ABSTRACT

Title of Dissertation:	DECIPHERING THE MOLECULAR MECHANISM BEHIND THE SARS-COV-2 FUSION DOMAIN
	Daniel Birtles, Doctor of Philosophy, 2024
Dissertation directed by:	Professor Jinwoo Lee, Department of Chemistry and Biochemistry

SARS-CoV-2 is an extremely infectious virus, yet despite a plethora of research the viral lifecycle is still not well understood, particularly the process of membrane fusion. The traditional means by which viral glycoproteins facilitate fusion is that of the six-helix bundle, within which a short, conserved sequence known as the fusion domain (FD) initiates the process as it embeds within and perturbs the target cell membrane, in turn lowering the energetic barrier necessary to coalesce two opposing membranes. Furthermore, the highly conserved coronavirus FD is found to be available on the SARS-CoV-2 spike protein surface, which along with its integral role within the viral lifecycle makes it an ideal therapeutic target. However, limited knowledge of the exact molecular mechanism by which the SARS-COV-2 FD conducts its role within the fusion process has prevented the production of antiviral treatments.

Here we describe the elucidation of key molecular details regarding how the SARS-CoV-2 FD initiates the process of membrane fusion. Firstly, the FD was found to contain a unique assembly of fusogenic regions, known as a fusion peptide (FP) and fusion loop (FL), which operate

in synergy to elicit efficient fusion. This was followed by the discovery of a preference for the FD to fuse within conditions akin to the late endosomal membrane, with both pH and lipid composition significantly impacting fusion. It was found that the endosomal resident anionic lipid BMP imparts a negative impact on lipid packing within the membrane, which positively correlates with fusion. The unique mechanism by which the coronavirus FD initiates fusion was cemented when we uncovered the importance of several positively charged residues towards the FDs function. This also led to unearthing a mutant of the FD (K825A), which if found to have naturally occurred within the full spike protein, has the potential to produce a more virulent strain of SARS-CoV-2.

DECIPHERING THE MOLECULAR MECHANISM BEHIND THE SARS-COV-2 FUSION DOMAIN

by

Daniel Birtles

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Dr. Jinwoo Lee
Dr. Lai-Xi Wang
Dr. John Orban
Dr. Wade Winkler
Dr. Margaret Scull

© Copyright by
Daniel Birtles
2024

Acknowledgements

Firstly, I would like to thank my PI, Jinwoo, for being a great mentor over the past five years, having provided me with the support, skillset and freedom that has allowed me to develop as an independent researcher. I would also like to thank all of the graduate students who have joined me on this journey from within the Lee lab: Shane, Patrick, Hallie, Nick, Kiruthika and Moses, for the time spent listening to practice talks, reading/editing my papers and your friendship. A special thanks goes to Anna, Wafa and Lijon, I have accomplished more than I thought possible thanks to their hard work and dedication. Outside of the lab, coffee mornings with Levi and Tanner, playing soccer with Reid, as well as the feedback from my committee members I view as equally important towards my success. Finally, thank you to my parents, siblings, and wife (and Liz) who without your constant love and support I could never have achieved all that I have.

Table of Contents

Acknowledgements ii
Table of Contents iii
List of Figures v
List of Abbreviations vii
Chapter 1: Introduction 1
1.1 History of the Coronavirus Family 1
1.2 Human Coronavirus Outbreaks: SARS-CoV, MERS-CoV and SARS-CoV-2. 2
1.3 SARS-CoV-2 Lifecycle 5
1.4 Molecular Mechanism of SARS-CoV-2 Membrane Fusion 7
1.5 Role of the Viral Fusion Domain 9
1.6 Outlook 11
Chapter 2: SARS-CoV-2 Contains a Structurally Unique Fusion Domain 13
2.1 Introduction 13
2.2 Results 16
2.2.1 Pre- and Post-Fusion States of the FD 16
2.2.2 The FP Alone Embeds Within the Membrane 21
2.3 Discussion 24
2.4 Conclusion 29
2.5 Materials and Methods 30
2.5.1 Expression 30
2.5.2 Purification 30
2.5.3 NMR Experiments 31
2.5.4 Paramagnetic Relaxation Enhancement 32
2.5.5 Circular Dichroism 33
2.5.6 Preparation of Small Unilamellar Vesicles 33
Chapter 3: The Fusion Peptide and Fusion Loop act in Synergy to Initiate Fusion 34
3.1 Introduction 34
3.2 Results 36
3.2.1 Minor pH Dependent Conformational Change 36
3.2.2 Environmental pH is Critical for FD Mediated Fusion. 40
3.2.3 The FP and FL are Both Integral for Fusogenicity. 43
3.3 Discussion 46
3.4 Conclusion 53
3.5 Materials and Methods 54
3.5.1 Expression and Purification 54
3.5.2 NMR Experiments 54
3.5.3 Preparation of Large Unilamellar Vesicles 55
3.5.4 Lipid Mixing Assay 56
3.5.6 Circular Dichroism 57
Chapter 4: A Preference for the Late Endosomal Pathway 58
4.1 Introduction 58
4.2 Results 60
4.2.1 The FD Preferentially Fuses with Membranes Containing BMP 60

4.2.2 Probing the Interactions of Different Anionic Lipids with the FD 63
4.2.3 BMP Negatively Impacts Lipid Packing 65
4.2.4 Limited Synergy Amongst Other Lipids Relevant to Viral Fusion 68
4.3 Discussion 70
4.4 Conclusion 75
4.5 Materials and Methods 75
4.5.1 Materials 75
4.5.2 Preparation of Small and Large Unilamellar Vesicles 76
4.5.3 Expression and Purification 77
4.5.4 Circular Dichroism 77
4.5.5 Lipid Mixing Assay 77
4.5.6 Zeta Potential Measurements 78
4.5.7 Isothermal Titration Calorimetry 78
4.5.8 C-Laurdan Experiments 79
Chapter 5: Basic Amino Acids are Key for the Molecular Mechanism of Fusion 80
5.1 Introduction 80
5.2 Results 83
5.2.1 Basic Amino Acids are Important for the Fusogenecity of the FD 83
5.2.2 Structurally Investigating the Role of the Basic Residues 86
5.3 Discussion 91
5.4 Conclusions 95
5.5 Materials and Methods 96
5.5.1 Materials 96
5.5.2 Expression of 19F-Phe Fluorinated Fusion Domain 96
5.5.3 Purification of the Fusion Domain 97
5.5.4 Preparation of Small and Large Unilamellar Vesicles 97
5.5.5 NMR experiments 98
5.5.6 Circular Dichroism 99
5.5.7 Lipid Mixing Assay 99
5.5.8 Isothermal Titration Calorimetry 99
Chapter 6: Concluding Remarks 102
Appendices 104
Bibliography 111

List of Figures

Figure 1.1 Classification of the Coronavirus family 2
Figure 1.2 Comparison of Coronavirus Infectivity and Mortality 4
Figure 1.3 Coronavirus Lifecycle 6
Figure 1.4 Coronavirus Virion and Spike Protein 7
Figure 1.5 Six-Helix Bundle Mechanism 8
Figure 1.6 SARS-CoV-2 Membrane Fusion 9
Figure 2.1 Functional Domains of the Spike Protein 14
Figure 2.2 Pre- and Post-Fusion States of the FD 16
Figure 2.3 Pre-Fusion Backbone Assignment 17
Figure 2.4 FD Purification 18
Figure 2.5 Confirmation of Disulfide Bond via Cysteine δCβ 19
Figure 2.6 Severing the Disulfide Bond Perturbs the FD Structure 20
Figure 2.7 Dynamics of Pre- and Post-Fusion States 21
Figure 2.8 Membrane Depth of the Post-Fusion State 22
Figure 2.9 Amphipathic Nature of Helix One and Helix Two 26
Figure 2.10 Structural Model of the pH7.4 FD Embedded in the Membrane 27
Figure 3.1 Minor pH Dependent Conformational Change 36
Figure 3.2 No pH Dependent Change detected via CD 37
Figure 3.3 Dynamics of Post-Fusion State at pH7.4 and pH5.0 38
Figure 3.4 Membrane Depth of Post-Fusion state at pH7.4 and pH5.0 39
Figure 3.5 FD Displays pH Dependent Fusogenecity 40
Figure 3.6 Positive Trend Between P/L Ratio and Fusion at pH7.4 41
Figure 3.7 The ‘LLF’ Motif and Disulfide Bond are Key for FD Fusogenecity at pH5.0 43
Figure 3.8 Ca^{2+} does not Impact the Structure or Function of the FD at pH7.4 46
Figure 3.9 Structural Model of the FD in a Membrane at pH7.4 and pH5.0 48
Figure 3.10 The ‘LLF’ Motif and Disulfide Bond are Key for FD Fusogenecity at pH7.4 50
Figure 4.1 FD Prefers to Fuse in Conditions Akin to the Endosomal Pathway 57
Figure 4.2 Preference for Endosomal Pathway is Regardless of pH and Negative Charge 59
Figure 4.3 A Positive Specificity for BMP over other Anionic Lipids 60

Figure 4.4 Negative Impact of POPE and Cholesterol on Fusion....60
Figure 4.5 General Interactions Present Between FD and Anionic Lipids....61
Figure 4.6 FD does not Interact with Non-Anionic membranes....62
Figure 4.7 BMP Decreases Lipid Packing and Positively Impact on Fusion....64
Figure 4.8 Chemical Structures of POPC, DOPC, DPPC, POPG, POPS and BMP Lipids....65
Figure 4.9 Lipid Packing is not Impacted by BMP in an Already Fluid Membrane....66
Figure 4.10 Synergistic Impact of POPE and Cholesterol....67
Figure 4.11 BMP Fusion Control....70
Figure 5.1 Basic Residues of the FD....80
Figure 5.2 Positive Amino Acids Within the FD are Important for Fusion....82
Figure 5.3 Impact of Basic Mutants on Kd with Anionic Liposomes....83
Figure 5.4 ^{19}F Wt Assignment....84
Figure 5.5 ^{19}F-Phe + Tyr labelling....85
Figure 5.6 Mutants ^{19}F Chemical Shift Perturbations....86
Figure 5.7 ^{19}F dynamics....87
Figure 5.8 Elongation of Helix One in K825A Increases Fusogenecity of the FD....88
Figure 5.9 Amphipathicity of Helix One in Wt and K825A at pH5.0....92

List of Abbreviations

3-^{19}F-Tyr	3-fluoro-tyrosine
4-^{19}F-Phe	4-fluoro-phenylalanine
6HB	Six-Helix Bundle
ACE2	Angiotensin Converting Enzyme 2
α-CoV	Alphacoronavirus
ANTS	8-aminonaphthalene-1,3,6-trisulfonic acid
ASLV	Avian Sarcoma Leukosis Virus
β-CoV	Betacoronavirus
BME	β-mercaptoethanol
BMP	Bis(monoacylglycero)phosphate
CD	Circular Dichroism
CHOL	Cholesterol
CFR	Case Fatality Rate
CMC	Critical Micelle Concentration
COVID-19	Coronavirus Disease 2019
CPMG	Carr-Purcell-Meiboom-Gill
CRYO-EM	Cryo-Electron Microscopy
CSP	Chemical shift perturbations
δ-CoV	Deltacoronavirus
DHPC	1,2-dihexanoyl-sn-glycero-3-phosphocholine
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DPC	Dodecylphosphocholine
DPPC	1,2-dipalmitoyl-sn-glycero-3-phosphocholine
DOPC	1,2-dioleoyl-sn-glycero-3-phosphocholine
DPX	N,N-*p*-xylylenebis-(pyridinium bromide)
DSA	Doxyl steric acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EBOV	Ebola Virus

EPR	Electron Paramagnetic Resonance
ERGIC	Endoplasmic Reticulum-Golgi Intermediary Complex
FD	Fusion domain
FL	Fusion Loop
FP	Fusion Peptide
FRET	Förster Resonance Energy Transfer
γ-CoV	Gammacoronavirus
Gd-[DTPA]	Gadolinium (III) diethylenetriaminepentaacetic acid
GP41	Glycoprotein 41
HA	Hemagglutinin
hCoV	Human Infecting Coronaviruses
HIV	Human Immunodeficiency Virus
HR1	Heptad Repeat One
HR2	Heptad Repeat Two
HSQC	Heteronuclear single quantum coherence
IFP	Internal Fusion Peptide
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LASV	Lassa Virus
LB	Luria-Bertani
LUV	Large Unilamellar Vesicles
MALDI-TOF	Matrix Assisted Laser Desorption/Ionization Time-Of-Flight
MD	Molecular Dynamics
MERS-CoV	Middle Eastern Respiratory Syndrome Coronavirus
MPER	Membrane Proximal External Region
mRNA	Messenger Ribonucleic Acid
MSP	Membrane Scaffold Protein
NBD-PE	1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(7-nitro-2-1,3-benzoxidiazol-4-yl)
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy

NSP	Non-Structural Protein
NUS	Non-uniform sampling
OD	Optical density
ORF	Open Reading Frame
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PE	Phosphatidylethanolamine
POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine
POPE	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
POPG	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-[1'-rac-glycerol]
POPS	1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine
PRE	Paramagnetic Relaxation Enhancement
PS	Phosphatidylserine
Rh-PE	1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl)
RBD	Receptor Binding Domain
RNA	Ribonucleic Acid
SARS-CoV	Severe Acute Respiratory Syndrome Coronavirus
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SEC	Size Exclusion Chromatography
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SM	Sphingomyelin
SUV	Small Unilamellar Vesicles
TCEP	Tris(2-carboxyethyl)phosphine
TMD	Transmembrane Domain
TMPRSS2	Transmembrane Serine Protease 2
WHO	World Health Organization

Chapter 1: Introduction

Some of this work has been published as a *Biochemistry Viewpoint*.(1)

Birtles D, Lee J. SARS-CoV-2 Fusion Domain Provides Clues toward the Molecular Mechanism for Membrane Fusion. Biochemistry. 2023(62):3033–5.

D.B. and J.L. reviewed the literature and prepared the manuscript

1.1 History of the Coronavirus Family

A virus is an infective agent that is required to invade other living cells in order to propagate and survive.(2) In 1965, Tyrrell and Bynoe passaged a virus they named B814 from the respiratory tract of a human adult with a common cold. (3) The morphological analysis of this virus revealed a medium size (80-150 nm), enveloped virion with club-shaped surface projections.(4) Over the next several years, other morphologically similar viruses were identified,(5, 6) eventually leading to the creation of a new viral family named *Coronaviridae* due to the crown-like appearance of the projections on the viral surface.(7) The family was eventually split into two subfamilies: *Letovirinae* and *Orthocoronivinae*, the latter of which are commonly referred to as coronaviruses (Figure 1.1).

Research into the human infecting members of the coronavirus (hCoV) family proceeded with the strains OC43 and 229E, due to their relative ease to culture.(8) Pathogenicity was initially thought to be low despite the families association with several respiratory illnesses such as asthma, and bronchitis. Coronaviruses were even thought to contribute to approximately 15% of the all common adult colds at the time.(9) (10) Meanwhile, knowledge of animal coronaviruses grew at a rapid rate, with these viruses were able to infect multiple species, displayed several different phenotypes and contained varying degrees of pathogenicity. As a result, there was a need to introduce further classification systems, which utilizing antigenicity and genetic makeup provided

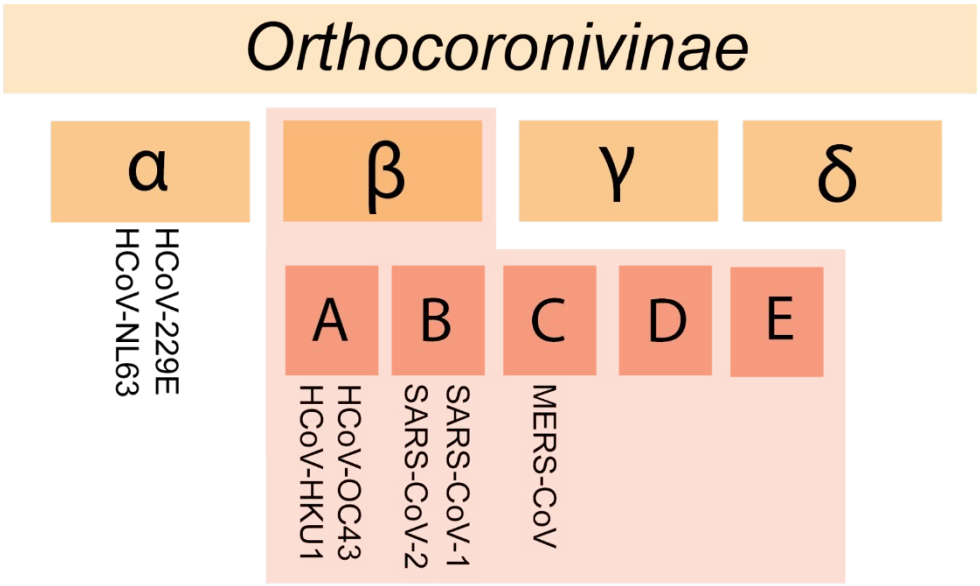

Figure 1.1 Classification of the coronavirus family, highlighting the seven known human infecting coronaviruses (HCoV).

the foundations for the four genera we have today: Alphacoronavirus (α-CoV), Betacoronavirus (β-CoV), Gammacoronavirus (γ-CoV) and Deltacoronavirus (δ-CoV).(11) Generally, α-CoV and β-CoV infect mammals, whilst γ-CoV and δ-CoV primarily infect avian species.

The β-CoV family is the largest and most clinically relevant genera, containing five of the seven known hCoVs (OC43, HKU1, SARS-CoV, SARS-CoV-2 and MERS-CoV) with the other two found within the α-CoV genera (229E and NL63) (Figure 1.1). Based on the phylogenetic analysis of conserved protein domains, β-CoVs were further divided into four main subgenera, with a fifth recently added(12): *Embecovirus* (Lineage A), *Sarbecovirus* (Lineage B), *Merbecovirus* (Lineage C), *Nobecovirus* (Lineage D) and *Hibecovirus* (Lineage E).

1.2 Human Coronavirus Outbreaks: SARS-CoV, MERS-CoV and SARS-CoV-2.

In 2003, the outbreak of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) put the coronavirus family into the global spotlight. A β-CoV in the S*arbecovirus* subgenera, SARS-CoV was reported across North America, South America, Europe and Asia, infecting over 8000

individuals and leading to nearly 800 deaths.(13) Within China, SARS-CoV was deemed an epidemic by the World Health Organization (WHO) in early 2003, making it the first viral outbreak of the 21st century. The virus is thought to have originated in Chinese horseshoe bats as natural reservoirs, with interspecies transmission allowing the virus to infect palm civets as intermediate hosts before the disease could spread to the human population via live animal markets.(14) A highly transmissible, airborne virus similar to influenza, SARS-CoV quickly disseminated with symptoms including a fever, cough, fatigue, headache, diarrhea and muscle soreness.(15) Despite the rapid emergence of SARS-CoV, a collaborative international effort from the scientific community helped control the disease outbreak within nine months after its initial occurrence.

Less than ten years later, another outbreak in humans was discovered in Saudia Arabia, called Middle East Respiratory Syndrome Coronavirus (MERS-CoV).(16) Also believed to have originated from a bat population, MERS-CoV infection was heavily implicated with dromedary camels as intermediate hosts.(17) Again a β-CoV, but this time from the *Merbecovirus* subgenera, MERS-CoV was significantly more lethal than SARS-CoV with a Case Fatality Rate (CFR) of ~35% compared to ~10%.(18) Since April 2012, ~2600 cases have been confirmed with ~940 deaths.(19) As there is no specific treatment or vaccine for MERS-CoV, it continues to be a problem today particularly in the Arabian Peninsula, with the last outbreak occurring in 2019 and several isolated cases noted since.(20)

Due the reoccurrence and severity of hCoV outbreaks, in 2018 the World Health Organization (WHO) declared both SARS and MERS as requiring urgent and prioritized research due to their pandemic potential.(21) This proved to be too little to late when at the end of 2019, SARS-CoV-2 emerged in Wuhan, China. Commonly referred to as the Coronavirus Disease 2019 (COVID-19), SARS-CoV-2 took just four months from the first reported case to be declared as a

global pandemic, and by July 2020 had already infected more than 26 million people, leading to approximately 865,000 deaths.(22) As of May 2024 with the pandemic over, SARS-CoV-2 is reported to have infected more than 775,000,000 people and caused over 7,000,000 deaths.(23) An incredibly infectious virus, as displayed by its increased R_0,(24) not only when compared to other members within the coronavirus family but amongst several other well known viruses such as influenza, and Ebola.(25) With a CFR of ~2%, the decreased lethality of the virus compared to previous coronaviruses has been proposed to have allowed it to infect more people on a dramatically increased scale, with the global spread and economic impact of SARS-CoV-2 unprecedented (Figure 1.2). In an astonishingly short period of time, just seven months after the first case of SARS-CoV-2 had been identified, two mRNA-based vaccines had been produced which significantly decreased sever clinical outcomes thanks to an incredible, global collaborative effort.(26-29) The rapid worldwide distribution of vaccines gradually decreased the number of

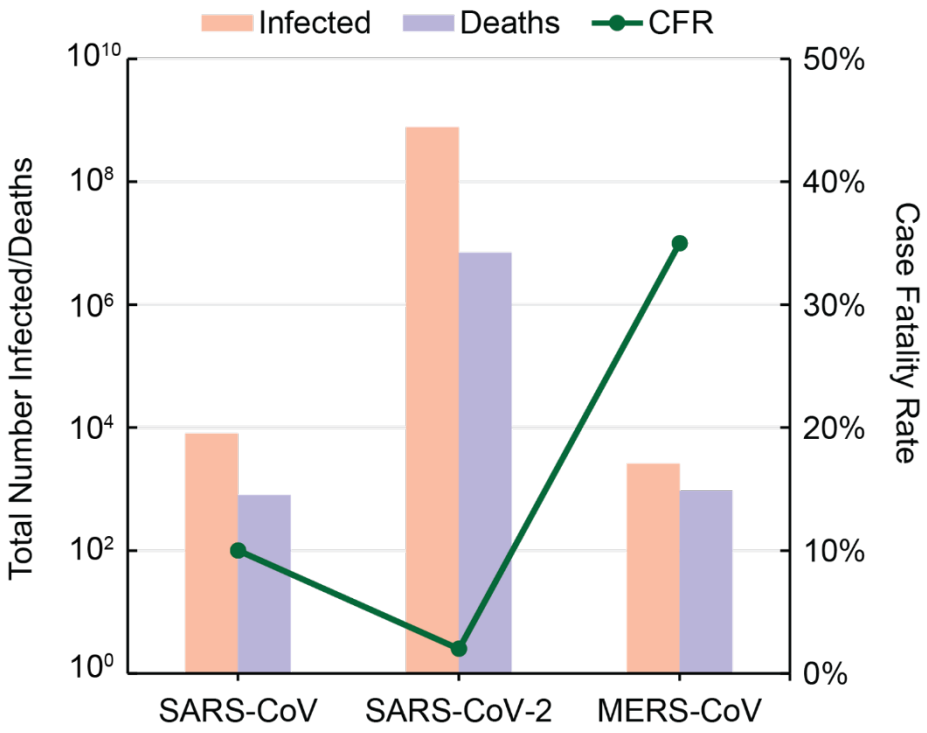

Figure 1.2 Comparing infection and death rates from SARS-CoV, SARS-CoV-2 and MERS-CoV, alongside the case fatality rate (CFR).

reported cases, and alongside the discovery of antiviral treatments significantly lowered the CFR in healthy, vaccinated individuals. On May 5th, 2023, the public health emergency caused by COVID-19 was declared over. Nevertheless, there remains a plethora of information to be uncovered regarding SARS-CoV-2 which can aid in the preparation for future viral pandemics.

1.3 SARS-CoV-2 Lifecycle

The SARS-CoV-2 lifecycle can be simplified into four distinct stages: (1) Infection, (2) Translation, (3) Assembly, (4) Release (Figure 1.3). Infection, also referred to as viral entry, is mediated by the spike (S) glycoprotein, which is one of only three proteins alongside the envelope (E) and matrix (M) proteins displayed on the mature virions surface (Figure 1.4). The S 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.(30, 31) The receptor binding domain (RBD) of S1 recognizes and interacts with the angiotensin-converting enzyme 2 (ACE2) receptor found ubiquitously throughout the body.(32) Following receptor binding there is a cleavage event at the S2' site facilitated by transmembrane serine protease 2 (TMPRSS2) at the plasma membrane or cathepsin L at the endosomal membrane. After cleavage has occurred the S2 domain conducts membrane fusion through significant structural changes that bring the opposing membranes within proximity to coalesce.

Once the viral RNA is released, it takes advantage of the host cells molecular machinery to translate the genomic information into functional proteins. Coronaviruses possess enormous RNA genomes, that include *cis*-acting secondary RNA structures essential for RNA synthesis.(33) As such the translation process for SARS-CoV-2 is extremely complex, consisting of multiple open reading frames (ORFs) with genomic RNA encoding for non-structural proteins (NSPs) and

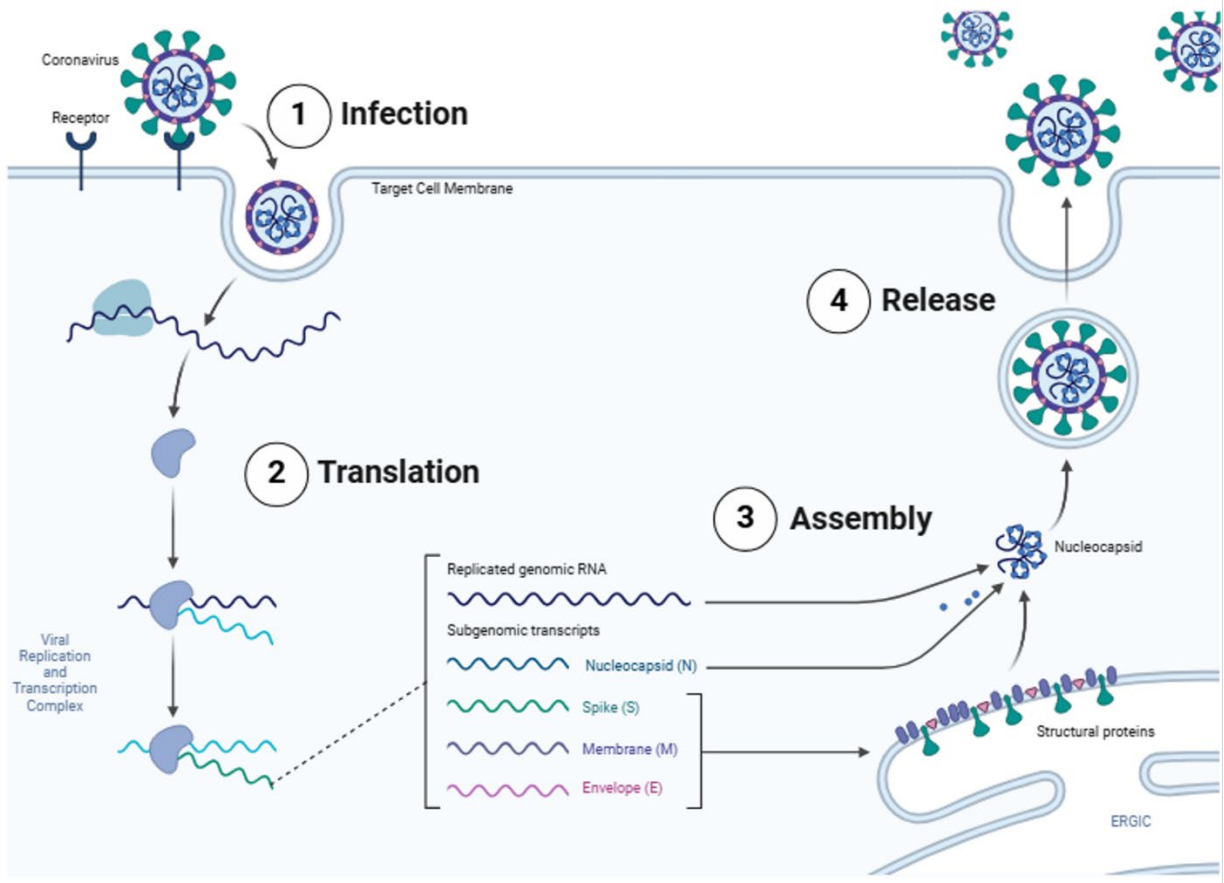

Figure 1.3 SARS-CoV-2 lifecycle. The virus infects the host by first binding to the receptor, before then fusing with the target cell membrane. Once the viral RNA is released non-structural proteins (NSPs) are translated facilitating the creation of the viral replication and transcription complex. In turn, this facilitates the production of the genomic RNA and the structural proteins: nucleocapsid (N), Spike (S), Membrane (M) and Envelope (E), in the endoplasmic reticulum-golgi intermediary complex (ERGIC). Within the ERGIC is where the proteins assemble, and pre-processing cleavage events occur to produce the mature virion which is subsequently exocytosed and released from the cell. Created with Biorender.com.

subgenomic RNA encoding for structural proteins (S, E and M proteins) and accessory proteins, all of which contain diverse and essential roles. NSPs are responsible for forming the viral replication and transcription complex, essential for producing genomic RNA for continuous viral propagation.(34) Accessory proteins appear to play key roles in both immune evasion and viral pathogenesis,(35) whilst the structural proteins are responsible for forming new virions that can go on to infect other cells. Following translation of the S, M, and E structural proteins in the endoplasmic reticulum-golgi intermediary complex (ERGIC), the S protein is cleaved at a S1/S2

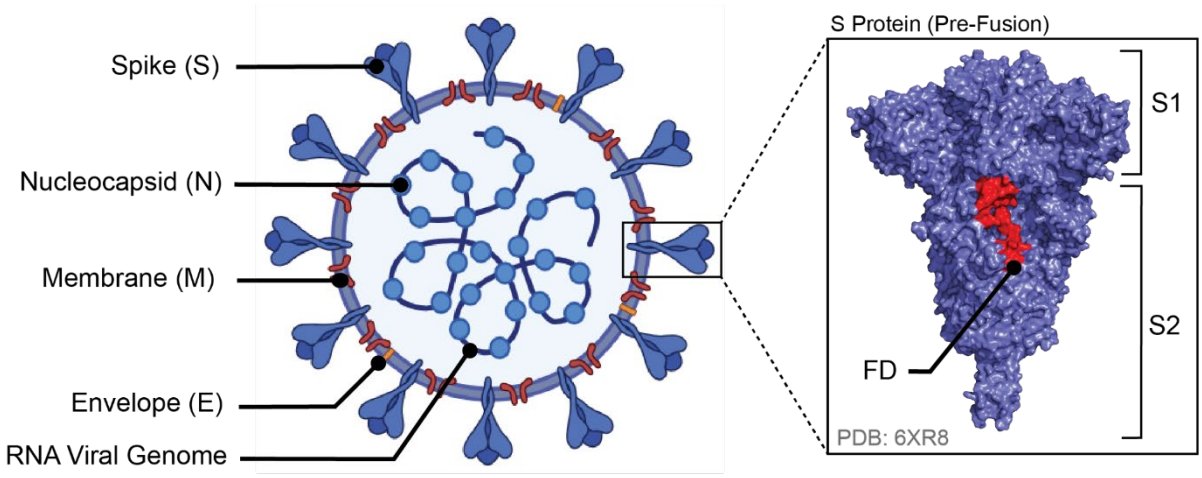

Figure 1.4 Overview of the mature SARS-CoV-2 virion displaying four structural proteins (S, N, M and E) along with the RNA viral genome. One of the first published structures of the S protein is displayed (PDB:6XR8) with the fusion domain (FD) highlighted in red. Made with Biorender.com

site by furin protease, covalently releasing the two subunits from one another.(36) Once the structural proteins come into contact with newly produced genomic RNA, the virion forms and eventually buds from the ERGIC into the lumen of secretory vesicles.(37) Finally, mature virions are secreted from infected cells following exocytosis.

In order to prevent viral infection, the virus in question must first be understood in great detail. By understanding the SARS-CoV-2 viral lifecycle we can not only determine what makes the virus so successful, but where it may be inhibited. This body of work is specifically focused on understanding the process of membrane fusion on a molecular level.

1.4 Molecular Mechanism of SARS-CoV-2 Membrane Fusion

The S glycoprotein is considered to be a class I fusion protein due to its proteolytic processing, oligomeric state, membrane orientation and fusion domain (FD) location.(31, 38) Other notable members of this class include glycoprotein 41 (GP41) of HIV and hemagglutinin (HA) of influenza.(39, 40) The classical model of fusion associated with all of these viruses is the six-helix

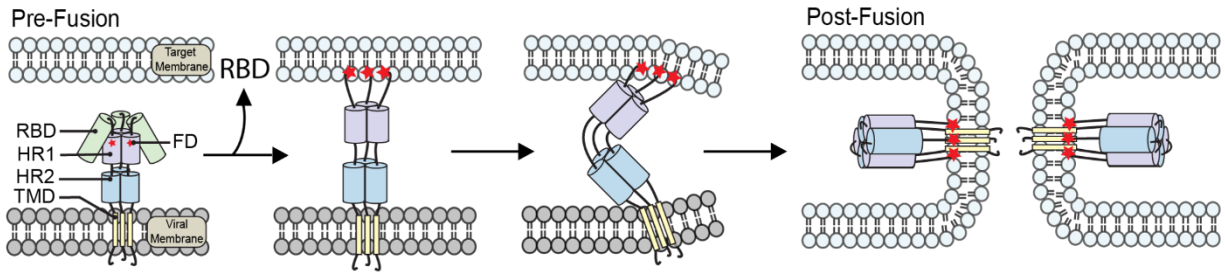

Figure 1.5 Six-Helix Bundle (6HB) mechanism for Class I viral fusion. In the Pre-Fusion state, the receptor binding domain (RBD) sits atop the fusogenic subunit until it is triggered through an event such as receptor binding, cleavage or decreasing pH. Then the RBD dissociates, releasing the fusion domain (FD) which is propelled into the target cell membrane following dynamic structural rearrangement. The protein is now anchored in both the viral membrane and target membrane, via the transmembrane domain (TMD) and FD respectively. Favorable interactions between the helices in heptad repeat 1 (HR1) and HR2 pull the membranes into proximity until they eventually coalesce, forming the fusion pore and allowing the viral genomic information to enter the cell.

bundle (6HB) mechanism, (Figure 1.5) which begins following the dissociation of the receptor binding subunit and in some cases, such as with SARS-CoV-2, a cleavage event.

Our current understanding of SARS-CoV-2 fusion suggests that following S2' cleavage, the FD is released and embeds itself within the target cell membrane to perturb the local lipid environment.(1) This action results in the S protein serving as a molecular bridge connecting the opposing membranes, as it is embedded in both the target membrane via the FD, and the viral membrane through its own transmembrane domain (TMD). One of many novel aspects of SARS-CoV-2 fusion is the association of heptad repeat 1 (HR1) and heptad repeat 2 (HR2) with the membranes, contributing towards this molecular bridge. HR1 binds the target membrane, whilst HR2 interacts with the viral membrane in a state what can be considered as a metastable intermediate state, serving as the initial interaction complex between the S protein and target cells lipid membrane.(41) Structural changes in S2' then drive HR1 and HR2 together into the thermodynamically favorable six-helix bundle, which is composed of a trimer of hetero-dimers, pulling the opposing membranes into close proximity.(38) Once the membranes coalesce, the FD and TMD come into close proximity and may interact with one another to aid in the transition from

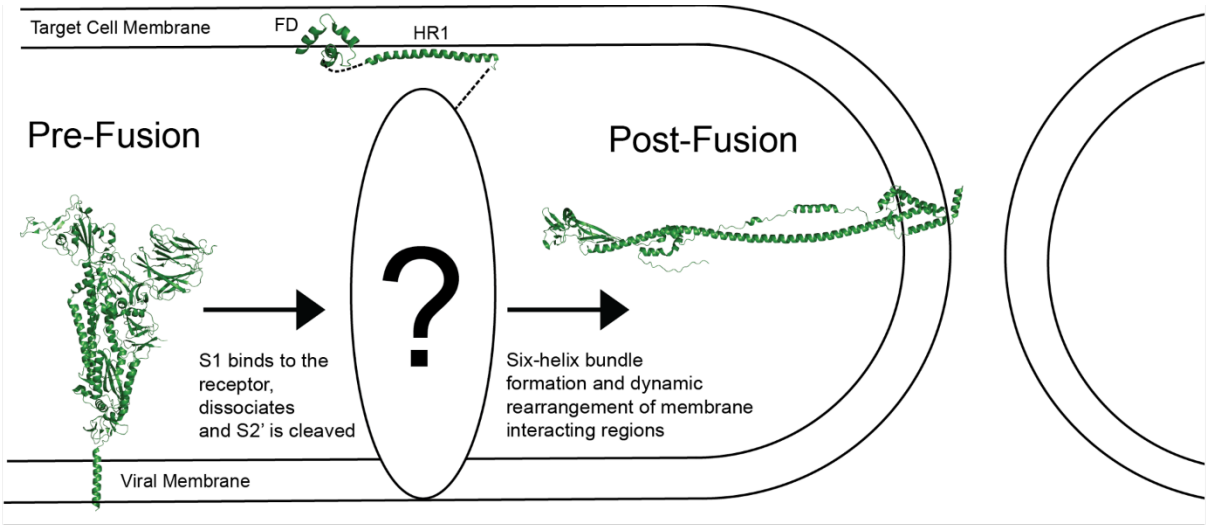

Figure 1.6 Current understanding of SARS-CoV-2 membrane fusion via atomic resolution structures. The Pre-Fusion (PDB:6XR8) SARS-CoV-2 spike protein binds to its target cell receptor, allowing S1 to dissociate and S2' cleavage to occur, releasing the fusion domain (FD). Once free, the FD embeds within 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 proximity before they eventually 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.

hemi-fusion to fusion pore.(42) The formation of the fusion pore then creates a passageway for the viral RNA to enter the host cell. Despite extensive research, several gaps within our knowledge of the SARS-CoV-2 fusion process still exist, particularly pertaining to the membrane interacting regions of the S protein such as the FD and the initiation of membrane fusion. (Figure 1.6)

1.5 Role of the Viral Fusion Domain

Every studied viral fusion protein contains an FD, whether that be in the form of a fusion peptide (FP), fusion loop (FL) or a combination of the two. Within Class I viruses HIV and influenza contain FPs,(43, 44) Avian Sarcoma Leukosis Virus (ASLV) and Ebola virus (EBOV) contain FLs,(45, 46) whilst SARS-CoV-2 and Lassa Virus (LASV) contain a combination of the two.(47)

The primary role of the viral FD is to anchor the viral fusion protein to the target cells membrane and subsequently perturb the local lipid environment to lower the energetic cost necessary to merge two opposing membranes.(48) To perform this function, FDs generally display rich conformational flexibility and certain degree of structural polymorphism that can be dictated by specific lipid, pH and protein environments. This was witnessed with the HIV FP, which displays different secondary structure character dependent on the amount of cholesterol present within the membrane.(49) Despite only moderate hydrophobicity amongst the majority of FDs, triggered structural changes result in the formation of hydrophobic patches that can facilitate membrane insertion.(46) Accordingly, small motifs consisting of bulky, hydrophobic amino acids are often extremely well conserved throughout viral families due to this crucial role.

