STRUCTURE AND FUNCTION STUDY OF HIV AND INFLUENZA FUSION PROTEINS

By

Shuang Liang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry—Doctor of Philosophy

2017

ABSTRACT

STRUCTURE AND FUNCTION STUDY OF HIV AND INFLUENZA FUSION PROTEINS

By

Shuang Liang

Human immunodeficiency virus (HIV) and influenza virus are membrane-enveloped viruses causing acquired immunodeficiency syndrome (AIDS) and flu. The initial step of HIV and influenza virus infection is fusion between viral and host cell membrane catalyzed by the viral fusion protein gp41 and hemagglutinin (HA) respectively. However, the structure of gp41 and HA as well as the infection mechanism are still not fully understood. This work addresses (1) full length gp41 ectodomain and TM domain structure and function and (2) IFP membrane location and IFP-membrane interaction. My studies of gp41 protein and IFP can provide better understanding of the membrane fusion mechanism and may aid development of anti-viral therapeutics and vaccine.

The full length ectodomain and transmembrane domain of gp41 and shorter constructs were expressed, purified and solubilized at physiology conditions. The constructs adopt overall α helical structure in SDS and DPC detergents, and showed hyperthermostability with T_m > 90 °C. The oligomeric states of these proteins vary in different detergent buffer: predominant trimer for all constructs and some hexamer fraction for HM and HM_TM protein in SDS at pH 7.4; and mixtures of monomer, trimer, and higher-order oligomer protein in DPC at pH 4.0 and 7.4. Substantial protein-induced vesicle fusion was observed, including fusion of neutral vesicles at neutral pH, which are the conditions similar HIV/cell fusion. Vesicle fusion by a gp41 ectodomain construct has rarely been observed under these conditions, and is aided by inclusion of both the FP

and TM, and by protein which is predominantly trimer rather than monomer. Current data was integrated with existing data, and a structural model was proposed.

Secondary structure and conformation of IFP is a helix-turn-helix structure in membrane. However, there has been arguments about the IFP membrane location. ^{13}C-^{2}H REDOR solid-state NMR is used to solve this problem. The IFP adopts major α helical, minor β strand secondary structure in PC/PG membrane. The α helical IFP's with respectively ^{13}CO labeled Leu-2, Ala-7 and Gly-16 all show close contacts with the lipid acyl chain tail, suggesting IFP has strong interaction with the membrane. By screening the current IFP topology models, it either has a membrane-spanning confirmation, or it promotes lipid trail protrusion.

IFP bounded lipid membrane structure was studied by paramagnetic relaxation enhancement (PRE) solid-state NMR to provide more information about the detailed IFP membrane location model. The T_2 relaxation time and rate were measured for membrane with or without IFP and with or without Mn^{2+}. Based on the results, it is concluded that IFP does not promote lipid protrusion at both gel phase and liquid phase, which is evidenced by that the R_2 difference with and without Mn^{2+} is smaller for IFP free membrane than IFP bounded membrane, meaning IFP does not induce a smaller average distance between lipid acyl chain and aqueous layer. By integrating these results, a IFP membrane spanning model was proposed, in which IFP N-terminal helix adopts a 45° angle with respect to membrane normal.

Dedicated to mom, dad and Luyao

ACKNOWLEDGEMENTS

I would like to take this chance to thank my advisor Dr. David P Weliky for giving me a chance to join this wonderful group, teaching me incredible things not only about science but also about life, and being patient and always helpful. I want to thank him for all the help, suggestion, support and courage he gave me when I faced any difficulties. He taught me NMR and biological knowledge, techniques and, most importantly, learning and problem-solving abilities in the past five years. I also want to thank my committee members: Prof. John McCracken, Prof. Heedeok Hong and Prof. Daniel Jones, for their guidance, support and valuable suggestions.

I am appreciative that I can join the Weliky group, thank Dr. Li Xie, Dr. Koyeli Banerjee, Dr. Punsisi Ratnayake, Dr. Ujjayini Ghosh, Lihui Jia, Robert Wolfe and Ahinsa Ranaweera for not only being colleagues that help and support me in my research, but also as my good friends that we share happiness and sorrow. I am also appreciative to have the Max T Rogers NMR facility staffs Dr. Dan Holmes and Dr. Li Xie. Thank them for all the help and suggestions whenever I have any NMR issues or questions throughout my graduate school.

I am grateful to for everybody in Hong's lab and Geiger's lab. Thank them for all the help, valuable suggestions, and for sharing instrument. I also got to know so many good friends in Michigan State University. I will never forget all the joy we had together and thank them for accompanying to be a better and mature person.

At the end, I want to express my gratefulness for my family. Thank my mom and dad for all the care, support and encouragement to let me study in the United States. Special thanks for my husband Luyao for all his love and the wonderful moments that we spent together. Word cannot express my love for my family and this love will last forever.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

AA	Amino acid
AB	Aqueous binding
AIDS	Acquired immunodeficiency syndrome
AUC	Analytical ultracentrifugation
CD	Circular dichroism
CDC	Centers of Disease Control and Prevention
Chol	Cholesterol
CHR	C-heptad repeat
CP	Cross polarization
DCM	Dichloromethane
DEPBT	3-(Diethylphosphoryloxy)-1,2,3-benzotriazin-4(3H)-one
DIEA	N,N-Diisopropylethylamine
DM	n-Decyl- β-D-Maltoside
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DPC	n-Dodecylphosphocholine
DPPG	1,2-Dipalmitoyl-sn-glycero-3-phosphocholine
DTPC	1,2-di-O-tetradecyl-sn-glycero-3-phosphocholine
E.coli	Escherichia coli
Endo	Endodomain
EPR	Electron paramagnetic resonance

FID	Free induction decay
Fmoc	Fluorenylmethyloxycarbonyl
FP	Fuison peptide
FWHM	Full width at half maximum
GuHCl	Guanidinium chloride
HA	Hemagglutinin
HBTU	O-(benzotriazol-1-yl)-N,N,N’,N’-tetramethyluronium hexafluorophosphate
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HIV	Human immunodeficiency virus
HOBt	1-Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
IFP	Influenza fusion peptide
IPTG	Isopropyl β-D-thiogalactopyranoside
LB	Luria-Bertani broth
MAS	Magic-angle spinning
MD	Molecular dynamics
MES	2-(N-morpholino)ethanesulfonic acid
MPAA	4-Mercaptophenylacetic acid
MPER	Membrane-proximal external region
NA	Neuraminidase
NDB-PE	*N*-(7-Nitrobenz-2-oxa-1,3-diazol-4-yl)-1,2-dihexadecanoyl-*sn*-glycero-3-phosphoethanolamine
NHR	N-heptad repeat
NMR	Nuclear magnetic resonance

NOE	Nuclear Overhauser effect
OC	Organic co-solubilization
PBS	Phosphate-buffered saline
POPC	1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPG	1-Palmitoyl-2-oleoyl-sn-glycero-3-[phospho-rac-(1-glycerol)]
PRE	Paramagnetic relaxation enhancement
Quecho	Quadrupolar echo
REDOR	Rotational-echo double-resonance
Rh-PE	N-(Lissamine rhodamine b sulfonyl)-1,2-dihexadecanoyl-*sn*-glycero-3-phosphoethanolamine
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
RP	Recombinant protein
Sarkosyl	Sodium lauryl sarcosinate
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SHB	Six helical bundle
SPR	Surface plasmon resonance
t-Boc	tert-Butyloxycarbonyl
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TM	Transmembrane

vRNP	Viral ribonucleoproteins
WHO	World Health Organization
DTPG	1,2-di-O-tetradecyl-sn-glycero-3-phospho-(1'-rac-glycerol)

Chapter 1

Introduction to HIV gp41 and influenza hemagglutinin fusion proteins

1.1 Introduction to HIV gp41 fusion protein

1.1.1 HIV virus structure and its infection pathway

The Acquired Immunodeficiency Syndrome (AIDS) is a disease that causes human immune system failure and it is caused by the Human Immunodeficiency Virus (HIV).[1] HIV/AIDS is one of the most significant infectious disease challenges in the world. According to World Health Organization (WHO), there were approximately 36.7 million people living with HIV and 1.1 million people died from HIV related diseases in 2015. There is no cure or vaccine for HIV. The HIV antiretroviral drugs can suppress the viral replication in human and allow human immune system to recover its capability to defend infection. However, only 46% of AIDS patients are using these antiretroviral drugs and the cost of the drug is around $20,000 per year.

HIV virus has two types: HIV-1 and HIV-2, and the current studies mainly focus on the HIV type 1 virus. In the following contents in this dissertation, HIV refers to HIV type 1 virus. HIV virus has a diameter about 100-120 nm and its structure is shown in Figure 1.1.[2] HIV is in the genus Lentivirus, which is part of the family Retroviridae. HIV virus is an enveloped virus, meaning the virion is wrapped by lipid membrane. There are ~72 "knob" on the viral membrane, which are the viral glycoprotein trimers.[3] These proteins are cleaved inside the host cell from a precursor called gp160 into two noncovalently bonded glycoproteins: external surface protein gp120 and transmembrane protein gp41 which are responsible for virus binding to host cell membrane and promoting viral and host cell membrane fusion, respectivley.[4] The matrix proteins forms a sphere underneath membrane to ensure the viral structure and integrity of the virion. There are two identical positive single-stranded RNAs enclosed by a conical shaped capsid. The reverse transcriptase and nucleocapsid proteins are also closely associated to the RNA.[3, 5, 6]

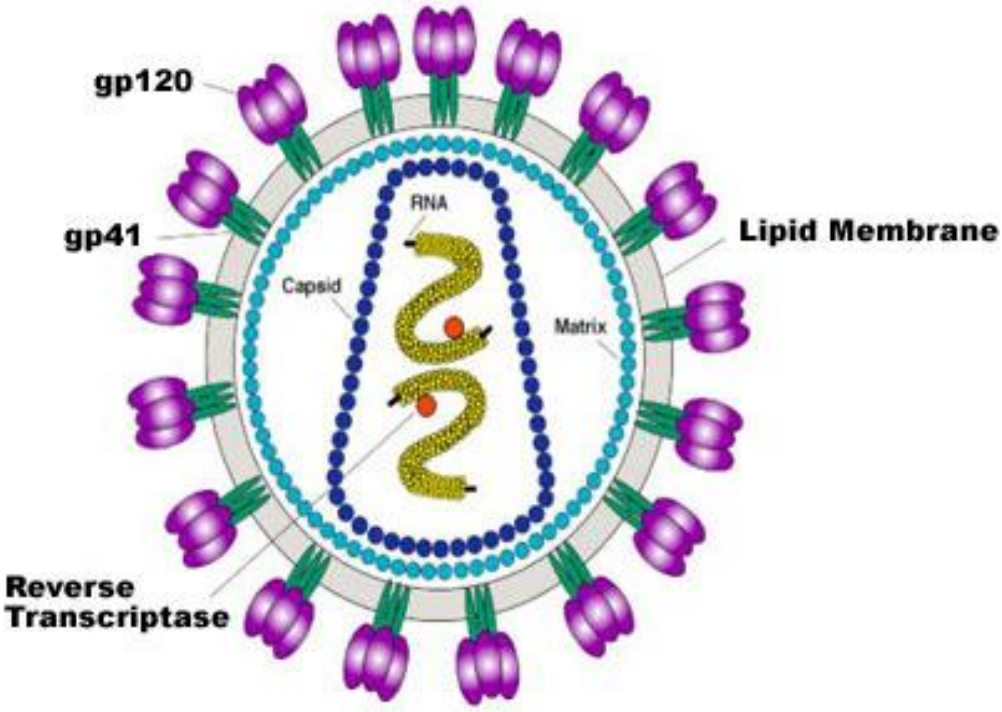

Figure 1.1 Cutaway schematic of the structure of an HIV virion.

