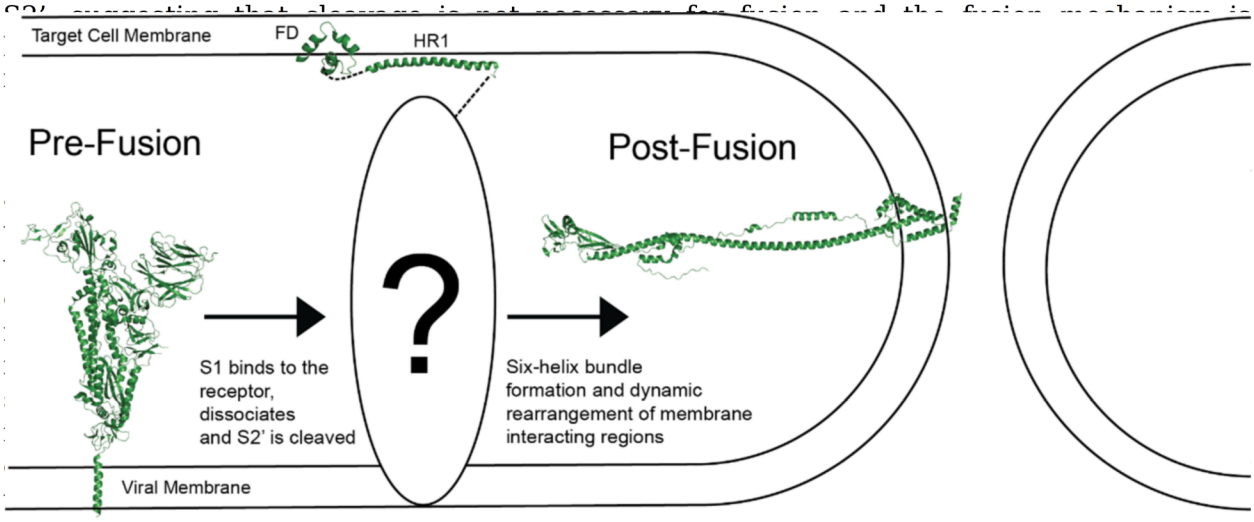

Figure 2: Molecular Mechanism of SARS-CoV-2 spike protein mediated membrane fusion with atomic resolution structures. The Pre-Fusion (PDB:6XR8) SARS-CoV-2 spike protein binds to its target cell receptor, allowing S1 to dissociate and S2' cleavage to occur, releasing the fusion domain (FD). Once free the FD embeds within the target cell membrane to initiate the fusion process (PDB:7MY8) and whilst the conformation of the remainder of the spike protein during this interaction is unknown, HR1 (PDB:7R95) and HR2 are thought to interact with the target and viral cell membranes respectively. As HR1/HR2 come together to form the six-helix bundle, the opposing membranes are also brought into close proximity before they coalesce. During this process, the FP region (S^{816}- I^{834}) leaves or is pushed out of the membrane to make way for the internal fusion peptide (IFP) which inserts itself and forms interactions with both the FL and transmembrane domain, producing the Post-Fusion conformation (PDB:8FDW). The spike protein is represented as a monomer throughout this schematic.

References

(1) 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* (24), 3875-3892. DOI: 10.1016/j.jmb.2017.10.017.
(2) Koppisetti, R. K.; Fulcher, Y. G.; Van Doren, S. R. Fusion Peptide of SARS-CoV-2 Spike Rearranges into a Wedge Inserted in Bilayered Micelles. *Journal of the American Chemical Society* **2021**, *143* (33), 13205-13211. DOI: 10.1021/jacs.1c05435
(3) Birtles, D.; Oh, A. E.; 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). DOI: 10.1002/pro.4390
(4) Chiliveri, S. C.; Louis, J. M.; 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. DOI: 10.1126/sciadv.abk2226.
(5) Shi, W.; Cai, Y.; Zhu, H.; Peng, H.; Voyer, J.; Rits-Volloch, S.; Cao, H.; Mayer, M. L.; Song, K.; Xu, C.; et al. Cryo-EM structure of SARS-CoV-2 postfusion spike in membrane. *Nature* **2023**, *619* (7969), 403-409. DOI: 10.1038/s41586-023-06273-4