Following the initial outbreak of SARS-CoV in 2002, a significant increase in research surrounding the coronavirus family took place to identify the FD. Several membrane interacting regions within the S protein were quickly identified, characterized, and proposed, with most evidence pointing towards the N-terminus of S2' as the most likely candidate. (50-53) 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.(54) Whilst this research was carried out in SARS-CoV, strong sequence conservation suggested that this was applicable to all members of the coronavirus family. This led to the coronavirus FD being labeled as a bipartite system that forms an extended fusion platform, due to it containing two regions that can elicit fusion.

Although extensive evidence supports the N-terminal region (S^{816}-F^{855} in SARS-CoV-2) region as the putative FD, debate is still present within the field. A Cryo-Electron Microscopy (CRYO-EM) structure of the S protein(42) suggested a different model with regards to the role of

the FD. Based on that structural data, at some point during the fusion process the FP portion of the FD (816-838) is thought to leave the membrane, making way for the internal fusion peptide (IFP; 867 - 909). Following the completion of the fusion event, the IFP is also found to interact with the TMD (1211-1234) and the FL (839-855) within the newly fused membrane. However, within the structure the Post-Fusion structure remains uncleaved at S2', suggesting that cleavage may not be necessary for fusion. This emphasizes the redundancy of the SARS-CoV-2 fusion mechanism, something that has been touted previously as being a key component of the viruses incredible infectivity. Moreover, this implies that with S2' cleaved a different structural state for the S protein may also exist, and therefore, a different role for the FD.

1.6 Outlook

The outbreak of SARS-CoV-2 and resulting COVID-19 pandemic, led to an unprecedented amount of scientific discoveries in a relatively short period of time. Due to this, a wealth of information has been discovered regarding not just SARS-CoV-2, but the coronavirus family as a whole due to strong sequence conservation and comparative studies. Even so, there is still much that we do not know regarding the virus, particularly regarding the molecular details behind its elevated levels of infectivity in comparison to other viruses.

Here, we have sought to elucidate the molecular mechanism behind the initiation of S mediated fusion via the FD of SARS-CoV-2. Not only will this aid in our understanding of the fundamental machinery necessary for SARS-CoV-2 fusion, but it may also serve as the foundation for therapeutic intervention against the coronavirus family, past, present, and future. Within the next chapter, I will describe how we found that the SARS-CoV-2 FD consisted of a novel assembly of a FP and a FL, alongside the distinct structural features of the two regions. In Chapter 3,

evidence is presented that confirms that the FP and FL, through a hydrophobic motif and disulfide bond respectively, display a synergistic relationship where neither can efficiently perform fusion without the other. Chapter 4 is focused on understanding the mechanism from a more lipid-based perspective, where we compare the impact of the plasma and endosomal membranes on FD mediated fusion. Then building on the knowledge gained, Chapter 5 seeks to explain the positive environmental relationships witnessed with FD mediated fusion at atomic resolution. Finally, Chapter 6 summarizes my findings, providing meaning to that which has been found within the grander scheme of SARS-CoV-2 and the implications towards future work.

Chapter 2: SARS-CoV-2 Contains a Structurally Unique Fusion Domain

This work has been published in *Biochemistry*.(47)

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.

D.B. performed research; D.B., and J.L. conceptualized research, analyzed data and prepared the manuscript.

2.1 Introduction

Viral entry of the coronavirus family is mediated by the 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.(55) The interaction between the receptor binding domain (RBD) in S1 and the host cell receptor, angiotensin converting enzyme 2 (ACE2), has been studied extensively.(56, 57) 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 lifecycle. The most well accepted mechanism for viral membrane fusion is that of the six-helix bundle (6HB), a process which is initiated by the fusion domain (FD) after it embeds itself in the host cell membrane.(58) The coronavirus family has been found to fuse at both the plasma and endosomal membranes, highlighting the remarkable adaptability of these viruses.(59) Accordingly the FD is found to be well conserved throughout, particularly in those most pathogenic amongst the human population

(Figure 2.1). 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), Lambda (C.37) and Omicron (B.1.1.529).(60)

Interestingly the FD has two regions which both closely resemble fusogenic structures witnessed in other viruses. Firstly, 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 both HIV(39, 43, 61) and influenza.(44, 62, 63) 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).(64) 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 revealed the structure of the spike protein in both Pre- and Post-Fusion states for SARS-CoV-2.(65) Yet at the time of this research, the Post-Fusion state was missing a large

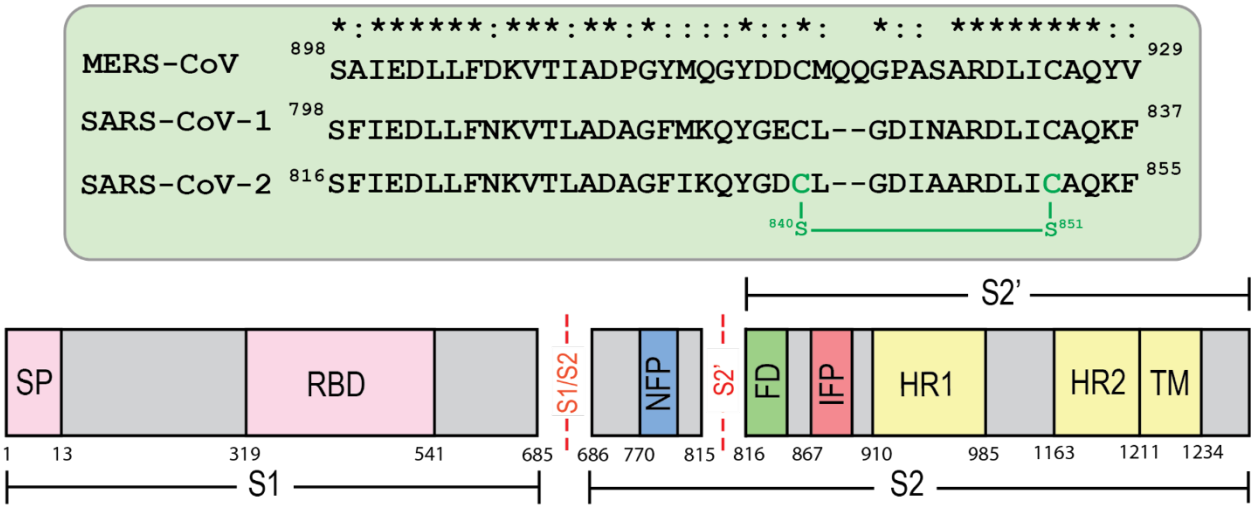

Figure 2.1 SARS-CoV-2 spike protein domains with fusion domain (FD) conservation. Three potential FDs were initially identified: N-terminal fusion peptide (NFP 770-788; Blue), the fusion domain (FD 816-855; Green) and the internal fusion peptide (IFP 867-909; Red). The highly conserved nature of the FD is one trait that suggests it is integral for s protein function and thus, the viral lifecycle. 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'.

disordered region (I^{770}-T^{912}) in proximity to the membrane, which included the FD (S^{816}-F^{855}). Previous work carried out on the FD within SARS-CoV, revealed that both the FP and FL contain independent fusogenic activity, but are most effective in synergy.(50, 51) This led to the development of a bipartite fusion platform model for how the FD interacts with the membrane.(54) Moreover, the SARS-CoV-2 FD has been suggested to contain a greater membrane perturbing ability than the FD of SARS-CoV through the use of electron paramagnetic resonance (EPR) and membrane ordering experiments.(66) The same study also displayed highly specific interactions between the FD and Ca^{2+}, which is thought to be an important co-factor within membrane fusion. 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.(54, 65) Throughout, the terminology FP and FL will be used exclusively as 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.

In the present 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 Post-Fusion 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, whilst 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 other well established viral fusion domains.

2.2 Results

2.2.1 Pre- and Post-Fusion States of the FD

The process of viral membrane fusion begins with an insertion in into the host cell membrane. Due to the hydration layer of water molecules and lipid packing, this is an energetically unfavorable process.(67) The FD is thought to undergo a significant structural change to overcome this barrier, embed within the membrane and disrupt the host cell lipid bilayer, initiating the process of membrane fusion. Here we show the presence of a Pre-Fusion and Post-Fusion state of the SARS-CoV-2 FD (Figure 2.2).

In solution the FD displays an almost completely random coil secondary structure, evident by the large negative dip at 200 nm when assessed through CD (Figure 2.2A). Following the introduction of a membrane mimic, whether that be DPC or small unilamellar vesicles (SUV), a change in global secondary structure is witnessed (Figure 2.2A). This can be characterized by the

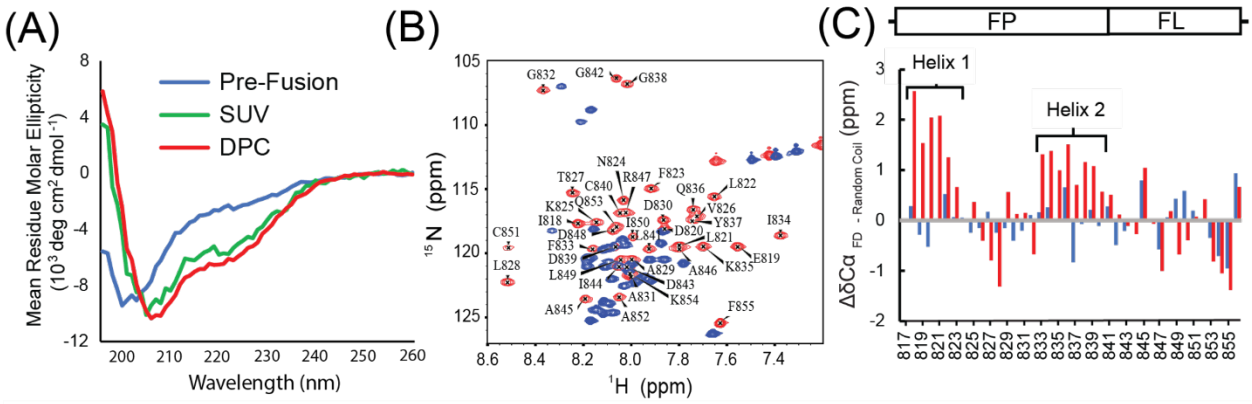

Figure 2.2 The Pre-Fusion (Blue) and Post-Fusion (Red) states of the SARS-CoV-2 FD. (A) CD highlights the secondary structure change between the two states, brought about by the introduction of 100 mM DPC micelles (Red) or 1 mM small unilamellar vesicles (SUVs) (Green). (B) An overlay of the ^{1}H-^{15}N HSQC spectra of the Pre-Fusion and Post-Fusion states with the backbone assignment shown for the latter. (C) Chemical shift index (CSI) using Cα values confirms the presence of two alpha helices in the Post-fusion conformation of the FD.

distinctive double-dip at 208 nm and 222 nm in the spectra, indicative of α-helical content. In the presence of POPC:POPG (50:50) SUVs, the FD displayed 19% α-helical content which agrees well with the 20% witnessed with 100 mM DPC. 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 Post-Fusion states can be witnessed (Figure 2.2B). These large chemical shift perturbations (CSPs) following the addition of the membrane mimic DPC, suggests 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 Pre-Fusion (Figure 2.3) and Post-Fusion states (Figure 2.2B), a more detailed investigation into 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

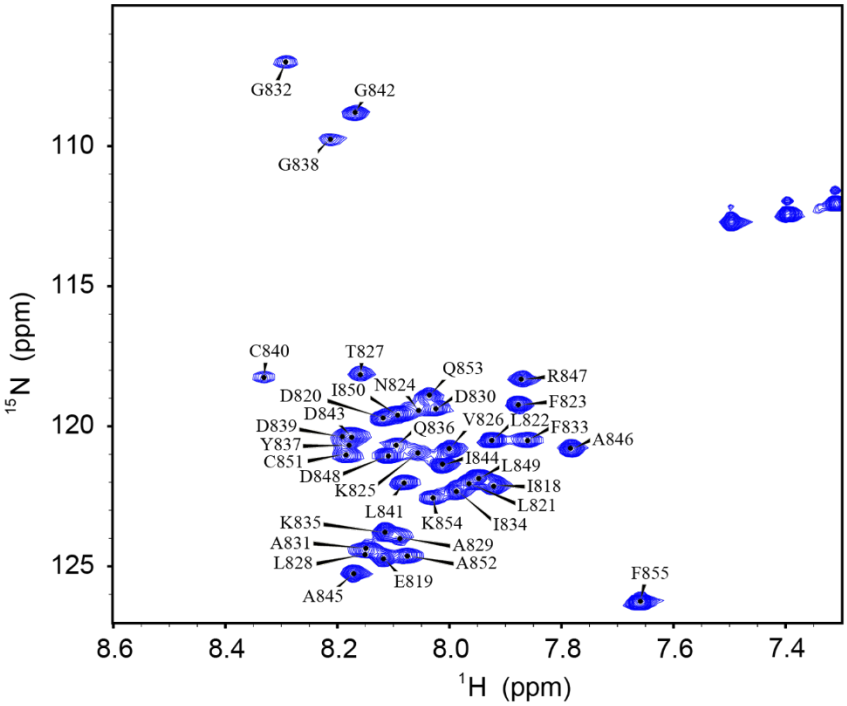

Figure 2.3 Pre-Fusion backbone assignment for the SARS-CoV-2 fusion domain (FD).

structure could be ascertained.(68) 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 Pre-Fusion state as all values are within 1 of a random coil (Figure 2.2C, Blue). However for the Post-Fusion state, it suggests that two α-helices separated by a short break (Figure 2.2C, Red) form within the FP region. Specifically, we see helix one from residues F^{817}-L^{822} and helix two 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, with an increase in α-helical character observed in the Post-Fusion conformation. Interestingly, large secondary structure changes are evident for the FP (S^{816}-G^{838}), but much less so towards the FL (D^{839}-F^{855}) (Figure 2.2C). The internal disulfide bond within this region (C^{840}-C^{851}) most likely plays a significant role in confining the FL to a similar structure within both the Pre- and Post-Fusion states. Presence of the FD monomer was confirmed by SEC and MALDI-TOF (Figure 2.4) and disulfide bond formation via C_β chemical shifts above 35 ppm representing

oxidized cysteines (Figure 2.5).(69) Furthermore, large CSPs are witnessed via ^{1}H-^{15}N HMQC after the disulfide bond is severed through addition of 5 mM TCEP to the NMR sample (Figure 2.6).

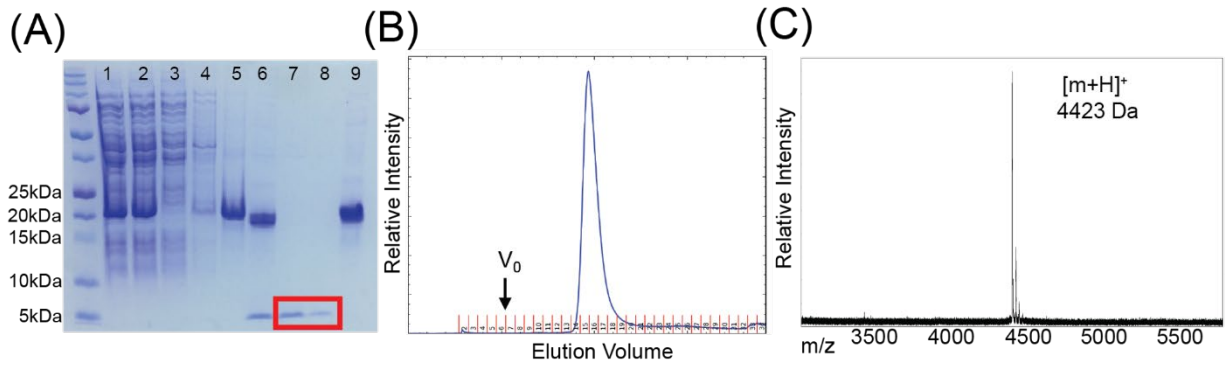

Figure 2.4 Expression and Purification of the SARS-CoV-2 fusion domain (FD). (A) The pET41-FD construct is purified from 1 L cell pellet which is resuspended, sonicated and centrifuged with a sample taken pre- (1) and post-centrifugation (2). The supernatant is then bound the Ni-NTA resin, and the flowthrough collected (3). The resin is washed (4) and then eluted (5) using 250 mM imidazole. Sumo protease is added to the elution and the cleavage reaction takes place overnight whilst in dialysis. The cleaved sample (6) is incubated with the Ni-NTA resin again, with the cleaved FD no longer able to bind the resin. Both the FT(7) and wash (8) both contain the FD (Red Box) whilst the elution (9) contains the cleaved tag. (B) A single peak is witnessed through fractions 15-17, indicating the peptide is in a monomeric state, when ran through a Superdex^{TM} 10/30 HR Peptide SEC column. (C) MALDI-TOF mass spectrometry for the SEC fractions 15-17 reveals a single species that matches the MW of the FD (4424 Daltons).

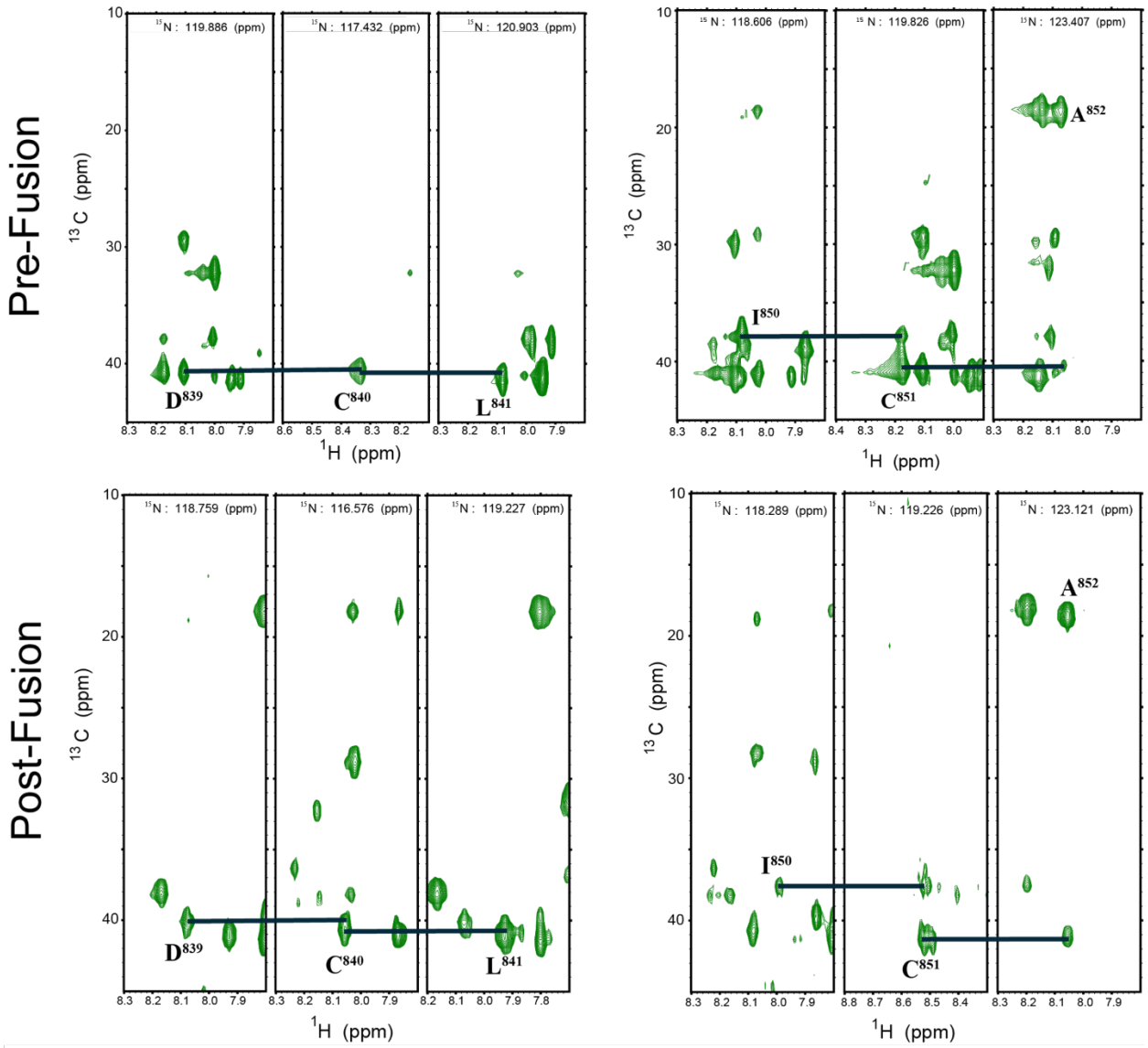

Figure 2.5. Correct formation of the internal disulfide bond within the fusion loop (FL) was confirmed via the C_β chemical shifts of the cysteine residues, all of which are over 35 indicating that they are oxidized.

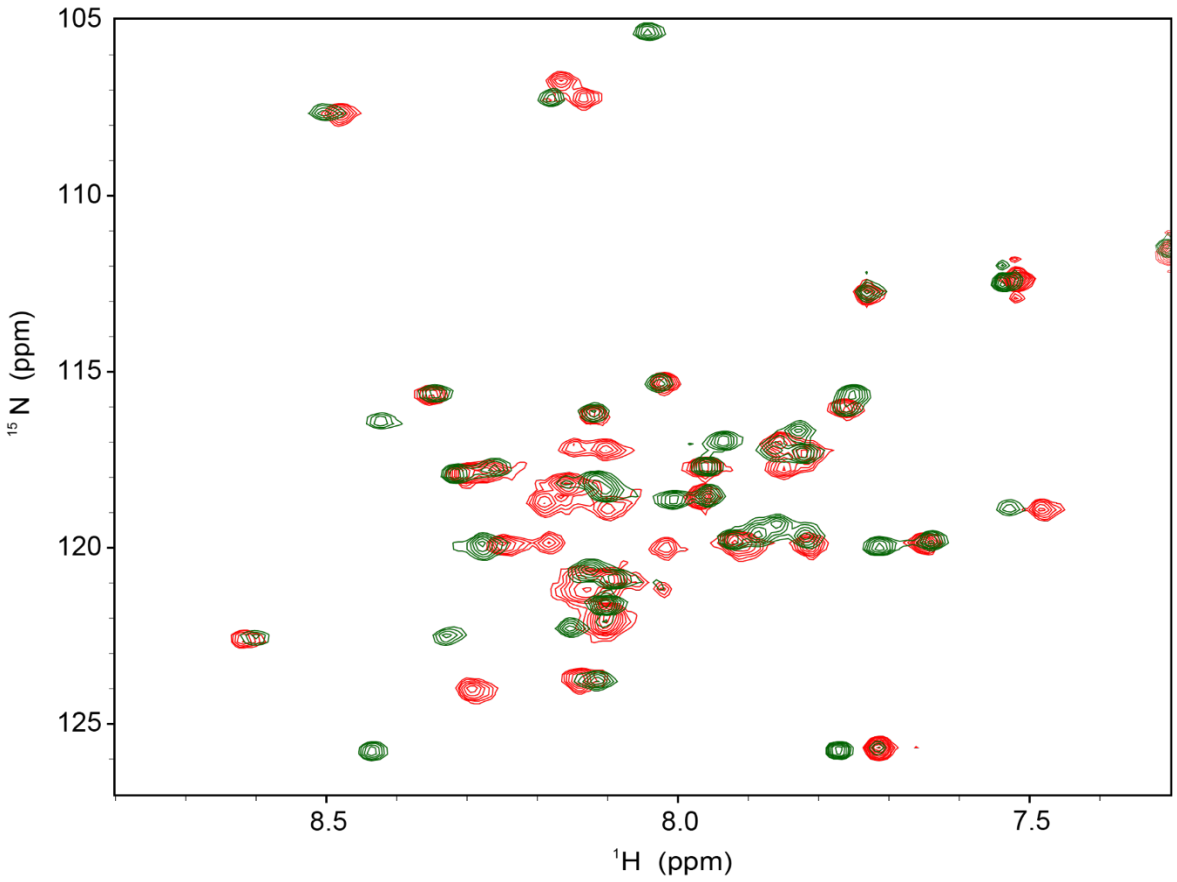

Figure 2.6 Chemical shift perturbations in the Post-Fusion conformation of the fusion domain (FD) (Red) following the addition of 5 mM TCEP (Green) highlights the correct formation of the internal disulfide bond within the fusion loop (FL).

2.2.2 The FP Alone Embeds Within the Membrane

To better understand how the Post-Fusion state of the FD interacts with the membrane, the dynamic behavior of the protein was investigated. It is to be expected that the Pre-Fusion state would be significantly more dynamic than the Post-Fusion state, with the latter restricted by its association with the membrane. Through spin-lattice (T1), spin-spin (T2) and ^{1}H-^{15}N NOE experiments, clear differences in the dynamics of the Pre- and Post-Fusion states can be witnessed (Figure 2.7).

The overall relaxation times of the FD in its Pre-Fusion state: ^{1}H-^{15}N NOEs = 0.207 ± 0.040 (Figure 2.7A), T1 = 0.563 s ± 0.010 (Figure 2.7B) and T2 = 0.272 s ± 0.008 (Figure 2.7C),

agree very well with the dynamics of a random coil. The Post-Fusion state on the other hand: ^{1}H-^{15}N NOEs = 0.625 ± 0.026 (Figure 2.7A), T1 = 0.808 s ± 0.007 (Figure 2.7B) and T2 = 0.090 s ± 0.005 (Figure 2.7C), represents that of a much more restricted structural conformation. These values correlate excellently with previous work carried out on the influenza FP which transitions to a more structured form when bound to DPC micelles.(70) Moreover, paramagnetic relaxation enhancement (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 which are solvent exposed and further emphasizes the difference between the Pre- and Post-Fusion

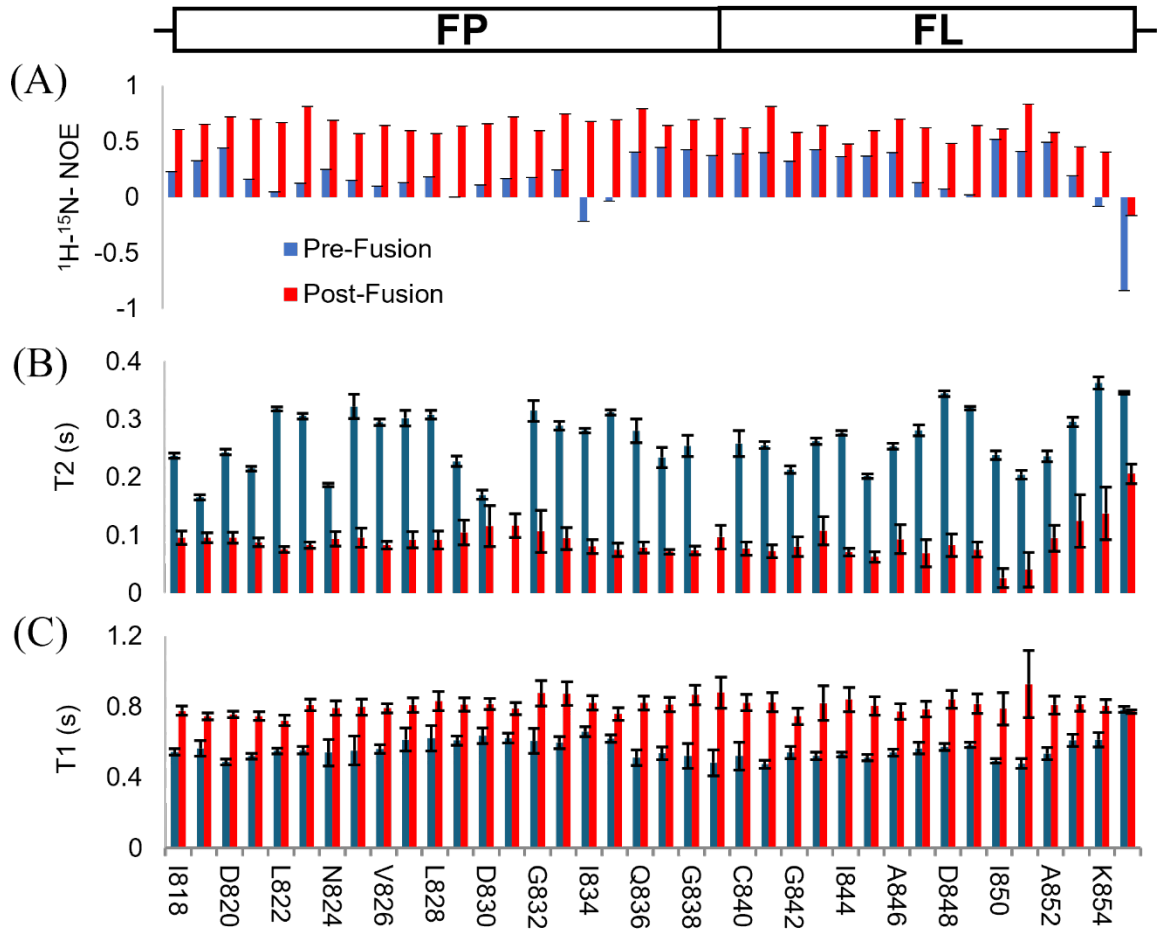

Figure 2.7 Measuring the (A) T1, (B) T2 and (C) ^{1}H-^{15}N NOEs relaxation rates for both the Pre-Fusion (Blue) and Post-Fusion (Red) states. In all measurements the pre-fusion state can be found to be significantly more dynamic than the post-fusion state due to the latter being restricted by its association with the membrane. Error bars are created from the signal to noise.

states (Figure 2.8A). In the Pre-Fusion 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 Post-

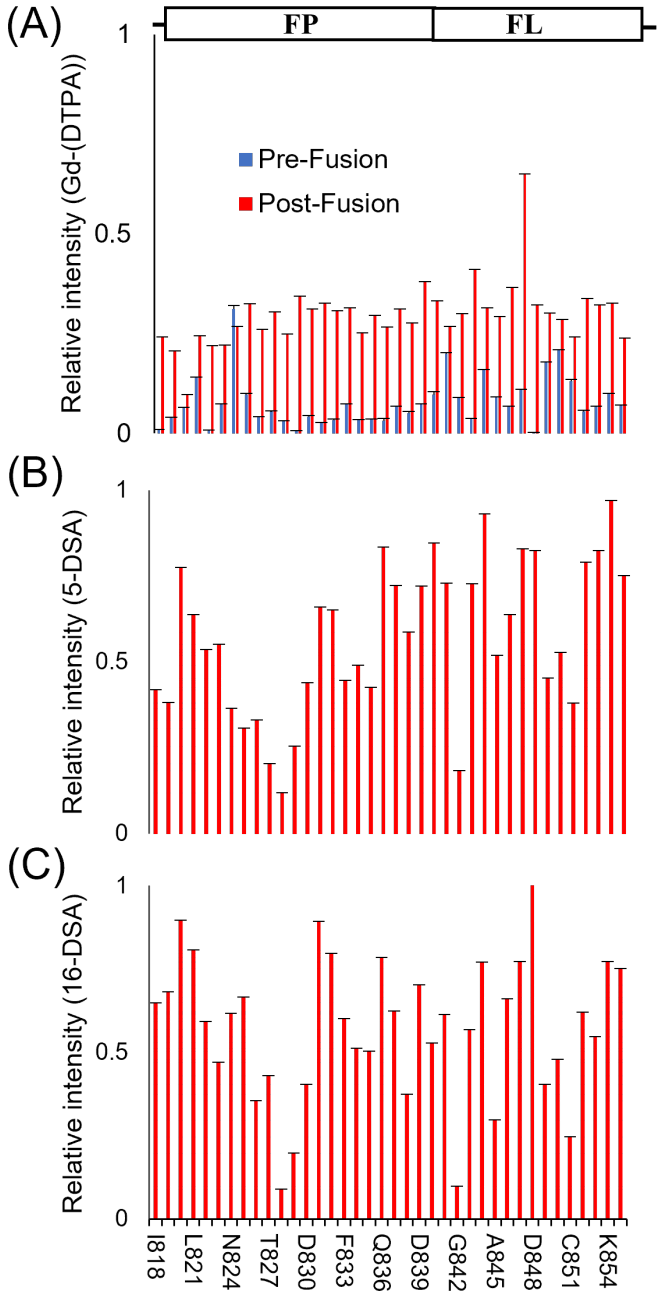

Figure 2.8 Assessing the depth of the fusion domain (FD) within the membrane .(A) With Gd(DTPA) significantly more quenching is witnessed for the Pre-Fusion state (Blue) than the Post-Fusion (Red), indicating it is more exposed to solvent. Relative intensities witnessed in the presence of (B) 1 mM 5-DSA 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 is capped at 1, although this is only applied to D848 in (C). Error bars shown are propagated from the signal to noise.

Fusion conformation also displays quenched signals throughout (Figure 2.8A). Although to a lesser extent than the Pre-Fusion state, this still suggests that no residues within the FD are found sufficiently deep enough 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 2.8B) and 16-DSA (Figure 2.8C). 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 2.8C) are remarkably 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 relative intensity for the FP; 0.507 ± 0.009, then we do for the FL; 0.685 ± 0.012. This can be attributed to the FP region existing in closer proximity to the 5-DSA probe than the FL. Furthermore, residues F823-D830, a linker region found in-between the α-helices of the FP, is the most severely quenched by both DSA probes. This suggests that F823-D830 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 just below the lipid headgroups, with the stretch of amino acids between the two α-helices F823-D830 the deepest point of insertion. The FL is not as severely quenched by 5-DSA, which leads us to believe that it does not insert itself within the membrane but rather interacts with the membrane surface in a more superficial manner.

2.3 Discussion

There has 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.(71-76) Whilst this has

facilitated the interactions between the spike protein and ACE2 to be investigated extensively, the molecular mechanisms behind membrane fusion remain elusive. Imperative towards this is the elucidation of the spike protein whilst associated with a lipid membrane, however this poses its own challenges. The structure of the membrane associated (Post-Fusion) SARS-CoV-2 spike protein currently available (PDB:6XRA)(65) is missing a large region in the S2 sub-unit (I^{770}-T^{912}), which contains the 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 towards understanding how the virus enters the cell. This is evident in previous research for influenza,(70, 77, 78) HIV,(39, 49) and EBOV,(46, 64, 79) where structural elucidation of the FD provided much needed molecular insight into the initiation 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 environment.

Research into the coronavirus FD has been hindered by debate regarding its location, with several candidates proposed.(52, 53, 80-82) However, the region S^{816}-F^{855} in SARS-CoV-2 appears to have the strongest case due to high sequence conservation, N-terminal location and high fusogenic activity, (50, 51) all of which are key requisites for the FDs of viral proteins. 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 fusion loop (FL: D^{839}-F^{855}). Both FP and FL have been shown to display independent fusogenic activity but are most effective in synergy.(54) Hence, we believe that the FP and FL work in concert to create a highly fusogenic structure, one which resembles two different classes of fusion machinery. The FP is akin to the N-terminal namesakes witnessed in HIV and influenza, whilst the internal FL resembles the structure seen in EBOV. This novel

arrangement of fusogenic machinery within a single virus glycoprotein is highly unique and could potentially explain the incredible fusogenecity of SARS-CoV-2.