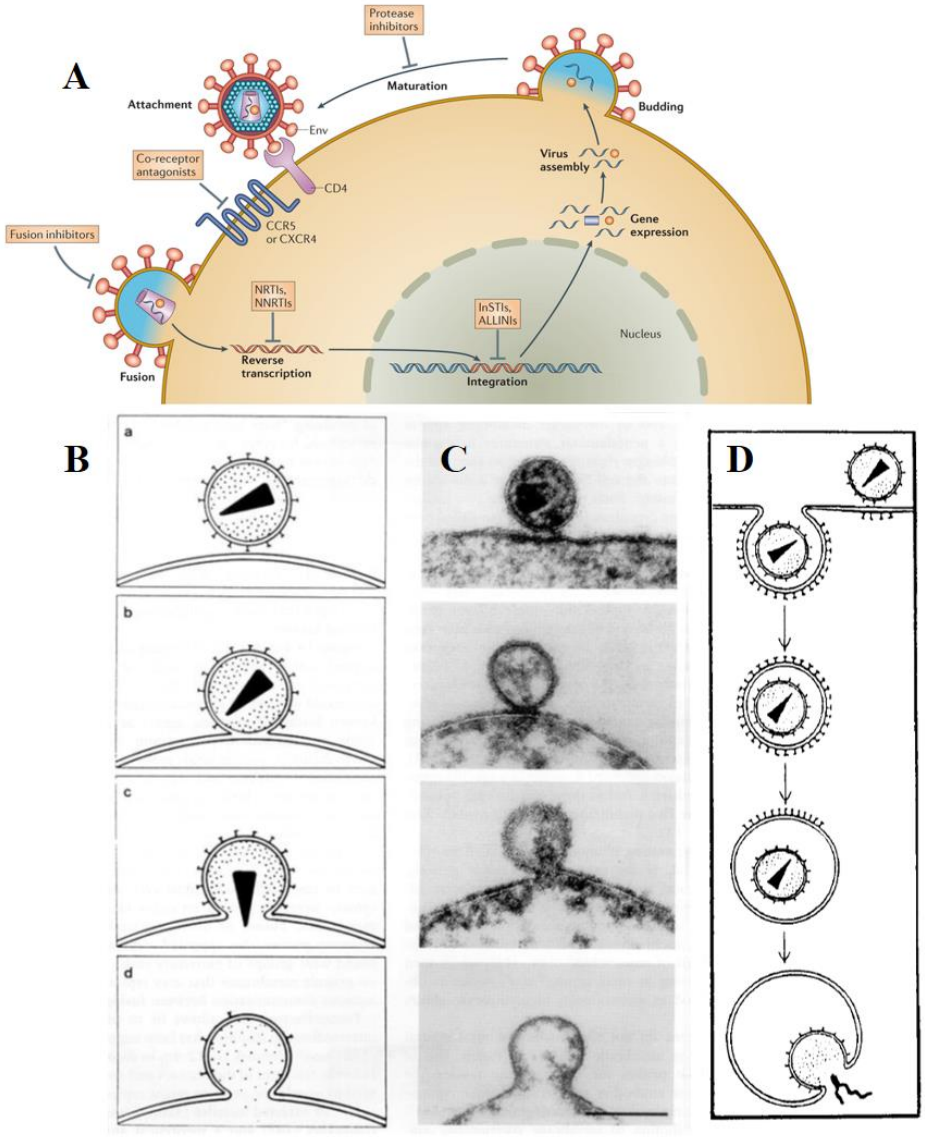

Figure 1.2 (A)Stages of the HIV-1 life cycle. The yellow boxes in the diagram show the possible antiretroviral therapy mechanism.[7] (B) Schematic infection model of HIV. The spikes on virion are gp120/gp41 complex. (a) Binding of HIV on host cell, (b) Hemifusion of viral and host cell membrane, (c) Pore formation, and (d) complete fusion and entry of viral genetic material into the host cell. (C) Electron microscopy pictures of HIV infection process correspond to panel B. [8] (D) Schematic infection model of an alternative HIV entry pathway via clathrin-mediated endocytosis.[8, 9]

The infection cycle of HIV is shown in Figure 1.2 panel A. This process is initiated by the binding of gp120 onto the host cell. HIV infects vital cells in the human immune system such as $CD4^+$ T cells and macrophages.[3, 10, 11] The gp120 typically binds to the CD4 molecules on the host cell where gp120 and CD4 forms a cavity-laden interface and this binding can cause structural changes in gp120. This conformation change facilitates the coreceptor, such as CXCR4 and CCR5, which are in the chemokine receptor family, to bind to a conserved binding site on gp120.[12, 13] After the binding, gp120 is displaced and results in the exposure of gp41.The function of gp41 is to catalysis the fusion between viral and host cell membrane and during the fusion process a large conformational change happens.[14, 15] At the end of the fusion, a fusion pore is formed and the nucleoprotein complexes can get into the host cell. Figure 1.2 panel B and C are the schematic diagram and electron microscopy pictures of the fusion process.[8] On the other hand, endocytosis is another alternative entry pathway for HIV. Electron microscopy results showed that HIV can be endocytosed via clathrin-coated pits (Figure 1.2 D).[8, 9] After fusion, the RNA is reverse transcribed into DNA catalyzed by reverse transcriptase and then the DNA can get into the host cell nucleus and be integrated into the host cell genome. Then the viral proteins can be expressed and finally new viruses bud from the host cell membrane and start the new infection cycle.[16] The antiretroviral drugs are designed based on the virus replication process. To prevent the drug resistance, antiretroviral drugs are usually administered in combination and help the patient to live a near-normal life span.[7]

1.1.2 Structure of gp41

The HIV envelop glycoproteins forms a “knob” shape complex on the viral membrane in a noncovalent manner where three molecules of gp120 form the “cap” and three molecules of gp41

form the "stem" (as shown in Figure 1.2 on the virus surface).[7, 14] The function of gp120 includes the target cell binding while gp41 fuses the viral and host cell membrane to release the RNA into host cell. The gp120 and gp41 are the product from a precursor gp160 which is expressed from the viral gene and be cleaved in the host cell .[3] Since gp41 plays a key role in membrane fusion, it has been a target protein in HIV vaccine and antiretroviral drug development. Gp41 is a transmembrane protein with 345 amino acids. The schematic diagrams of full-length HIV gp41 and the corresponding amino acid sequence are shown in Figure 1.3. Numerous of previous work illustrates that the gp41 undergoes a huge conformational change before and after fusion. X-ray crystallography and cryo-electron microscopy are used to study the structure of gp41 and big progresses has been made for reveal the structure and its possible fusion mechanism.[14, 15, 17-19]

Chan and his coworkers[14] found out the high resolution X-ray crystallographic structure using a protein-dissection approach by which they were able to obtain a water soluble, highly stable complex consists of two peptide fragments denoted N36 and C34. The N36 peptide corresponds to residue 546 – 581, which is in the NHR region, while C34 peptide are residue 628 – 661 and it is in the CHR region. The complex is a six-helical bundle (SHB) whose center consists of a parallel trimeric coiled-coil of three N36 helices and three C34 helices wrap antiparallel outside of the center trimer (Figure 1.4). This complex has a ~35 Å diameter and ~55 Å height. The N36 coiled-coil adopts the "knobs-into-holes" packing, which allows β-branched residues in the complex center to pack into the cavities in the nearby helices. The three C34 helices pack into the highly conserved and hydrophobic grooves on the N36 trimer surface. The analog single chain protein N34(L6)C28 (residue 546 – 579 in NHR, six residue linker SGGRGG and residue 628 – 655 in CHR) also has confirmed the formation of α-helix trimer and showed high termostability.[20] This SHB structure is believed to be the structure of gp41 at post-fusion state.[21, 22]

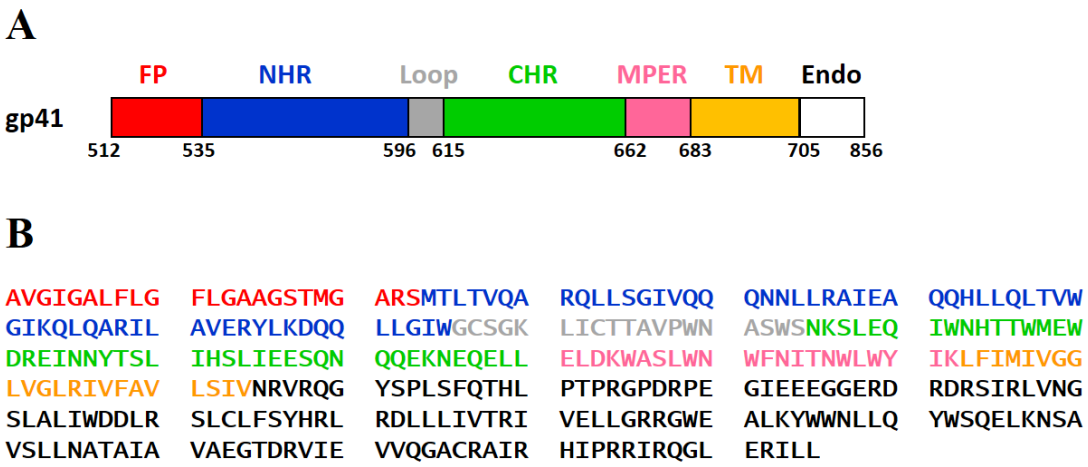

Figure 1.3 (A) Schematic diagrams of full-length HIV gp41 and corresponding colors: FP ≡ fusion peptide, red; NHR ≡ N-helix region, blue; Loop, grey; CHR ≡ C-helix region, green; MPER ≡ membrane-proximal external-region, pink; TM ≡ transmembrane domain, orange; and endo ≡ endodomain, white. (B) Amino acid sequences with colors matching segments in panel A except for the endo domain in black. The sequence is from the HXB2 laboratory strain of HIV and have the gp160 precursor residue numbering, 1-511 for gp120 and 512-856 for gp41.