Structural plasticity is a key feature of N-terminal FPs which allows them to insert themselves within and perturb host cell membranes.(38) Our data shows that the N-terminal FP undergoes structural rearrangement in the presence of a membrane (Figure 2.2). A transition from an unstructured (Pre-Fusion) to a structured (Post-Fusion) state as the FP interacts with the membrane results in the reorganization of lipids, decreasing the activation energy necessary for membrane fusion. Additionally, CSI revealed that the FP forms two α-helices upon interaction with a lipid membrane (Figure 2.2C). 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 when separated from the FL.(51, 54) Between helix one (F^{817}- L^{822}) and helix two (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. In agreement with models created through computational studies, the FP likely plays an integral role in perturbing the host cell membrane.(83, 84) We are proposing that the FP transitions into an obtuse helix-turn-helix motif upon interacting with a lipid membrane. This conformation is often witnessed in membrane proteins and a similar structural transition is witnessed in the boomerang model of the influenza FP.(70, 85) Previous work had identified the hydrophobic motif L^{821}-F^{823} as essential for the FDs membrane interaction and perturbing effects.(51) Our structural date places this motif at the end of helix one and 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}-L^{822} (Figure 2.9), whereas a turn may allow all three residues to face into the

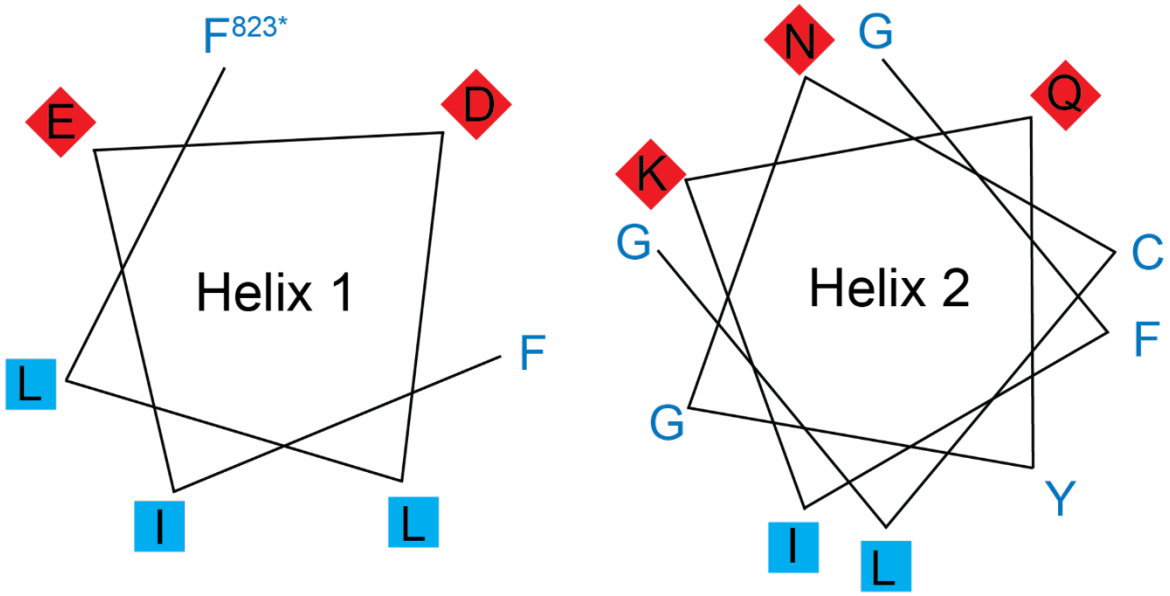

Figure 2.9 Amphipathic nature of helix one (F^{817}-L^{822}) and helix two (G^{832}-G^{838}). Helix one is displayed with F^{823} incorporated to highlight how it would interrupt the amphipathicity if incorporated.

membrane core. Deeper insertion within the membrane often suggests increased perturbing effects and fusion activity, due to greater destabilization of lipid packing and subsequent lowering of the activation energy necessary for membrane fusion.(86) 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.

Opposingly, in the disulfide bonded FL we see that the secondary structure remains undefined in both Pre- and Post-Fusion states (Figure 2.2C). This disulfide bond is highly conserved throughout the entire coronavirus family, 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.(57) This region found at the spike-interdomain interface within the spike protein trimer on the viral surface, was observed to undergo dramatic conformational change at endosomal pH (pH5.5). Resulting in coordinated movements of the receptor binding domain (RBD) and potential involvement in evasion of the immune response. Whilst this suggests that the FL may play an important role in the conformation of the glycoprotein

trimer, it also points towards functional implications within the endocytic pathway, similar to those witnessed within the internal FL of the EBOV.(64) An interesting thought on the FL's function surrounds cleavage by endosomal cathepsins. An auxiliary cleavage site has been identified between A^{829}-D^{830} which is in the middle of helix two of the FP and before the FL.(87) 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.

Based on the data, we have created a structural model for the FD embedded within the membrane (Figure 2.10). When looking into how deeply the FD inserts itself into the membrane, it was 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 FP and 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 head groups. 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

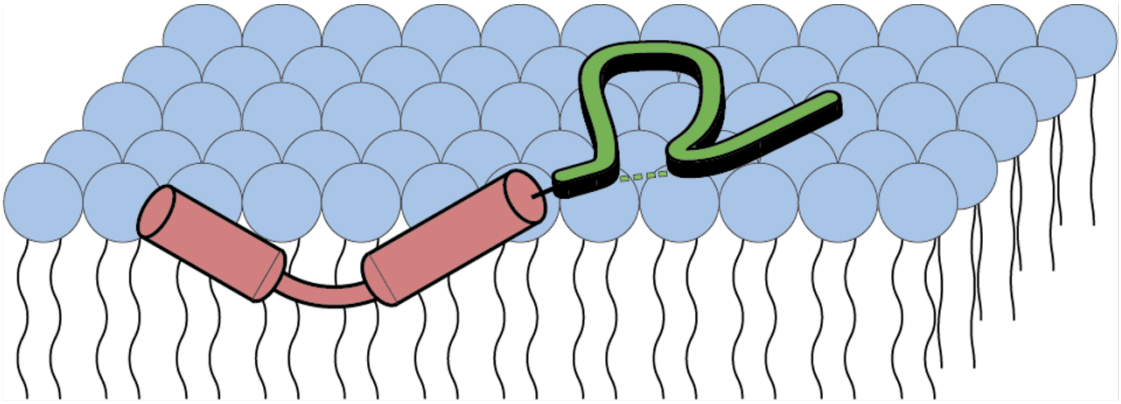

Figure 2.10 Structural model of the fusion domain (FD) of SARS-CoV-2 embedded in the host cell membrane. The fusion peptide (FP: S^{816}-G^{838}) shown in red, with a helix-turn-helix motif, embeds itself within the membrane in a shallow manner, just below the lipid head groups. Alternatively, the fusion loop (FL: D^{839}-F^{855}) displayed in green, lays upon the surface of the membrane.

quenching data witnessed with 16-DSA may be due to both regions laying at the limit of the probes paramagnetic effect. Therefore, within our model the FP is embedded within the membrane just below the lipid head groups, whilst the FL lays upon the lipid surface in a perpendicular orientation to the membrane. 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. This hypothesis is reinforced by the implication of Ca^{2+} as a cofactor in the membrane perturbing activity of the FD with binding sites found within both the FP and FL.(54) Additionally, specific lipids have been shown previously to have dramatic effects on the structure and activity of FDs in multiple other viruses.(78, 88)

2.4 Conclusion

In conclusion, we have provided structural evidence towards a working model for the FD of SARS-CoV-2 when associated with a lipid membrane (Figure 2.10). 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. Whilst 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 fusogenecity. The structural insights provided here can allow a better understanding the mechanistic process of SARS-CoV-2 membrane fusion and further work to solve the tertiary structure of the FD will build upon this work to provide a more in depth understanding of how the FD interacts with the host cell membrane. Comprehensive structural elucidation of the FD could serve as a platform for the development of novel antiviral therapeutics to target the FD and prevent SARS-CoV-2 fusion.

2.5 Materials and Methods

2.5.1 Expression

The SARS-CoV-2 fusion domain construct ^{816}(SFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKF)^{855} was designed with an N-terminal 6x His-tag, followed by a small ubiquitin like-modifier (SUMO) tag from *Saccharomyces cerevisiae* (*SMT3* gene) to aid with solubility and expression.(89) 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 the 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, 225 rpm then added to 1 L LB media. The large culture was grown to an OD_{600} of between 0.6-0.8 and induced with 1 mM IPTG final concentration for 4 hrs at 37 ºC. For isotope labelling (^{13}C ^{15}N), M9 minimal media was used with 1 g of ^{15}N ammonium chloride and 2 g ^{13}C glucose. For triple labeled samples (^{2}H ^{13}C ^{15}N), 1L of D_2O was used to make the minimal media instead of H_2O. A 25 mL starter culture in the appropriate minimal media was grown overnight at 37 ºC 225 rpm then added to a 1L volume. Following induction with final concentration of 0.1 mM IPTG at an OD_{600} of 0.6-0.8, cells were grown overnight at 18ºC, 225 rpm. Cells were then pelleted at 4 °C for 30 min at 4,500 xG and stored at -80 ºC or purified immediately.

2.5.2 Purification

The cell pellet was resuspended in lysis buffer (8 M Urea, 20 mM Tris, 300 mM NaCl, 10 mM BME, pH7.5) and sonicated for 10 minutes on ice (40% power, 1 s on, 1 s off). The clear solution was then centrifuged for 30 mins at 40,000 xG 4 °C. The supernatant was added to Ni-NTA resin

and incubated for at least 2 hrs at 4 °C. Following incubation, the resin was washed (8 M Urea, 20 mM Tris, 300 mM NaCl, 10 mM Imidazole, pH7.5) and then the bound construct was eluted (8 M Urea, 20 mM Tris, 300 mM NaCl, 250 mM Imidazole, pH7.5). SUMO protease ULP1 was added to the elution at a ratio of ~1:10,000, then dialyzed overnight against 4 L of cleavage buffer (20 mM Tris, 300 mM NaCl, 5 mM BME, pH 7.5) at 4 °C. The protease vector pFGET19_Ulp1 was a gift from Hideo Iwai (Addgene plasmid # 64697).(90) The following day the sample was rebound to the Ni-NTA resin with at least a 2 hr incubation at 4 °C. The flowthrough was collected, and the resin washed (20 mM Tris, 300 mM NaCl, 10 mM Imidazole, pH 7.5) to obtain all the cleaved construct. The rebind flowthrough and wash were then pooled and dialyzed against 4 L volumes of folding buffer (20 mM Tris, 300 mM NaCl, pH 7.5) at 4°C. Two buffer exchanges were made, both after 3 hr intervals before the sample was left overnight to allow for correct disulfide bond formation. The dialyzed sample was then concentrated with a Amicon® Ultra Centrifugal Filter (3kDa MWCO) prior to further purification via size exclusion chromatography (SEC) using a Superdex™ Peptide 10/30 HR column with 25 mM Na_2HPO_4 150 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.

2.5.3 NMR Experiments

NMR spectra were acquired using a shigemi NMR tube with a sample volume of ~300 µL in 25 mM Na_2HPO_4, 150 mM NaCl, pH7.4 buffer with 90%H_2O/10%D_2O. 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% non-uniform sampling (NUS). All data was processed using the software NMRPipe(91) and NMRFAM-SPARKY(92) via NMRBox(93). Triple labeled (^{2}H ^{13}C ^{15}N) samples were produced solely for the HN(CA)CB experiments, to improve the signal to noise. Dodecyl phosphocholine (DPC) was applied as a membrane mimic at 100 mM for the backbone assignment. Accurate DPC concentrations were gathered via phosphorous 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) in order to ascertain the DPC concentration. T1, T2 and heteronuclear NOE experiments were conducted to determine the dynamic properties of both Pre- and Post-Fusion states of the FP. T1 and T2 data was processed using the Bruker Topspin Dynamics Center, whilst heteronuclear NOEs were processed via NMRFAM-SPARKY. Error shown is propagated from the signal to noise or standard error of the mean (SEM).

2.5.4 Paramagnetic Relaxation Enhancement

For probing 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 paramagnetic relaxation enhancement (PRE) experiments. Gd(DTPA) was solubilized in autoclaved H_2O at a concentration of 100 mM and titrated into the NMR sample to reach the following concentrations: 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 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 conducted for concentrations of 0.5, 1, 2, 3 and 5 mM. All data was collected through ^{1}H-^{15}N HSQC at 296K, and none of the paramagnetic agents altered the structure of the peptide as evident by the lack of chemical shift perturbations.

For all non-overlapping signals at each titration point, the relative intensity was calculated as $\frac{I_0-I_n}{I_0}$. 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_o was the first titration point due to increasing signal intensity as a result of the paramagnetic effect.(94) Error shown is propagated from the signal to noise or SEM.

2.5.5 Circular Dichroism

CD spectra were collected on a Jasco J810 Spectro-Polarimeter using a quartz cuvette with a 1 mm path length. All experiments were conducted at 23 °C in 2.5 mM Na_2HPO_4, 15 mM NaCl, pH 7.4 with 20uM protein. Concentrations of 1 mM SUVs and 100 mM DPC was used where applicable. Data was collected from 260 nm to 190 nm with a step size of 1 nm at 20 nm/min and averaged over 3 accumulations. The spectra were converted from millidegrees (mdeg) to mean residue molar ellipticity (θ) in order to calculate the percentage helix at 222 nm using the following equation: $\frac{\theta_{222}-\theta_C}{\theta_H-\theta_C}$, where $\theta_H = (250T - 44{,}000)\left(1-\frac{3}{n}\right)$ and $\theta_C = 2{,}220 - 53T$.(95) T is the temperature in °C and *n* is the number of residues in the peptide.

2.5.6 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 whilst 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 15 mM NaCl pH7.4 and sonicated for 15 mins at 10% duty cycle (1 s on/1 s off) with the sample sat in ice water, using a Branson® ultrasonicator microtip.

Chapter 3: The Fusion Peptide and Fusion Loop act in Synergy to Initiate Fusion

This work has been published in *Protein Science*.(96)

Birtles, D., A. E. Oh, and J. Lee. 2022. Exploring the pH dependence of the SARS-CoV-2 complete fusion domain and the role of its unique structural features. *Protein Science*. 31(9):e4390

D.B. and J.L. conceptualized research; D.B. and A.E.O. performed research; D.B., and J.L. analyzed data and all authors prepared the manuscript.

3.1 Introduction

Although FDs are generally well conserved within a viral family, they can vary greatly across different families. For example, HIV and influenza both contain N-terminal fusion peptides (FPs) but with different structures when embedded within the membrane.(39, 49, 70) Ebola virus (EBOV) contains an internal fusion loop (FL) that is held together by a disulfide bond and undergoes a large conformation change at low pH,(64) whilst the coronavirus FD contains both an N-terminal FP and an internal FL that has only ever been witnessed previously in the arenavirus family.(47, 97) Nonetheless, all of the aforementioned FD's are responsible for the initiation of viral fusion through embedding within and perturbing their respective target membranes.

Previous work on the FD of SARS-CoV 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.(54) More recently, structural information gathered for the SARS-CoV-2 FD in dodecyl phosphocholine (DPC) micelles at pH7.4(47) and 1,2-dimyristoyl- sn -glycero-3-phosphocholine (DMPC): 1,2-diheptanoyl-sn-glycero-3-phosphocholine (DH^7PC) bicelles at pH5.0(98) has revealed further details. The FP was found to contain a helix-turn-helix motif that buries into the membrane, whilst the FL is more superficially

associated with the lipid headgroups (Figure 2.10). 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.(57, 99) Fusion at the endosomal membrane is often associated with low pH as it is commonly found to be a major determinant for viral fusion.(100) 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, suggesting a possible preference for the virus to enter the cell via the endocytic pathway. Through solution NMR only a slight change in the conformation of the FD was observed between pH7.4 and pH5.0. However, utilizing a FRET-based lipid mixing assay the SARS-CoV-2 FD displayed significantly more efficient fusion at pH5.0 than at pH7.0, 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 new model to explain the structure and function relationship involved in how the SARS-CoV-2 FD initiates membrane fusion in different pH environments.

3.2 Results

3.2.1 Minor pH Dependent Conformational Change

The majority of viruses enter the target cell via a single pathway, yet SARS-CoV-2 has been shown to fuse at both the plasma and endosomal membrane,(57, 99) 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.(62, 64) Lowering the local pH environment of the SARS-CoV-2 FD from pH7.4 to pH5.0 induces a slight conformational change within the membrane embedded peptide (Figure 3.1).

We previously showed that when associated with either a small unilamellar vesicle (SUV) or DPC micelle at pH7.4, the conformation of the FD (S^{816}-F^{855}) undergoes a drastic conformational change compared to in solution (Figure 2.2).(47) When the FD is embedded within a DPC micelle, decreasing the pH to that present in the acidic environment of the late endosome (pH5.0) appears to induce a minor structural rearrangement when compared to neutral pH (Figure 3.1). This lack of global structural change is reiterated by CD, utilizing the more physiologically relevant bilayer environment of small unilamellar vesicles (SUVs) (Figure 3.2). Comparing the ^{1}H-^{15}N HSQC spectra at pH7.4 and pH5.0 we witness some chemical shift perturbations, suggestive of a change in tertiary conformation (Figure 3.1A). 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 FD to embed within the DPC micelle at pH 7.4, the helix-

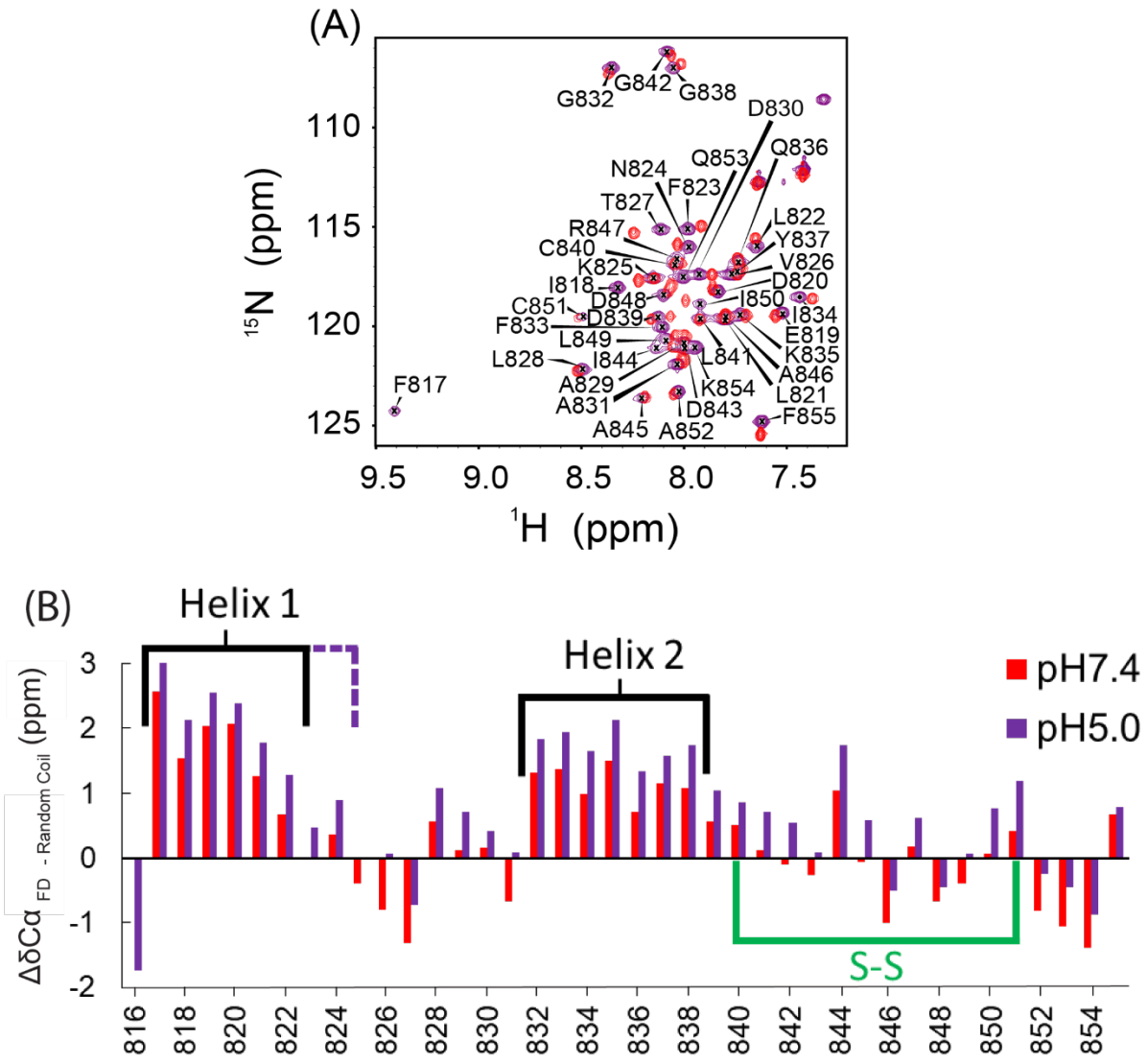

Figure 3.1 A slight conformational change is witnessed for the SARS-CoV-2 FD embedded within the membrane between pH7.4 (Red) and pH5.0 (Purple). (A) A minor change in the tertiary conformation of the FD was detected via ^{1}H-^{15}N HSQC chemical shift perturbations, with the backbone assignment shown at pH5.0. (B) Chemical shift indexing revealed the same secondary structure elements at both pH7.4 and pH5.0, with a slight elongation of helix one at pH5.0.

turn-helix, exist at pH7.4 and pH5.0 (Figure 3.1B). Within helix one, a slight elongation was detectable from L^{822} to N^{824}; this correlating with the published structure of the FD embedded in a lipid bilayer at pH5.0 (PDB: 7MY8)(98). A potential elongation of helix two from G^{838} to G^{842} was also detected in the pH5.0 CSI (Figure 3.1B). 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 pH5.0 structure (PDB:7MY8) would be unable to form, offering an explanation as to why helix two ends at G^{838} in all available experimental structural data of the FD.(47, 98) We also compared the dynamic nature and depth of insertion within the DPC micelle of the FD at pH7.4 and pH5.0. Both T1 and HN-NOEs revealed no significant changes in the dynamics of the peptide, yet the T2 data displayed more dynamic behavior when embedded in the membrane at pH5.0 than at pH7.4 (Figure 3.3). The 16-DSA paramagnetic probe also displayed only minor deviations based on changing pH alone, but the 5-DSA probe did suggest that the FD embeds deeper into the membrane at pH5.0 (0.20±0.02) than at pH7.4 (0.32±0.03) (Figure 3.4).

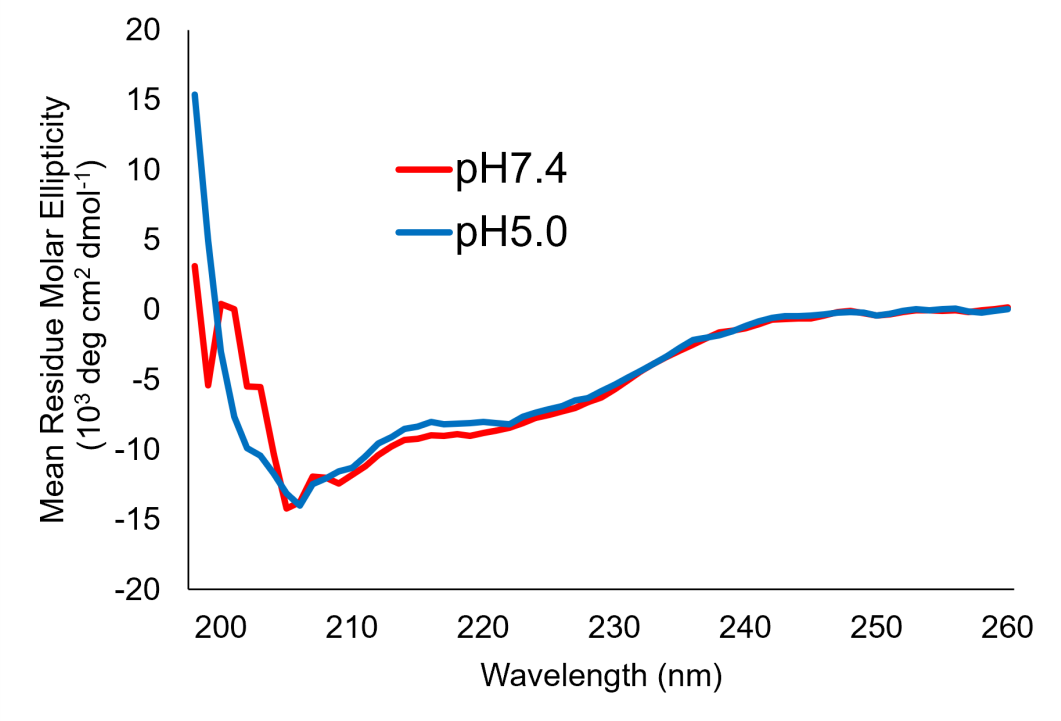

Figure 3.2 CD spectrum for the SARS-CoV-2 fusion domain (FD) embedded in the membrane of 75:25 POPC:POPG small unilamellar vesicles (SUVs) at pH7.4 (Red) and pH5.0 (Blue).

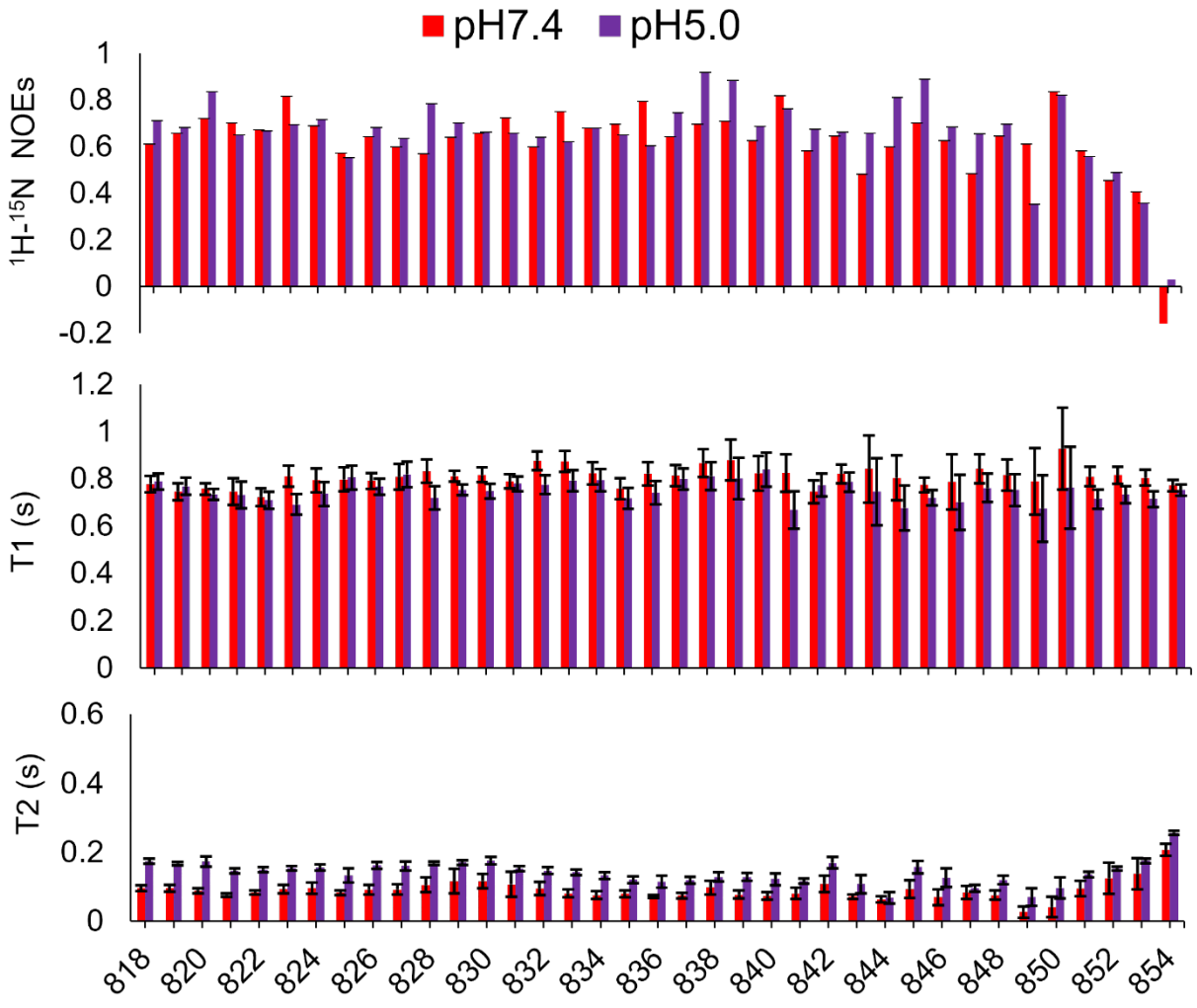

Figure 3.3 Relaxation experiments comparing the fusion domain (FD) embedded in the membrane at pH7.4 (Red) and pH5.0 (Purple).

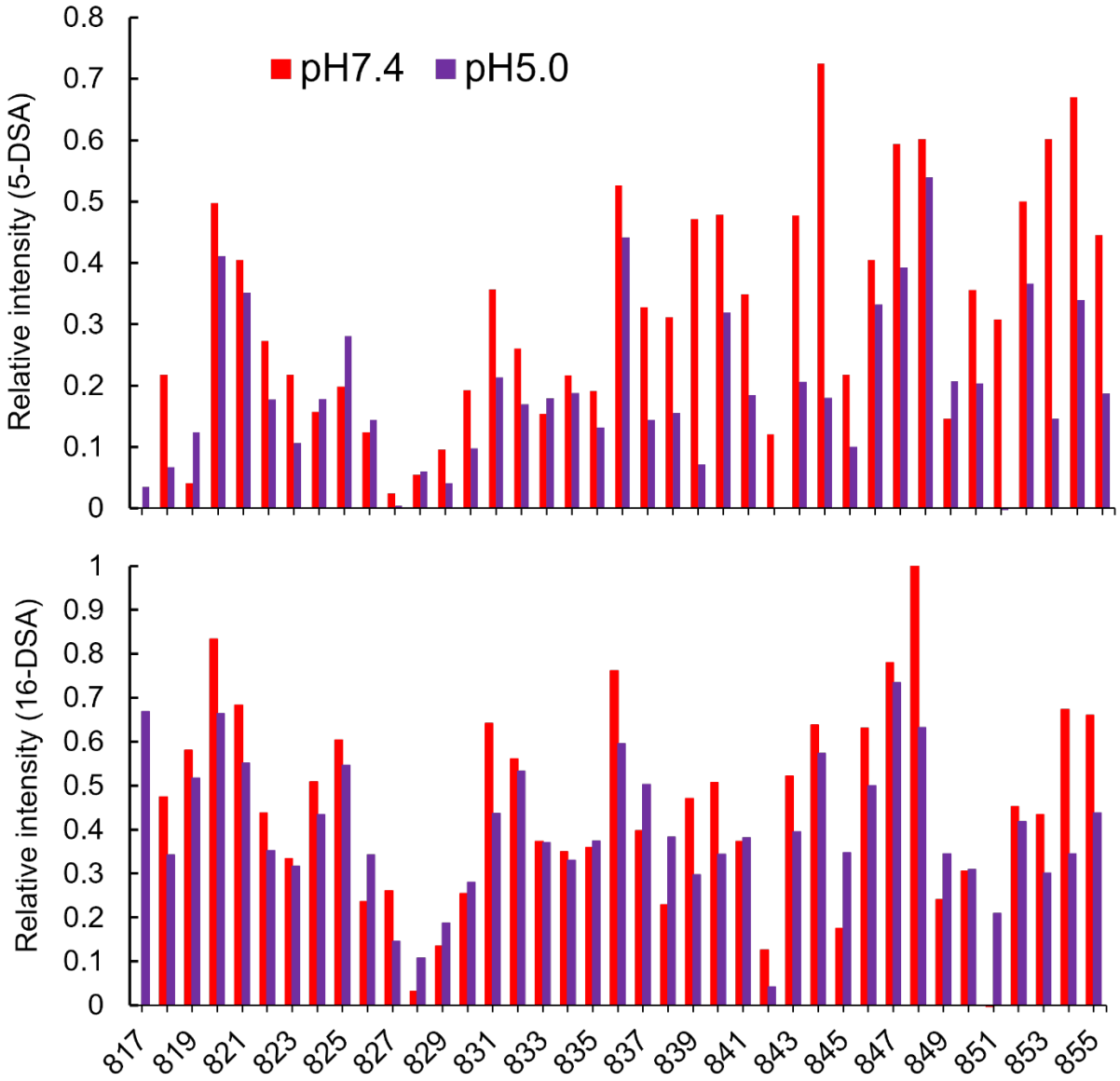

Figure 3.4 Paramagnetic relaxation enhancement (PRE) experiments with 2 mM (A) 5-DSA and (B) 16-DSA comparing how deeply the FD embeds itself within the membrane at pH7.4 (Red) and pH5.0 (Purple).

3.2.2 Environmental pH is Critical for FD Mediated Fusion.

Functionally, environmental pH can be associated with the fusogenicity of the FD at the plasma membrane (~pH7.4-7.0), as well as distinct stages of the endosomal membrane (~pH6.5-5.0). The endocytic pathway sees a large decrease in pH as the lysosomal compartment forms, ranging

from pH6.5-6.0 in the early endosome to pH5.5-5.0 in the late endosome.(100) 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 3.5).

At pH5.0 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 3.5A). A P/L ratio of 0.01 is expected to be close to that of physiological conditions based on electron tomography analysis of the viral surface.(74, 101) However, at pH7.0 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 pH5.0 (Figure 3.5A). A linear correlation up to a P/L ratio of 0.2 is still observed, signifying a positive relationship between the concentration of peptide and the amount of lipid mixing witnessed (Figure 3.6). This relationship eventually tails off at higher P/L ratios (>0.2), most likely due to oversaturation of the peptide on the liposome surface.

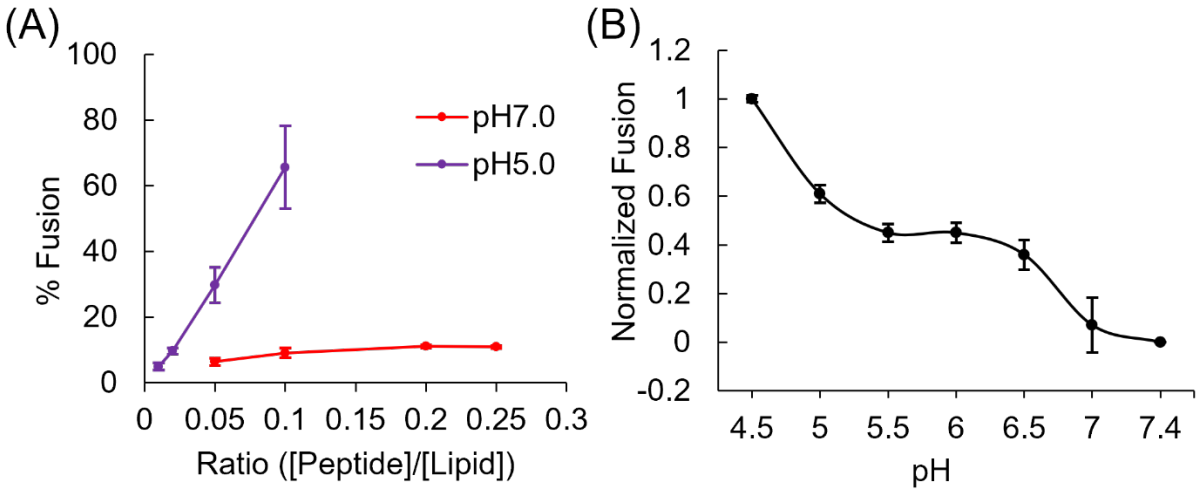

Figure 3.5 The SARS-CoV-2 fusion domain (FD) displays a large pH dependence to fusion in simple lipid compositions of 3:1 POPC:POPG. (A) A linear relationship can be seen with the P/L ratio at pH5.0 (Purple), but drastically reduced fusion is present at pH7.0 (Red) even at much higher peptide concentrations. (B) A positive, biphasic relationship is witnessed between decreasing environmental pH and lipid mixing with a P/L ratio of 0.05.

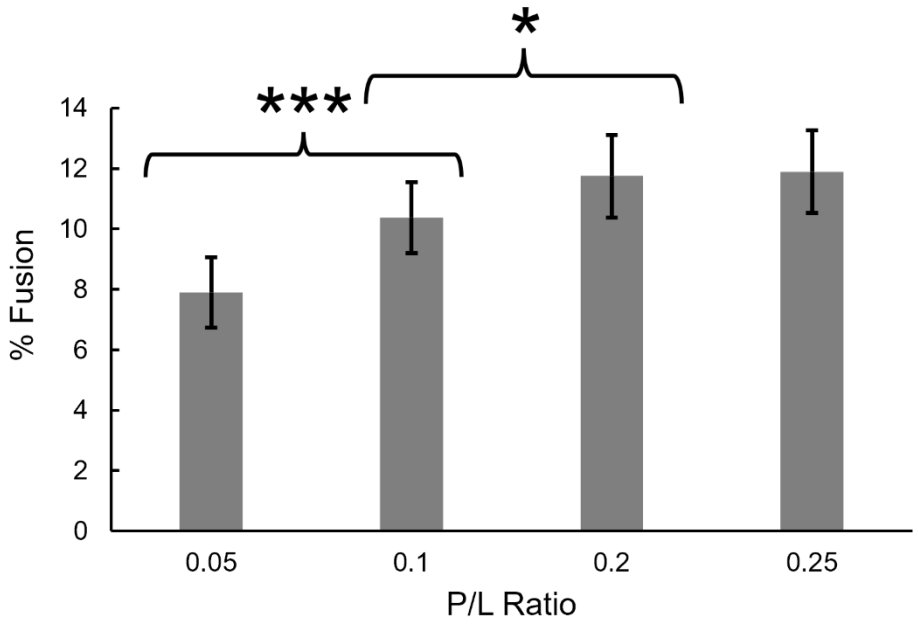

Figure 3.6 A linear increase in fusogenecity is witnessed at pH7.0 with increasing P/L ratios that are statistically significant. Through an independent T-test it was found that the lipid mixing values between P:L ratios of 0.05 (M = 7.90, SD = 1.16) and 0.1 (M = 10.38, SD = 1.17) were significant t(14)=-4.23, p<.001, as well as between 0.1 and 0.2 (M = 11.74, SD = 1.37) t(14)=-2.15, p = .049. No significant difference was found between P:L ratios of 0.2 and 0.25 (M = 11.60, SD = 1.37) with t(14)= 0.21, p = 0.83, likely due to oversaturation at such high peptide concentrations. *** = p<.001, * = p<.05

To probe the relationship between pH and fusogenicity further, we conducted the same functional assay over a range of several different environmental pH's that are relevant within the endocytic pathway (Figure 3.5B). A biphasic relationship was revealed where we see an initial increase in fusogenicity at a pH corresponding to that of the early endosome (pH6.5-6.0) followed by a brief plateau. The jump in fusion from pH7.0 to pH6.0 is approximately 6-fold, 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 (pH5.0), and when we compare the ability of the FD to initiate fusion at pH5.0 to that at pH7.0, we see an almost 9-fold increase in activity. Our results suggest that the FD can initiate fusion in the pH environment of the plasma membrane (~pH7.0) and early endosome

(pH6.5-6.0), yet as the pH decreases in the later stages of the endocytic pathway (pH5.5-5.0) 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 (<pH7.0). Moreover, the two-stage fusion profile witnessed (Figure 3.5B) demonstrates that the FD can initiate fusion in the early endosome (~pH6.0), yet the potency of the FD increases by approximately 50% at a pH corresponding to that of the late endosome (~pH5.0). This indicates 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.

3.2.3 The FP and FL are Both Integral for Fusogenicity.

We earlier found that the FD consists of two structurally defined regions, the FP and the FL, with the former proving to be the primary site of interaction with the DPC micelle.(47) Previous work on the FP of SARS-CoV found that a 'hydrophobic motif' consisting of L^{821}, L^{822}, and F^{823} played an important role in perturbing the membrane.(51, 54) On the other hand, although the FL is known to be important for the activity of the FD, how it contributes to the initiation of membrane fusion at different pH's is still not well understood.(50, 51) 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.

Firstly, we mutated each individual amino acid in the 'LLF' motif to alanine creating three single mutant constructs as well as a triple mutant; all of which displayed a remarkable loss of

function (Figure 3.7A). We then explored whether or not the aromatic functional group of F^{823} plays a key role by replacing this phenylalanine with a tyrosine ($F^{823}Y$) or tryptophan ($F^{823}W$) residue(Figure 3.7B). The $F^{823}Y$ mutant recovered little functionality and was equivalent to $F^{823}A$.

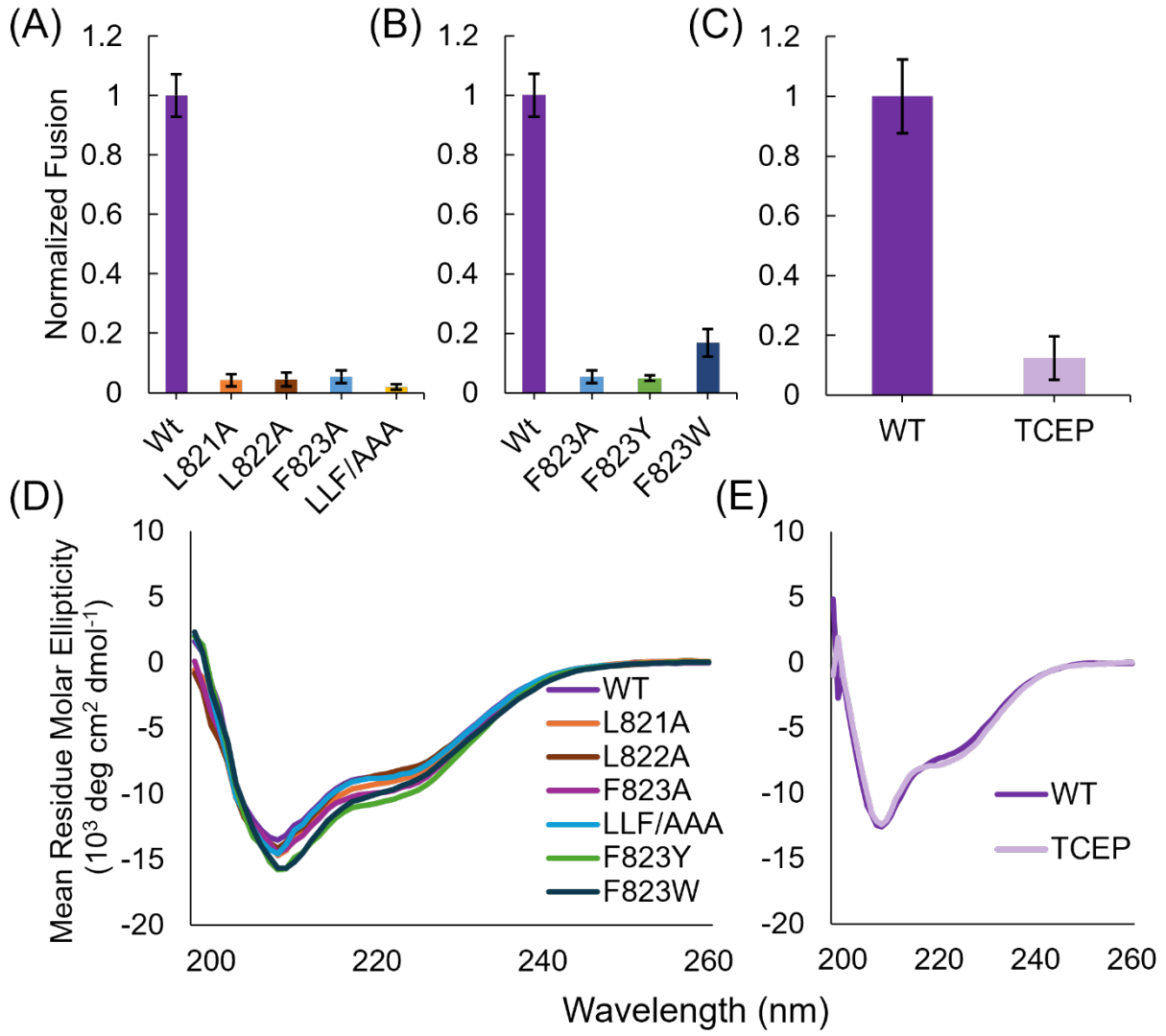

Figure 3.7 The hydrophobic motif 'LLF' and the fusion loop (FL) play a key role in the fusogenicity of the fusion domain (FD) at pH5.0. (A) Mutating the LLF motif residues to alanine individually, or all three simultaneously, leads to an almost complete loss of function (n≥12). (B) Conserving the aromatic 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.

Opposingly, $F^{823}W$ recovered more fusogenic activity than both $F^{823}A$ and $F^{823}Y$, although this was still approximately 5-fold lower than the WT (Figure 3.7B). These results echo that witnessed in SARS-CoV, which also found each residue in the 'LLF' motif to be critical towards membrane perturbation via viral infectivity assays.(51) All of the LLF mutants displayed the same global secondary structure as the WT (Figure 3.7D), 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 towards the function of the FD, with a potential specific role for the phenylalanine residue at position 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 conserved internal disulfide bond. In the presence of 1 mM TCEP a significant decrease in fusogenic activity is witnessed (Figure 3.7C), 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 FD as witnessed through CD (Figure 3.7E). This suggests that the helix-turn-helix motif of the FP remains intact even after the disulfide bond of the FL is lost. Interestingly, CSPs in the FL were found when comparing the Wt FD to that present under reducing conditions via ^{1}H-^{15}N HSQC, indicating large tertiary structure changes (Figure 2.6).(47) 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.

3.3 Discussion

Much of the research involving the activity of the coronavirus FD has been conducted in SARS-CoV, 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 pH5.0 or assessed the FP and FL independently.(51, 98, 102, 103) Initially, both domains were found to contain some fusogenic properties in isolation, with the FP even displaying a preference for fusion at low pH.(51) 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.(54) Thus, in order to fully understand the molecular mechanisms behind how the FD of SARS-CoV-2 interacts with the membrane to initiate 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 towards 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 towards membrane fusion.(49, 66, 78, 79, 88, 104-107) For the FD of SARS-CoV-2, limited research has been conducted regarding the impact of lipid compositions on membrane perturbation, yet the role of Ca^{2+} has been shown to be significant at low pH.(66) Interestingly when investigating the role of Ca^{2+} at neutral pH, we found that no significant impact was witnessed either structurally or functionally (Figure 3.8). 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 differing protein: Ca^{2+} and

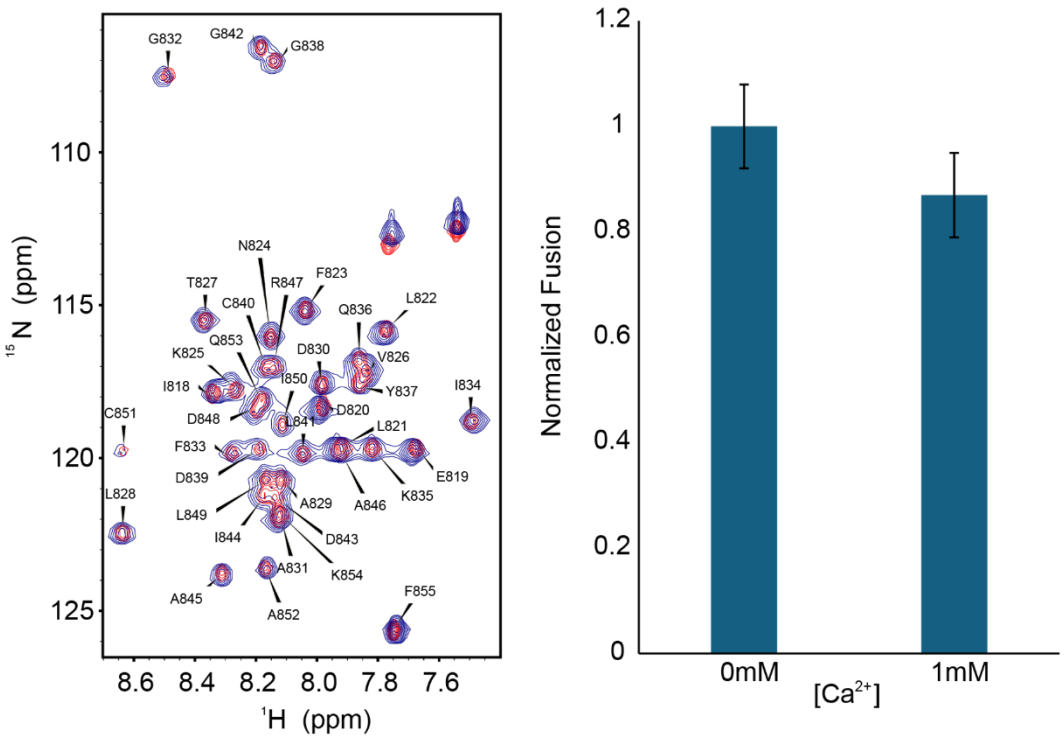

Figure 3.8 At pH7.4 physiological concentrations of Ca^{2+} have no significant impact on the structure or function of the SARS-CoV-2 fusion domain (FD). (A) ^{1}H-^{15}N HSQC of the FD with 1.5 mM Ca^{2+} (Blue) and without (Red) (B) Lipid mixing displayed no significant difference when 1mM Ca^{2+} was introduced to the buffer environment (n = >8).

protein: lipid interactions, we opted to use no Ca^{2+} and a simple lipid composition of POPC:POPG 75:25, which is akin to that used previously in the study of several different viral FDs. (46, 64, 108)

SARS-CoV-2 is well established as being able to fuse at both the plasma membrane(99) and endosomal membrane(57) with the route taken thought to be due to the cell type being infected and the presence of particular proteases.(109) 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 FD of SARS-CoV-2 has been shown here to contain a clear pH dependency in order to initiate membrane fusion (Figure 3.5). 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 (pH6.5-6.0)

but may more readily occur in the late endosome as the pH continues to decrease (~pH5.0) (Figure 3.5B). A similar profile has been witnessed previously for the internal fusion peptide (IFP) of the avian sarcoma/leukosis virus subtype A (ASLV-A) virus.(110) Here the two-step process of fusion were attributed to the more complete transition from hemi-fusion to full fusion at lower pH, which is consistent with other studies on the pH requirements for fusion.(45) 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 to prove this theory. Furthermore, a preference for SARS-CoV and SARS-CoV-2 to fuse in the late endosome (~pH5.0) has also been shown *in vivo*.(111, 112) Whilst 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 a lower pH.

Structurally, we witnessed little change in secondary structure in the FD between pH7.4 and pH5.0, 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 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 pH5.0 (PDB:7MY8)(98) and MD simulations,(84, 113) 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 FD witnessed at low pH (Figure 3.9). In this model at pH5.0, the slight elongation of helix 1 sees a minor change in the orientation of the helix-turn-helix motif where the entire FP

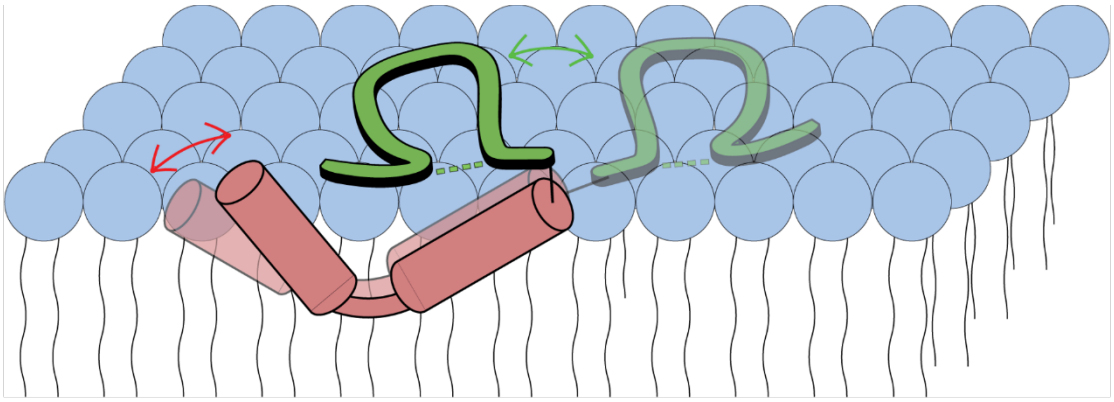

Figure 3.9 Structural model of the fusion domain (FD) at pH7.4 (Opaque) and pH5.0 (Full). The fusion peptide (FP - Red) embeds in the membrane with a helix-turn-helix conformation regardless of pH, yet at pH5.0 we see a slight elongation of helix one resulting in the turn region burying deeper into the membrane. The Fusion loop (FL - Green) interacts more superficially with the surface of the membrane and may display more dynamic motions as the pH is decreased. 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.

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(98) and thus as the FP embeds deeper, the FL is also drawn further into the membrane. This would result in a much greater local membrane perturbing effect, greatly impacting the lipid packing in the immediate membrane environment. This increased depth of insertion throughout the FD within the lipid membrane at lower pH is also corroborated by our PRE data with 5-DSA (Figure 3.4). 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 support the FP inside the membrane and increase the force needed to remove it.(113) 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 spike protein carried out at a series of low pH's.(57) 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 whilst 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 the hydrophobic motif LLF in the FP was shown to be integral to the regions capacity to initiate fusion.(54) This motif also appears to be critical for fusion in SARS-CoV-2, at both pH5.0 (Figure 3.7) and pH7.0 (Figure 3.10), although the mechanistic details are still unclear. LLF is found immediately preceding the turn region, which is the deepest embedded component of the FD. Hence, this hydrophobic motif may be integral towards 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 point of insertion within the membrane.(70, 114) Crucial to the turn region within influenza are two aromatic sidechains that flank either side.(77) 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 3.7), whilst 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 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 FD.(84) Within our

model (Figure 3.9) the ‘LLF’ motif plays an integral role in the pH dependent rearrangement of helix one, 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

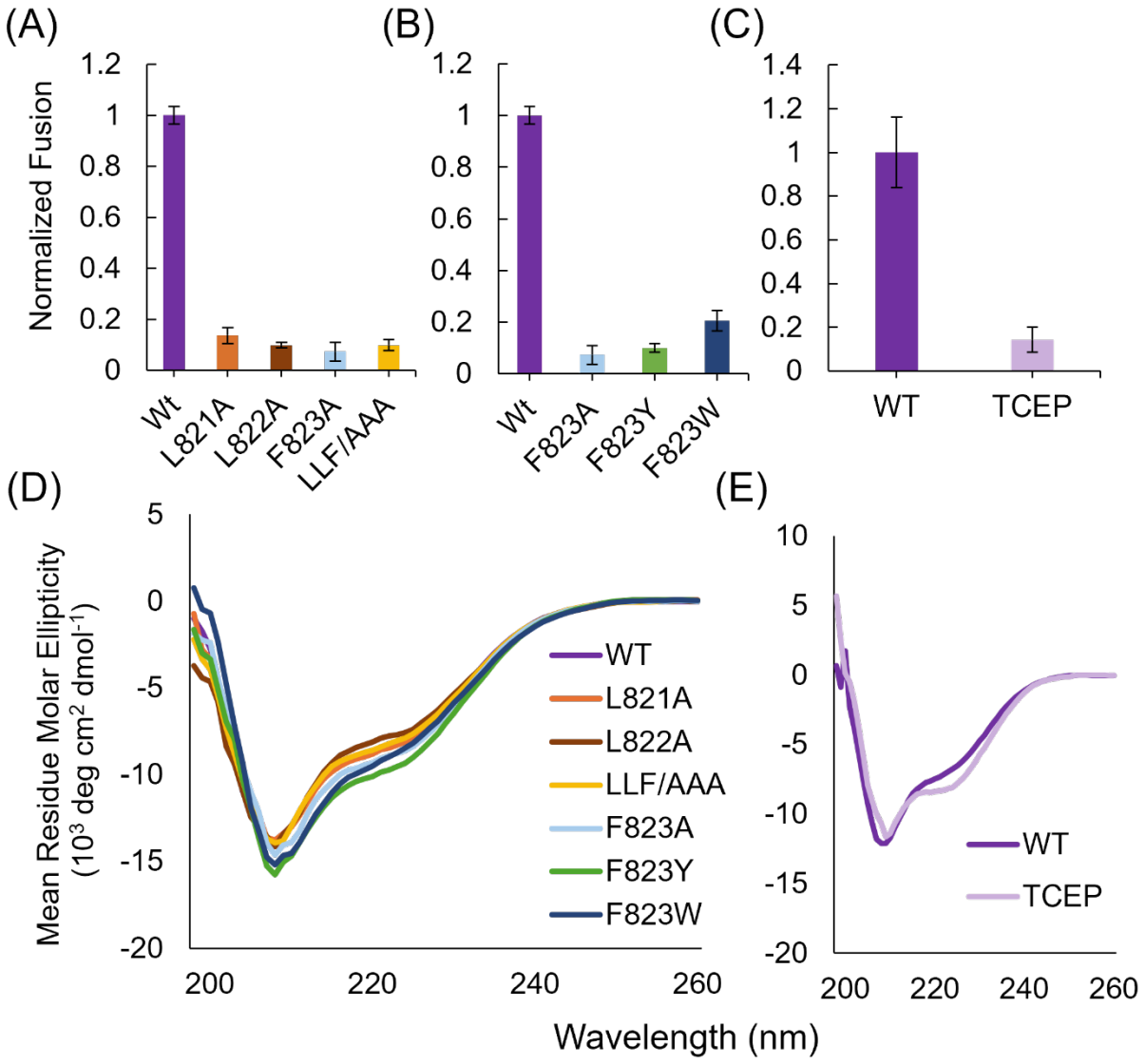

Figure 3.10 The hydrophobic motif 'LLF' and the fusion loop (FL) play a key role in the fusogenicity of the complete fusion domain (FD) at pH7.0. (A) Mutating the LLF motif residues to alanine individually, or together results in an almost complete loss of function. (B) Conserving the aromatic functional group at position 823 maintained some activity only when tryptophan was present. (C) The presence of 1 mM TCEP in the environment decreases the amount of lipid mixing witnessed substantially. CD confirmed the global secondary structure of all (D) LLF mutants and the (E) WT in the presence of 500 μM TCEP exhibited no significant changes in the global secondary. n ≥ 8 for all lipid mixing experiments.

but several domains in perturbing the membrane.(41) 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 internal fusion peptide (IFP). (52, 81, 102, 115) 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.(47, 98, 102) 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 FD and IFP, 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.(41) Whilst 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.

3.4 Conclusion

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 secondary structure 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 towards 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.

3.5 Materials and Methods

3.5.1 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 (SUMO) tag from *Saccharomyces cerevisiae* (*SMT3* gene) to aid with solubility and expression.(90) 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.(47) All mutants in this study were purified in the same manner as the wild type.

3.5.2 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, pH7.4 or pH5.0 buffer with 90%H_2O/10%D_2O. All experiments were conducted on the Bruker Ultrashield^{TM} 600 MHz magnet with a CPTXI 600S3 H-C/N-D-05 Z Cryoprobe at a temperature of 23 $^{\circ}$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 was processed using the software NMRPipe and

NMRFAM-SPARKY via NMRBox.(91-93) Dodecyl phosphocholine (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/15N/D Z-GRD Cryoprobe. Spin-lattice (T1), spin-spin (T2) and heteronuclear NOE experiments were conducted to determine the dynamic properties of the FD at both pH7.4 and pH5.0. T1 and T2 data was processed using the Bruker Topspin Dynamics Center, whilst 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.(47) 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).

3.5.3 Preparation of Large Unilamellar Vesicles

Large unilamellar vesicles (LUVs) were produced by mixing specified amounts of lipid stock solutions in glass test tubes. Chloroform was removed via gentle vortexing whilst 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, pH7.4 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^{TM}), a total of 21 times through 2 polycarbonate membranes with a 100nm pore size. LUVs were either used immediately or stored for a maximum of 72 hrs at 4 ºC prior to use.

3.5.4 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.(49, 64, 116) 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 5µM FD and 100µM LUVs (protein/lipid ratio of 0.05) unless stated otherwise, with all conditions containing at least 4 replicas (n≥4). Both liposomes and FD were suspended in 10 mM HMA, 100 mM NaCl, pH7.4 prior to the experiment, with acidification to the described pH controlled by addition of a specific amount of 1M HCl. Fluorescence was recorded using a SpectraMax M5 microplate with all readings taken at room temperature (~22 °C). Percent lipid mixing/fusion was calculated as $\frac{(I_F - I_B)}{(I_{100} - I_B)} x\ 100$, 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 pH7.4 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 standard error of the mean (SEM).

3.5.6 Circular Dichroism

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 $^{\circ}$C) in 1 mM HMA, 10 mM NaCl, 50 mM DPC, pH7.4 or pH5.0 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 3 accumulations, with all smoothing carried out using the program CDToolX.(117)

Chapter 4: A Preference for the Late Endosomal Pathway

This work has been published in *The Journal of Physical Chemistry B*.(118)

Birtles, D., W. Abbas, and J. Lee. 2024. Bis(Monoacylglycero)Phosphate Promotes Membrane Fusion Facilitated by the SARS-CoV-2 Fusion Domain. *The Journal of Physical Chemistry B*. 128(11)

D.B. and J.L. conceptualized research; D.B. and W.A. performed research; D.B., and J.L. analyzed data and all authors prepared the manuscript.

4.1 Introduction

The exact mechanism by which the SARS-CoV-2 FD initiates membrane fusion becomes even more intriguing when considering that it can utilize the plasma and endosomal membrane pathways for fusion.(30) This is a rather uncommon characteristic for a virus, likely improving the overall success rate of infection. As a result, the FD is expected to embed within two different lipid environments, given that the plasma membrane and endosomal membrane lipid compositions are distinct from one another. In the plasma membrane the predominant lipids are POPC:POPE:POPS:CHOL:SM at an approx. molar ratio of 30:15:5:30:15, whilst the late endosomal membrane contains POPC:POPE:BMP:CHOL:SM at approx. 45:15:20:10:5 (Figure 4.1A).(119) The remaining 5 mol% of both membranes is attributed to 'other' lipids, as a diverse array can be present in the membrane dependent on several environmental factors such as cell type and stress.(120) POPS and BMP are anionic lipids that both contain a single negative charge within their phosphate headgroup. However, due to the rising concentration of BMP as the endocytic pathway progresses, a substantial difference in abundance is present, equating to an approx. four-fold increase in negative charge in the late endosomal membrane when compared to the plasma membrane.(121, 122) With this in mind, it comes as no surprise that BMP has been more heavily implicated than POPS in viral fusion, particularly with viruses that fuse through the endocytic

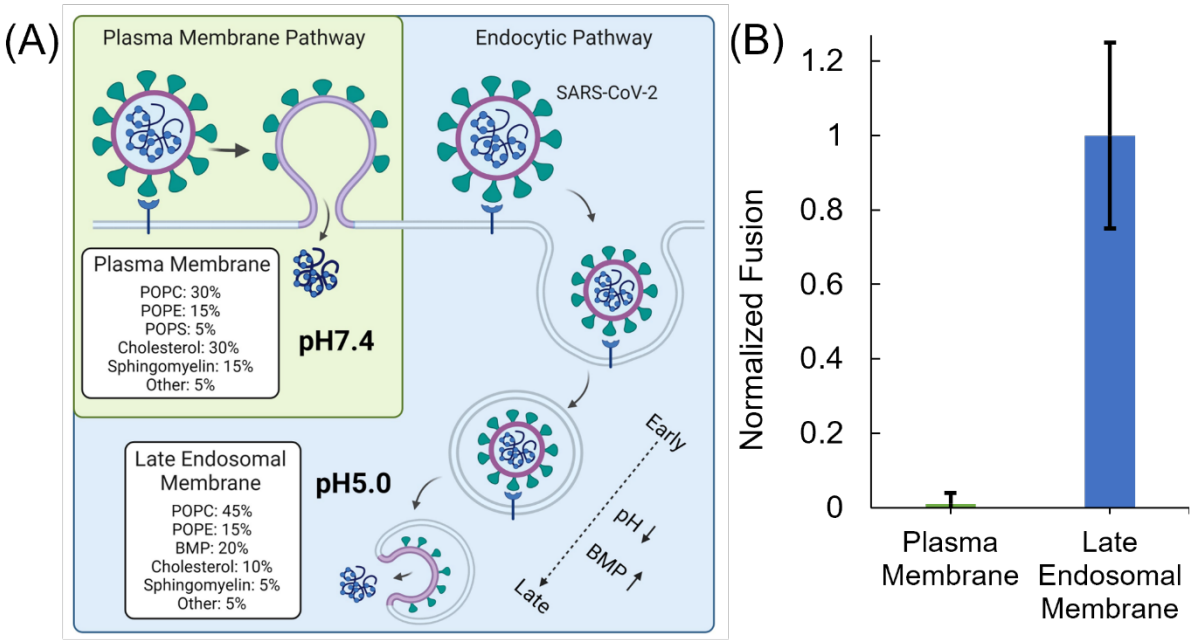

Figure 4.1 The SARS-CoV-2 fusion domain (FD) preferentially initiates fusion with a lipid composition resembling that of the late endosomal membrane. (A) SARS-CoV-2 can enter the target cell via both plasma and endosomal routes. (B) When comparing the fusogenecity of the FD utilizing the respective pH and lipid environments displayed in (A) a clear preference for the late endosomal membrane is observed ($n \leq 4$). The mol% of 'Other' lipids was replaced by POPC in the fusion assay. Created (A) through biorender.com.

pathway such as influenza,(123) Lassa Virus (LASV),(124) and Vesicular Stomatitis Virus (VSV).(88) It should be noted that the exact concentration of BMP throughout the endocytic pathway is not certain, though it is established to increase as the pathway progresses, reaching a high of 15-20% in the late endosomal membrane.(119, 121) Another notable difference in the two membrane's lipid compositions is the cholesterol and sphingomyelin content, with both concentrations higher in the plasma membrane.(119) These two lipids are the major constituents of lipid rafts, which are often involved with viral fusion due to receptor localization, however specific impacts on viral FD's have been noted previously.(49, 104, 119, 125, 126) Following extensive research into the proteins involved in SARS-CoV-2 membrane fusion, it is long overdue that the impact of membrane lipids is further investigated to fully understand the fusion process.

Here we found that the SARS-CoV-2 FD more readily elicits fusion with a lipid composition resembling that of the endosomal membrane when compared to the plasma membrane. Moreover, an anionic lipid dependence for FD mediated fusion was discovered, with a significant preference for the endosomal specific lipid BMP. To greater understand the molecular intricacies associated with this preference we conducted a biophysical analysis of how the anionic lipids properties may impact fusion. Despite slight variation in binding affinity and secondary structure of the FD amongst different anionic lipids, BMP was discovered to negatively impact lipid packing when compared to POPS. This effect on lipid packing appears to be the main deviation between the two lipids, and due to a negative relationship with the fusogenicity of the FD, could account for the preference for BMP witnessed. Together, the data suggests that the SARS-CoV-2 FD preferentially initiates fusion with a membrane representing the late endosomal membrane, due to the properties conferred by BMP on to the lipid bilayer.

4.2 Results

4.2.1 The FD Preferentially Fuses with Membranes Containing BMP

With the effect of pH on SARS-CoV-2 FD initiated fusion already established,(96) the next environmental factor to assess was that of the target cells lipid membrane composition. Utilizing an *in vitro* FRET-based fusion assay, the fusogenic ability of the SARS-CoV-2 FD in different physiologically relevant membrane environments has been determined. When presented with a membrane mimicking the plasma or endosomal membrane, we observed a clear and substantial preference for the latter. (Figure 4.1B) Even when the POPS content of the plasma membrane mimic is increased to 20 mol% and the pH dropped to 5, creating the same net anionic charge and pH environment as seen in the late endosomal membrane, the amount of fusion witnessed is still

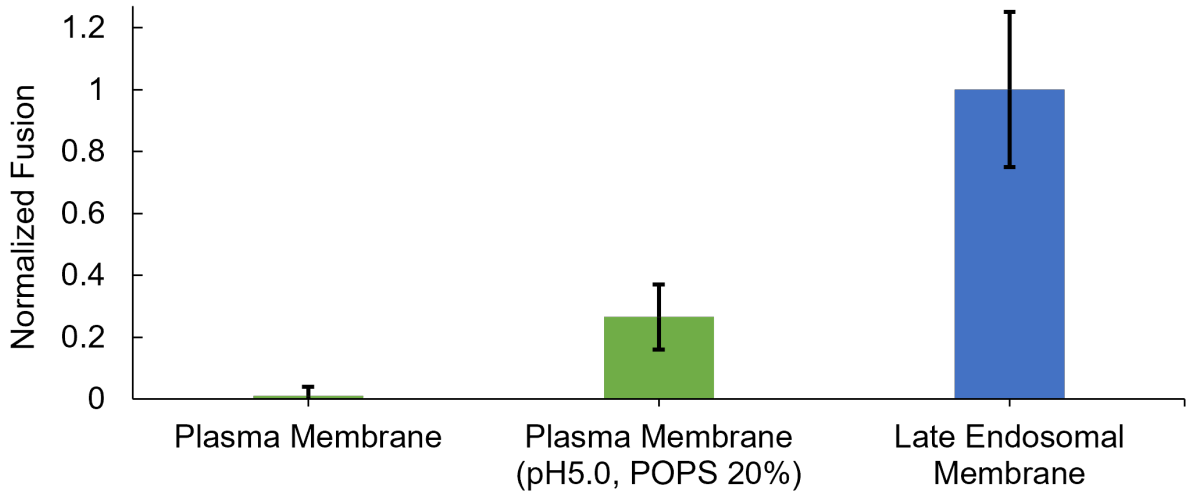

Figure 4.2 Even when the anionic lipid component of the plasma membrane and environmental pH matches that of the late endosomal membrane, the amount of fusion elicited by the SARS-CoV-2 fusion domain (FD) is still approx. 4-fold lower.

~4-fold decreased in comparison. (Figure 4.2) To better understand this preference, a bottom-up approach was employed to assess whether individual lipids may contribute towards fusion. Interestingly, a significant anionic lipid dependence was observed for both POPS (plasma membrane) and BMP (endosomal membrane), with increasing either lipid resulting in a greater amount of fusion observed (Figure 4.3). The role of the negative charge as the primary reason for this positive relationship is corroborated further by the same relationship witnessed with the primarily prokaryotic anionic lipid POPG (Figure 4.3A), as well as two uncharged lipids, POPE and CHOL, having a negative rather than positive impact on fusion (Figure 4.4). Alongside this general and positive relationship with negatively charged lipids, a specificity for BMP is also apparent through the use of a bias plot that aids in identifying bias towards a given effector molecule (Figure 4.3B).(127) The graph is generated by plotting the normalized SARS-CoV-2 FD fusion witnessed in BMP membranes against that in POPS, with anionic lipid concentration increasing from left to right. Bias is present if the trajectory does not align with the inserted dashed line and thus, a preference for the FD to initiate fusion with membranes containing BMP over

POPS exists. The concentrations of anionic lipids utilized here approximately mimic the increasing concentrations seen throughout the endocytic pathway; intriguingly, the bias plot highlights how the specificity for BMP increases at higher anionic lipid concentrations such as those seen at the late endosomal membrane.

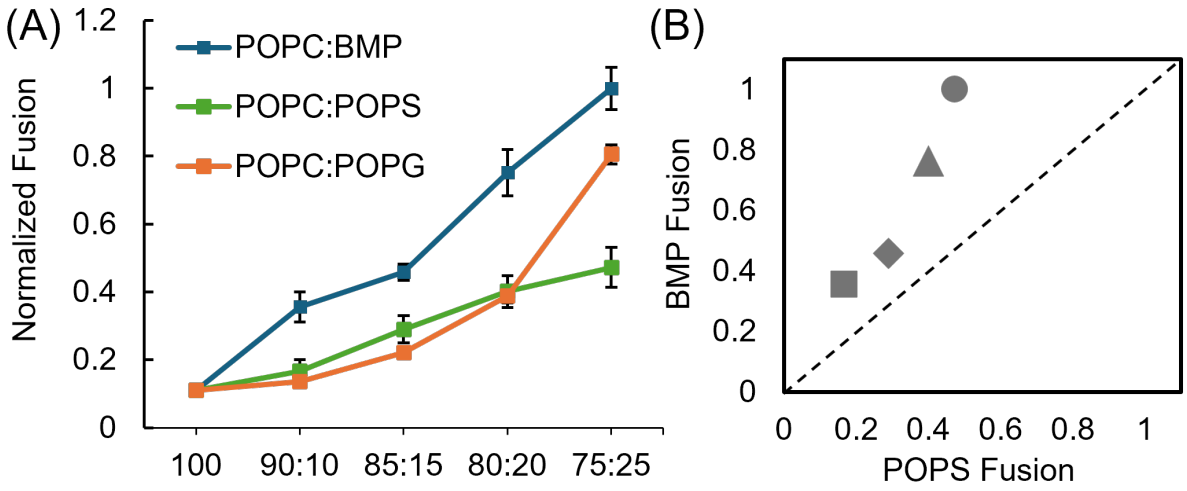

Figure 4.3 BMP elicits more fusion domain (FD) mediated fusion than other anionic lipids. (A) When increasing the concentration of anionic lipid present, we observe a positive relationship with fusion for both POPG, POPS and BMP, yet more fusion is witnessed for the latter. (B) The bias plot highlights how a specificity for BMP exists over POPS and increases alongside anionic lipid concentration. The Y axis represents normalized fusion for BMP, whilst the X axis for POPS. The four data points plotted correspond to the four different anionic lipid concentrations displayed in (A), which are increasing from left to right. ■=10 mol%, ♦=15 mol%, ▲=20 mol%, ●=25 mol%. n ≥ 17.