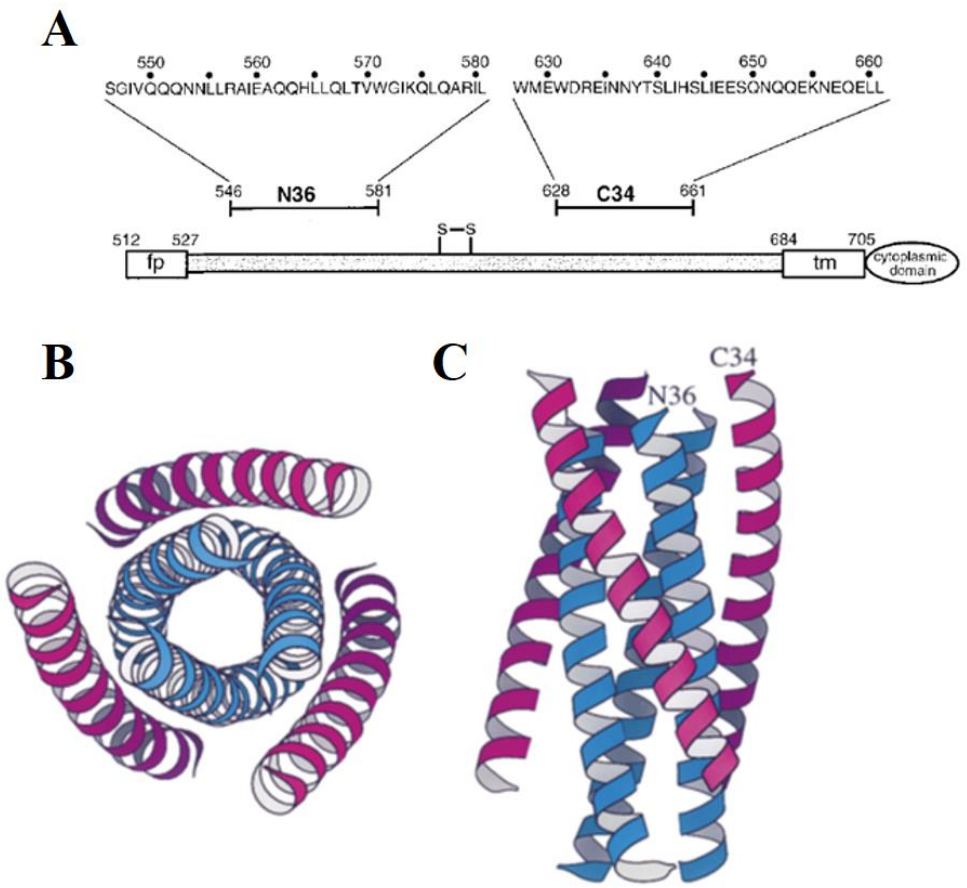

Figure 1.4 (A)Sequence of the N36/C34 complex. (B) The end-on view of the N36/C34 complex looking down the three-fold axis of the trimer. The N36 residues are in blue and the C34 residues are in red. (C)The side view of the N36/C34 complex. The amino termini of the N36 helices point toward the top of the page, while those of the C34 helices point toward the bottom.[14]

Other researchers[15, 19, 22] studied the structure of soluble, cleaved HIV envelope glycoprotein trimer by X-ray crystallography and cryo-electron microscopy with ~ 5 Å resolution. In these studies, the gp120/gp41 complex were obtained from HIV virion and stabilized by antibodies, and the structure was at a prefusion state of gp41. The schematic diagram and the structure for the complex is shown in Figure 1.5. In this study[15], gp41 NHR formed a ~50 Å helix trimer bundle and an additional short helix extended but was kinked away from the trimer axis. CHR had a ~60° angle relative to NHR and it wrapped around central trimer bottom (Figure 1.5 B and C). This NHR/CHR interaction in pre-fusion state was completely different compare to the post fusion state, where NHR and CHR formed closely packed SHB.[14] Similar structures were observed from cryo-electron microscopy results.[19, 22] The difference of the NHR coiled-coil structure in pre- and post- fusion stated is shown Figure 1.6.[22] There was a 15° angle change in each N-terminal helices from the pre-fusion state (Figure 1.6 A) to the post-fusion state (Figure 1.6 B). In addition, in the post-fusion state, the N-terminal helices became more compactly packed and the NHR/CHR SHB was formed.

Although the X-ray crystallography and cryo-electron microscopy results illustrated a lot details about the core domain of gp41 before and after fusion, the membrane-associate MPER and TM domain were truncated in gp140 proteins, and the FP structure was missing.

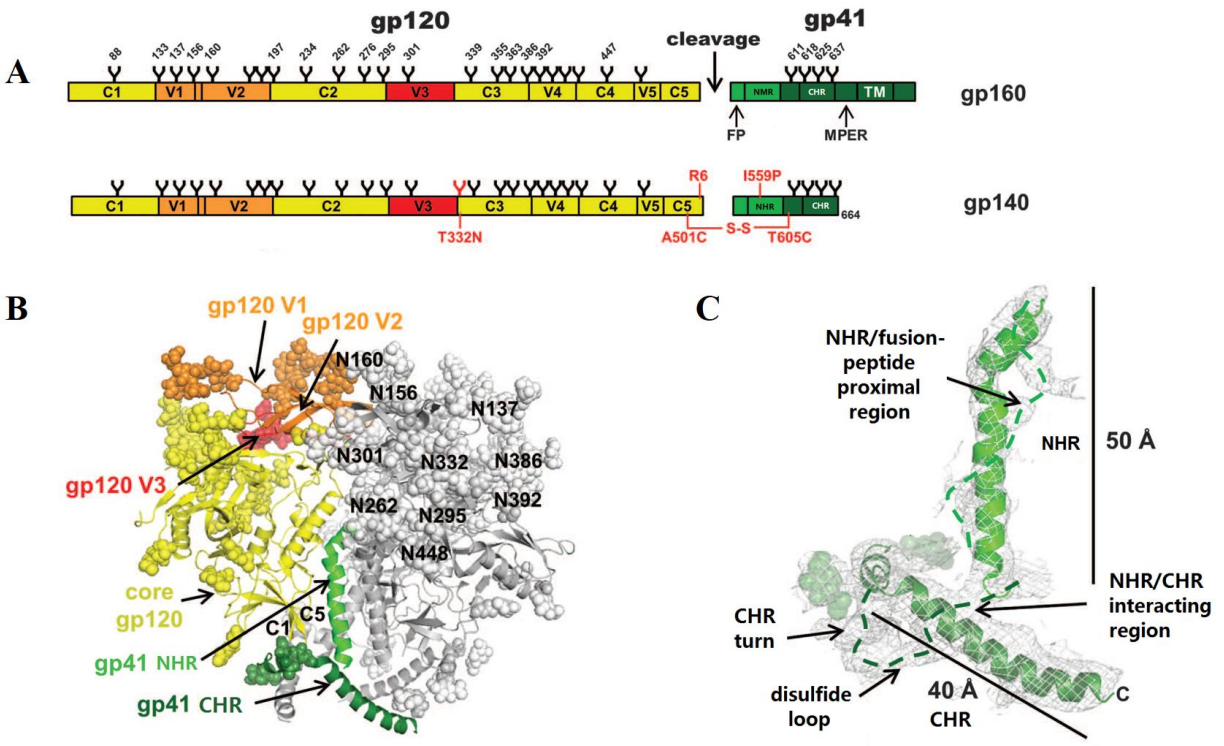

Figure 1.5 (A) Schematic of the HIV gp140 construct studied[15] in comparison to full-length gp160. N-linked glycans are shown and numbered on their respective Asn residues. The FP, NHR, CHR, MPER, transmembrane (TM), and endodomain elements in gp41 are indicated. The mutations are shown in red, as well as the added N332 glycan site. The color coding is preserved in (B) and (C). (B) Side view of the gp140 trimer. The main domains are labeled correspond to panel A and glycans are shown as spheres. (C) gp41 in gp140 complex. Secondary structure determination was ambiguous at the dashed line areas.

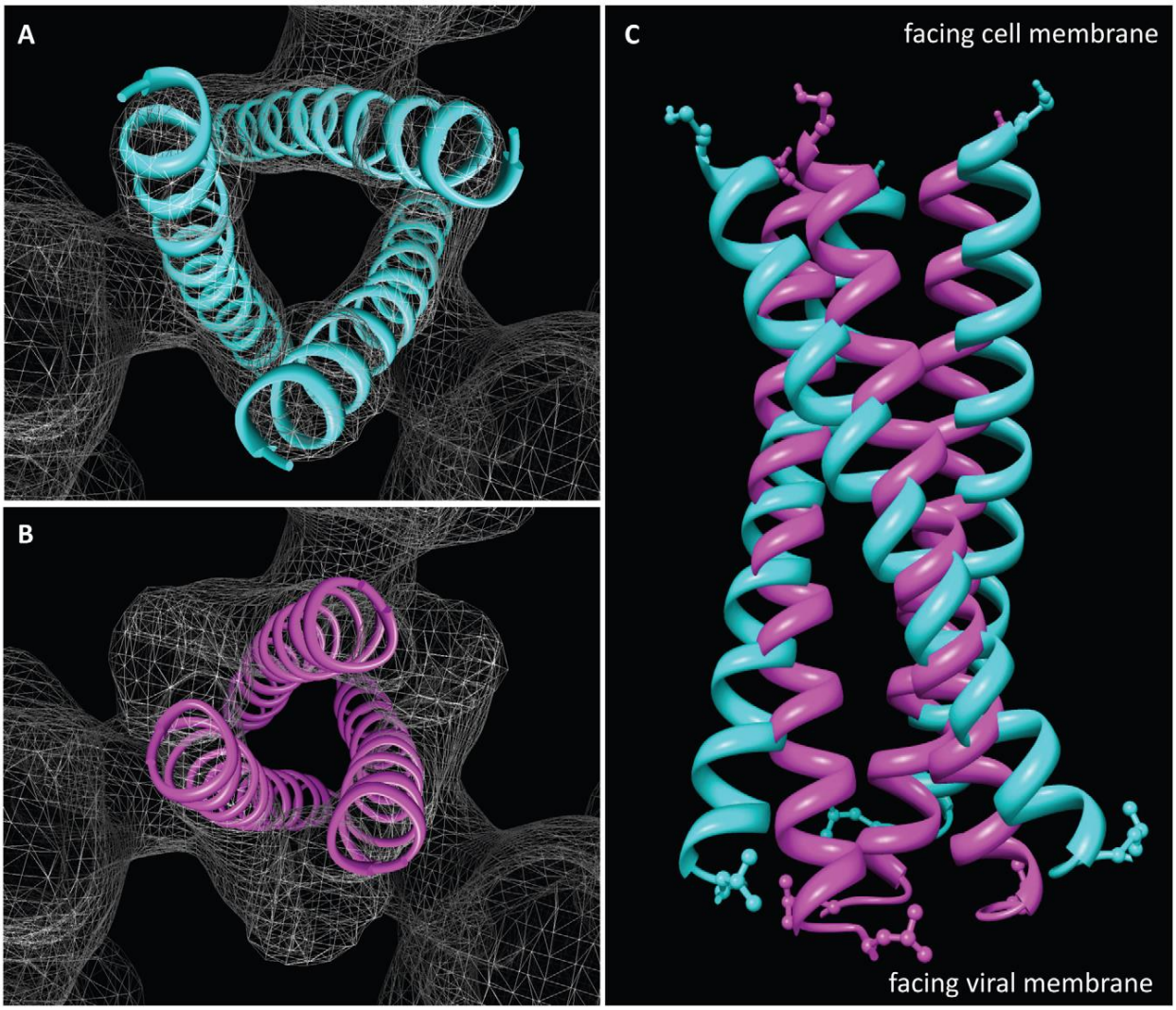

Figure 1.6 Compare of NHR coiled-coil trimer at pre- and post-fusion states. (A) Top view of the locations of the three N-terminal helices at pre-fusion state derived from cryo-electron microscopy. (B) Top view of the locations of the same helices in the post-fusion state derived by X-ray crystallography. (C) Superposition of the arrangement of the three N-terminal helices in pre-(cyan) and the post-fusion (magenta) conformation.

The fusion peptide (FP) is first 23 amino acid in the N-terminus of gp41 with sequence: AVGIGALFLGFLGAAGSTMGARS. FP is a key domain for fusion, and the FP alone can promote lipid vesicle fusion.[23] Even the structure of FP was not clear in gp41 crystal or cryo-electron microscopy results, there were a lot of NMR research on the peptide alone in both detergent micelles and lipid membrane. The FP formed α-helical structure in the SDS or DPC detergent micelles[24-26], however, other studies showed that FP formed antiparallel β-sheet structure in physiologically relevant membrane environment.[27-30] In membrane containing ~30 mol % cholesterol, the first 16 residues of FP had 0.30 fraction of antiparallel β sheet secondary structure with residue 16→1/1→16 and 17→1/1→17 registries for adjacent strands (Figure 1.7 A) and parallel β sheet with up to 0.15 fraction. The antiparallel conformation was also confirmed in a larger gp41 construct. [27, 30] FP had two membrane locations: major deeply-inserted and minor shallowly-inserted (Figure 1.7 B). [29] By studying the membrane location and fusogenicity of FP wild type and V2E mutation, a positive correlation between the insertion depths and the fusion activities was observed.[28]

Membrane proximal external region (MPER) is the ectodomain region that is closest to the viral membrane. The MPER is 22 amino acid tryptophan-rich domain and the sequence is ELDKWASLWNWFNITNWLWYIK. Liquid state NMR was used to study the MPER peptide in DPC micelle and the result demonstrated a well-defined α-helical peptide (Figure 1.8 A).

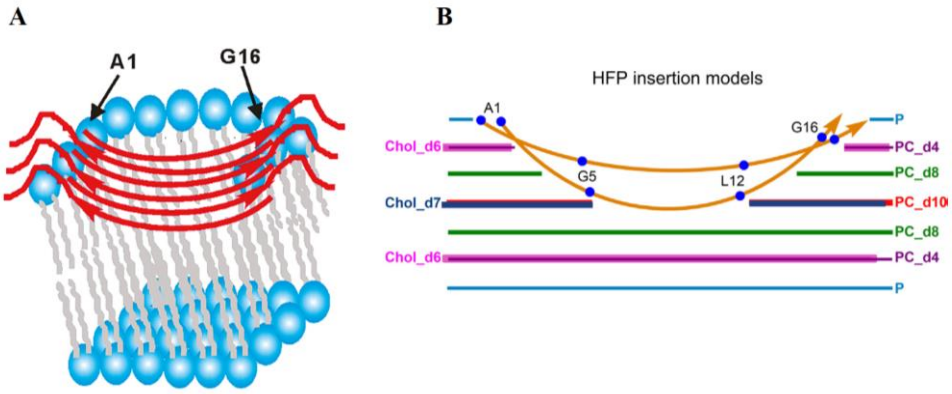

Figure 1.7 Models of FP in membrane. (A) Antiparallel β sheet FP with residue 16→1/1→16 or 17→1/1→17 registries for adjacent strands binding to membrane.[27] The FP is shown as red lines and the lipids are the blue sphere with gray lines.(B) Two membrane locations of FP studied by ^{13}C-^{2}H REDOR: majorly deep and minorly shallow insertion. The FP is ^{13}CO labeled on Gly_5, Leu_12 and Gly_16 residues. The labels on both side of the diagram are the approximate membrane locations of the ^{2}H's and ^{31}P's where P ≡ phosphorous in lipid; Chol_d6, Chol_d7, PC_d4, PC_d8 and PC_d10 are deuterium in cholesterol or phosphocholine lipids. [29]

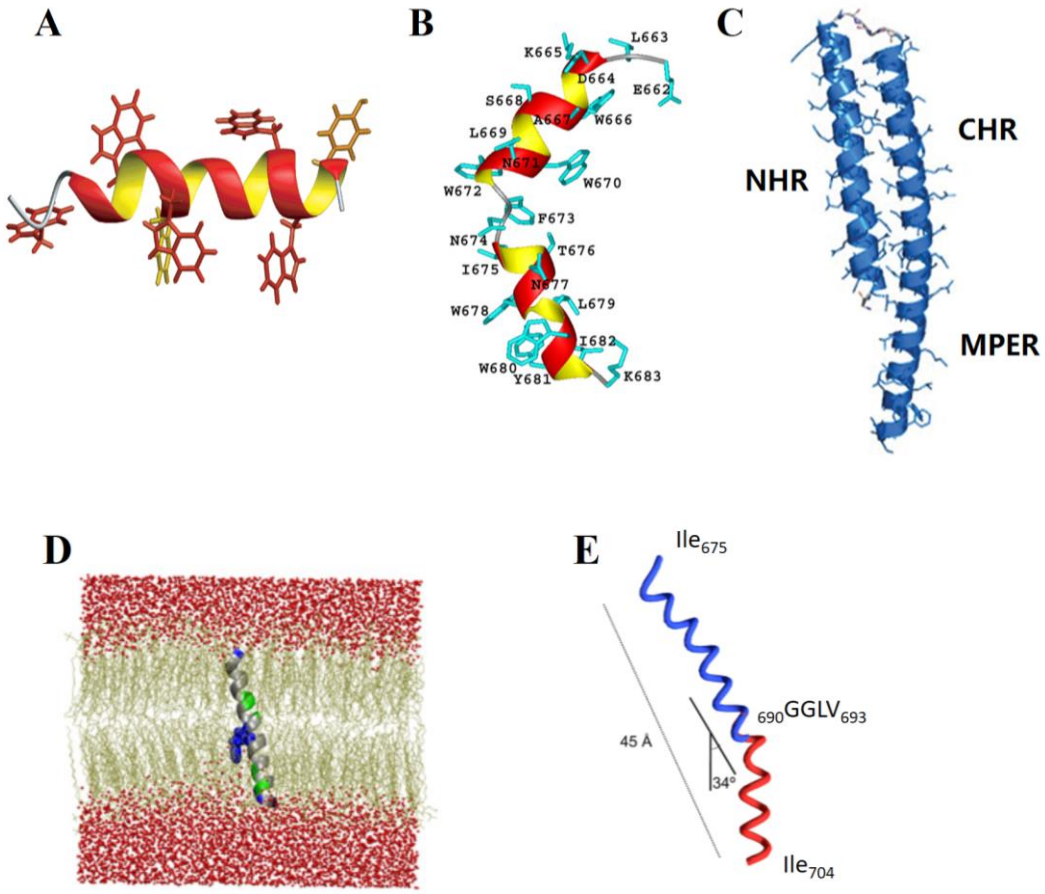

Figure 1.8 Models of MPER and TM. (A) $MPER_{665-683}$ in DPC micelle at pH 3.5 showing a straight α-helix;[31] (B) $MPER_{662-683}$ in DPC micelles at pH 6.6 showing a L-shape kink between two α-helices;[32] (C) Crystallography result of $NHR_{547–575}$-$CHR_{630-662}$-$MPER_{663-675}$ monomer[33]. The CHR and MPER forms a continuous helix and the MPER in this structure ends at I675. (D) Molecular dynamics (MD) simulation result of $TM_{681-707}$ in membrane with Arg694 snorkeling toward the membrane surface. TM peptide is shown in grey, Arg residues in blue, lipid molecules in yellow and water molecules in red.[34] (E) Liquid-state NMR result of $MPER_{675-683}$-$TM_{684-704}$ peptide. The MPER and TM form a continuous α-helix with a turn at $_{690}GGLV_{693}$ residues.[35]

Additionally, the MPER peptide showed an amphipathic structure where the aromatic residues were positioned on one side of the helix while the polar residues were on the other side, suggesting the MPER was attached to the viral membrane.[31] A newer study using liquid-state NMR, electron paramagnetic resonance (EPR), and surface plasmon resonance (SPR) techniques revealed that MPER had two discrete helical segments with a central hinge, forming an L-shape (Figure 1.8 B).[32] A crystal structure of gp41 containing NHR, CHR and MPER region revealed a trimeric, coiled-coil SHB and the MPER extended from CHR in continuation of a slightly bent long helix (Figure 1.8 C).[33]

The transmembrane (TM) domain is a 22-residue domain that anchors the gp41 in the viral membrane. Based on computer simulations data, it acquired a tilted α-helical conformation in membrane and had tendency to form a trimer (Figure 1.8 D).[34, 36] Liquid state NMR also revealed that MPER and TM forms an uninterrupted long α-helix (Figure 1.8 E).[35] Even there were some differences in the experimental conditions such as pH and the construct length, the MPER and TM domains all showed high helicity. Although there is no high-resolution structure for the whole gp41 protein, these isolated ectodomain or membrane segments structures provide a lot of information and help further studies.

1.1.3 Possible membrane fusion mechanism

As concluded in the previous sections, after the gp120 binds to CD4 and coreceptor, gp41 is exposed and the fusion process starts. The gp41 undergoes a large conformational change: from a NHR/60°-turn/CHR structure to become a NHR/180°-turn/CHR SHB structure. Like gp41, distinct change also happens in membrane structures during fusion (Figure 1.9).[37] The fusion starts with the lipid contact between viral and host cell membranes outer leaflet followed by stalk

formation where the outer leaflets fuse while the inner leaflets stay unfused. Hemifusion intermediate characterized by intermembrane lipid mixing and no contents mixing. This is followed by breaking the hemifusion barrier and formation of a fusion pore.

There are several models being proposed to explain the mechanism of gp41 catalyzed fusion.