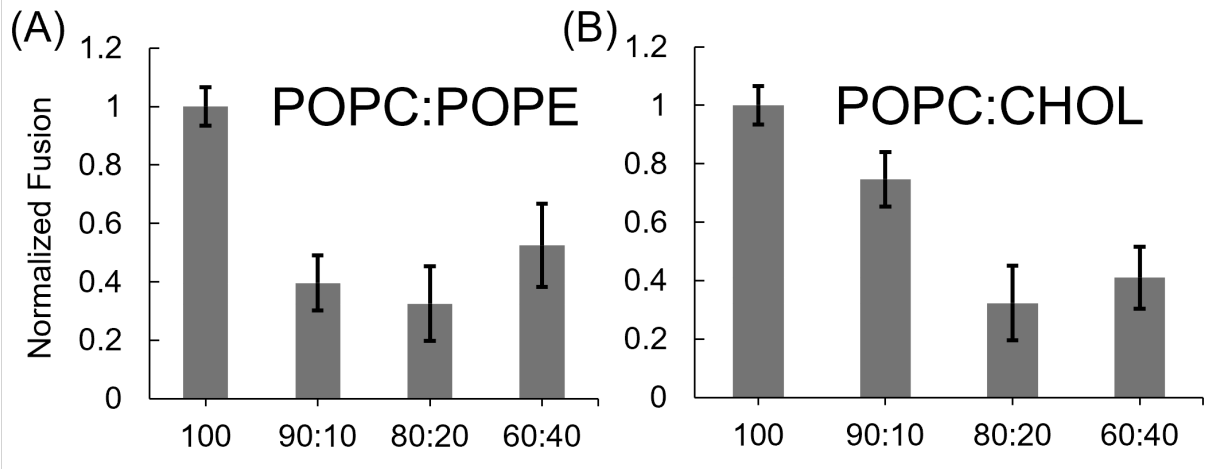

Figure 4.4 Increasing concentrations of (A) POPE and (B) Cholesterol had a negative impact on SARS-CoV-2 fusion domain (FD) mediated fusion.

4.2.2 Probing the Interactions of Different Anionic Lipids with the FD

To further understand how different anionic lipids impact the interaction of the FD with the membrane, we quantified the strength of this interaction via isothermal titration calorimetry (ITC) (Figure 4.5). A Kd of 1.2 ± 0.1μM, 3.0 ± 0.1μM, and 2.6 ± 0.1μM was found for the FD associating

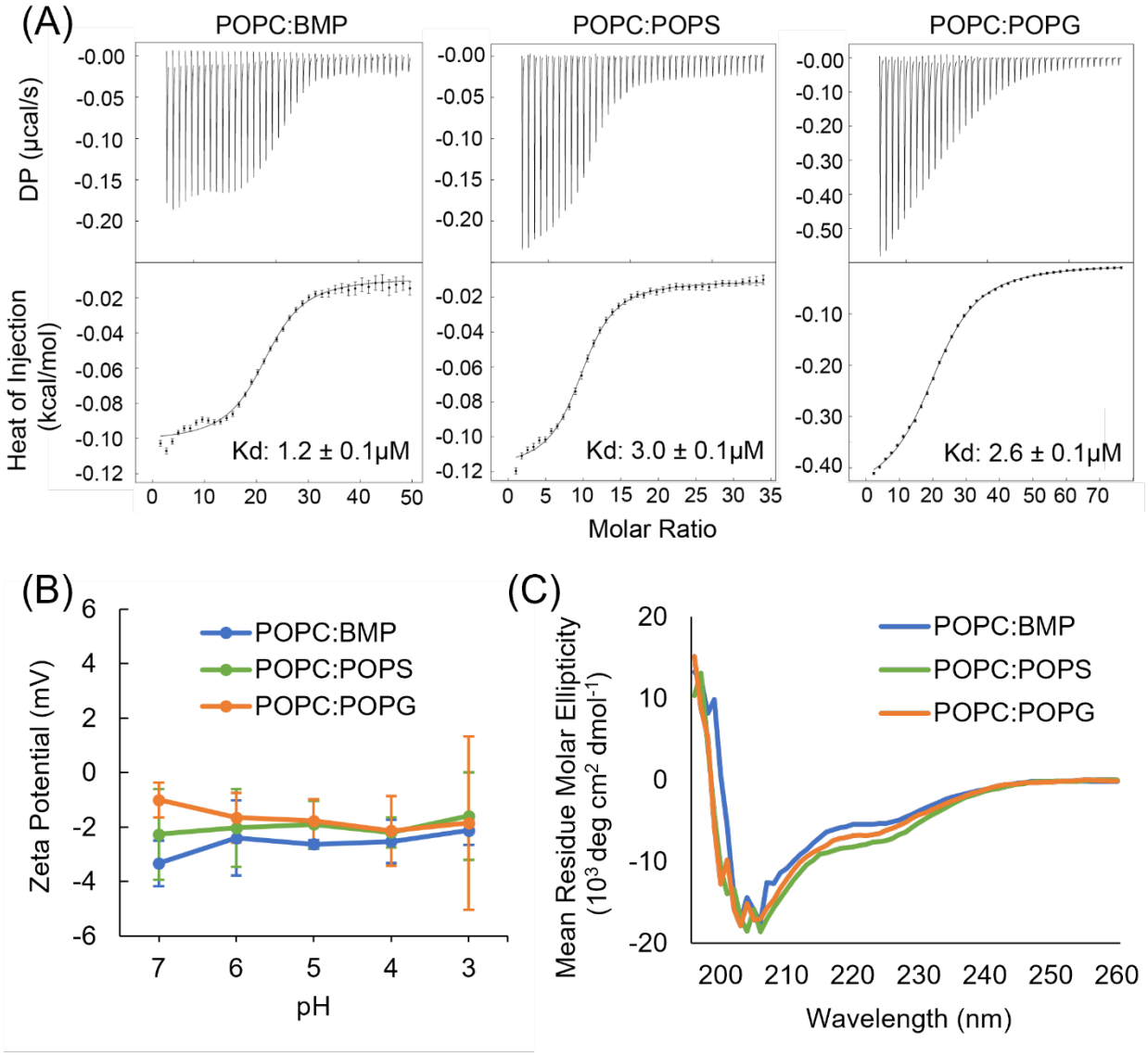

Figure 4.5 The initial interaction between the fusion domain (FD) and lipid vesicles containing BMP, POPS or POPG. (A) Regardless of which anionic lipid is present, the dissociation constants (Kd) found through ITC are similar. Vesicles contained 50 mol% anionic lipid with a total concentration of 10mM, and a protein concentration ~20 μM was utilized in all ITC experiments. (B) Zeta potential measurements displayed no significant change in charge from pH7 to pH3 with 20 mol% anionic lipid present. (C) No major perturbations in global secondary structure were observable via CD based on the presence of different anionic lipids at 2 0mol%.

with membranes containing BMP, POPS, and POPG respectively. It should also be noted that the presence of anionic lipids was found to be critical for the interaction between the FD and lipid vesicles, with a ~1000-fold increase in Kd present (4.2 ± 0.7 mM) with only POPC in the membrane (Figure 4.6). Despite being negatively charged at neutral pH, it is unknown whether this charge is maintained for all three anionic lipids under more acidic conditions, such as those found in the endocytic pathway. To experimentally probe the relationship between anionic lipid charge and pH we measured the zeta potential of LUVs from pH7.0 to pH3.0. Put simply, a zeta potential is the measurement of how the electric potential found to surround a nanoparticle due to the attraction of counterions, changes as the particle moves in solution.(128) By measuring the zeta potentials of LUVs containing BMP, POPS and POPG, we see that no significant alteration in charge occurs for any of the anionic lipids from pH7.0 through to pH3.0 (Figure 4.5B). Thus, environmental pH does not appear to significantly impact the net negative charge of the constructed lipid vesicles. We also assessed the structural impact of differing anionic lipids on the

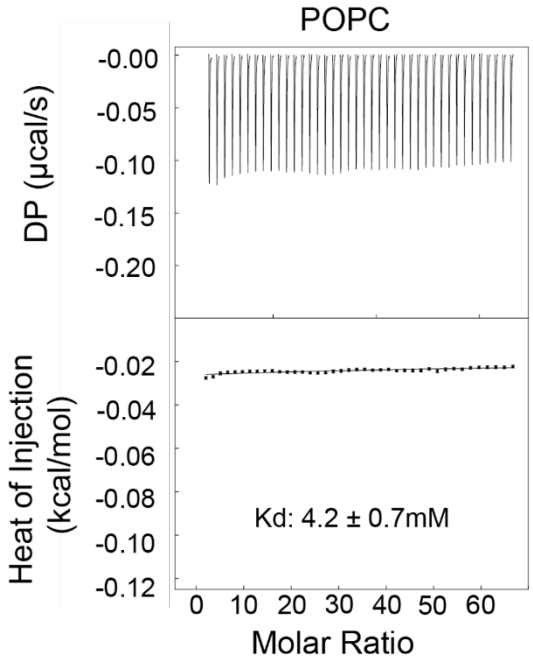

Figure 4.6 SARS-CoV-2 fusion domain (FD) binding to 100% POPC vesicles with a dissociation constant of 4.2 ± 0.7 mM.

embedded FD through circular dichroism (CD). An α-helical conformation for the FD was found in all three lipid environments (Figure 4.5C). This suggests that the presence of different lipid headgroups has a negligible impact on the global secondary structure of the FD. Taken together, this data supports a conclusion where a negative charge is essential for the interaction between the SARS-CoV-2 FD and lipid vesicles, however, the different headgroups of the anionic lipids involved have minimal impact on this interaction.

4.2.3 BMP Negatively Impacts Lipid Packing

Unable to discern any substantial differences between the anionic lipid headgroups via their interaction with the FD, we next turned our focus to how POPS and BMP may affect the packing of the membrane, and in turn, fusion. C-Laurdan is a fluorescent probe that is sensitive to the local environment and has been used previously to study lipid packing.(129, 130) The probe contains two emission peaks at ~440 nm (Increased lipid packing) and ~490 nm (Decreased lipid packing), with the relative intensity changing in the presence of different degrees of lipid packing. When first comparing POPC:POPS to POPC:BMP liposomes we see that the presence of POPS results in a slightly more packed membrane, with a λmax of 478 nm and 484 nm, respectively (Figure 4.7A&B). To probe this effect further we combined the anionic lipids with DPPC and DOPC, with the hypothesis that they would impact lipid packing based primarily on chain saturation. (Figure 4.8) No change in packing was witnessed for DOPC:BMP (λmax = 484 nm), but a minor shift for DOPC:POPS (λmax = 483 nm) towards a less packed membrane was observed (Figure 4.9). Moreover, a clear difference in packing can be seen for both anionic lipids in the presence of DPPC when compared to POPC (Figure 4.7A&B). DPPC:POPS (λmax = 444 nm) and DPPC:BMP (λmax = 450 nm) displayed a marked shift, with a large decrease in λmax indicative of an increase

in lipid packing. (Figure 4.7A&B) Intriguingly, DPPC:BMP still displayed less lipid packing than DPPC:POPS, the same trend as was witnessed with POPC (Figure 4.7B). Furthermore, a negative relationship was found between increased lipid packing and the fusogenicity of the SARS-CoV-2 FD (Figure 4.7C). Comparing DPPC:POPS to POPC:POPS and DPPC:BMP to POPC:BMP, an increase in fusion was witnessed with both POPC compositions. In summary, a negative relationship between lipid packing and SARS-CoV-2 FD fusogenecity has been found, with BMP displaying a negative impact on lipid packing in comparison to POPS.

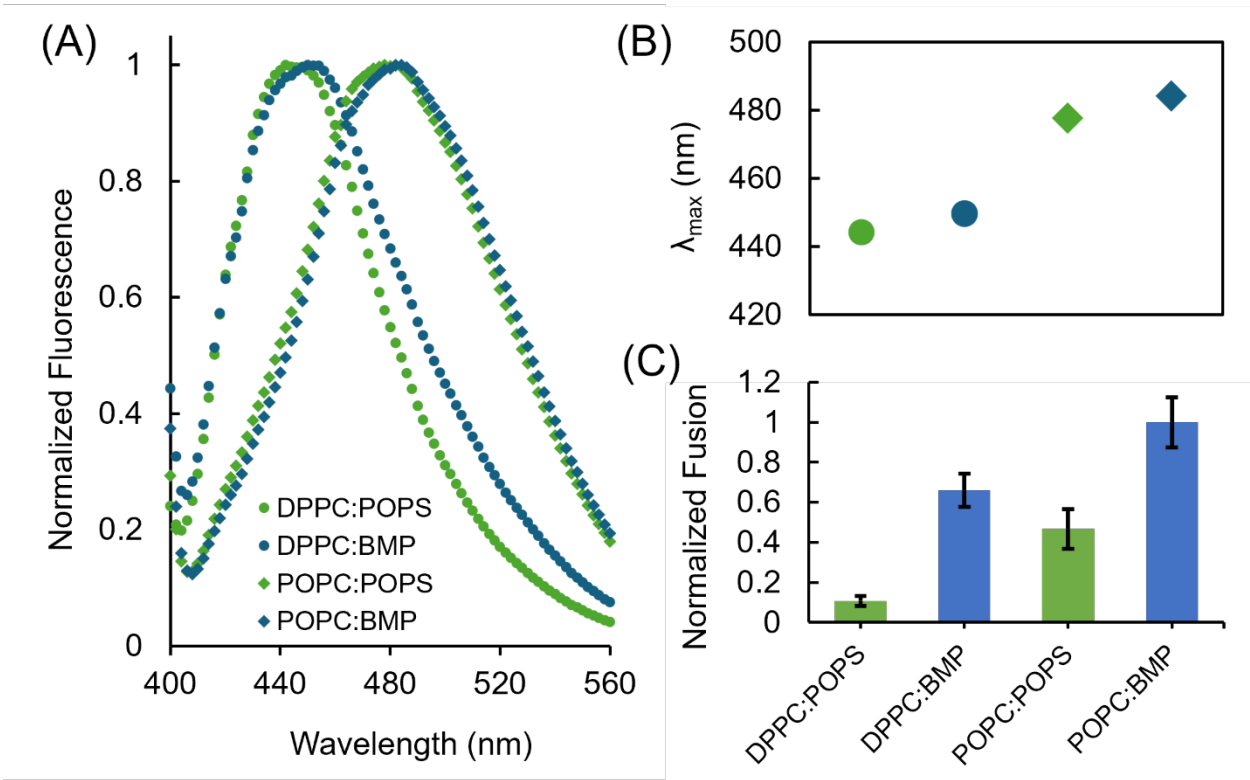

Figure 4.7 BMP decreases lipid packing, which has a positive impact on fusion. (A) DPPC:POPS and DPPC:BMP show increased lipid packing in comparison to POPC:POPS and POPC:BMP with a molar ratio of 80:20 present. (B) Plotting the lambda max displays a trend where BMP elicits a negative effect on lipid packing in comparison to POPS. (C) When comparing like headgroups, the liposomes containing lower levels of lipid packing allowed the fusion domain (FD) to elicit increased amounts of fusion, with a molar ratio of 90:10 and n ≥ 12.

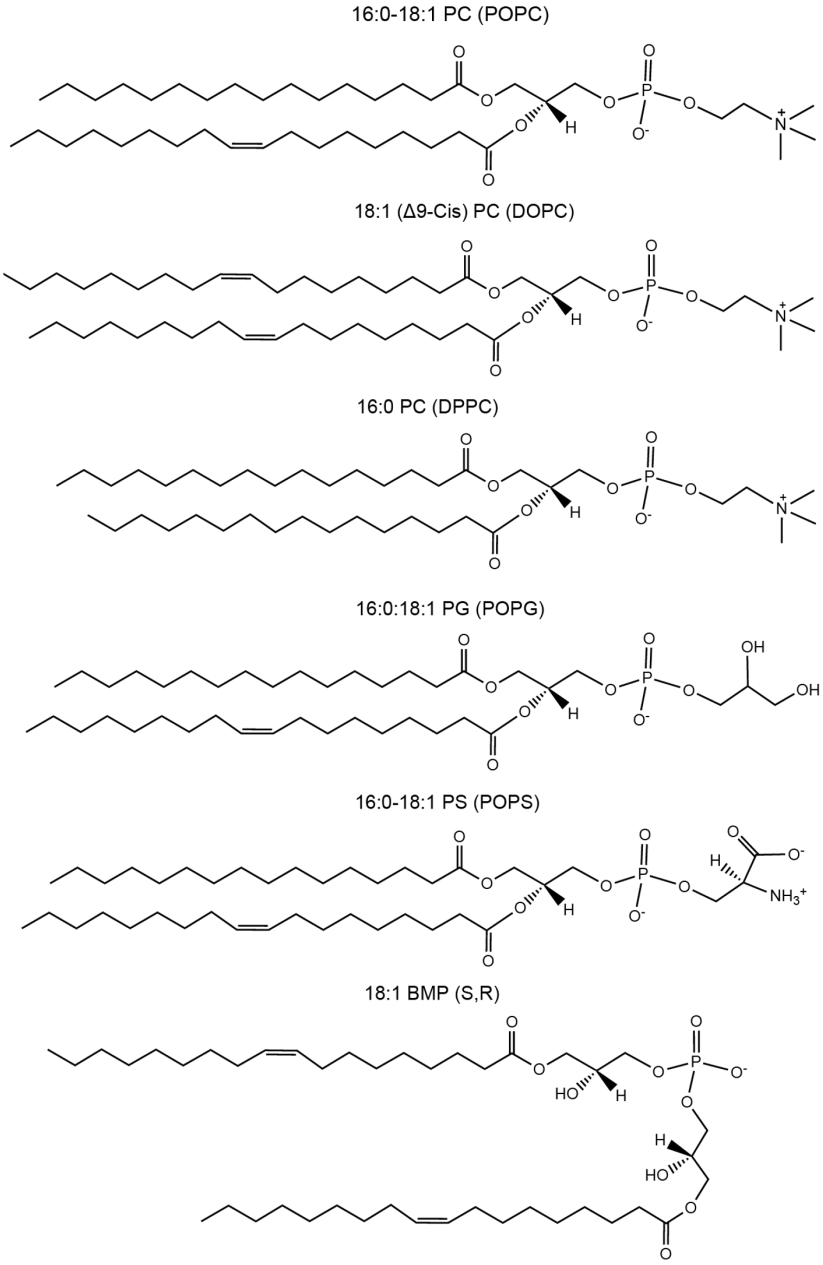

Figure 4.8 Chemical structures of POPC, DOPC, DPPC, POPG, POPS and BMP lipids.

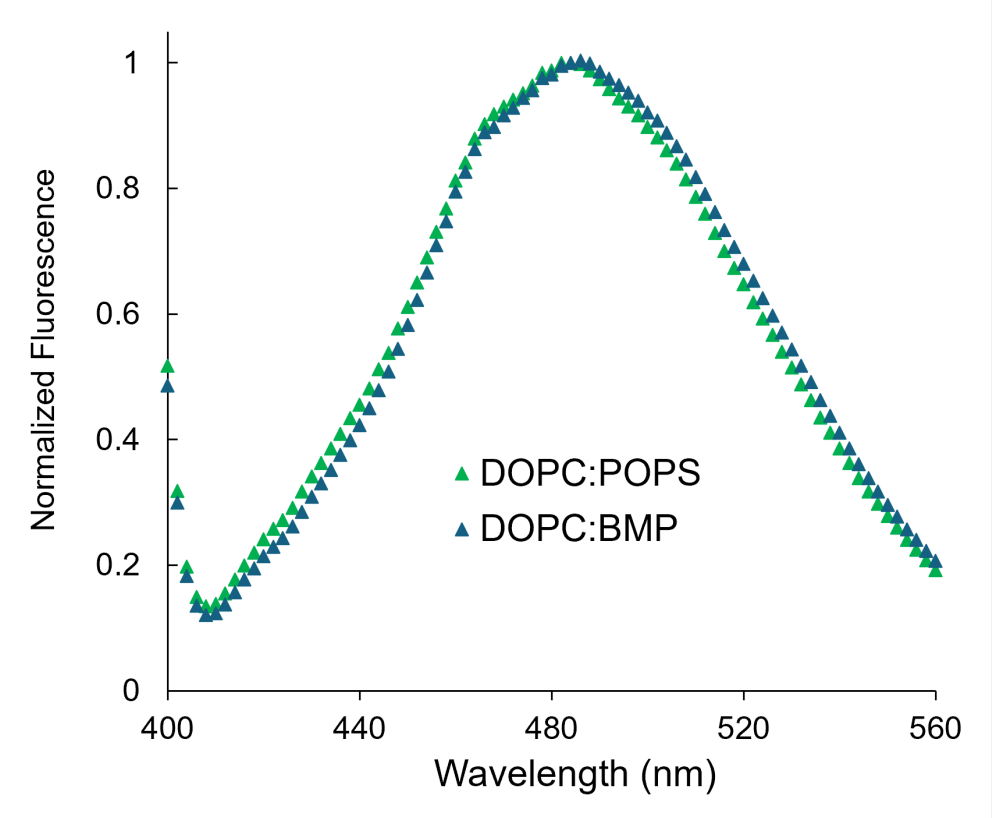

Figure 4.9 Using C-Laurdan, the lipid packing of DOPC:POPS (λmax = 483 nm) and DOPC:BMP (λmax = 484 nm) suggests a similar, less packed membrane.

4.2.4 Limited Synergy Amongst Other Lipids Relevant to Viral Fusion

After assessing the anionic lipids in unison, we finally wanted to investigate whether they displayed any synergistic effects alongside other physiologically relevant lipids. Due to their relatively high abundance in both the plasma and endosomal membranes as well as their previous implications in viral fusion, we focused on POPE and Cholesterol. Intriguingly, the two lipids showed some synergy together (Figure 4.10A), although at relatively low levels of fusion in comparison to the anionic lipids. When individually incorporating POPE or CHOL into LUVs alongside BMP no significant impact on fusion is observed (Figure 4.10B). On the other hand, a significant decrease in fusion is witnessed when POPS is incorporated alongside POPE or CHOL

(Figure 4.10C). Hence, whilst neither anionic lipid displays positive synergistic properties alongside POPE or CHOL, both lipids have a detrimental impact on POPS mediated fusion.

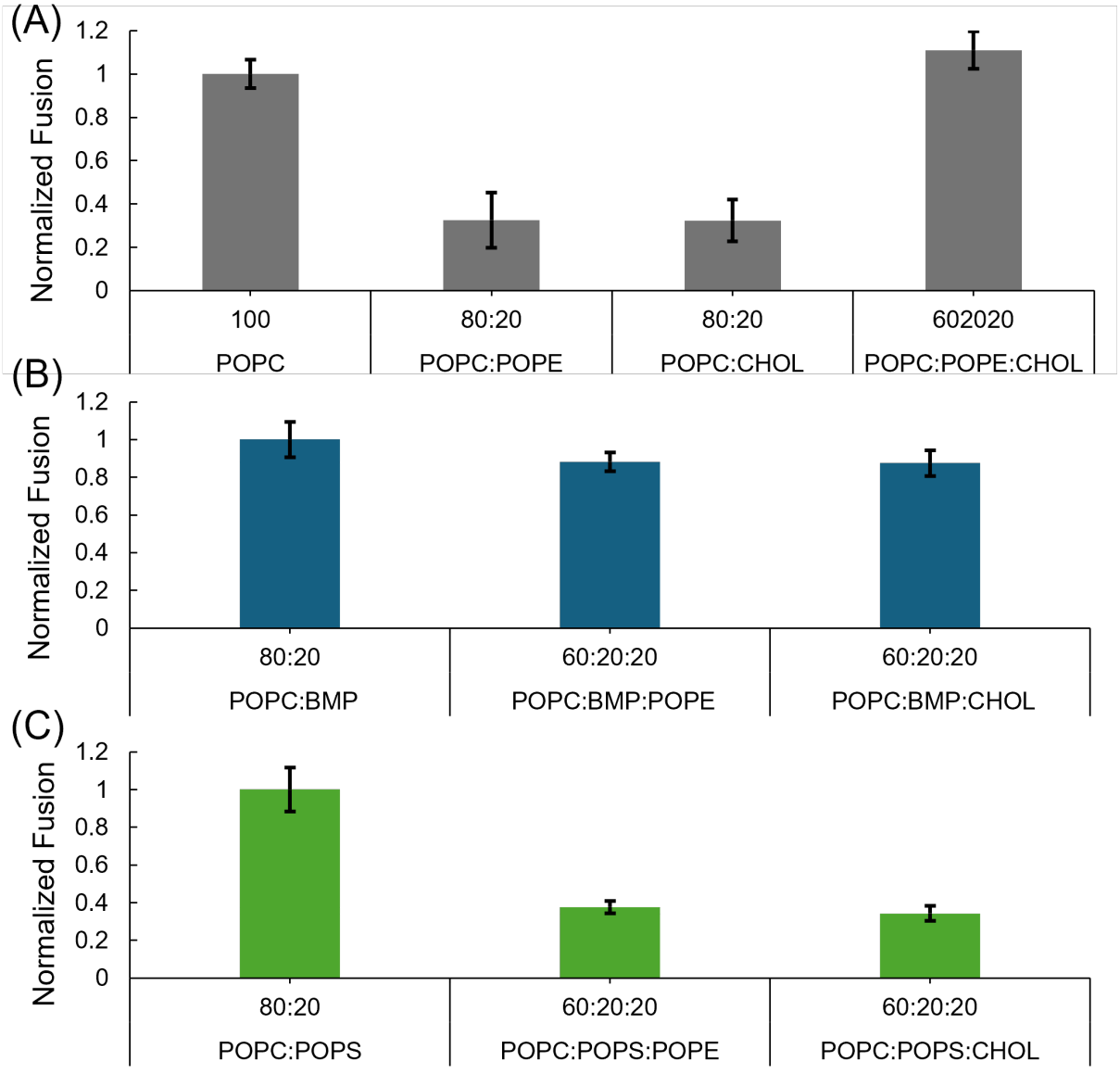

Figure 4.10 The impact of POPE and Cholesterol. (A) POPE and cholesterol show a synergistic relationship in SARS-CoV-2 fusion domain (FD) mediated fusion, but at relatively low levels in comparison to anionic lipids. (B) No difference in fusion is observed with vesicles containing BMP when in the presence of POPE or Cholesterol. (C) FD mediated fusion is significantly decreased when vesicles containing POPS are accompanied with POPE or Cholesterol. n ≤ 10.

4.3 Discussion

SARS-CoV-2 is an incredibly infectious virus, in part due to its ability to enter the target cell through either the plasma or endosomal membrane.(57, 99, 131) Whilst most viruses can only access one pathway, the ability to utilize both increases the likelihood of infection. One key difference between the two pathways is environmental pH: the plasma membrane exists at a neutral pH, whilst the pH within the endocytic pathway gradually decreases throughout. Both *in vitro(96)* and *in vivo(131)* SARS-CoV-2 has been shown to preferentially initiate fusion at a pH akin to that witnessed in the endocytic pathway. However, neither study considered the lipid composition of the target cell membrane; another critical environmental factor that differs between the plasma and endosomal routes of fusion. It should be noted that previous work has also found Ca^{2+} to be involved within the initiation of membrane fusion. However, the exact role of Ca^{2+} is yet to be elucidated within this process and we have thus decided to omit the divalent ion from all of our experiments allowing us to focus on the impact of lipid composition. Here we have shown that the FD of SARS-CoV-2 displays a preference to initiate fusion with the lipid composition found in the late endosomal membrane (Figure 4.1B), regardless of the environmental pH (Figure 4.2). Hence, alongside pH, lipid composition could be another major determining factor that dictates the likelihood of success for SARS-CoV-2 fusion.

Within the human body, the sum of POPS (~5 mol%) in the plasma membrane is generally thought to be significantly lower than that of BMP (~20 mol%) in the late endosomal membrane.(119, 121) Additionally, POPS is thought to be more commonly found on the inner leaflet of the plasma membrane rather than the outer leaflet where it would be exposed to the external environment.(132, 133) So, not only is POPS at a lower concentration than BMP in their respective membranes, but it is also less accessible to interact with the FD. Such reasoning could

provide a physiological explanation as to why a preference for the SARS-CoV-2 FD to initiate fusion with membranes containing BMP over POPS exists, as the former is the predominant anionic lipid that comes into contact with the viral glycoprotein (Figure 4.3). Furthermore, several other viruses which utilize the endocytic pathway (influenza, LASV, and VSV) have all displayed a positive fusogenic relationships with BMP. (88, 123, 124) It is possible that a trend could be emerging, where enveloped viruses which travel through the endocytic pathway take advantage of the increasing concentrations of BMP as a potential trigger to initiate their fusion processes.

We sought to clarify where this preference for BMP arises from by looking at how the anionic lipids impact the membrane from a biophysical perspective. It has previously been stated that BMP elicits fusion in a pH dependent manner,(121) however this was not witnessed within our experimental set-up as controls without FD displayed no fusion regardless of lipid composition (Figure 4.11). A third anionic lipid was introduced, POPG, which despite being a structural isomer and precursor of BMP retains more similarities with POPS due to headgroup stoichiometry, chain length and saturation.(123) Similarly to POPS, POPG was also not as efficient at eliciting fusion when compared to BMP (Figure 4.3A). Through the use of zeta potential measurements, the overall charge of the lipid membrane was assessed across a wide pH range. The thought process behind these measurements includes the possibility that the negative charge could be lost as the lipid headgroup becomes protonated at low pH, with the isoelectric points for each functional group not well established. Yet, all three anionic lipids displayed no significant variation between pH7 and pH3, with the negative charge maintained throughout (Figure 4.5B). Membranes containing BMP, however, were found to be slightly more negative than POPS and POPG. Interestingly, a similar trend was witnessed through ITC, where a slight increase in binding affinity was also discovered for BMP (Figure 4.5A). We hypothesize that both of these results may be due

to a greater preference for BMP to partition into the outer leaflet of LUVs, as the zeta potential measurement only detects the outer surface charge of a given nanoparticle and the FD is only thought to interact with the lipid headgroups of the membranes outer leaflet.(98, 134) Further studies involving the impact of lipid asymmetry may prove beneficial in further elucidating the preference for BMP as well as how the complexity of the membrane impacts fusion. Additionally, the same alpha helical conformation found previously (Figure 3.2) was also present in all three different lipid environments (Figure 4.5B). Hence, the FD appears to directly interact with anionic

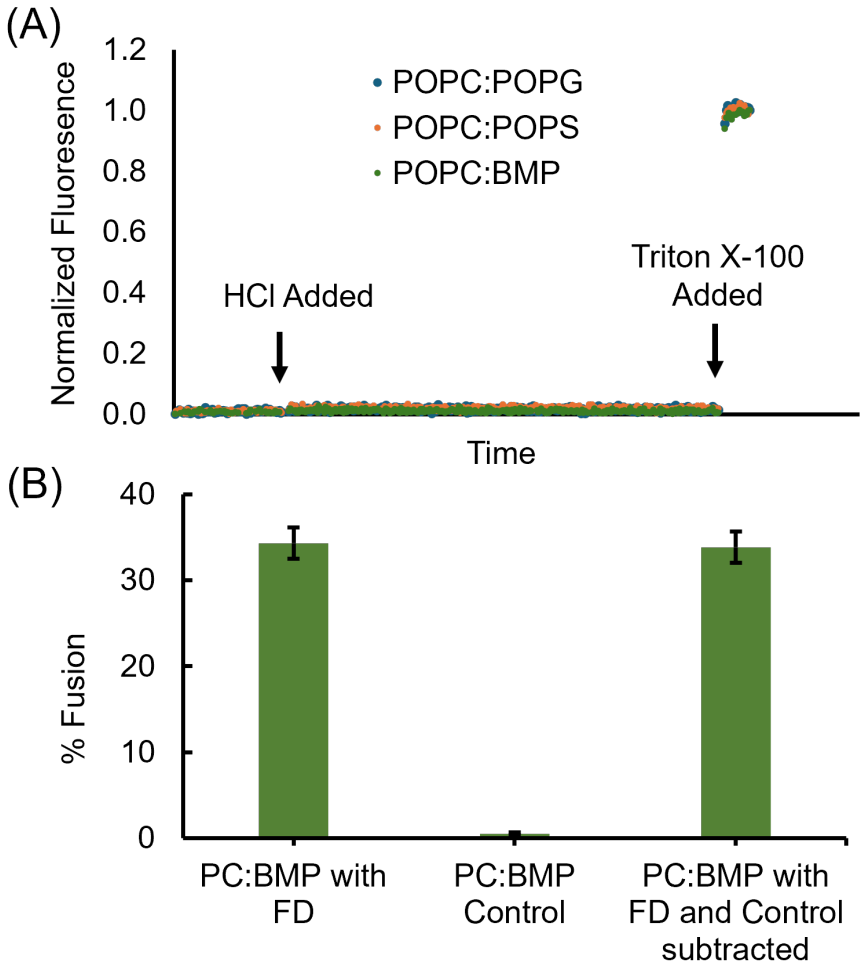

Figure 4.11 POPC:BMP LUVs do not elicit fusion independent of the SARS-CoV-2 Fusion Domain (FD) at pH5.0. (A) Raw data displaying no increase in fluorescence upon addition of HCl with vesicles containing BMP showing the same inert behavior as those containing POPS or POPG. (B) Example of data processing workflow. The % fusion is calculated for the FD and the controls independently (N=4) for each condition, allowing the standard deviation to be gathered for both. The control is then subtracted from the FD fusion and the errors propagated. All lipid compositions shown are 80:20 mol%.

lipids in order to initiate fusion, which explains the positive relationship witnessed. With four positively charged residues within the FD (K825, K835, R847 and K854), an increase in anionic lipid present in the membrane could facilitate more ionic interactions regardless of lipid structure, which based upon the lack of binding for a membrane containing 100 mol% POPC, appears critical. However, such interactions do not explain the preference for the FD to fuse at membranes containing BMP over other anionic lipids.

Alongside anionic lipids, both POPE and CHOL have been implicated in viral fusion previously with the lipids found to impact membrane curvature and fluidity, respectively. (86, 105, 135, 136) Unexpectedly, neither POPE nor CHOL displayed any positive impact on the ability of the FD to elicit fusion (Figure 4.4). This was a particular surprise as *in vivo* based studies have shown cholesterol to positively influence SARS-CoV-2 fusion.(137, 138) The discrepancy in results could be due to a greater complexity of membrane lipid compositions involved in such studies, or the more likely scenario, by where the formation of lipid rafts results in the localization of the ACE2 receptor on the cell surface. Hence, rather than direct protein-lipid interactions between the SARS-CoV-2 spike protein and the target cell membrane, it is in fact the membranes impact on the ACE2 receptor that may promote fusion in such cases.

The two main characteristics of lipids are a hydrophilic headgroup and two hydrophobic fatty acid tails, both of which play a key role in modulating the biophysical make-up of the membrane.(119, 139, 140) Headgroups often dictate interactions with the soluble environment and are distinguishable by traits such as shape, size and charge. Alternatively, aliphatic chains are more directly involved within the hydrophobic core of the membrane, relying on varying degrees of saturation and length for differentiation. . Two key structural differences exist between BMP and POPS which may help to explain the significant differences within fusion, one in the lipid

headgroup and another in the fatty acid tails (Figure 4.8). Firstly, BMP is a structural isomer of phosphatidyl glycerol (PG) that contains a rare sn-1;sn-1' configuration that is not found in any other mammalian glycerophospholipids.(141) This unique configuration has been found to participate in less hydrogen bonding than the traditional sn-3:sn-1' orientation that is found in both POPS and POPG, thought to be due to the orientation of available hydrogen bond donors.(142-144) Secondly, BMP most commonly contains two unsaturated tails (18:1) whilst POPS and POPG have one saturated and one unsaturated chain (16:0-18:1). Saturated fatty acids allow lipids to pack tightly alongside one another, however double bonds introduce kinks to the acyl chains and prevent efficient lipid packing. A simplified example is lipids with saturated acyl chains being thought of as having an overall rectangular shape, so the headgroup and acyl chains are a similar width. Alternatively, unsaturated lipids have a more triangle like shape with their lipid tails consuming a wider area than there headgroup counterpart (Figure 4.8). Through MD simulations, it has been suggested that the SARS-CoV-2 FD can more readily insert itself into endosomal membranes as opposed to the plasma membrane due to a decrease in lipid packing.(113) Our experimental results further indicate that BMP has a subtle, negative impact on lipid packing when compared to POPS (Figure 4.7). Put another way, BMP increases the fluidity of the membrane, a finding that has also been corroborated previously,(145) and appears to directly correlate with an increase in fusion elicited by the SARS-CoV-2 FD (Figure 4.7C). Increasing polyunsaturated lipids in the membrane has been shown to favor fusion through decreasing the overall activation energy required previously.(146) Thus, greater membrane fluidity via the presence of BMP as a result of less hydrogen bonding and increased unsaturation, could aid the FD in reducing the energetic barrier necessary to elicit fusion.

4.4 Conclusion

The SARS-CoV-2 FD contains a preference to elicit fusion in membranes containing the endosomal lipid BMP. Alongside its negative charge, which is maintained across a wide pH range encompassing the endosomal pathway, BMP is thought to have a subtle, negative impact lipid packing, likely due to its unique structural characteristics. Together, these properties are deemed to be critical for the positive and specific impact on membrane fusion witnessed when BMP is compared to POPS. The negative charge shared by both lipids is critical for the initial interaction between the FD and the lipid membrane, whilst the minor decrease in lipid packing imparted by BMP results in increased FD mediated fusion. Intriguingly, whilst fusion with membranes containing POPS was negatively impacted by the presence of POPE or cholesterol, BMP showed neither positive nor negative synergistic effects alongside other lipids (Figure 4.10). Further work should prioritize understanding the interplay of such lipids within the membrane and how that may directly impact SARS-CoV-2 mediated fusion in more complex systems incorporating features such as lipid asymmetry and phase separation. This knowledge could be applied to other viruses, aiding in the understanding of viral tropism through different membrane lipid compositions.

4.5 Materials and Methods

4.5.1 Materials

All lipids were acquired from Avanti Polar Lipids in chloroform and their complete chemical names are provided below. POPC: 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine, POPG: 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol), POPS: 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine, POPE: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine,

SM: Sphingomyelin, BMP: bis(monooleoylglycero)phosphate (S,R Isomer), DPPC: 1,2-dipalmitoyl-sn-glycero-3-phosphocholine, DOPC: 1,2-dioleoyl-sn-glycero-3-phosphocholine, LissRhod-PE: 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl) and NBD-PE: 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(7-nitro-2-1,3-benzoxadiazol-4-yl). Cholesterol (Chol) was purchased from Anatrace in powder form, then later dissolved with chloroform and aliquoted when needed. C-Laurdan was acquired from Tocris Bioscience, dissolved in DMSO and stored at -80ºC.

4.5.2 Preparation of Small and Large Unilamellar Vesicles

Vesicles were assembled from lipid stock solutions in chloroform, with specified amounts added to glass test tubes using Hamilton syringes. The chloroform was then removed by applying a stream of nitrogen whilst gently vortexing the sample to create a lipid film, before being put into a desiccator under vacuum overnight to remove any residual solvent. For large unilamellar vesicles (LUVs), the lipid film was resuspended in 10 mM HEPES/MES/Sodium acetate (HMA), 100 mM NaCl, pH7.4 buffer through extensive vortexing. When using DPPC, a 5-minute incubation at 55 ºC prior to vortexing was necessary, due to the lipids transition temperature. The lipid suspensions were then subjected to 10 freeze-thaw cycles between liquid nitrogen and water bath. Liposomes were extruded using a liposofast extrusion kit (Avestin) a total of 21 times through 2 polycarbonate membranes with a 100 nm pore size. For small unilamellar vesicles (SUVs), the lipid film was resuspended in the desired buffer and sonicated for 15 mins at 10% duty cycle (1 s on/1 s off) with the sample sat in ice water, using a Branson ultrasonicator microtip. After sonication the SUVs were centrifuged at 20,000 xG for 10 mins to remove residual metal particulates. All vesicles were either used immediately or stored for a maximum of 72 hrs at 4 ºC prior to use.

4.5.3 Expression and Purification

The SARS-CoV-2 FD is made up of 40 amino acids (^{816}SFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKF^{855}) and designed with an N-terminal 6x His-tag, followed by a SUMO tag to aid with solubility and expression. The expression and purification of the SARS-CoV-2 FD has been described in detail previously.(47) Briefly, 8 M urea is used to solubilize the cell pellet followed by Ni-NTA affinity chromatography and cleavage via SUMO protease in dialysis. The cleaved FD is then isolated through the use of Ni-NTA affinity chromatography once again and dialyzed to remove reducing agents, ensuring correct formation of the disulfide bond. Following dialysis, the sample is purified further via a Superdex30 size exclusion chromatography (SEC) column with 10 mM HMA 100 mM NaCl pH7.4 as the mobile phase. Resulting fractions are pooled, concentration determined via A_{280}, and stored at 4 ºC.

4.5.4 Circular Dichroism

All CD data was acquired via the Jasco J810 Spectro-Polarimeter using a quartz cuvette with a 2 mm path length. Each experiment was conducted at room temperature (~22 °C) in 1 mM HMA, 10 mM NaCl, pH7.4 with a protein concentration of ~20 μM and a liposome concentration of 800 μM. Data was collected from 260 nm to 190 nm with a step size of 1 nm at 20 nm/min and averaged over three accumulations. Baselines were acquired without any protein present and subtracted from all data.

4.5.5 Lipid Mixing Assay

LUVs composed of specified lipids were mixed with labeled LUVs containing the same lipid composition as well as 1 mol% of the fluorescent labels: LissRhod-PE and NBD-PE. Experiments were conducted at a ratio of 9:1 unlabeled:labeled in Corning Costar black walled, clear bottom

96 well plates with excitation and emission wavelengths at 460 nm and 538 nm respectively, with a 530 nm cut-off. Fluorescence was recorded using a SpectraMax M5 microplate at room temperature (~22 ˚C). Percent fusion was calculated as $\frac{(I_F - I_B)}{(I_{100} - I_B)}$, where I_B is the initial background fluorescence, I_F is fluorescence intensity measured after decreasing the pH, and I_{100} is the 100% fluorescence intensity value gathered after complete vesicle rupture following the addition of 1% Triton X-100. All experiments were conducted at pH5.0 unless stated otherwise and contained a peptide/lipid ratio of 0.05 (5 μM and 100 μM). Controls containing no protein were also ran alongside all conditions, subtracted from the final values and errors propagated from the SEM.

4.5.6 Zeta Potential Measurements

Measurements were acquired on a Horiba Scientific nanoPARTICA Nanoparticle Analyzer SZ-100V2 in a glass capillary cell, with a nominal volume of ~400 μL. Each sample was measured in triplicate and the SD is provided. For all pH measurements, a known amount of 1 M HCl was added to LUVs in 10 mM HMA 100 mM NaCl and final values verified via pH meter.

4.5.7 Isothermal Titration Calorimetry

The FD sample was prepared through dialysis against 4 L of 10 mM Sodium Acetate, 100 mM NaCl at pH5.0 for 2hrs at 4ºC. After this time, the dialysis buffer was used to resuspend the lipid film prior to forming SUVs, with both protein and SUV samples centrifuged at 20,000 xG for 10 mins and degassed prior to each experiment. All ITC measurements were taken using a Malvern VP-ITC microcalorimeter with the following experimental parameters in place: initial delay: 2400 s, 41 injections (1x2 μL and 40x6 μL), spacing: 300 s, duration: 14.4 s, Stir Speed: 270 rpm, ref power: 2 μcal/sec and temperature: 22 ºC. The concentration of the protein sample was taken following dialysis via A_{280}, and the SUV concentration was known from drying down specific

volumes of lipid stocks. All processing was conducted through NITPIC and SEDPHAT, with final figures generated through GUSSI.(147, 148) Errors for the simulated fit were generated via Montecarlo analysis within SEDPHAT, with 10,000 iterations and a confidence interval of 0.68.

4.5.8 C-Laurdan Experiments

All experiments were performed with a C-Laurdan concentration of 500 nm, and 100 μM of LUVs of a known lipid composition for a final ratio of 1:200. Each composition was run in at least triplicate using a quartz cuvette with a 1 cm path length and fluorescence recorded using a SpectraMax M5 microplate at room temperature (~22 $^{\circ}$C). Excitation wavelength was set to 385 nm with the emission spectrum captured from 400-600 nm in 2 nm steps. Unless stated otherwise, the lipid composition utilized in this assay was 80:20 mol% (POPC: Anionic lipid).

Chapter 5: Basic Amino Acids are Key for the Molecular Mechanism of Fusion

This work has been submitted for publication.

Birtles, D., Guiyab, L., W. Abbas, and J. Lee. 2024. Positive residues of the SARS-CoV-2 fusion domain are key contributors to the initiation of membrane fusion

D.B. and J.L. conceptualized research; D.B., L.G., W.A. performed research; D.B., and J.L. analyzed data and all authors prepared the manuscript.

5.1 Introduction

SARS-CoV-2 is understood to be able to enter host cells via membrane fusion at either the plasma membrane or the endosomal membrane, a unique trait that is thought to aid in the success of viral infection.(99, 109) One key contributing factor to this ability is the presence of the proteolytic enzymes transmembrane serine protease 2 (TMPRSS2) in the plasma membrane and Cathepsin L in the endocytic pathway, both of which can cleave S2' to create a fusion-ready glycoprotein on the viral surface. Furthermore, a preference for the SARS-CoV-2 FD to fuse under environmental conditions similar to that witnessed in the endocytic pathway has been established. Both *in vitro* and *in vivo* work exhibited an increase in fusogenecity at a pH below neutral (≤pH7), (96, 131) as well as a positive and specific relationship with the endosomal resident, anionic lipid bis(monoacylglycero)phosphate (BMP).(118) Specific pH environment, enzyme localization, and membrane lipid composition, are all common triggers associated with the initiation of the fusion process for several different viruses.(100) A greater understanding of the SARS-CoV-2 FD's molecular mechanism regarding the initiation of fusion via its unique FD is necessary to fully understand how such triggers lead to a successful fusion event.

The coronavirus FD has previously been shown to contain two distinct regions: a fusion peptide (FP) and a fusion loop (FL), with both capable of fusing independently but significantly

more effective in synergy.(47, 54) This concept of two synergistic domains within a single fusion subunit is unique and suggests a novel molecular mechanism. Although initially thought to only exist within the coronavirus family, a very similar FD with seemingly two distinct regions was recently discovered in Lassa virus, a prominent member of the Arenavirus family deemed to have pandemic potential by the World Health Organization (WHO).(97) Alongside the unique impact of the two fusogenic domains, the SARS-CoV-2 FD also contains several charged residues spread throughout, four positive and six negative. This is in stark contrast to other viral FDs such as those found in influenza and HIV which contain two and zero charged residues respectively.(44, 61) Whilst the formation of an internal salt bridge has been ruled out,(97) and several acidic residues have been implicated in binding Ca^{2+},(103) little research has focused on the significance of the positively charged residues. A structure-based approach using a combination of NMR and molecular dynamics found that K825 and K835, found in the two helices of the FP, were buried within a zwitterionic membrane, snorkeling to place their charges amongst the lipid headgroups.(134) Alternatively, while the positive charge of R847 was also found within the headgroup region, the position of the backbone and aliphatic chain was dictated by dynamic movements of the FL.(134) As anionic lipids have been found to be crucial for the fusogenecity of the FD,(118) it is reasonable to assume that complimentary interactions between those lipids and basic amino acids within the FD could be crucial for the underlying molecular mechanism to initiate fusion.

Here we have sought to elucidate the importance of four well-conserved(149), positive amino acids found within the SARS-CoV-2 FD (Figure 5.1). Within the 40-amino acid FD construct, both alanine and charge conserving mutants for all positive amino acids (K825, K835, R847 and K854) were created and expressed, with the exception of K854A. Each mutant perturbed

the function of the FD, with both R847 mutants displaying the greatest loss, whilst K825A increased fusogenecity. Moreover, an increase in α-helicity for K825A was found via CD, with all other mutants displaying the same global secondary structure as Wt. To understand the role of the basic amino acids further, we employed a ^{19}F-Phe labelling approach to gather atomic resolution structural data. After assigning the ^{19}F-Phe Wt peaks through a mutagenesis-based approach, chemical shift perturbations (CSPs) were plotted for each mutant against Wt. K825A displayed the only significant CSP, corresponding to a ^{19}F-Phe labelled residue (F823) in helix one and this was confirmed as the site of the α-helical extension via traditional ^{13}C and ^{15}N backbone assignment for K825A. In summary, all four basic residues play a part within the fusion process, however K825 in helix one of the FP, and R847 within the FL, appear to have the most important roles. Together, the data suggests that electrostatic interactions between positive amino acids within the FD and negatively charged lipids in the target membrane, could be important for the molecular mechanism behind the initiation of SARS-CoV-2 fusion.

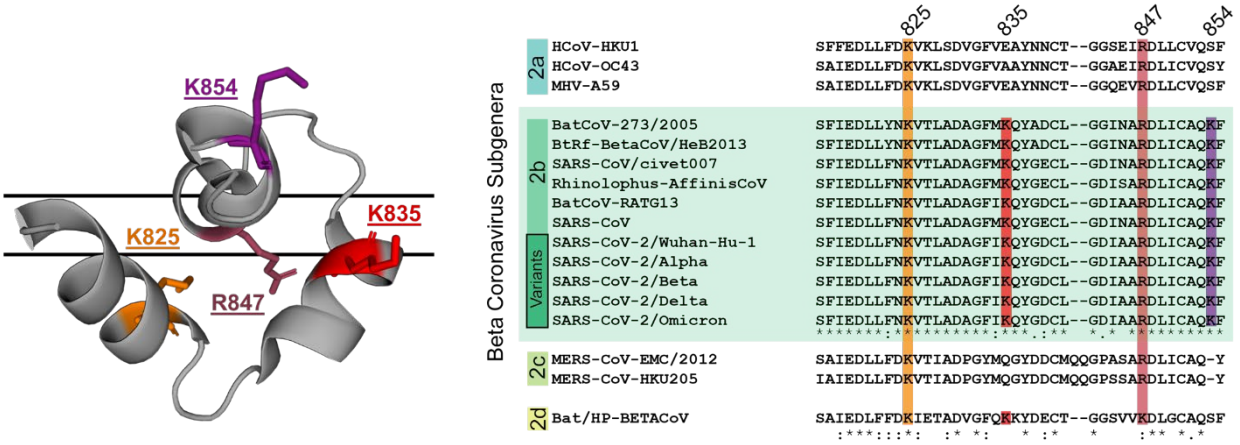

Figure 5.1 The structure of the SARS-COV-2 fusion domain (FD) embedded in a DMPC:DH^{7}PC bicelle (PDB:7MY8), with the basic amino acids highlighted and sequence conservation across the four beta coronavirus subgenera.

5.2 Results

5.2.1 Basic Amino Acids are Important for the Fusogenecity of the FD

A preference for the SARS-CoV-2 FD to initiate fusion with membranes containing anionic lipids has implicated four positively charged amino acids within the FD's molecular mechanism. A total of eight mutants were created, consisting of changing a positive residue to alanine and also maintaining the positive charge: K825A/R, K835A/R, R847A/K, and K854A/R. All mutants were expressed via an *E.coli* system except K854A which did not express at all. The location of the basic residues' side chains was identified within a zwitterionic bicelle and are displayed in Figure 5.1, along with the high level of conservation for each residue. Within this structure, K825, K835, and R847 sidechains can all be found inside the membrane in proximity to the lipid headgroups, potentially implicating them in crucial interactions for the structure and/or function of the FD.(134) On the other hand, K854 is found in solution, facing away from the membrane.

To understand the role of individual positive charges on the fusogenic properties of the FD, we first sought to measure the impact of each mutant via an *in vitro* fusion assay with membranes containing either BMP or 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine (POPS). The reason for independently assessing the two anionic lipids was to understand whether the previously discovered preference for BMP was due to any specific interactions with a particular basic residue. When standardizing the results to Wt with the respective lipid composition, the same trend is witnessed for both anionic lipids suggesting no preferential interactions take place (Figure 5.2A). Within these trends, we see that K825A results in a gain of function/fusogenecity (GoF), whilst all other mutants resulted in a significant loss of function/fusogenecity (LoF), with R847 mutants displaying the least amount of fusion (Figure 5.2A). Alongside fusion, the affinity of the initial interaction between the FD and anionic membrane was perturbed, with K825 and R847 again the

most significantly impacted (Figure 5.3). Intriguingly, when analyzing the structure of the mutants via CD, K825A displayed a significant increase in α-helicity compared to all other mutants which were similar to Wt (Figure 5.2B). Hence, K825A displays an increase in α-helicity that appears to directly correlate with an increase in fusogenecity, whilst a lack of change in the CD spectra from Wt suggests changes in secondary structure are not responsible for the LoF witnessed across all other mutants.

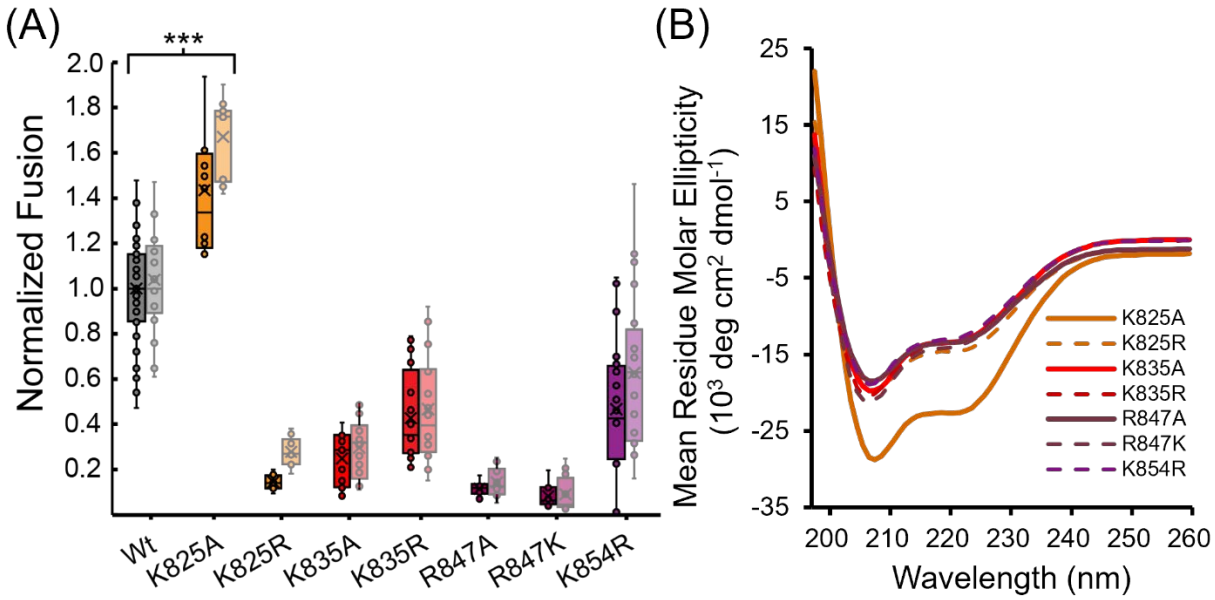

Figure 5.2 Functional and global secondary structure analysis of the fusion domain (FD) mutants. (A) The same general relationship is witnessed for the fusogenecity of each mutant when normalized to Wt, regardless of the anionic lipid present (80:20 POPC:BMP = Full, 80:20 POPC:POPS = Opaque). K825A results in a significant gain of function ($t(24)=-4.75$, $p=7.75\times10^{-5}$ and $t(12)=-3.75$, $p=0.02$ Wt against BMP and POPS respectively) whilst all other mutants result in a significant loss of function ($p<0.05$) as found through an independent T-test. (B) Through CD, K825A displays a large increase in α-helicity whilst all other mutants show a similar global secondary structure to Wt.

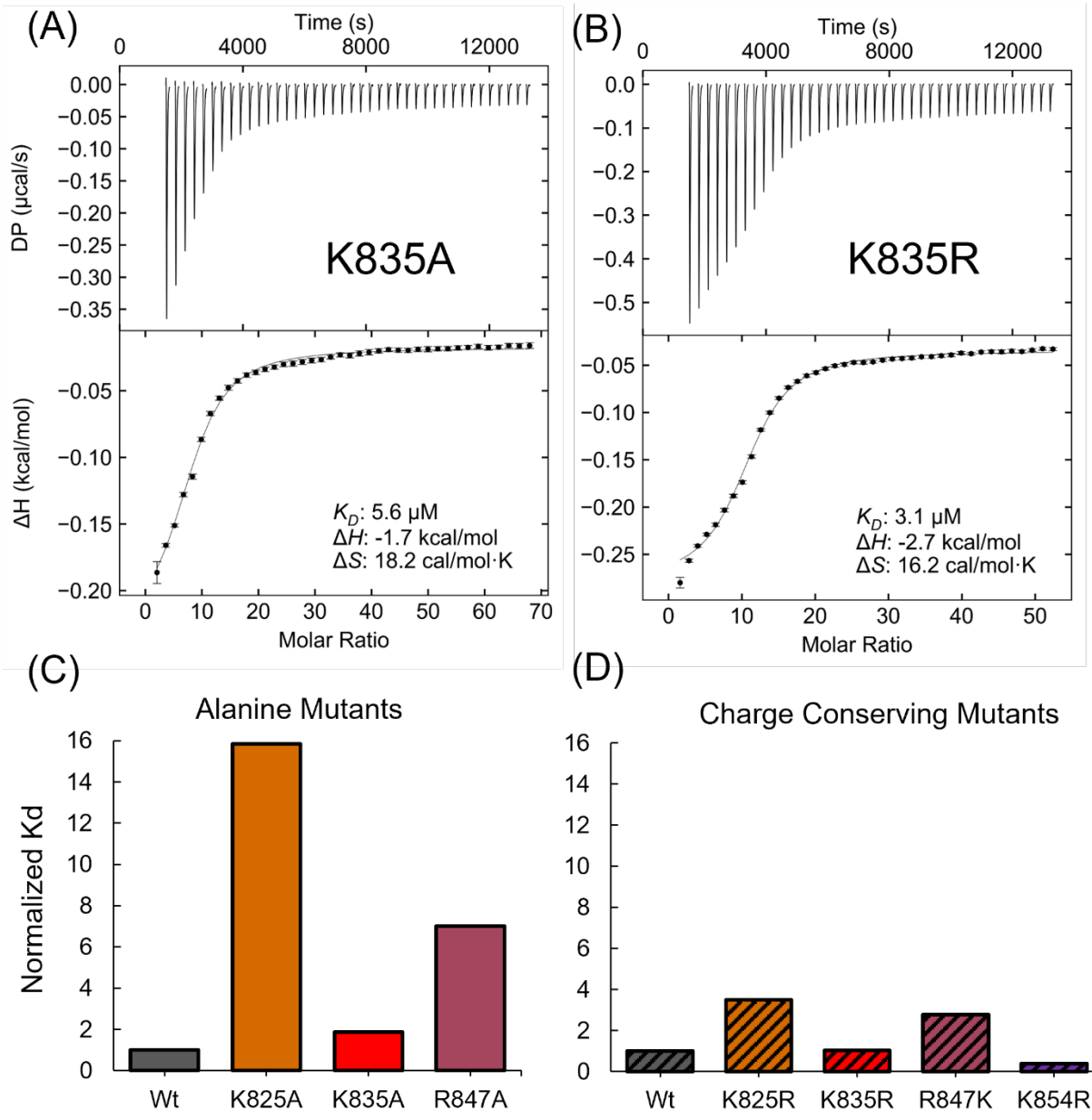

Figure 5.3 Fusion Domain (FD) mutants interacting with 50:50 POPC:POPS liposomes in comparison to Wt. Example titrations for (A) K835A and (B) K835R. (C) Perturbations in affinity are observable for all alanine mutants with 16-fold, 2-fold and 7-fold increases in Kd observed for K825A, K835A and R847A, respectively. (B) A similar trend can be witnessed for the charge conserving mutants but to a lesser extent, with K825R 3.5-fold and R847K 2.8-fold the most significantly impacted. K835R displays the same Kd as Wt and K854R shows an increased affinity for the liposomes.

5.2.2 Structurally Investigating the Role of the Basic Residues

To probe the structural nuances of the basic FD mutants further, a ^{19}F NMR approach was utilized, through a ^{19}F-Phe labelling scheme.(150) The rationale being that through a deconvoluted ^{19}F spectra, the local site of any structural perturbation could easily be identified. There are four phenylalanine residues within the FD (Figure 5.4A), one near the N-terminus (F817), one at the C-terminus (F855) and two conveniently located in the helix-turn-helix motif found in the FP; F823 in helix one and F833 in helix two. The only other aromatic amino acid within the FD is a single tyrosine (Y837), which was successfully labeled alongside the four phenylalanine residues in the ^{19}F-Phe+Tyr labelled Wt FD (Figure 5.5). Unfortunately, no

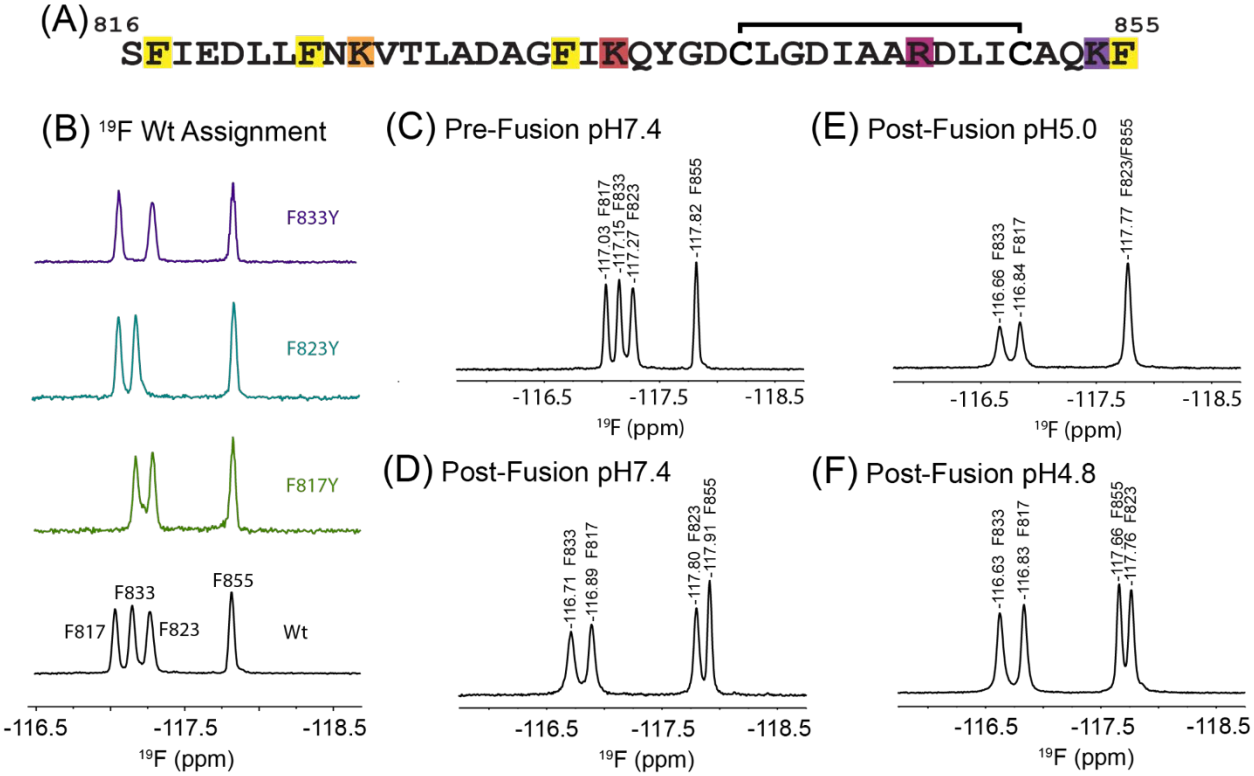

Figure 5.4 ^{19}F-Phe Wt FD assignment and sensitivity to local environment. (A) The SARS-CoV-2 fusion domain (FD) sequence with ^{19}F labelled phenylalanine's (Yellow) and positively charged residues highlighted. (B) A mutagenesis approach was utilized to assign the four ^{19}F peaks, with each mutant resulting in the loss of a single peak corresponding to the missing phenylalanine. Clear chemical shifting is present as the FD goes from (C) a random coil in solution, (D) to a helix-turn-helix in the presence of DPC. The slight shifting from (D) pH7.4 to (E) pH5.0 correlates well with the minor conformational changes witnessed previously.[8] (E) F823 and F855 overlap at pH5.0 , however, (F) a minor drop in pH can facilitate peak separation.

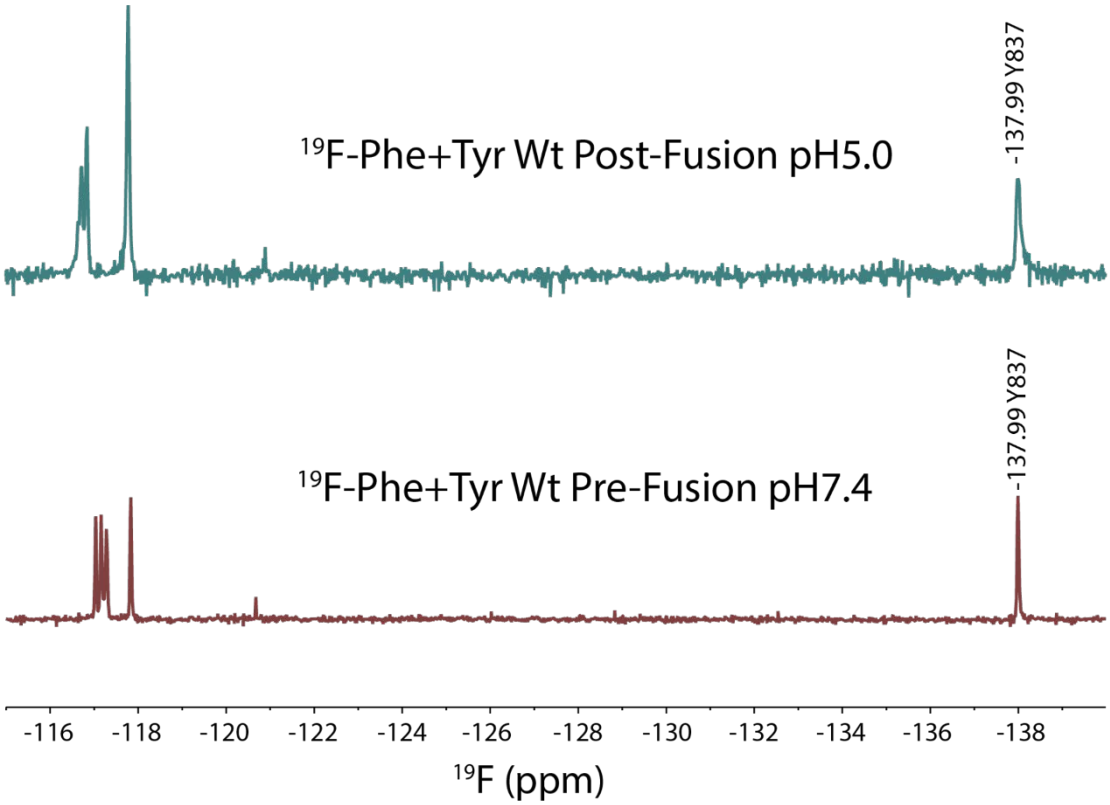

Figure 5.5 Phenylalanine and Tyrosine residues can be ^{19}F labelled simultaneously. However, ^{19}F-Y837 displayed no change in chemical shift and thus could not provide any further structural information.

shifting for ^{19}F-Y837 was witnessed between the functional states of the FD, likely due to the unique positioning of Y837 between the FP and FL. Each chemical shift was assigned to the correct corresponding residue via a mutagenesis approach, where three mutants were created (F817Y, F823Y and F833Y), expressed with ^{19}F-Phe and analyzed (Figure 5.4B). For example, with the F817Y mutant (Figure 5.4B – Green) only three ^{19}F-Phe peaks are observable such that the missing peak corresponds to F817 when compared to Wt. In the pre-fusion state, where the FD is a completely random coil,(47) four individual and well dispersed peaks are observed (Figure 5.4C). Following the addition of DPC, CSPs are witnessed for all peaks as the FD embeds within the membrane and forms the helix-turn-helix structure(Figure 5.4D).(98) As expected F823 and F833 undergo the largest chemical shifts, with absolute CSPs of 0.53 ppm and 0.44 ppm respectively, as helix one and helix two are embedded within the membrane in the

post-fusion state. With CSPs of 0.15 ppm and 0.09 ppm, F817 and F855 undergo shifting to a much lesser extent indicating less significant structural changes in those regions. Minor shifts are then observable as we decrease the pH from pH 7.4 to pH 5.0 (Figure 5.4D-E), with F817 and F833 shifting 0.05 ppm, F823 0.03 ppm and F855 0.14 ppm. This agrees well with the modest structural change discovered previously from the backbone assignment of the pH7.4 and pH5.0 Post-Fusion states.(96) Due to the overlapping of residues F823 and F855 at 117.77 ppm in the post-fusion pH5.0 state, we decreased the pH to pH4.8 in order to separate the two. F855 is particularly sensitive to pH likely due to its c-terminal carboxylic acid group, displaying a CSP of 0.14 ppm from pH7.4 to pH5.0 and 0.10 ppm from pH5.0 to pH4.8, allowing us to discern the two peaks. The data shown in Figure 5.4 verifies the use of ^{19}F-Phe labelling to detect structural changes within the FD at specific regions of interest such as the two helices in the helix-turn-helix motif.

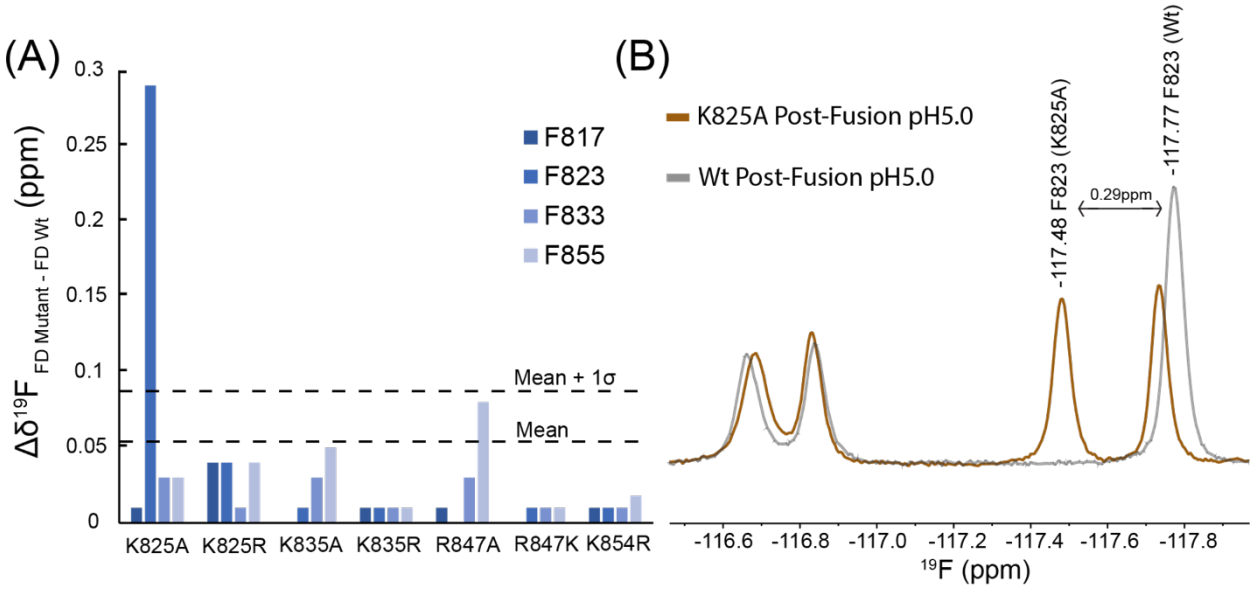

Figure 5.6 ^{19}F-Phe structural analysis of fusion domain (FD) mutants. (A) Only K825A showed any significant CSPs, specifically for residue F823 found in helix one. (B) Overlaying the Wt and K825A ^{19}F-Phe 1D spectra exhibits clear shifting of F823, with only marginal shifting for all other peaks.

Every basic amino acid mutant was labelled with ^{19}F-Phe, expressed, and analyzed. All three states were captured (Pre-Fusion, Post-Fusion pH7.4 and Post-Fusion pH5.0) and compared to Wt to observe any structural perturbations. Our primary focus was the Post-Fusion pH5.0 state due to the increased functional relevance of the FD structure at low pH. CSPs compared to Wt were plotted (Figure 5.6A) and displayed significant perturbations for F823 within the K825A mutant (Figure 5.6B), suggesting the mutation impacts helix one. With little structural perturbations present we turned to investigate dynamics, with spin-lattice (T1) and spin-spin (T2) experiments performed in the ^{19}F dimension. Whilst little deviation from Wt was observed across each mutant, differences in the dynamics of individual residues were detected. (Figure 5.7) Together, this data supports the use of ^{19}F-Phe to identify the location of structural perturbations such as the local impact of K825A, from a simplified, deconvoluted spectra at atomic resolution.