Model 1[18]: (1) gp120 binding and gp41 exposure; (2) formation of an "prehairpin intermediate" (PHI) state followed by FP insertion into host cell membrane. The PHI is hypothetically a NHR/0°-turn/CHR SHB structure and the NHR and CHR regions stay as trimer bundle; (3) NHR and CHR in PHI state fold into SHB which consequently brings the two membranes closer; (4) viral and host cell membrane hemifusion; (5) initial fusion pore formation; and (6) fusion pore expansion (Figure 1.10). It was also proposed that the free energy released by PHI folding to SHB is used to overcome the energy barrier during fusion process. However, there are some evidence that is contradictory to this hypothesis. Either NHR, CHR, CHR+MPER or NHR+loop+CHR peptides can inhibit fusion since these peptides were presumed to bind to the PHI and stop the formation of SHB.[20, 38, 39] These functional studies showed that these peptides stopped the fusion pore expansion which demonstrated that the pore formation happened before PHI → SHB folding. There were also evidence supporting that the hemifusion happens before SHB forms.[40] Thus **Model 2** was proposed:

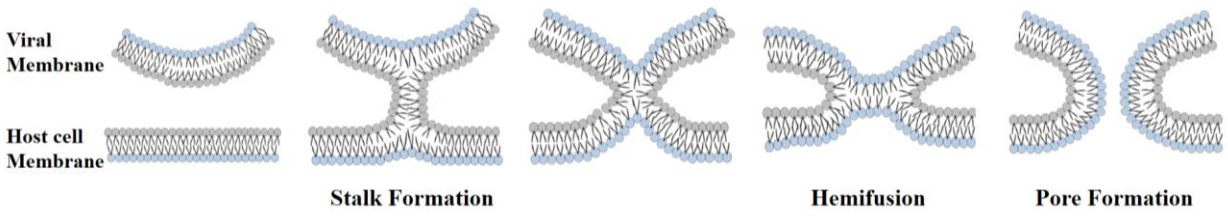

Figure 1.9 Stages of membrane fusion.

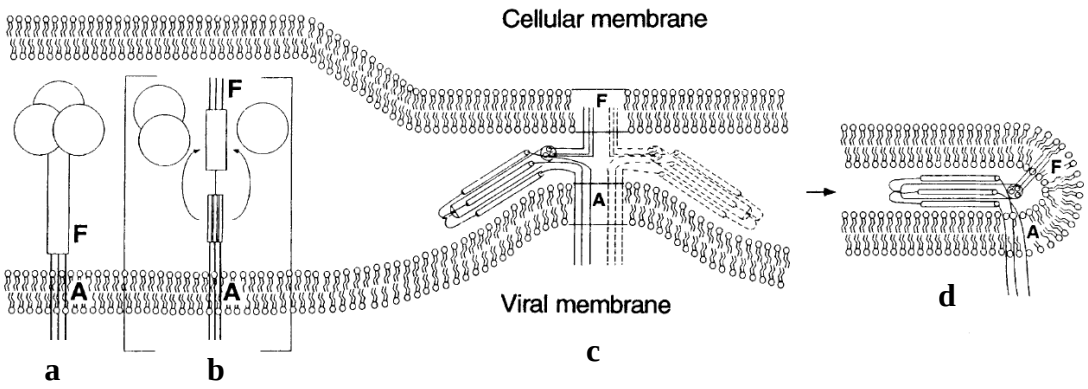

Figure 1.10 Schematic diagram of **Model 1**.[18] (a) Trimer gp120 and gp41 at pre-fusion state; (b) displacement of gp120 and PHI formation; (c) SHB formation; (d) gp41 at the SHB post-fusion state. 'A' represents the transmembrane domain and 'F' represent the FP domain.

Model 2[30]: (1) gp120 binding and gp41 exposure; (2) formation of PHI followed by FP insertion into host cell membrane; (3) Viral and host cell membrane hemifusion; (4) initial fusion pore formation; (5) PHI → SHB trimer folding; and (6) fusion pore expansion (shown in Figure 1.11). However, the oligomeric state studies at low pH or neutral pH with 6M GuHCl showed that gp41 ectodomain formed stable hairpin monomers as well as stable hexamers, which is contradictory with the trimeric PHI and SHB state in **Model 2**.[41]

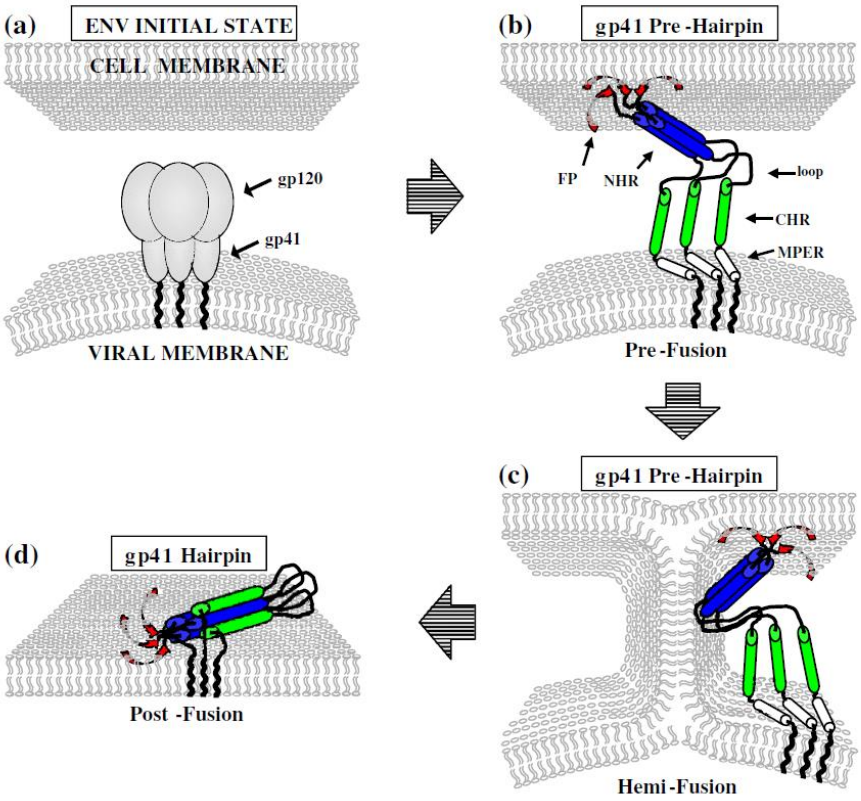

Figure 1.11 Schematic diagram of **Model 2**.[30] (a) Trimer gp120 and gp41 at pre-fusion state; (b) PHI formation and IFP inserted in host cell membrane; (c) Viral and host cell membrane hemifusion; (D)gp41 at the SHB post-fusion state. FP is shown in red, NHR in blue, CHR in green and MPER in white.

Model 3[41]: (1) gp120 binding and gp41 exposure; (2) dissociation of gp41 ectodomain into monomers and formation of extended PHI followed by FP insertion into host cell membrane; (3) PHI → hairpin monomer folding that brings the two membranes close together; (4) hemifusion; (5) initial pore formation; (6) hairpin monomer → SHB trimer → hexamer ectodomain assembly; and (7) fusion pore expansion (shown in Figure 1.12)

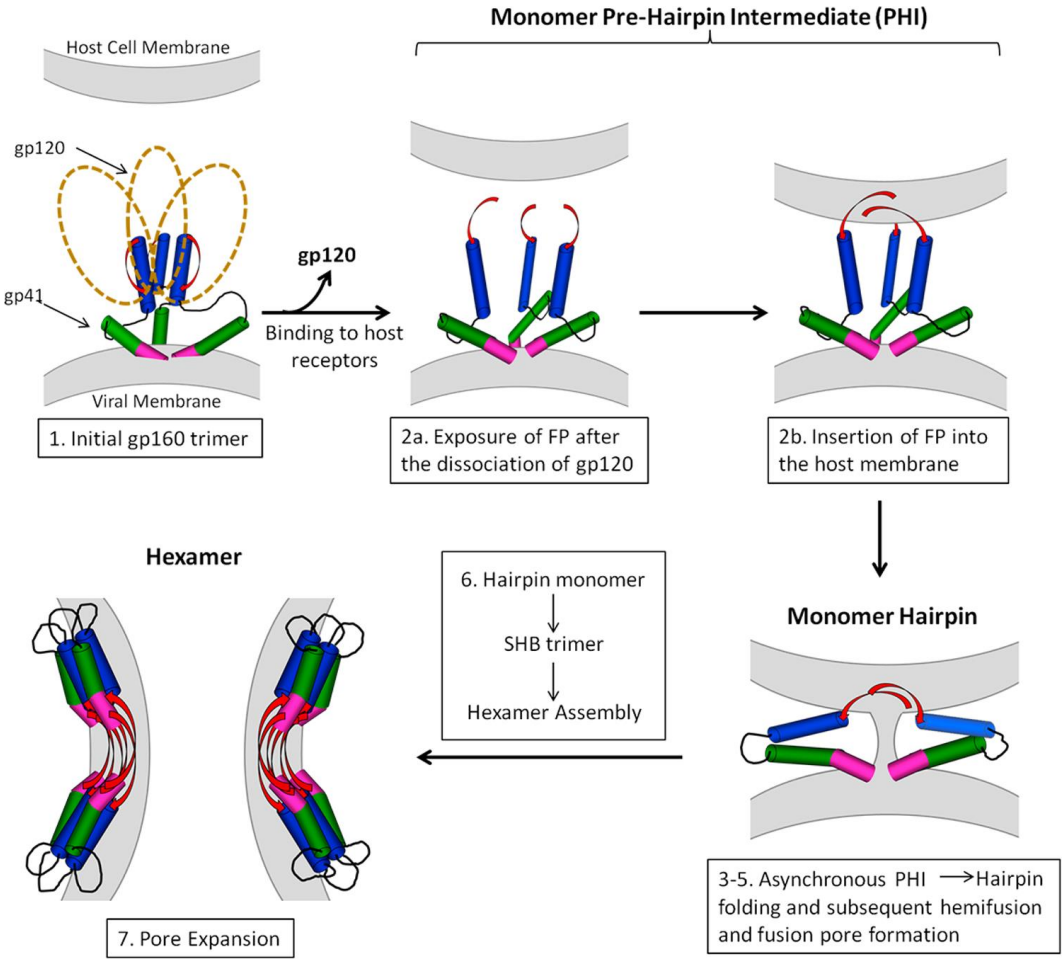

Figure 1.12 Membrane fusion **Model 3** of gp41 ectodomain monomer and hexamer. The different domains of gp41 are color coded the same as Figure 1.3 and the TM and endodomain are not shown. One of the monomers is not displayed in steps 3−5. The initial gp41 structure of step 1 and the final SHB structure of step 7 are based on high-resolution structures.[14, 15, 41]

As shown in **Model 3**, the PHI → SHB transition initializes the following hemifusion and fusion pore formation. The PHI → SHB monomer folding was also supported by hyper-thermostable ectodomain monomer. Besides, the stable hexamer is consistent with the requirement of multiple gp160 trimers for membrane fusion and HIV infection.[41, 42]

1.2 Introduction to influenza hemagglutinin fusion peptide

1.2.1 Influenza virus structure and its infection pathway

Influenza virus causes contagious respiratory illness that infects the nose, throat, bronchi and lungs. Not only for human, influenza viruses also infect other mammals and birds.[43] According to the World Health Organization's report in 2009, even though influenza vaccine is the most effective way to avoid infection and severe consequences, influenza virus still resulted in about 3-5 million cases of severe illness, and about 250,000 to 500,000 deaths worldwide. There were three other major influenza pandemics the past 100 years: 1918, 1957 and 1968, and each pandemic causes over one million of deaths. For example, in the 1918 Spanish flu, 500 million people were infected and 50 million of deaths across the world. Flu vaccines are widely used in the United States. However, people need to get vaccinated each year because of the influenza virus mutates rapidly. This variability results from different mechanisms including point mutations (antigenic drift) and gene reassortment (gene shift).[44] Therefore, even with having a flu vaccine, influenza viruses continue to pose a significant health risk to public worldwide.