To confirm the conclusions gathered from the ^{19}F data regarding K825A, a more traditional ^{13}C ^{15}N NMR labelling approach was utilized. All forty residues within K825A were assigned and compared to Wt at pH5.0 (Figure 5.8A) where several CSPs were found, particularly around the site of the mutation. (Figure 5.8B) This indicates that the mutation only significantly impacted

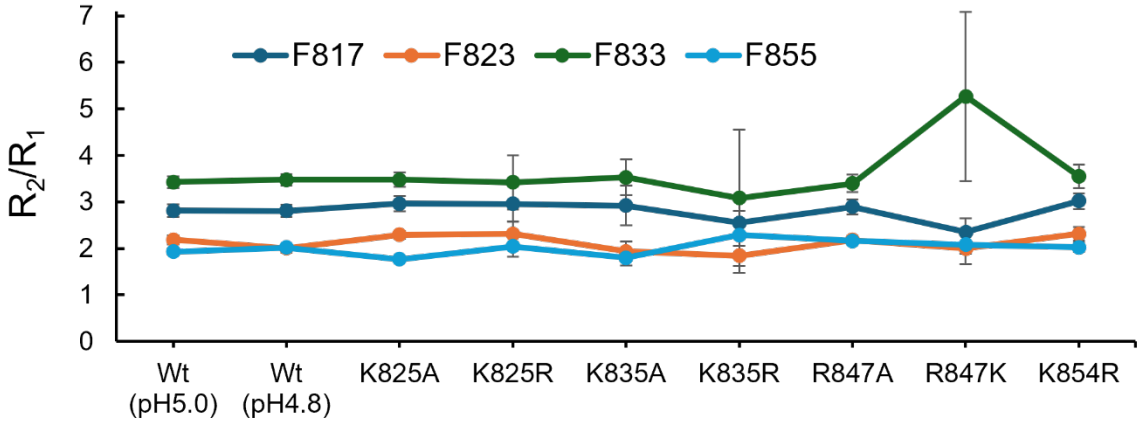

Figure 5.7 ^{19}F dynamics. Spin-spin (T1) and Spin-lattice (T2) measurements were taken for the Wt and all mutants, however little deviation was observed. Wt pH4.8 was analyzed due to the splitting of F823 and F855 allowing the analysis of individual residues.

helix one, in terms of the secondary structure, as predicted via the ^{19}F data (Figure 5.6). When plotting the chemical shift index (CSI) for K825A, a clear elongation of helix one is observable (Figure 5.8C). Intriguingly, when comparing the length of helix one via CSI across Post-Fusion pH7.4 (6 residues), Post-Fusion pH5.0 (8 residues) and K825A (12 residues) pH5.0 to fusogenic activity, a positive correlation can be observed. (Figure 5.8D) In summary, helix one within the FP region of the FD has been elucidated as a key component in the perturbation of the local membrane to initiate membrane fusion.

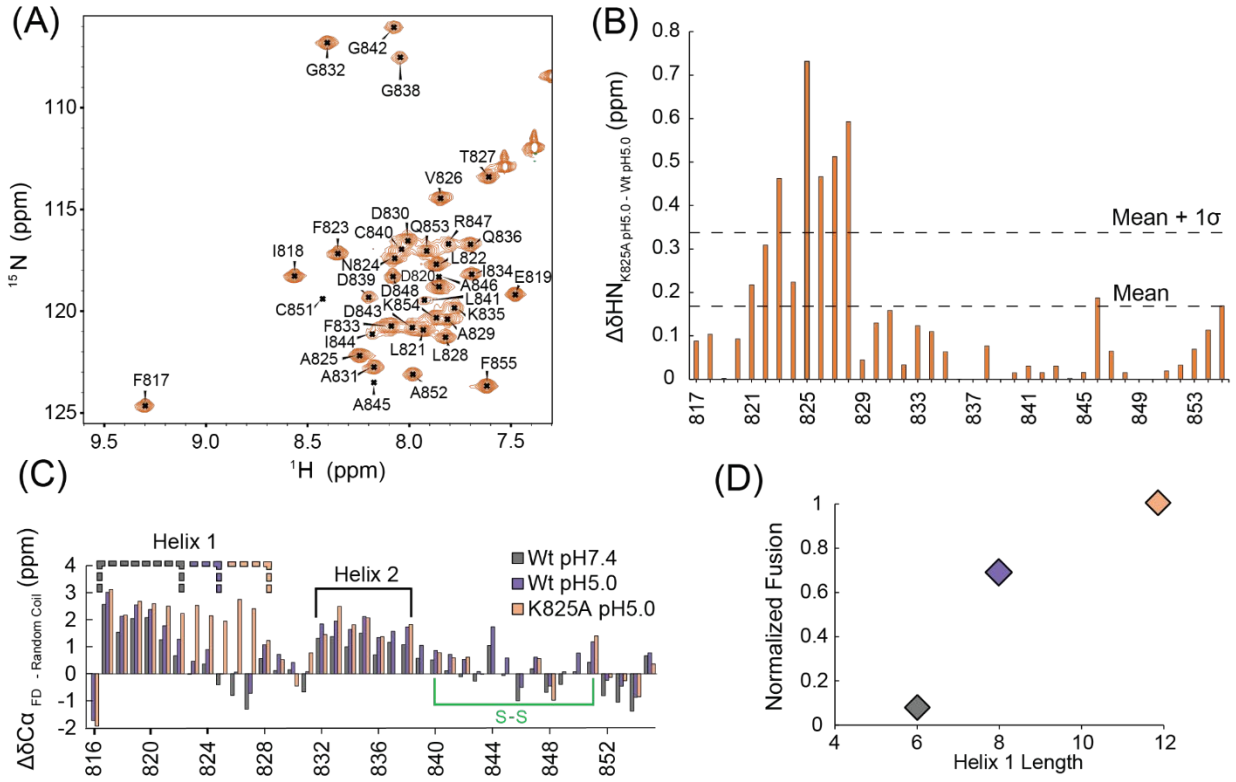

Figure 5.8 K825A has an extended helix one which is heavily implicated in fusion domain (FD) fusogenecity. (A) The backbone of K825A was assigned (B) and CSPs plotted, displaying perturbations within helix one only. (C) K825A shows a helical elongation within helix one compared to the Wt Post-Fusion state at pH7.4 and 5.0, (D) which correlates with an increase in fusion.

5.3 Discussion

During the process of viral entry, the SARS-CoV-2 spike protein displays multiple traits that stand out in comparison to other enveloped viruses, such as a secondary cleavage event(30, 151) and the ability to fuse within either the plasma or endosomal pathway.(57, 109) Intriguingly, both of these traits are linked to the FD, a well conserved stretch of forty amino acids at the N-terminus of the cleaved S2 subunit. Previous work has implicated pH, and membrane lipid composition as key modulators of FD mediated fusion, with the FD demonstrating a positive fusogenic relationship with low pH and increasing anionic lipid concentrations; two characteristics of the endosomal membrane.(57, 96, 118, 131) With this in mind, the relationship between the FD and negatively charged lipids is one that deserves further attention in order to elucidate the molecular details by which the FD initiates viral fusion. Through a mutagenesis-based approach the mechanism of SARS-CoV-2 FD initiated fusion has been further elucidated, with the importance of basic amino acids (K825, K835, R847 and K854) clarified, and helix one of the FP region proven to be a key constituent within the fusion process.

Electrostatic interactions are not heavily associated with the role of viral FDs, predominately due to the necessity for such fusogenic regions to bury within the membrane, making charged residues rather counterintuitive due to the increased energetic cost.(152, 153) Astonishingly, the FD of SARS-CoV-2 contains ten such charged residues (four positive and six negative), with high sequence conservation indicative of a critical role, which to date remains unclear. There is some evidence to suggest that each negative residue participates in the coordination of two Ca^{2+} ions.(66) For clarity, Ca^{2+} was omitted from this study in order to focus on the interactions between the individual positive amino acids and membrane lipids, particularly as the role of the divalent cation is still unclear within the FDs fusion mechanism. Moreover, the

possibility of a potential intra or inter-molecular salt bridge was also ruled out in a comparative study alongside the LASV FD without Ca^{2+} present.(97) In a high salt environment it was shown that the affinity and fusogenicity of the LASV FD with anionic membranes decreased significantly more so than that of the SARS-CoV-2 FD due to the loss of an intra-molecular salt bridge. However, a decrease in the initial binding affinity and fusogenecity for the SARS-CoV-2 FD in a high salt environment was witnessed, suggesting that weaker electrostatic interactions may have been perturbed. The decreased binding affinity when the basic residues are mutated to alanine as opposed to conserving their positive charge (Figure 5.3) and the lack of interaction witnessed in the absence of anionic lipid,(118) suggests that transient electrostatic interactions could well be taking place between the SARS-CoV-2 FD and anionic membrane.

As the FD previously displayed a fusogenic preference for the endosomal resident anionic lipid BMP,(118) it was pertinent to establish whether this was due to any specific interactions on an individual amino acid basis. Regardless of whether the POPS or BMP was present, the same functional trends were witnessed amongst all mutants (Figure 5.2), suggesting that any specific protein-lipid interactions were independent of the anionic headgroup. Looking at the mutants individually, K835R and K854R retain the most fusogenecity compared to Wt. The fact that K835R displays more fusion than K835A, could indicate potential electrostatic interactions are maintained by the arginine mutant. This theory is reinforced by K835R maintaining the same affinity to anionic liposomes as Wt. (Figure 5.3) With regards to K854, as it exists so close to the c-terminus, the dynamic nature of the region could be augmenting the gathered data and thus, it is tough to predict a mechanistic role for this residue. K825 and R847 are highly conserved throughout the β-Coronavirus sub-group, with R847 even found to be exclusive to this branch of the viral family (Figure 5.1). R847 was found to be essential for both fusion and the initial

interaction of the FD with anionic liposomes, suggesting a significant role within the FDs molecular mechanism. Simulations of the SARS-CoV-2 FD have provided evidence towards an ability to thin the local membrane environment, and even implicated charged residues within the mechanism of action.(134) Perhaps the interactions between basic residues and the anionic membrane allow the FD to modulate the lipid headgroups to induce negative curvature. This action previously labelled as polar head intrusion(154), was proposed to catalyze fusion by reducing the repulsion of opposing membrane headgroups and has since been elucidated as an important component in the fusogenecity of the influenza fusion peptide.(48, 155) This theory is further strengthened by the specific, positive relationship witnessed between the SARS-CoV-2 FD and the endosomal resident lipid BMP.(118) The inverted conical shape of the lipid favors local negative curvature and promotes membrane fluidity, likely lowering the energetic barrier necessary for the FD to modulate the local membrane environment. Further research into the interactions of individual basic amino acids with negatively charged lipids and their resulting impact on the shape of the membrane is essential to prove this hypothesis, with R847 expected to play a key role.

In a simple model, we anticipate that the binding affinity of the FD to a membrane would positively correlate with fusion. This is intuitive as FD mediated fusion cannot occur unless the FD comes into contact with the membrane in question. Curiously, the K825A mutant displayed a significant decrease in binding affinity (Figure 5.3) but an increase in fusion (Figure 5.2). Any GoF mutant found within the FD alone could elicit the same increased fusogenecity in the complete spike protein, and although further research is required, the mutant K825A may produce a more virulent strain of SARS-CoV-2.

To try and understand the unique properties of K825A, we investigated the structure of the mutant at atomic resolution. The increased α-helicity of K825A was found to be a four amino acid elongation of helix one, through a combination of ^{19}F-Phe CSPs and a more traditional ^{13}C ^{15}N backbone assignment. In the Wt at pH7.4, helix one terminates at L822, at pH5.0 the helix contains the full hydrophobic motif '^{821}LLF^{823}' and terminates at N824, whilst in K825A the helix extends to L828. The highly conserved 'LLF' motif has been shown on multiple occasions to be critical for the fusogenecity of the FD, playing a key role in membrane insertion and ordering. (51, 54, 84, 113) The modulation of 'LLF' within helix one appears to play a critical role in fusion and incorporation at pH5.0 is proposed to be a significant factor for the increased fusogenecity witnessed for the FD at low pH. It is less clear how 'LLF' may contribute to the increased fusion witnessed in K825A, however CSPs are witnessed for the motif suggesting the chemical environment could be altered in a manner that is positive for fusogenecity. This could be in the form of a novel orientation, formation of new interactions, or an increased depth of insertion. Moreover, the positive relationship found between increasing α-helicity and fusion witnessed

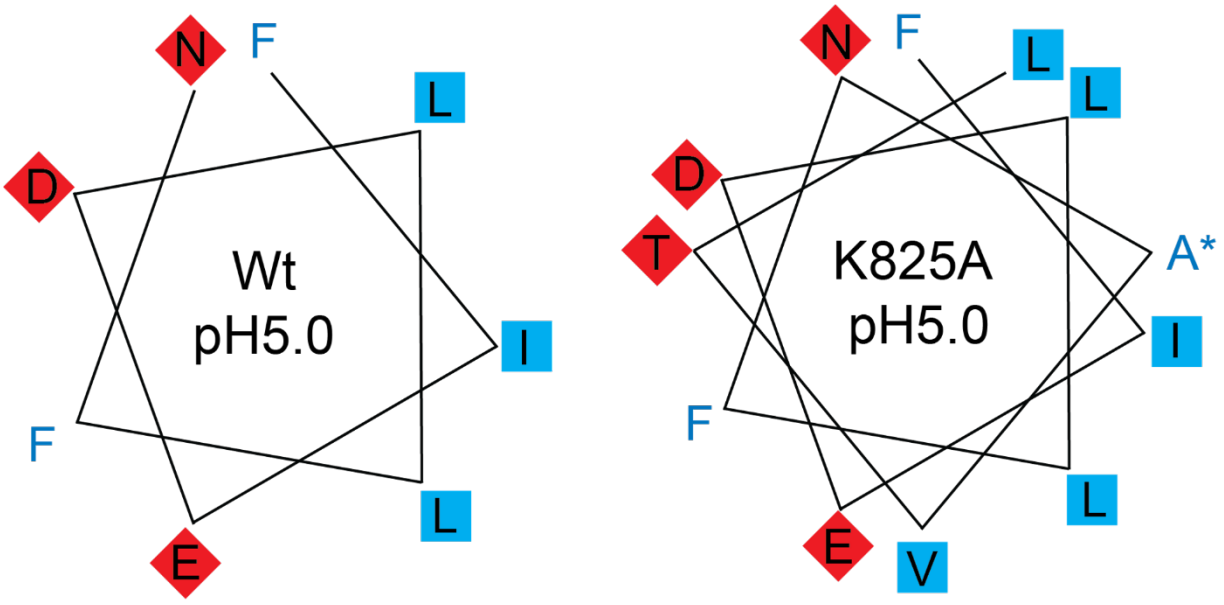

Figure 5.9 Helical wheel projection for helix one in Wt at pH5.0 and K825A at pH5.0.

(Figure 5.8D) has been found on several occasions previously, with α-helices capable of embedding within and destabilizing membrane environments.(39, 64, 114, 156). Using a helical wheel projection, an amphipathic character can be witnessed in helix one when K825 is not incorporated such as in the Wt at pH5.0. However, replacing this residue with a non-polar residue such as alanine allows the helix to elongate whilst maintaining amphipathicity (Figure 5.9). With polar residues energetically unfavorable within the hydrophobic core of the membrane, snorkeling is a common trait witnessed with charged residues. Snorkeling allows the aliphatic chain of the residue to exist within the membrane, whilst the charge group is placed amongst the polar lipid headgroups.(157) Therefore in the Wt, K825 could be snorkeling to form electrostatic interactions in the presence of anionic lipid, which in turn may prevent the elongation of helix one. Similar interactions were predicted in K825R, which maintains the same secondary structure as Wt, (Figure 5.2B) however a decreased binding affinity and fusogenecity compared to Wt suggest a more complicated explanation for the role of K825 than solely electrostatic interactions.

5.4 Conclusions

In conclusion, basic amino acids in the SARS-CoV-2 FD are important towards the mechanism of membrane perturbation that leads to the initiation of viral fusion. Even when the positive charged is conserved through mutagenesis, a significant loss of function remains present, suggesting a highly specialized role for those original residues. In particular, K825 and R847 have been elucidated as playing key roles within the fusion mechanism and are intriguingly found to be very well conserved throughout the β-coronavirus family. Furthermore, we have provided functional evidence regarding a potential mutation (K825A) that could lead to a more fusogenic spike protein and thus, a novel strain of SARS-CoV-2 with increased pathogenicity. Future work should focus on understanding the GoF witnessed for K825A in the context of the full spike protein and the

mature virion. Finally, it is important to note that although a lysine to alanine mutation may be unlikely to randomly occur as two base pair mutations are necessary, other non-polar mutations at position 825 could elicit the same effect and should also be considered for further investigation.

5.5 Materials and Methods

5.5.1 Materials

All lipids were acquired from Avanti Polar Lipids in chloroform (POPC: 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine, POPS: 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine, BMP: bis(monooleoylglycero)phosphate (S,R Isomer), LissRhod-PE: 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl) and NBD-PE: 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(7-nitro-2-1,3-benzoxadiazol-4-yl), except for the detergent Dodecyl phosphocholine (DPC) which was purchased from Anatrace. 4-Fluoro-DL-Phenylalanine (^{19}F-Phe), 3-Fluoro-DL-Tyrosine (^{19}F-Tyr), and glyphosate were acquired from Fisher Scientific, with traditional NMR isotopes (^{13}C-glucose and ^{15}N-ammonium chloride) purchased from Cambridge Isotopes.

5.5.2 Expression of 19F-Phe Fluorinated Fusion Domain

The SARS-CoV-2 FD is made up of forty amino acids (^{816}SFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKF^{855}), within which exists four phenylalanine, one tyrosine and zero tryptophan. The following ^{19}F expression protocol was initially adapted from Mandala and co-workers,(150) with unlabeled, and ^{13}C ^{15}N Expression described previously. A single colony from an LB plate was used to inoculate 25 mL of minimal media (pH7.3, 56 mM Na_2HPO_4, 15 mM KH_2PO_4, 9 mM NaCl, 8 mM NH_4Cl, 55 mM Glucose, 1

mM $MgSO_4$, 0.3 mM Na_2SO_4 0.2 mM $CaCl_2$, trace amounts of thiamine and biotin, 50mg kanamycin) and grown overnight (~20 hrs) at 37 °C. The next morning the confluent starter culture was added to 1 L of minimal media and grown at 37 °C until an OD_{600} of 0.5 was achieved (approx. 6 hrs). Subsequently, the temperature was lowered to 18 °C and 1.5 g of glyphosate added to prevent aromatic amino acid biosynthesis through arresting the pentose phosphate pathway.(158) This was followed by addition of 115 mg L-Tryptophan, 115 mg L-Tyrosine (or 400 mg of 3-Fluoro-DL-Tyrosine) and 400 mg 4-Fluoro-DL-Phenylalanine. After 30-60 minutes, 1 mM IPTG was added, and protein expression proceeded at 18 °C overnight (~20 hrs). Cells were then collected by centrifugation at 4,000 xG for 30 mins at 4 °C, with the resulting pellet either stored at -80 °C or purified immediately.

5.5.3 Purification of the Fusion Domain

The purification of the SARS-CoV-2 FD has been described in detail previously and remains consistent regardless of the expression/labelling system.(47) Briefly, 8 M urea was used to solubilize the cell pellet followed by Ni-NTA affinity chromatography and cleavage via SUMO protease in dialysis. The cleaved FD is then isolated through the use of Ni-NTA affinity chromatography once again and dialyzed to remove reducing agents, ensuring the correct formation of the disulfide bond. Following dialysis, the sample was purified further via size exclusion chromatography (SEC) with 10 mM HMA 100 mM NaCl pH7.4 as the mobile phase. Resulting fractions are pooled, concentration determined via A_{280}, and stored at 4 °C.

5.5.4 Preparation of Small and Large Unilamellar Vesicles

Specific amounts of lipids solubilized in chloroform were added to glass test tubes via a Hamilton syringe. Chloroform was then removed via gentle vortexing whilst applying a continuous stream

of nitrogen to create a lipid film, followed by storage in a desiccator under vacuum overnight. For small unilamellar vesicles (SUVs), the lipid film was resuspended in the chosen buffer via vortex then sonicated for 15 mins at 10% duty cycle (1s on/1s off) using a Branson ultrasonicator microtip with the sample sat in ice water. The SUVs were then centrifuged at 20,000 xG for 10 mins to remove residual metal particulates. For large unilamellar vesicles (LUVs), the lipid film was resuspended in buffer and vortexed extensively. The lipid suspension was subjected to 10 freeze-thaw cycles between liquid nitrogen and a 42 ºC water bath. Liposomes were extruded using a liposofast extrusion kit (Avestin) a total of 21 times through 2 polycarbonate membranes with a 100 nm pore size. All vesicles were either used immediately or stored for a maximum of 72 hrs at 4ºC prior to use.

5.5.5 NMR experiments

NMR spectra were acquired using a Shigemi NMR tube in 10 mM Hepes/MES/Sodium Acetate (HMA), 100 mM NaCl, pH7.4 or pH5.0 buffer with 90%H2O/10%D2O. DPC was present in all Post-Fusion states as a membrane mimic at a concentration of ~100 mM. All experiments were conducted on the Bruker Ultrashield^{TM} 600 MHz magnet with a CP2.1 TCI 600S3 H&F/C/N-D-05 Z XT at a temperature of 23 °C. Typical ^{19}F experiments consisted of 2000 scans, an acquisition time of 0.15 s, D1 of 0.3 s and a sweep width of 20 ppm. All ^{19}F data was processed using the software Mestrenova,(159) and corrected to an internal TFA standard at -76.55 ppm. For the backbone assignment of K825A, HNCA, HN(CO)CA, HNCO and HN(CA)CO experiments were performed, and processing was carried out through SPARKY((92) via NMR Box.(93)

5.5.6 Circular Dichroism

CD data was acquired via a Jasco J810 Spectro-Polarimeter using a quartz cuvette with a 2 mm path length. All experiments were conducted at room temperature (~23 °C) in 1 mM HMA, 10 mM NaCl, 50 mM DPC, pH5.0 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. Baselines were acquired in the exact same buffer, without any protein present and subtracted from all data.

5.5.7 Lipid Mixing Assay

Unlabeled LUVs composed of specified lipids were mixed with labeled LUVs containing the same composition with the addition of 1 mol% of the fluorescent labels: LissRhod-PE and NBD-PE. Experiments were conducted at a ratio of 9:1 unlabeled:labeled liposomes in Corning Costar black walled, clear bottom 96 well plates with excitation and emission wavelengths at 460 nm and 538 nm respectively, with a 530 nm cut-off. A SpectraMax M5 microplate at 23 °C was used to record the fluorescence measurements. Percent fusion was calculated as $\frac{(I_F - I_B)}{(I_{100} - I_B)}$, where I_B is the initial background fluorescence, I_F is fluorescence intensity measured after decreasing the pH, and I_{100} is the 100% fluorescence intensity value gathered after complete vesicle rupture following the addition of 1% Triton X-100. Experiments were performed at pH5.0 with a peptide/lipid ratio of 0.05 (5 μM and 100 μM). Negative controls identical to the experimental wells except no protein was present, were ran alongside all conditions and subtracted from the final values.

5.5.8 Isothermal Titration Calorimetry

To ensure sufficient buffer matching the FD sample was dialyzed against 4L of 10 mM Sodium Acetate, 100 mM NaCl at pH5.0 for 2 hrs at 4 ºC, with the same dialysis buffer used to resuspend

the lipid film prior to forming SUVs. All samples were centrifuged at 20,000 xG for 10 mins and degassed prior to each ITC experiment. All measurements were taken using a Malvern VP-ITC microcalorimeter with the following experimental parameters in place: initial delay: 1200s, 41 injections (1x2 µL and 40x6 µL), spacing: 300 s, duration: 14.4 s, Stir Speed: 270 rpm, ref power: 2 µcal/sec and temperature: 22 ºC. All processing was conducted through NITPIC,(148) and SEDPHAT, (147) with thermograms generated through GUSSI.(160)

Chapter 6: Concluding Remarks

Although there remains many outstanding questions regarding SARS-CoV-2 infection, this research has contributed towards our understanding of the membrane fusion process. Specifically, my work has provided molecular details regarding how the SARS-CoV-2 FD initiates fusion via embedding within and perturbing the target cell membrane in different environments.

We have shown that the SARS-CoV-2 FD only displays structural characteristics upon interacting with a membrane environment where a unique conformational ensemble of a FP and FL are present. Moreover, these two structurally distinct regions display a complex, synergistic relationship necessary for efficient fusion which is further impacted by environmental pH and lipid composition. All the data acquired points towards a novel molecular mechanism for the coronavirus FD that may have evolved to facilitate fusion with increasing efficiency. Whilst critical factors within the FD have been elucidated, such as the hydrophobic motif 'LLF' and the conserved disulfide bond, little is still understood regarding how exactly the FP and FL work in unison to facilitate fusion. Further atomic resolution structural data would be hugely beneficial in our understanding, especially within differing pH and lipid environments that reflect that of the endosomal membrane. To overcome the problems associated with structural elucidation of proteins in physiologically relevant membrane environments, a combined structural and MD-based approach could be helpful in providing the answers to such intricate questions.

The rather unique ability of SARS-CoV-2 to enter the host cell via both plasma membrane and endosomal membrane routes of fusion suggested a certain amount of redundancy within the fusion process. Naturally, it was assumed that this redundancy would be present in the FD, however it quickly became clear that the FD specifically prefers an environment similar to that found in the endosomal membrane to elicit fusion. The discovery of a positive, biphasic

relationship with decreasing pH was followed by finding a positive, specific relationship with the anionic lipid BMP, implicating two hallmarks of the endocytic pathway in FD mediated fusion. This information has provided further insight into the viral infection process, suggesting that the FD would rather enter cells via the endocytic pathway, possibly hinting at different driving force or membrane interacting region responsible for initiating fusion at the plasma membrane. The possibility of multiple fusogenic domains within a single viral protein that are environment specific may sound unlikely, but until the viral infection process is fully elucidated it remains a possibility.

Finally, one of the largest problems faced during the COVID-19 pandemic was the emergence of increasingly infectious variants of SARS-CoV-2. As mutations occur at random, variants can be a common characteristic with viruses that infect a vast amount of people as with every person infected the likelihood of a mutation beneficial to the virus increases. With the FD completely conserved across all known SARS-CoV-2 variants, it came as a surprise when we found that K825A resulted in a gain of fusogenicity. Whilst this mutation needs to be explored in the context of the complete spike protein, if not the mature virion, there is data to suggest that if this random mutation ever occurred then the resulting variant could be increasingly infectious. A more thorough investigation regarding the impact of other amino acid mutations on K825 is warranted, as any non-polar amino acid in place of the charged lysine could potentially increase fusogenecity. The early identification of increasingly pathogenic SARS-CoV-2 variants is essential to prevent future outbreaks and minimize the spread of infection.

Appendices

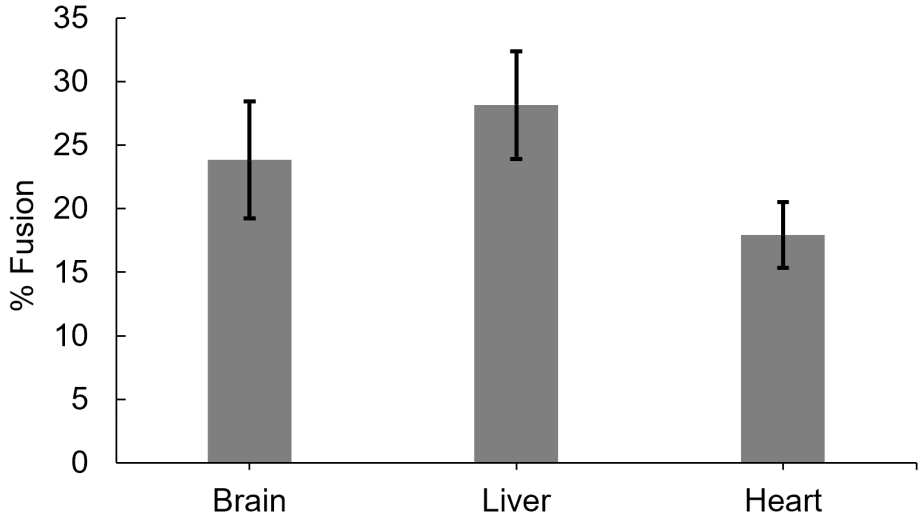

Figure A1 Fusion witnessed when making LUVs with complex cell extracts. Error shown is the SEM. n= 20

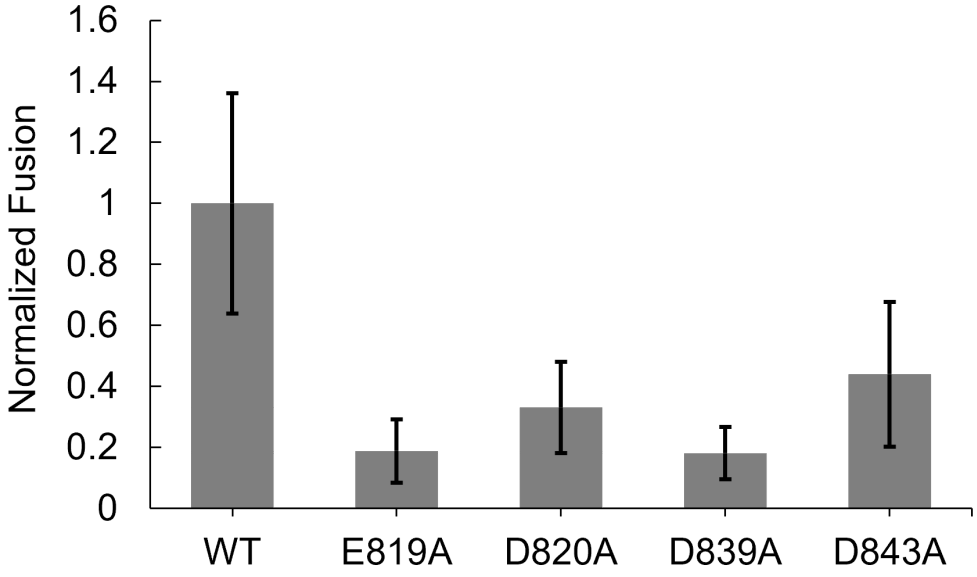

Figure A2 Negatively charged lipids are important for fusion with POPC:POPG 75:25. Other negative residues within the FD, D830A and D848A, are not shown as they did not express. n= 4

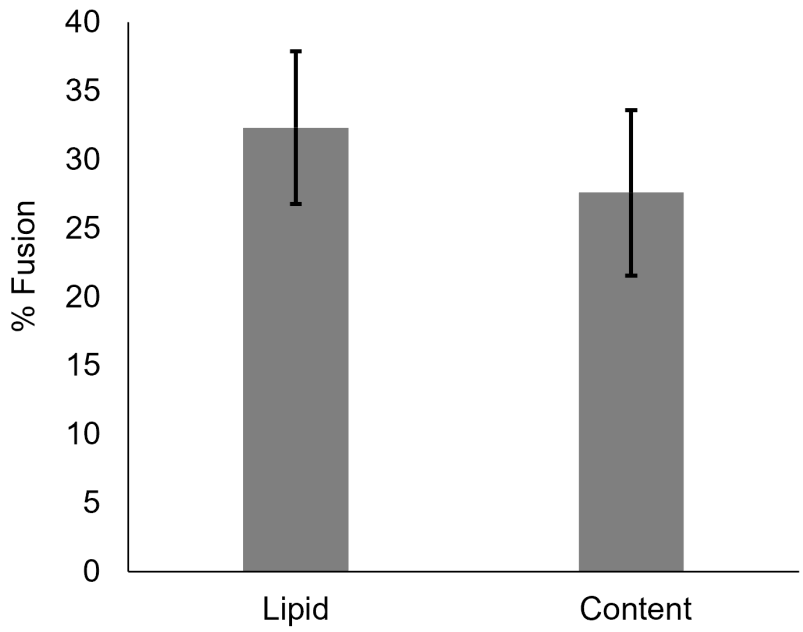

Figure A3 The fusion witnessed within the lipid mixing assay is confirmed to be most significantly fusion as opposed to bursting or hemi-fusion via the content mixing assay. Approximately 5% of the result via lipid mixing is though to be bursting. n= 4

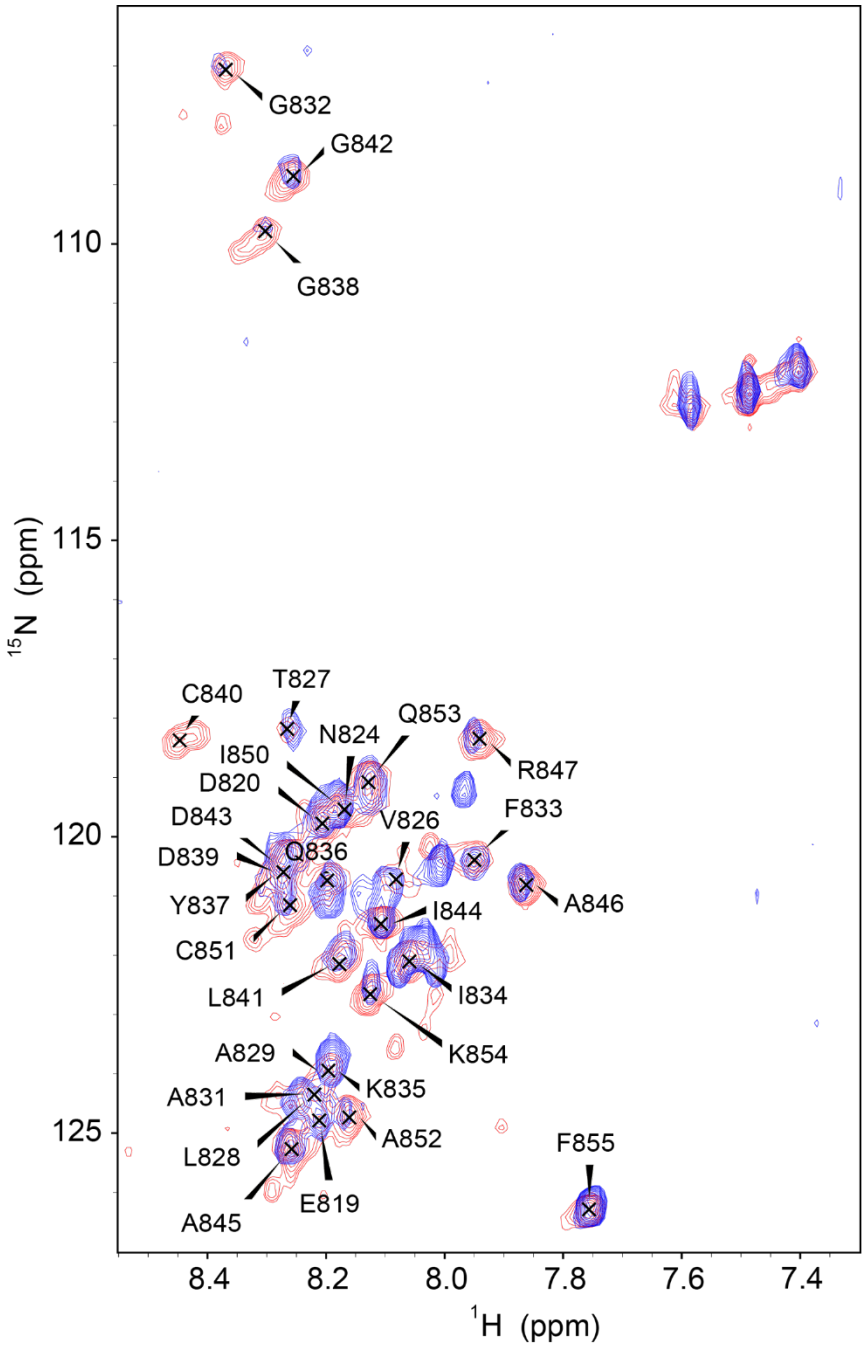

Figure A4 Pre-Fusion Wt pH7.4 (Blue) overlayed with FD with 5050 PCPG nanodiscs pH5.0 (Red) which are assigned based on proximity to Wt peaks. Whilst several residues did undergo significant enough perturbations to not be assigned, most of the peaks overlap suggesting lack of association between the FD and nanodiscs. Missing residues: S816, F817, I818, L821-F823, K825, D830, D848, L849.

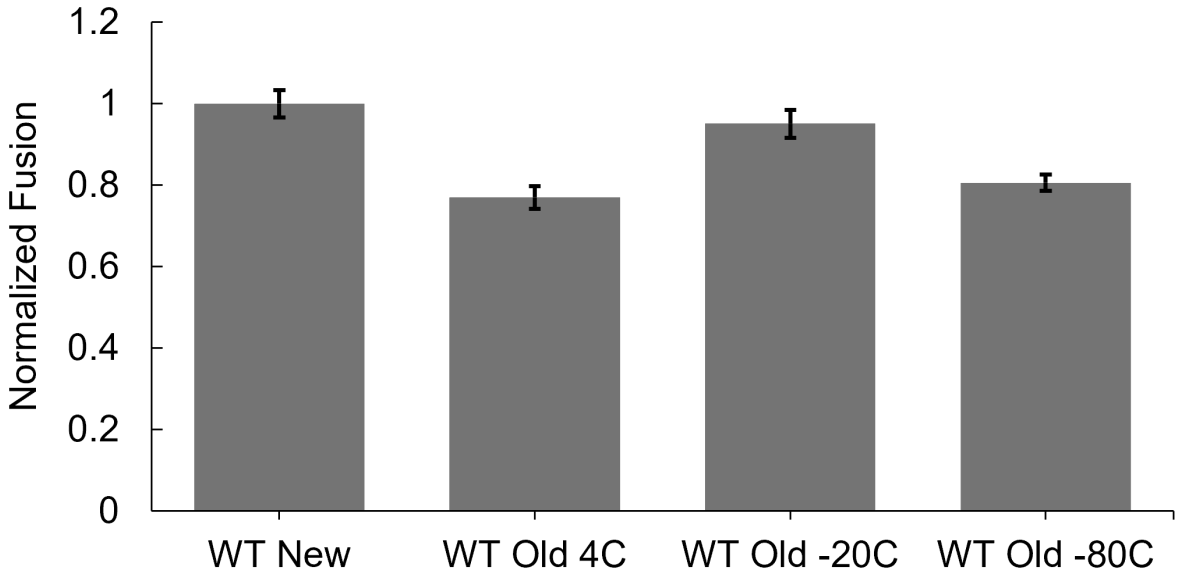

Figure A5 Analyzing storage conditions for SARS-COV-2 FD. All old samples were stored for one month prior to use, with the new sample purified the day before the assay and stored at 4C overnight.