Influenza virus has four types: A, B, C and D. Human influenza A virus is the most widespread type and it is the cause influenza pandemics.[45] According to Centers of Disease Control and Prevention (CDC), Influenza A and B viruses cause seasonal epidemics, and influenza C only causes a mild respiratory illness and are not thought to cause epidemics, while influenza D viruses primarily affect cattle and are not known to infect or cause illness in people. The following discussion in this dissertation will be focused on influenzas type A virus and will be abbreviated as "influenza virus".

Influenza virus is a member of orthomyxovirus family and it is an enveloped virus. The viral membrane comes from the budding when the virion leaves the host cell. There are three

membrane proteins on the viral membrane: hemagglutinin (HA), neuraminidase (NA), and M2 ion channel. HA is an integral membrane protein and there are ~ 400 copies of HA trimers per virion.[46, 47] HA is responsible for binding of virus to host cell and is involved in membrane fusion. HA is first expressed in the host cell as a precursor called HA0. The HA0 is cleaved into two subunits, HA1 and HA2 by a protease on a single Arginine residue. They are linked by a disulfide bond. HA1 has ~328 residues, including a receptor-binding site which can bind to sialic acids on carbohydrate side chains of cell-surface glycoproteins and glycolipids. HA2 has ~211 residues and the first ~20 residues at the N-terminus of HA2 is the influenza fusion peptide (IFP) and its function is essential for the membrane fusion.[48-50] NA is also an integral membrane glycoprotein with ~ 100 copies of tetramer per virion. The function of NA is to cleave terminal sialic acid from glycoproteins or glycolipids and free the virion form host cell receptors then the virus can spread and infect other cells.[44] HA and NA are highly mutable proteins and influenza A type virus is divided into strains based on the subtype of these two proteins. There are 18 different HA subtypes (H1 to H18) and 11 different NA (N1 to N11) subtypes and the strain of influenza is defined by the combinations of HA and NA, for example: H1N1, H3N1 etc.[48] M2 is a tetramer integral membrane protein act as proton ion channel. It lowers the pH to uncoat the virus.[51] The M1 protein is the most abundant protein in virion and it forms a shell surrounding the virion nucleocapsids inside of the envelope. The influenza viruses have eight unique segments of single-stranded RNA and the RNAs are loosely encapsidated by NP, which is the second most abundant protein. Viral polymerase proteins (PA, PB1 and PB2) are attached at the end of the genomic segments and provides RNA-dependent RNA polymerase activity. The complex of viral RNA, NP and polymerase proteins are called viral ribonucleoproteins (vRNP). The NS1(host antiviral response antagonist) and NS2 (nuclear export protein) are not incorporated into progeny virions, but they

present in infected cell nucleus (NS1) and cytoplasm (NS2). NS2 is responsible to export viral RNAs from the nucleus into the cytoplasm and pack viral RNAs into newly formed viruses.[44, 52, 53] Figure 1.13 demonstrates the structure of a virion.

The expression and replication of influenza viruses is a multi-step process, shown in Figure 1.14. An influenza virion with cleavage-activated HA recognizes and binds to the host cell through the binding of subunit on HA1 onto the terminal sialic acid of the host cell surface glycoproteins and glycolipids. Then the virion is endocytosed by the host cell. Subsequent lowering of pH to 5-6 in the endosome causes conformational changes of HA and the fusion peptide is exposed to trigger the fusion between viral membrane and endosomal membrane.[54-56] The pH in virion is lowered by M2 ion channel, removing the M1 proteins from vRNPs, which is also called uncoating. The genome of the virus is released into the cytoplasm and then imported into the nucleus for viral gene transcription and replication.[57] RNAs are transcribed into mRNA and the translation happens in the host cell cytosol. Newly synthesized viral proteins are assembled into RNPs in the nucleus and transported to the host cell membrane, while synthesized membrane proteins are integrated into the membrane. Finally, budding of new virus happens where host cell plasma membrane becomes the viral envelope.[44, 53]

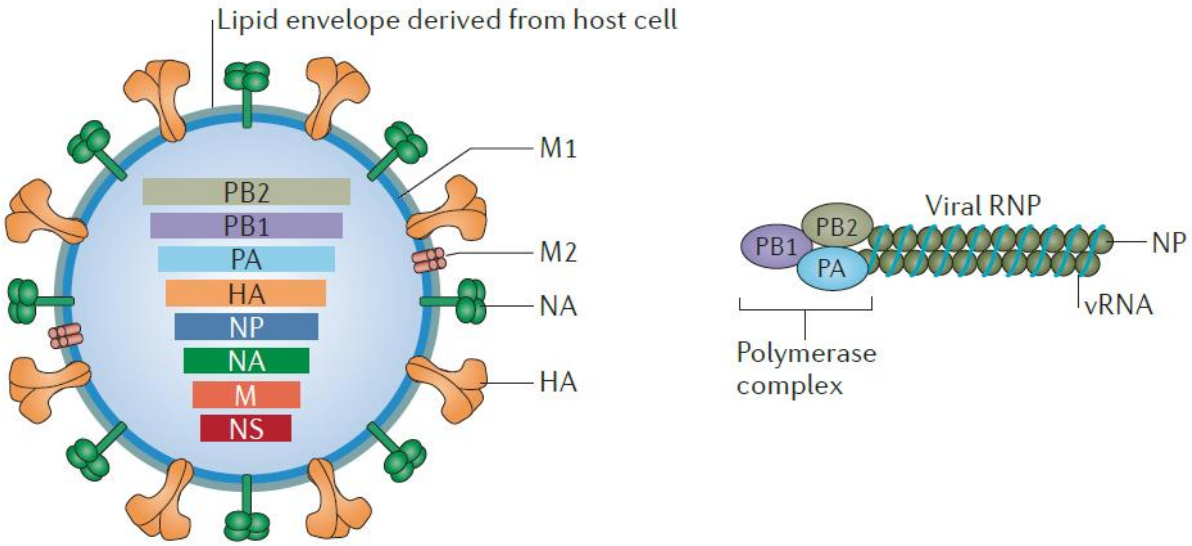

Figure 1.13 Schematic representation of the structure of influenza virus.[53] Viral envelope is shown as gray sphere and the membrane proteins (HA, NA and M2) are shown as spikes on the membrane. Inside the membrane is M1 that surrounds the viral ribonucleoproteins (vRNP). There are eight single-stranded, negative-sense RNA segments and each encoded one or two proteins.

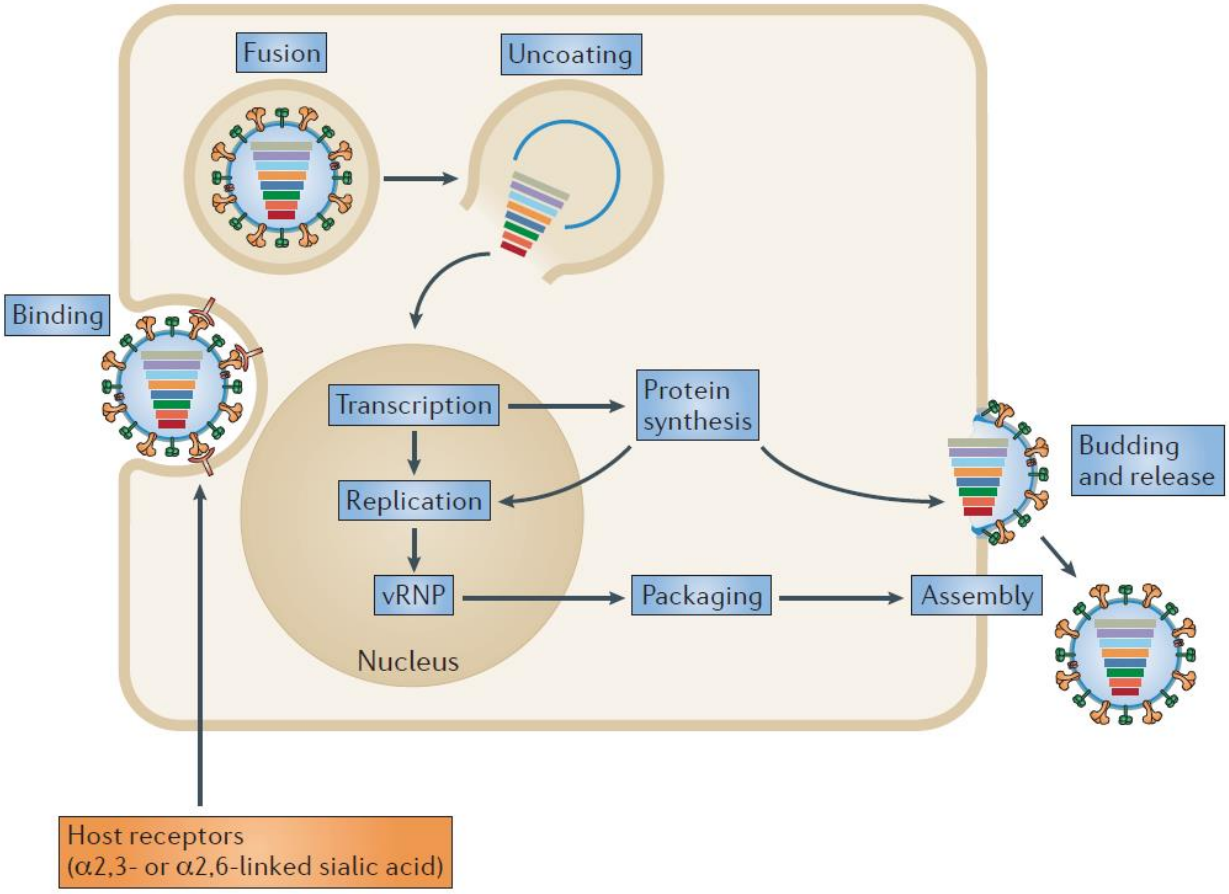

Figure 1.14 Influenza viral life cycle.[53]

1.2.2 Structure of HA2

HA is synthesized as a single polypeptide chain called HA0 and cleaved to HA1 and HA2 two subunits, which are linked by a disulfide bond. There are six glycosylation sites on HA1 and one on HA2 with ~13,000 MW.[56] The HA1 has ~328 residues and HA2 has ~211 residues, and the glycosylated site on HA2 is Asn154.[58] HA1 has a sialic acid bind site while HA2 is a transmembrane protein and it is responsible for catalyzing membrane fusion.[57] The structure of HA2 and its fusion mechanism is not fully understood. By studying HA2 may provide more information for vaccine and drug development to combat the virus.

As discussed in section 1.2.1, the infection process of influenza virus undergoes an endocytosis pathway. The pH of the endosome decreases from pH 7.4 to 5 or 6 and triggers the conformation change of HA2. Extensive studies have been made to elucidate the structure of HA at these two pHs, or in other words, the pre- and post-fusion state (Figure 1.15). At pre-fusion state, the oligomeric state of HA is trimer. The crystal structure of HA2 ectodomain containing residue 38 – 170 is shown in Figure 1.15 panel A. The N-terminal α-helix (29 Å) extended away from the membrane and connected by an extended chain to a very long α-helix (76 Å) that stretched back towards the membrane. In the trimer, the long helix formed the coiled-coil core with two identical helices from other subunits.[47] At post-fusion state, HA2 also formed three-stranded α-helix coiled-coil, approximately. The monomer HA2 from residue 38 - 175 was shown in Figure 1.15 panel B. The center helix including residue 40 – 105 followed by a 180˚ turn and helix including residue 113 - 129 packed antiparallelly outside.[49, 55] There was a big conformational change between these two conformations (Figure 1.15 C). From pre- to post-fusion state, helix A was relocated on top of the triple-stranded α-helix coiled-coil. Loop B adopted an α-helix conformation and region A – C formed a ~100 Å long helix. A portion of helix CD (residue 106 - 112) refolded to form a loop

and region D was antiparallel to the center bundle. Thus, region E, F and G were packed in an inverted orientation.[55] The region E and F were no longer antiparallel β-strands in the post fusion state.[49] Details about the fusion mechanism will be discussed in section 1.2.3.

The polypeptide chain of ~ 25 N-terminal residues of HA2 is called influenza fusion peptide (IFP). The sequence of IFP is highly conserved and hydrophobic.[59] IFP plays a vital role in fusion activity. Even a single site mutation may eliminate the fusion activity.[60] However, in these crystal structures, the IFP was not shown. Isolated IFP has been widely studied to understand its structure and function in fusion process. Liquid-state NMR, solid-state NMR as well as electron spin resonance (EPR) have been used to illustrate IFP structure in detergent micelle or lipid membrane at both physiology pH and low pH.[61-65] These studies showed that IFP adopted a helix-turn-helix structure in micelle or membranes without cholesterol. Two IFP subtypes have been extensively studied and their sequence are shown as below:

H1: GLFGAIAGFIE**G**GW**T**GMIDGWYG

H3: GLFGAIAGFIE**N**GW**E**GMIDGWYG

The difference in these two sequences are residue 12 and 15 (Labeled as bold). In H3 subtype, the residue 12 is Asparagine instead of Glycine, and the residue 15 is Glutamic acid instead of Threonine compared to H1 subtype.

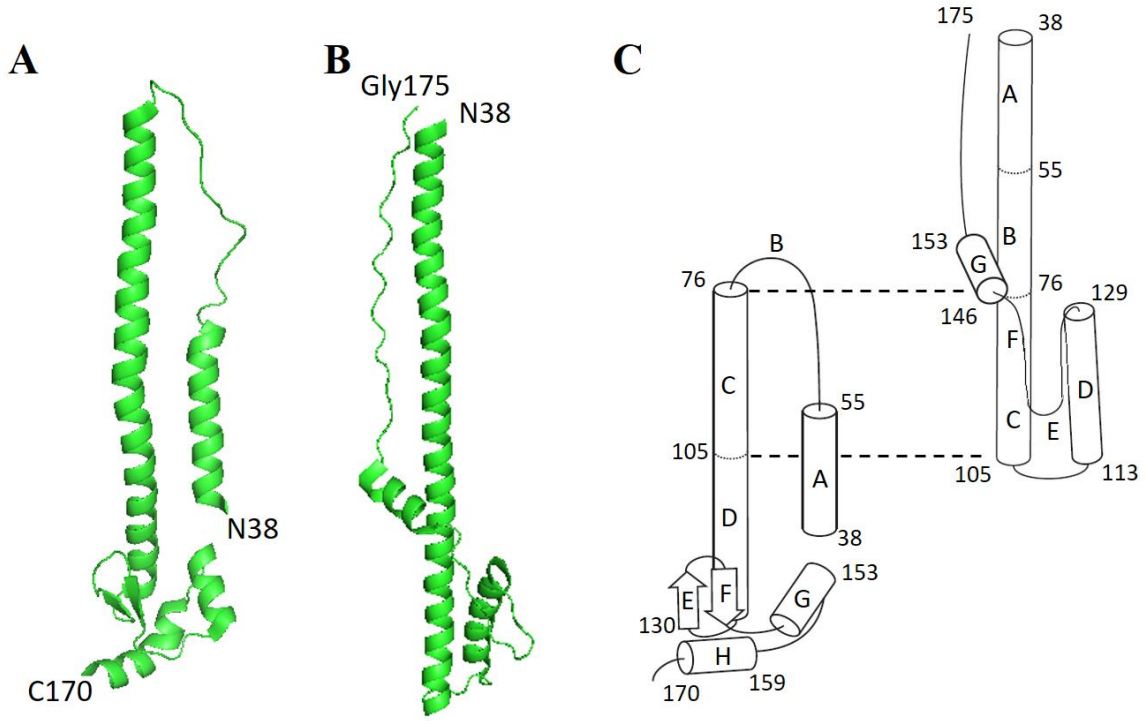

Figure 1.15 (A) Pre-fusion structures of HA2 ectodomain (Protein Data bank entries 1RD8). This structure includes residues 38 - 170 of HA2 at pH 7.5;[47, 66] (B) Post-fusion structures of HA2 (Protein Data bank entries 1QU1). This structure includes residues 38 - 175 of HA2 at pH 5.[49] (C) The structure of pre-fusion (left) and post-fusion (right). The HA2 is divided into A – H regions and the residue numbering is labeled. The β-strand of HA1 is also shown as "1" and the disulfide bond between 14_1 and 137_2 is indicated.[55]

Tamm's group studied the structure of the first 20 residues of H3 (denoted as H3_20 IFP) in DPC micelle using liquid-state NMR.[64] At pH 5 (Figure 1.16 A) IFP adopted a N-terminal helix (residues Leu-2 to Ile-10) / turn (residue Glu-11 to Gly-13)/C-terminal helix structure (residues Trp-14 to Ile-18) and this structure was referred as "open structure" with an interhelical angle ~ 100°. At pH 7.4, it formed a N-terminal helix (residues Leu-2 to Ile-10) / turn (residue Glu-11 to Gly-13)/C-terminal extended structure (residues Trp-14 to Gly-20) (Figure 1.16 B). The membrane location of H3_20 IFP was also studied by EPR. The 20 residues were replaced by Cystine individually and labeled with nitroxide groups. The result showed that H3_20 IFP was immerged in an inverted V-shaped manner into lipid hydrophobic core, where the N- and C-terminus deeply inserted while the turn is close to the aqueous surface. At pH 7.4, the N-terminal was ~ 5 Å from the lipid phosphate group, and C-terminal was ~ 2 Å. When at pH 5, the N-terminal has ~ 11 Å distance from the lipid phosphate group, while C-terminal ~ 10 Å. Additionally, H3_20 IFP is inserted deeper at pH 5 compare to at pH 7.4. The result also suggests that a greater perturbation of lipid bilayer at pH 5 and facilitate the fusion.

Another very different structure was proposed by Bax's group when they studied the H1_23 IFP in DPC micelle by liquid-state NMR.[61] H1_23 IFP had a tightly packed anti-parallel N-helix/turn/C-helix structure with a 158° interhelical angle at both pH 7.0 and pH 4 (Figure 1.16 C and D) and this structure was referred to "closed structure". The closed structure also showed an amphipathic property: the hydrophobic side chains were interacting with the hydrophobic core of the micelle and the hydrophilic side was exposed to the solvent (Figure 1.16 D). The such significant difference may result from the three additional C-terminal residues: Trp, Tyr and Gly.

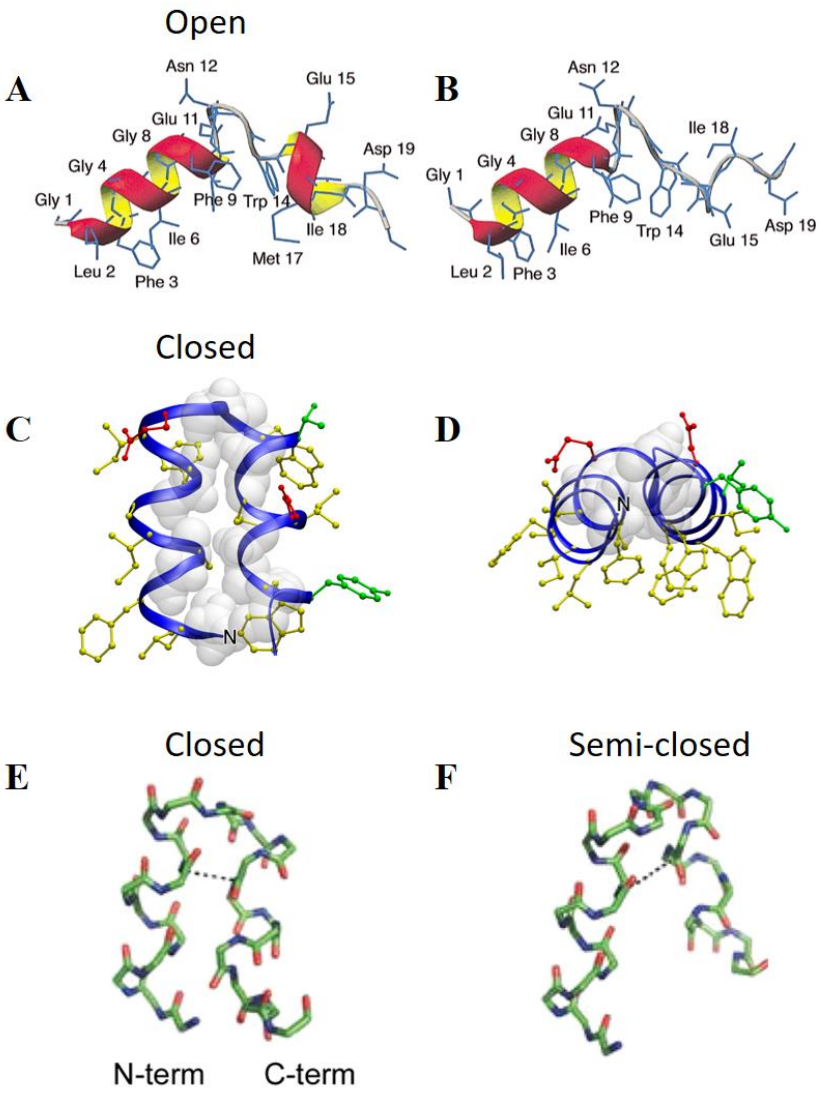

Figure 1.16 NMR structures of IFP. (A) (B) H3_20 in DPC micelle at pH5 and pH 7.4 respectively. Amino acids are labeled in the diagram. Backbone of the peptide is shown as ribbon representations and side chains shown as stick models.[64] (C) (D) Longitudinal view and lateral view of H1_23 in DPC micelle at pH 7 and pH 4, with the hydrophobic side chains in yellow, polar side chains in green, acidic side chains in red, and Gly residues shown in white van der Waals representation;[61] (E) (F) H1_23 and H3_20 in lipid membrane at pH 5 or pH 7. Carbon, nitrogen and oxygen atoms in green, blue and red vertices. The dashed lines are between F9 N and G16 CO with distances 3.9 Å and 5.5 Å respectively.[62, 63]