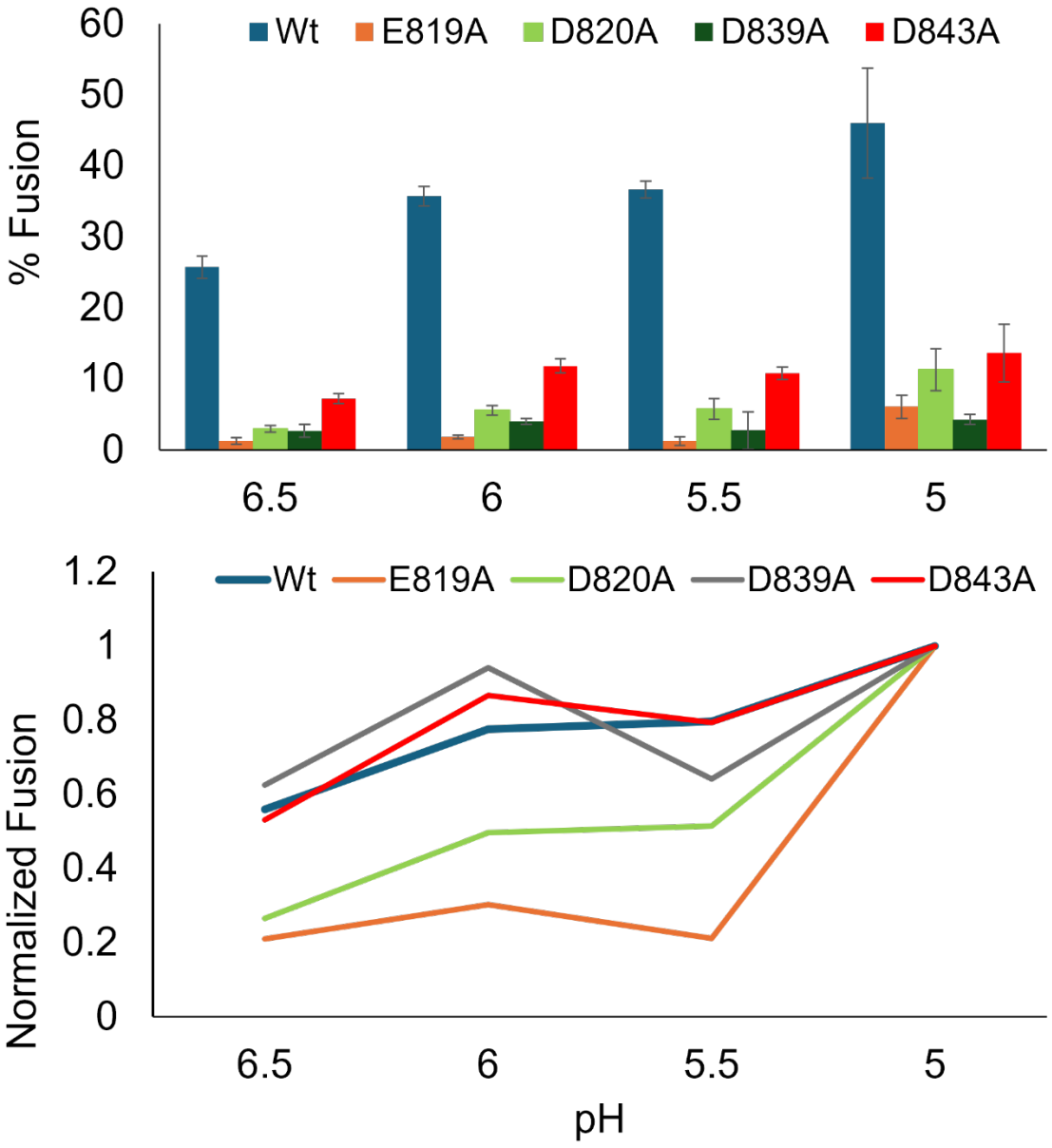

Figure A6 Relationship between pH and fusion for negative amino acid mutants compared to Wt.

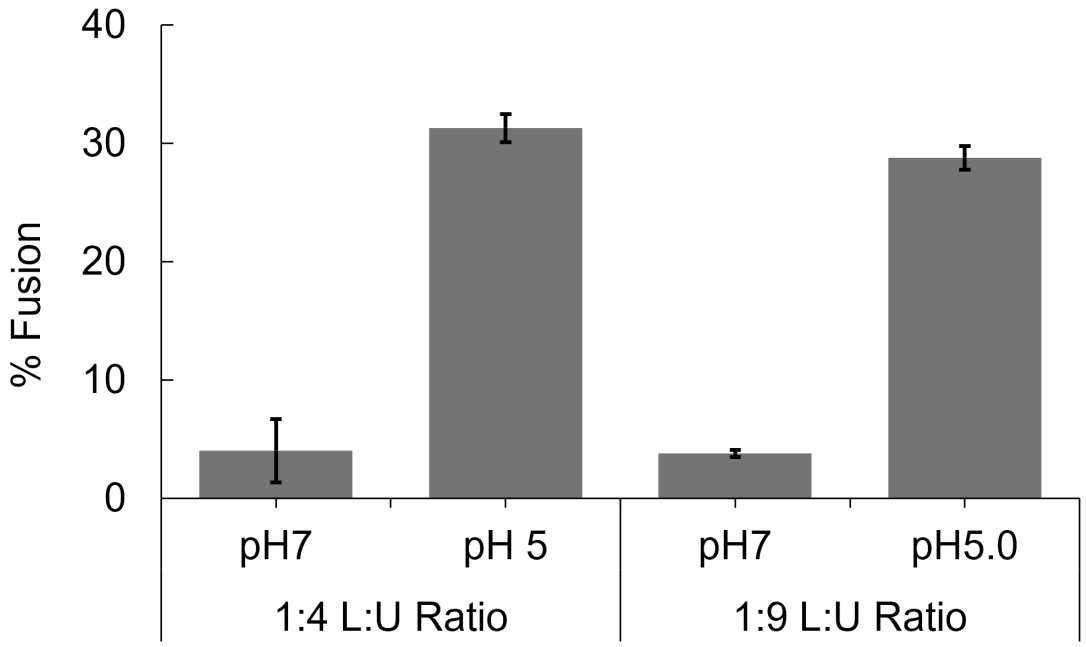

Figure A7 The ratio of unlabeled to labelled liposomes within the lipid mixing assay has little impact on the overall result.

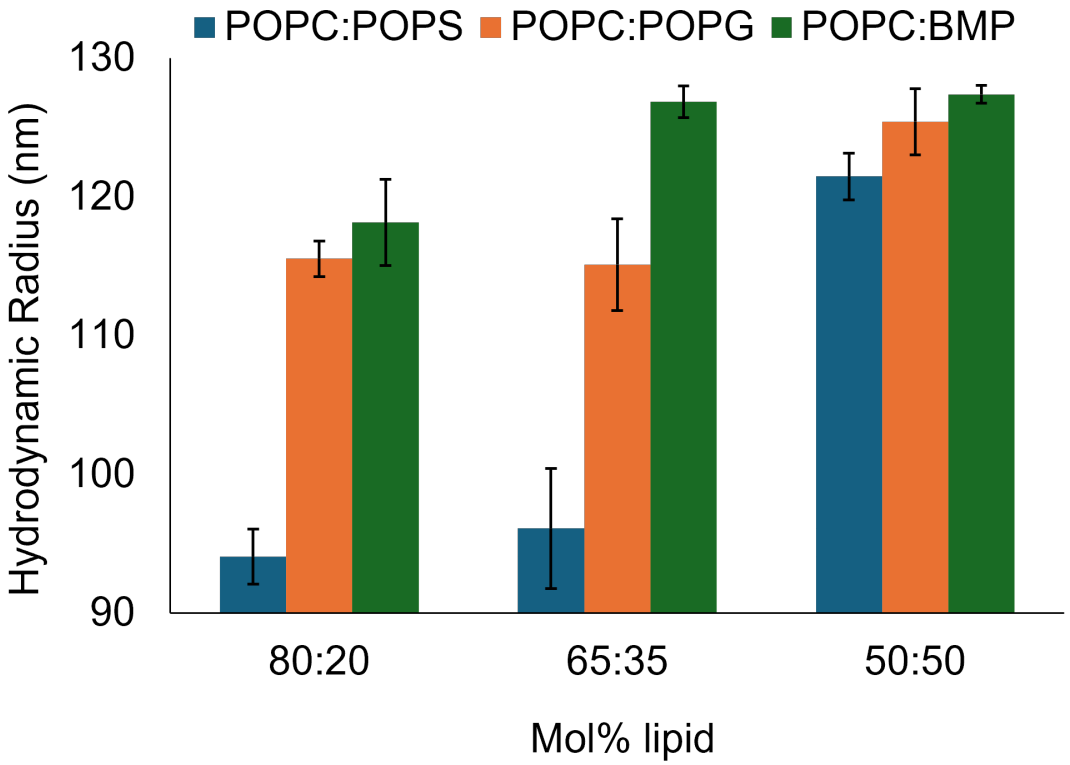

Figure A8 Dynamic light scattering (DLS) measurements for large unilamellar vesicles (LUVs) extruded via 100nm polycarbonate membranes. Data shown is the average peak height and the error bars are the standard deviation.

Bibliography

1. Birtles D, Lee J. SARS-CoV-2 Fusion Domain Provides Clues toward the Molecular Mechanism for Membrane Fusion. Biochemistry. 2023(62):3033–5.
2. Modrow S, Falke D, Truyen U, Schätzl H. Viruses: Definition, Structure, Classification. Molecular Virology: Springer Berlin Heidelberg; 2013. p. 17-30.
3. Tyrrell DA, Bynoe ML. Cultivation of viruses from a high proportion of patients with colds. Lancet. 1966;1(7428):76-7.
4. Almeida JD, Tyrrell DA. The morphology of three previously uncharacterized human respiratory viruses that grow in organ culture. Journal of General Virology. 1967;1(2):175-8.
5. Hamre D, Procknow JJ. A new virus isolated from the human respiratory tract. Proceedings of the Society for Experimental Biology and Medicine. 1966;121(1):190-3.
6. McIntosh K, Dees JH, Becker WB, Kapikian AZ, Chanock RM. Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. Proceedings of the National Academy of Sciences. 1967;57(4):933-40.
7. Tyrrell DA, Almeida JD, Cunningham CH, Dowdle WR, Hofstad MS, McIntosh K, et al. Coronaviridae. Intervirology. 1975;5(1-2):76-82.
8. Bradburne AF. Antigenic relationships amongst coronaviruses. Archiv für die gesamte Virusforschung. 1970;31(3-4):352-64.
9. Mcintosh K, Kapikian AZ, Turner HC, Hartley JW, Parrott RH, Chanock RM. Seroepidemiologic studies of coronavirus infection in adults and children. American Journal of Epidemiology. 1970;91(6):585-92.
10. Bradburne AF, Somerset BA. Coronavirus antibody titres in sera of healthy adults and experimentally infected volunteers. Journal of Hygiene. 1972;70(2):235-44.
11. Coronaviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ, editors. Virus Taxonomy2011. p. 806-28.
12. Wu Z, Yang L, Ren X, He G, Zhang J, Yang J, et al. Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. The ISME Journal. 2016;10(3).
13. WHO. Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003 2015 [Available from: https://www.who.int/publications/m/item/summary-of-probable-sars-cases-with-onset-of-illness-from-1-november-2002-to-31-july-2003.
14. Lau SKP, Woo PCY, Li KSM, Huang Y, Tsoi H-W, Wong BHL, et al. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proceedings of the National Academy of Sciences. 2005;102(39):14040-5.
15. WHO. Severe Acute Respiratory Syndrome (SARS) 2024 [Available from: https://www.who.int/health-topics/severe-acute-respiratory-syndrome.
16. Zaki AM, Van Boheemen S, Bestebroer TM, Osterhaus ADME, Fouchier RAM. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. New England Journal of Medicine. 2012;367(19):1814-20.
17. Azhar EI, El-Kafrawy SA, Farraj SA, Hassan AM, Al-Saeed MS, Hashem AM, et al. Evidence for Camel-to-Human Transmission of MERS Coronavirus. New England Journal of Medicine. 2014;370(26):2499-505.
18. Pustake M, Tambolkar I, Giri P, Gandhi C. SARS, MERS and CoVID-19: An overview and comparison of clinical, laboratory and radiological features. Journal of Family Medicine and Primary Care. 2022;11(1).
19. WHO. MERS situation update, March 2024 2024 [Available from: https://www.emro.who.int/health-topics/mers-cov/mers-outbreaks.html.
20. Ahmadzadeh J, Mobaraki K. Epidemiological status of the Middle East respiratory syndrome coronavirus in 2019: an update from January 1 to March 31, 2019. International Journal of General Medicine. 2019;12:305-11.
21. Mehand MS, Al-Shorbaji F, Millett P, Murgue B. The WHO R&D Blueprint: 2018 review of emerging infectious diseases requiring urgent research and development efforts. Antiviral Research. 2018;159:63-7.
22. WHO. Situation Report - 181 2020 [Available from: https://www.who.int/publications/m/item/situation-report---181.
23. WHO. WHO COVID-19 Database 2024 [Available from: https://data.who.int/dashboards/covid19/deaths?n=o.
24. Manathunga SS, Abeyagunawardena IA, Dharmaratne SD. A comparison of transmissibility of SARS-CoV-2 variants of concern. Virology Journal. 2023;20(1):59.

25. Warwick McKibbin RF. The global economic impacts of the COVID-19 pandemic. Economic Modelling. 2023;129.
26. Jackson LA, Anderson EJ, Rouphael NG, Roberts PC, Makhene M, Coler RN, et al. An mRNA Vaccine against SARS-CoV-2 — Preliminary Report. New England Journal of Medicine. 2020;383(20):1920-31.
27. Walsh EE, Frenck RW, Falsey AR, Kitchin N, Absalon J, Gurtman A, et al. Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates. New England Journal of Medicine. 2020;383(25):2439-50.
28. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. New England Journal of Medicine. 2020;383(27):2603-15.
29. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. New England Journal of Medicine. 2021;384(5):403-16.
30. 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. Molecular Cell. 2020;78(4):779-84.e5.
31. Bosch BJ, Zee Rvd, Haan CAMd, Rottier PJM. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. J Virol. 2003;77(16):8801-11.
32. Verdecchia P, Cavallini C, Spanevello A, Angeli F. The pivotal link between ACE2 deficiency and SARS-CoV-2 infection. European Journal of Internal Medicine. 2020;76:14-20.
33. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V, V'kovski P, et al. Coronavirus biology and replication: implications for SARS-CoV-2. Nature Reviews Microbiology. 2020;19(3).
34. Polatoğlu I, Oncu-Oner T, Dalman I, Ozdogan S. COVID-19 in early 2023: Structure, replication mechanism, variants of SARS-CoV-2, diagnostic tests, and vaccine & drug development studies. MedComm. 2023;4(2).
35. Zandi M, Shafaati M, Kalantar-Neyestanaki D, Pourghadamyari H, Fani M, Soltani S, et al. The role of SARS-CoV-2 accessory proteins in immune evasion. Biomedicine & Pharmacotherapy. 2022;156.
36. Takeda M. Proteolytic activation of SARS-CoV-2 spike protein. Microbiology and Immunology. 2022;66(1):15-23.
37. Pinto AL, Rai RK, Brown JC, Griffin P, Edgar JR, Shah A, et al. Ultrastructural insight into SARS-CoV-2 entry and budding in human airway epithelium. Nature Communications. 2022;13(1).
38. White JM, Delos SE, Brecher M, Schornberg K. Structures and Mechanisms of Viral Membrane Fusion Proteins: Multiple Variations on a Common Theme. Critical Reviews in Biochemistry and Molecular Biology. 2008;43(4):189-219.
39. Li Y, Tamm LK. Structure and plasticity of the human immunodeficiency virus gp41 fusion domain in lipid micelles and bilayers. Biophysical Journal. 2007;93(3):876-85.
40. Wilson IA, Skehel JJ, Wiley DC. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. Nature. 1981;289(5796):366-73.
41. 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. Science Advances. 2021;7(41):eabk2226.
42. Shi W, Cai Y, Zhu H, Peng H, Voyer J, Rits-Volloch S, et al. Cryo-EM structure of SARS-CoV-2 postfusion spike in membrane. Nature. 2023;619(7969):403-9.
43. Peisajovich SG, Epand RF, Pritsker M, Shai Y, Epand RM. The polar region consecutive to the HIV fusion peptide participates in membrane fusion. Biochemistry. 2000;39(7):1826-33.
44. Gray C, Tatulian SA, Wharton SA, Tamm LK. Effect of the N-terminal glycine on the secondary structure, orientation, and interaction of the influenza hemagglutinin fusion peptide with lipid bilayers. Biophysical Journal. 1996;70(5):2275-86.
45. 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. Journal of Virology. 2004;78(7):3753-62.
46. 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. Journal of Virology. 2014;88(12):6636-49.
47. Birtles D, Lee J. Identifying Distinct Structural Features of the SARS-CoV-2 Spike Protein Fusion Domain Essential for Membrane Interaction. Biochemistry. 2021;60(40):2978-86.
48. White JM, Ward AE, Odongo L, Tamm LK. Viral Membrane Fusion: A Dance Between Proteins and Lipids. Annual Review of Virology. 2023;10(1):139-61.
49. 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. Journal of Molecular Biology. 2012;418(1-2):3-15.

50. 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):265-71.
51. 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. Journal of Virology. 2009;83(15):7411-21.
52. 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. Journal of Virology. 2005;79(11):7195-206.
53. Guillén J, Pérez-Berná AJ, Moreno MR, Villalaín J. A Second SARS-CoV S2 Glycoprotein Internal Membrane-Active Peptide. Biophysical Characterization and Membrane Interaction. Biochemistry. 2008;47(31):8214-24.
54. 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. Journal of Molecular Biology. 2017;429(24):3875-92.
55. Wang Q, Zhang Y, Wu L, Niu S, Song C, Zhang Z, et al. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. Cell. 2020;181(4):894-904.
56. 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(6485):1444-8.
57. Zhou T, Tsybovsky Y, Gorman J, Rapp M, Cerutti G, Chuang GY, 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):867-79.e5.
58. Skehel JJ, Wiley DC. Coiled Coils in Both Intracellular Vesicle and Viral Membrane Fusion. Cell. 1998;95(7):871-4.
59. Belouzard S, Millet JK, Licitra BN, Whittaker GR. Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. Viruses. 2012;4(6):1011-33.
60. Elbe S, Buckland-Merrett G. Data, disease and diplomacy: GISAID's innovative contribution to global health. Global Challenges. 2017;1(1):33-46.
61. Dimitrov AS, Rawat SS, Jiang S, Blumenthal R. Role of the Fusion Peptide and Membrane-Proximal Domain in HIV-1 Envelope Glycoprotein-Mediated Membrane Fusion. Biochemistry. 2003;42(48):14150-8.
62. Han X, Bushweller JH, Cafiso DS, Tamm LK. Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. Nature Structural Biology. 2001;8(8):715-20.
63. Gething MJ, Doms RW, York D, White J. Studies on the mechanism of membrane fusion: site-specific mutagenesis of the hemagglutinin of influenza virus. Journal of Cellular Biology. 1986;102(1):11-23.
64. 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. Proceedings of the National Academy of Sciences. 2011;108(27):11211-6.
65. Cai Y, Zhang J, Xiao T, Peng H, Sterling SM, Walsh RM, et al. Distinct conformational states of SARS-CoV-2 spike protein. Science. 2020;369(6511):1586-92.
66. Lai AL, Freed JH. SARS-CoV-2 Fusion Peptide has a Greater Membrane Perturbating Effect than SARS-CoV with Highly Specific Dependence on Ca2+. Journal of Molecular Biology. 2021;433(10):166946.
67. Epand RM. Fusion peptides and the mechanism of viral fusion. Biochimica et Biophysica Acta. 2003;1614(1):116-21.
68. Wishart DS, Sykes BD, Richards FM. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry. 1992;31(6):1647-51.
69. Sharma D, Rajarathnam K. 13C NMR chemical shifts can predict disulfide bond formation. Journal of Biomolecular NMR. 2000;18(2):165-71.
70. Smrt ST, Draney AW, Lorieau JL. The Influenza Hemagglutinin Fusion Domain Is an Amphipathic Helical Hairpin That Functions by Inducing Membrane Curvature. Journal of Biological Chemistry. 2015;290(1):228-38.
71. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell. 2020;181(2):281-92.e6.
72. Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. Science. 2020;369(6501):330-3.
73. Grant OC, Montgomery D, Ito K, Woods RJ. Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. Scientific Reports. 2020;10(1):14991.

74. Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, et al. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. Nature Communications. 2020;11(1):5885.
75. Woo H, Park S-J, Choi YK, Park T, Tanveer M, Cao Y, et al. Developing a Fully Glycosylated Full-Length SARS-CoV-2 Spike Protein Model in a Viral Membrane. The Journal of Physical Chemistry B. 2020;124(33):7128-37.
76. Zhang X-Y, Guo J, Wan X, Zhou J-G, Jin W-P, Lu J, et al. 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 & Infections. 2020;9(1):2653-62.
77. Lai AL, Tamm LK. Locking the kink in the influenza hemagglutinin fusion domain structure. Journal of Biological Chemistry. 2007;282(33):23946-56.
78. Takeda M, Leser GP, Russell CJ, Lamb RA. Influenza virus hemagglutinin concentrates in lipid raft microdomains for efficient viral fusion. Proceedings of the National Academy of Sciences. 2003;100(25):14610-7.
79. Nathan L, Lai AL, Millet JK, Straus MR, Freed JH, Whittaker GR, et al. Calcium Ions Directly Interact with the Ebola Virus Fusion Peptide To Promote Structure–Function Changes That Enhance Infection. ACS Infectious Diseases. 2020;6(2):250-60.
80. 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. Biochimica et Biophysica Acta (BBA) - Biomembranes. 2018;1860(2):407-15.
81. 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. Biochimica et Biophysica Acta. 2015;1848(2):721-30.
82. Guillén J, Kinnunen PKJ, Villalaín J. Membrane insertion of the three main membranotropic sequences from SARS-CoV S2 glycoprotein. Biochimica et Biophysica Acta. 2008;1778(12):2765-74.
83. Khelashvili G, Plante A, Doktorova M, Weinstein H. Ca2+-dependent mechanism of membrane insertion and destabilization by the SARS-CoV-2 fusion peptide. Biophysical Journal. 2021;120(6):1105-19.
84. Gorgun D, Lihan M, Kapoor K, Tajkhorshid E. Binding mode of SARS-CoV-2 fusion peptide to human cellular membrane. Biophysical Journal. 2021;120(14):2914-26.
85. Oh KI, Smith-Dupont KB, Markiewicz BN, Gai F. Kinetics of peptide folding in lipid membranes. Biopolymers. 2015;104(4):281-90.
86. Qiang W, Sun Y, Weliky DP. A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide. Proceedings of the National Academy of Sciences. 2009;106(36):15314-9.
87. Bollavaram K, Leeman TH, Lee MW, Kulkarni A, Upshaw SG, Yang J, et al. Multiple sites on SARS-CoV-2 spike protein are susceptible to proteolysis by cathepsins B, K, L, S, and V. Protein Science. 2021;30(6):1131-43.
88. Roth SL, Whittaker GR. Promotion of vesicular stomatitis virus fusion by the endosome-specific phospholipid bis(monoacylglycero)phosphate (BMP). FEBS Letters. 2011;585(6):865-9.
89. Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD, Butt TR. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. Journal of Structural and Functional Genomics. 2004;5(1/2):75-86.
90. Guerrero F, Ciragan A, Iwaï H. Tandem SUMO fusion vectors for improving soluble protein expression and purification. Protein Expression and Purification. 2015;116:42-9.
91. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. Journal of Biomolecular NMR. 1995;6(3):277-93.
92. Lee W, Tonelli M, Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics. 2015;31(8):1325-7.
93. Maciejewski MW, Schuyler AD, Gryk MR, Moraru II, Romero PR, Ulrich EL, et al. NMRbox: A Resource for Biomolecular NMR Computation. Biophysical Journal. 2017;112(8):1529-34.
94. Kuriashkin IV, Losonsky JM. Contrast enhancement in magnetic resonance imaging using intravenous paramagnetic contrast media: A review. Veterinary Radiology. 2000;41(1):4-7.
95. Han X, Tamm LK. A host-guest system to study structure-function relationships of membrane fusion peptides. Proceedings of the National Academy of Sciences. 2000;97(24):13097-102.
96. 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.
97. Pennington HN, Lee J. Lassa virus glycoprotein complex review: insights into its unique fusion machinery. Bioscience Reports. 2022;42(2).

98. Koppisetti RK, Fulcher YG, Van Doren SR. Fusion Peptide of SARS-CoV-2 Spike Rearranges into a Wedge Inserted in Bilayered Micelles. Journal of the American Chemical Society. 2021;143(33):13205-11.
99. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen 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):271-80.e8.
100. White JM, Whittaker GR. Fusion of Enveloped Viruses in Endosomes. Traffic. 2016;17(6):593-614.
101. Ke Z, Oton J, Qu K, Cortese M, Zila V, McKeane L, et al. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. Nature. 2020;588(7838):498-502.
102. Santamaria A, Batchu KC, Matsarskaia O, Prévost SF, Russo D, Natali F, et al. Strikingly Different Roles of SARS-CoV-2 Fusion Peptides Uncovered by Neutron Scattering. Journal of the American Chemical Society. 2022;144(7):2968-79.
103. Lai AL, Freed JH. Negatively charged residues in the membrane ordering activity of SARS-CoV-1 and -2 fusion peptides. Biophysical Journal. 2022;121(2):207-27.
104. Yang S-T, Kiessling V, Simmons JA, White JM, Tamm LK. HIV gp41–mediated membrane fusion occurs at edges of cholesterol-rich lipid domains. Nature Chemical Biology. 2015;11(6):424-31.
105. Biswas S, Yin S-R, Blank PS, Zimmerberg J. Cholesterol Promotes Hemifusion and Pore Widening in Membrane Fusion Induced by Influenza Hemagglutinin. Journal of General Physiology. 2008;131(5):503-13.
106. 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. Biochimica et Biophysica Acta. 2008;1778(4):945-53.
107. Dubé M, Rey FA, Kielian M. Rubella Virus: First Calcium-Requiring Viral Fusion Protein. PLoS Pathogens. 2014;10(12):e1004530.
108. 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. Molecular Membrane Biology. 2010;27(1):31-44.
109. Jackson CB, Farzan M, Chen B, Choe H. Mechanisms of SARS-CoV-2 entry into cells. Nature Reviews Molecular and Cellular Biology. 2022;23(1):3-20.
110. 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. Journal of Virology. 2008;82(6):3131-4.
111. Mingo RM, Simmons JA, Shoemaker CJ, Nelson EA, Schornberg KL, D'Souza RS, 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. Journal of Virology. 2015;89(5):2931-43.
112. Zang R, Case JB, Yutuc E, Ma X, Shen S, Gomez Castro MF, et al. Cholesterol 25-hydroxylase suppresses SARS-CoV-2 replication by blocking membrane fusion. Proceedings of the National Academy of Sciences. 2020;117(50):32105-13.
113. Schaefer SL, Jung H, Hummer G. Binding of SARS-CoV-2 Fusion Peptide to Host Endosome and Plasma Membrane. Journal of Physical Chemistry B. 2021;125(28):7732-41.
114. Lai AL, Park H, White JM, Tamm LK. Fusion Peptide of Influenza Hemagglutinin Requires a Fixed Angle Boomerang Structure for Activity. Journal of Biological Chemistry. 2006;281(9):5760-70.
115. 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):559-62.
116. DK S, D H, RE P. Use of Resonance Energy Transfer to Monitor Membrane Fusion. Biochemistry. 1981;20(14).
117. Miles AJ, Wallace BA. CDtoolX, a downloadable software package for processing and analyses of circular dichroism spectroscopic data. Protein Science. 2018;27(9):1717-22.
118. Birtles D, Abbas W, Lee J. Bis(Monoacylglycero)Phosphate Promotes Membrane Fusion Facilitated by the SARS-CoV-2 Fusion Domain. The Journal of Physical Chemistry B. 2024;128(11).
119. Van Meer G, Voelker DR, Feigenson GW. Membrane lipids: where they are and how they behave. Nature Reviews Molecular Cell Biology. 2008;9(2):112-24.
120. Symons JL, Cho K-J, Chang JT, Du G, Waxham MN, Hancock JF, et al. Lipidomic atlas of mammalian cell membranes reveals hierarchical variation induced by culture conditions, subcellular membranes, and cell lineages. Soft Matter. 2021;17(2):288-97.
121. Kobayashi T, Beuchat MH, Chevallier J, Makino A, Mayran N, Escola JM, et al. Separation and characterization of late endosomal membrane domains. Journal of Biological Chemistry. 2002;277(35):32157-64.

122. Kobayashi T, Beuchat M-H, Lindsay M, Frias S, Palmiter RD, Sakuraba H, et al. Late endosomal membranes rich in lysobisphosphatidic acid regulate cholesterol transport. Nature Cell Biology. 1999;1(2):113-8.
123. Mannsverk S, Villamil Giraldo AM, Kasson PM. Influenza Virus Membrane Fusion Is Promoted by the Endosome-Resident Phospholipid Bis(monoacylglycero)phosphate. The Journal of Physical Chemistry B. 2022;126(49):10445-51.
124. Markosyan RM, Marin M, Zhang Y, Cohen FS, Melikyan GB. The late endosome-resident lipid bis(monoacylglycero)phosphate is a cofactor for Lassa virus fusion. PLOS Pathogens. 2021;17(9):e1009488.
125. Silvius JR. Role of cholesterol in lipid raft formation: lessons from lipid model systems. Biochimica et Biophysica Acta. 2003;1610(2):174-83.
126. Kulkarni R, Wiemer EAC, Chang W. Role of Lipid Rafts in Pathogen-Host Interaction - A Mini Review. Frontiers Immunology. 2021;12:815020.
127. Karl K, Paul MD, Pasquale EB, Hristova K. Ligand bias in receptor tyrosine kinase signaling. Journal of Biological Chemistry. 2020;295(52):18494-507.
128. Clogston JD, Patri AK. Zeta Potential Measurement. Methods in Molecular Biology: Humana Press; 2011. p. 63-70.
129. Sezgin E, Sadowski T, Simons K. Measuring Lipid Packing of Model and Cellular Membranes with Environment Sensitive Probes. Langmuir. 2014;30(27):8160-6.
130. Kim HM, Choo HJ, Jung SY, Ko YG, Park WH, Jeon SJ, et al. A Two-Photon Fluorescent Probe for Lipid Raft Imaging: C-Laurdan. ChemBioChem. 2007;8(5):553-9.
131. Kreutzberger AJB, Sanyal A, Saminathan A, Bloyet L-M, Stumpf S, Liu Z, et al. SARS-CoV-2 requires acidic pH to infect cells. Proceedings of the National Academy of Sciences. 2022;119(38).
132. Bretscher MS. Asymmetrical Lipid Bilayer Structure for Biological Membranes. Nature New Biology. 1972;236(61):11-2.
133. Op Den Kamp JAF. Lipid Asymmetry in Membranes. Annual Review of Biochemistry. 1979;48(1):47-71.
134. Van Doren SR, Scott BS, Koppisetti RK. SARS-CoV-2 fusion peptide sculpting of a membrane with insertion of charged and polar groups. Structure. 2023;31(10):1184-99.e3.
135. Lee J, Kreutzberger AJB, Odongo L, Nelson EA, Nyenhuis DA, Kiessling V, et al. Ebola virus glycoprotein interacts with cholesterol to enhance membrane fusion and cell entry. Nature Structural & Molecular Biology. 2021;28(2):181-9.
136. Lee M, Morgan CA, Hong M. Fully hydrophobic HIV gp41 adopts a hemifusion-like conformation in phospholipid bilayers. Journal of Biological Chemistry. 2019;294(40):14732-44.
137. Niort K, Dancourt J, Boedec E, Al Amir Dache Z, Lavieu G, Tareste D. Cholesterol and Ceramide Facilitate Membrane Fusion Mediated by the Fusion Peptide of the SARS-CoV-2 Spike Protein. ACS Omega. 2023;8(36):32729-39.
138. Sanders DW, Jumper CC, Ackerman PJ, Bracha D, Donlic A, Kim H, et al. SARS-CoV-2 requires cholesterol for viral entry and pathological syncytia formation. eLife. 2021;10.
139. Ding W, Palaiokostas M, Wang W, Orsi M. Effects of Lipid Composition on Bilayer Membranes Quantified by All-Atom Molecular Dynamics. The Journal of Physical Chemistry B. 2015;119(49):15263-74.
140. Ballweg S, Sezgin E, Doktorova M, Covino R, Reinhard J, Wunnicke D, et al. Regulation of lipid saturation without sensing membrane fluidity. Nature Communications. 2020;11(1).
141. Akgoc Z, Sena-Esteves M, Martin DR, Han X, D'Azzo A, Seyfried TN. Bis(monoacylglycero)phosphate: a secondary storage lipid in the gangliosidoses. Journal of Lipid Research. 2015;56(5):1005-6.
142. Hayakawa T, Hirano Y, Makino A, Michaud S, Lagarde M, Pageaux J-F, et al. Differential Membrane Packing of Stereoisomers of Bis(monoacylglycero)phosphate. Biochemistry. 2006;45(30):9198-209.
143. Dickey A, Faller R. Examining the Contributions of Lipid Shape and Headgroup Charge on Bilayer Behavior. Biophysical Journal. 2008;95(6):2636-46.
144. Mukhopadhyay P, Monticelli L, Tieleman DP. Molecular Dynamics Simulation of a Palmitoyl-Oleoyl Phosphatidylserine Bilayer with Na+ Counterions and NaCl. Biophysical Journal. 2004;86(3):1601-9.
145. Holopainen JM, Söderlund T, Alakoskela JM, Säily M, Eriksson O, Kinnunen PK. Intermolecular interactions of lysobisphosphatidic acid with phosphatidylcholine in mixed bilayers. Chemistry and Physics of Lipids. 2005;133(1):51-67.
146. François-Martin C, Bacle A, Rothman JE, Fuchs PFJ, Pincet F. Cooperation of Conical and Polyunsaturated Lipids to Regulate Initiation and Processing of Membrane Fusion. Frontiers in Molecular Biosciences. 2021;8:763115.
147. Zhao H, Piszczek G, Schuck P. SEDPHAT – A platform for global ITC analysis and global multi-method analysis of molecular interactions. Methods. 2015;76:137-48.

148. Scheuermann TH, Brautigam CA. High-precision, automated integration of multiple isothermal titration calorimetric thermograms: New features of NITPIC. Methods. 2015;76:87-98.
149. Lassmann T. Kalign 3: multiple sequence alignment of large datasets. Bioinformatics. 2020;36(6).
150. Mandala VS, Mckay MJ, Shcherbakov AA, Dregni AJ, Kolocouris A, Hong M. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. Nature Structural & Molecular Biology. 2020;27(12):1202-8.
151. Belouzard S, Chu VC, Whittaker GR. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. Proceedings of the National Academy of Sciences. 2009;106(14):5871-6.
152. Parsegian A. Energy of an Ion crossing a Low Dielectric Membrane: Solutions to Four Relevant Electrostatic Problems. Nature. 1969;221(5183):844-6.
153. Hessa T, Kim H, Bihlmaier K, Lundin C, Boekel J, Andersson H, et al. Recognition of transmembrane helices by the endoplasmic reticulum translocon. Nature. 2005;433(7024):377-81.
154. Légaré S, Lagüe P. The influenza fusion peptide promotes lipid polar head intrusion through hydrogen bonding with phosphates and N-terminal membrane insertion depth. Proteins: Structure, Function, and Bioinformatics. 2014;82(9):2118-27.
155. Rice A, Haldar S, Wang E, Blank PS, Akimov SA, Galimzyanov TR, et al. Planar aggregation of the influenza viral fusion peptide alters membrane structure and hydration, promoting poration. Nature Communications. 2022;13(1).
156. Lorin A, Charloteaux B, Fridmann-Sirkis Y, Thomas A, Shai Y, Brasseur R. Mode of Membrane Interaction and Fusogenic Properties of a de Novo Transmembrane Model Peptide Depend on the Length of the Hydrophobic Core. Journal of Biological Chemistry. 2007;282(25):18388-96.
157. Strandberg E, Killian JA. Snorkeling of lysine side chains in transmembrane helices: how easy can it get? FEBS Letters. 2003;544(1-3):69-73.
158. de Freitas-Silva L, Rodríguez-Ruiz M, Houmani H, da Silva LC, Palma JM, Corpas FJ. Glyphosate-induced oxidative stress in Arabidopsis thaliana affecting peroxisomal metabolism and triggers activity in the oxidative phase of the pentose phosphate pathway (OxPPP) involved in NADPH generation. Journal of Plant Physiology. 2017;218:196-205.
159. Willcott MR. MestRe Nova. Journal of the American Chemical Society. August 24, 2009;131(36).
160. Brautigam CA. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. Methods in Enzymology. 2015;562.