Due to the high curvature of detergent micelle, the IFP structures determined in DPC are less physiologically relevant. Solid-state NMR techniques has been used to determine the IFP structure in physiologically relevant lipid membrane. Sun from Weliky's group studied the structure of H3_20 IFP in lipid membrane and the result illustrated that at pH5.0, membrane-associate IFP adopted two distinct helix/turn/helix structure.[65] Ghosh from Weliky's group studied the distance between H3_20 IFP Phe19N and Gly16CO at pH 5. There was ~ 0.6 mole fraction of "semi-close" structure with 5.2 Å distance and ~ 0.4 mole fraction of "close" structure with 3.6 Å distance.[62] Further the interhelical geometry study of both H1_23 and H3_20 in lipid membrane at pH 5 or pH 7 showed a mixture of "closed" and "semi-closed" structures (Figure 1.16 E and F). Both structures showed N-helix/ turn/C-helix with different interhelical angel and they were also amphipathic. It was proposed that IFP interacted with the membrane hydrophobic core with its hydrophobic face while the hydrophilic face interacted with water. Semi-close structure had a bigger hydrophobic area than close structure and larger population of semi-close structure presents at pH 5 compared to pH 7. This is consistent with higher fusion activity at pH 5. H1_23 also showed a higher fusion activity compared to H3_20. The reason was because that H1_23 had additional WYG residues at C-terminal, and these hydrophobic residues increased the hydrophobic surface area of H1_23 and lead to greater membrane perturbation and membrane fusion.[63]

1.2.3 Proposed membrane fusion mechanism

Figure 1.17 shows the sequence of events in membrane fusion promoted by HA2.[67, 68] The crystal structure of HA1 and HA2 in a prefusion conformation was determined by Wiley, Wilson and Skehel in 1981(Figure 1.15 A).[47, 69] This crystal structure of HA at pH 7.5 showed that HA is a trimer. Three HA2 subunits anchored on the viral membrane and forms the coiled-coil trimer while three HA1 subunits binds outside the HA2 core. After the HA1 sialic acid binding site binds to the glycolipids or glycoprotein on host cell membrane, the virion is endocytosed and the pH drops to ~5 in the endosome. The low pH triggers a large-scale conformational rearrangement, resulting in HA1 separates from HA2 and HA2 is exposed to the target endosomal membrane. It is proposed that during the conformation change, HA2 forms an extended intermediate and the IFP on the N-terminus of HA2 interacts with endosomal.[68, 70] Then the intermediate starts to collapse and energy is released during the refolding process causing the membrane to contact each other. The formation of the hemifusion stalk happens afterwards and finally the protein refolds to form the stable post-fusion conformation. The post-fusion conformation was visualized in 1999 (Figure 1.15 B).[49, 55] There is no crystal structures for stage B to D, neither experimental evidence for energy releasing due to the protein refolding process. But some biochemical studies support many of the proposed steps. Oligomeric state studies of HA2 ectodomain or full length HA2 showed predominant trimer.[71, 72] When destabilized with either heat or urea at neutral pH, HA2 had a conformational change which showed strong fusion activity and was believed to be the evidence of the extended intermediate state.[73] HA2 without transmembrane domain caused lipid mixing but no contents mixing, which meant the fusion stopped at hemifusion stalk and transmembrane domain is crucial for the formation of fusion pore.[74]

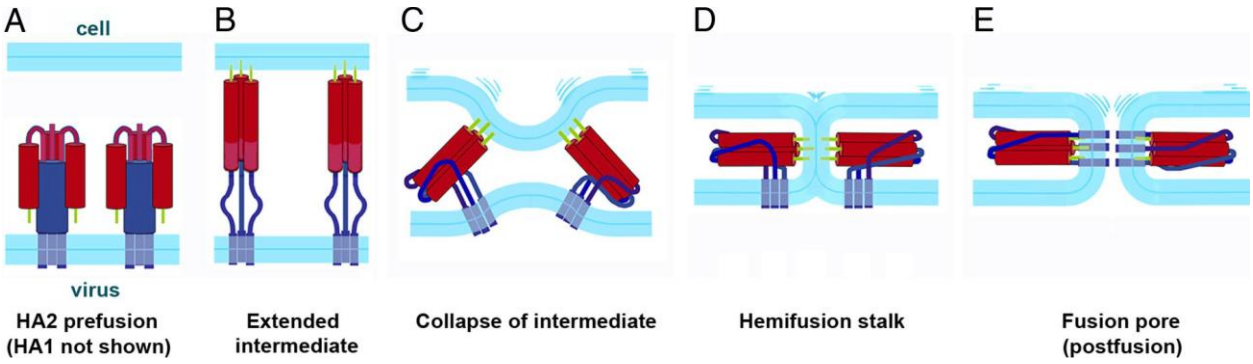

Figure 1.17 Mechanism of membrane fusion promoted by HA2.[67, 68] Host cell membrane is the blue bilayer on top in each diagram and viral membrane on bottom. The IFP is in green and N-terminal part of ectodomain of HA2 is in red while the C-terminal domains are in navy. The HA1 is not shown in this figure. (A) In the prefusion state, the protein is anchored to the viral membrane by a C-terminal transmembrane domain. (B) pH decreasing to pH 5 in endosome triggers a conformational change resulting in an extended intermediate and exposing IFP to the target membrane. (C) The intermediate collapses. (D) Hemifusion stalk. (E) Fusion pore formation. As the hemifused bilayers open into a fusion pore, the final zipping up of the C-terminal ectodomain segments snaps the refolded trimer into its post-fusion conformation, preventing the pore from resealing.

1.3 Introduction solid-state NMR techniques for membrane proteins

Membrane proteins are the proteins that interact with biological membrane and they have crucial functions such as membrane receptors and enzymes. However, due to the difficulty of growing protein crystals, only a small fraction of protein structure is solved.

Solid-state nuclear magnetic resonance (NMR) is a powerful technique to determine high resolution structure and function of large biomolecules. Comparing to widely used X-ray crystallography and liquid-state NMR, solid-state NMR is specifically suitable for membrane proteins, large proteins, protein aggregates and nucleic acids that cannot be crystallized or that are too large for solution NMR spectroscopy.[75, 76]

1.3.1 Magic-Angle Spinning (MAS)

In solution-state NMR spectra, anisotropic effects are rarely observed because of the rapid tumbling of the molecules in solution. Thus, the orientation of the molecules with respect to the external magnetic field $\mathbf{B_0}$ is rapidly averaged out and results in sharp peaks.[a] For solid samples or macro biomolecules, the tumbling is much slower which results in broad lines in NMR spectra since all the orientations in the sample contribute to different spectral frequencies.[77, 78] Magic-angle spinning (MAS) is a routinely used techniques to achieve high resolution spectra. It can remove the effects of chemical shift anisotropy and to assist in the removal of heteronuclear dipolar-coupling effects. It is also used to narrow lines from quadrupolar nuclei and removing the effects of homonuclear dipolar coupling.[78]

[a] In this dissertation, letters or symbols representing vectors are displayed in bold, and letters or symbols representing quantum mechanical operators have a "^" above them.

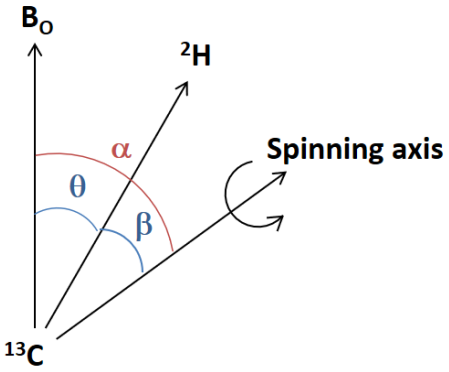

Figure 1.18 Geometry of the geometry of the ^{13}C – ^{2}H vector in solid state NMR sample under MAS. The sample is spun rapidly in a cylindrical rotor about a spinning axis oriented at the magic angle (α = 54.7°) with respect to external magnetic field $\mathbf{B_0}$.

Figure 1.18 shows the geometry of the MAS. In solid-state NMR MAS experiments, samples are packed in a cylindrical rotor and spun at high speed by an angle α with respect to $\mathbf{B_0}$. This angle is also called the magic angle and it equals to 54.7°. Angle θ is the angle between the ^{13}C – ^{2}H internuclear vector and $\mathbf{B_0}$, and β is the angle between the ^{13}C – ^{2}H internuclear vector and the spinning axis. When the sample is spin at α = 54.7°, then θ varies with time as the molecule rotates with the sample. The average of ($3\cos^2\theta$ -1) over each rotor period is:

$$\langle 3cos^2\theta - 1\rangle = \frac{1}{2}(3cos^2\alpha - 1)(3cos^2\beta - 1) \quad 1.1$$

In equation 1.1, β is fixed for a given nucleus in a rigid solid, and θ can be all possible values in a powder sample. The α is fixed and set to 54.7°, so over each rotor period:

$$\langle 3cos^2\theta - 1\rangle = 0 \quad 1.2$$

The secular Hamiltonian for ^{13}C-^{2}H dipolar coupling can be expressed as:

$$\hat{H}_{hetero} = \frac{\mu_0 h \gamma_C \gamma_D}{16\pi^2 r^3}(3cos^2\theta - 1)\left(2\hat{C}_z\hat{D}_z\right) \quad 1.3$$

Where C represents ^{13}C spin and D represents ^{2}H spin,

μ_0 = permeability of the free space,

$\hat{C}_z$ and $\hat{D}_z$ = the z component of spin operator $\hat{\boldsymbol{C}}$ and $\hat{\boldsymbol{D}}$ in a direction parallel to $\mathbf{B_0}$,

θ is the angle between the spin vector and $\mathbf{B_0}$,

r is the distance between spin *C* and *D*,

γ is gyromagnetic ratio.

Therefore, the ^{13}C-^{2}H dipolar coupling is averaged out by MAS.

1.3.2 ^{13}C-^{2}H Rotational-Echo Double-Resonance NMR (REDOR)

REDOR was developed by Gullion and Schafer and it is widely used MAS NMR techniques for studying molecular structure in solid-state.[29, 79-84] As discussed previously, MAS can average out the heteronuclear dipolar interactions. However, in REDOR experiments, by using simple rotor-synchronized π pulses, the heteronuclear dipolar interactions can be recovered. Since the dipolar interaction is inversely proportional to the cube of the internuclear distance, the distances information can be easily obtained. Another advantage of REDOR is the simplicity of its pulse sequence and data analysis.[79]

^{13}C-^{2}H REDOE pulse sequence is shown in Figure 1.19. ^{13}C-^{2}H REDOE is a three channel experiments. At the beginning of the sequence, a $\pi/2$ pulse is applied to rotate the ^{1}H magnetization from the $\mathbf{B_0}$ direction to the transverse plane. Then a ^{1}H-^{13}C cross polarization (CP) pulse sequence is applied to transfers ^{1}H transverse magnetization to ^{13}C nucleus and to enhance the ^{13}C signal. Due to the various orientation of molecules, chemical bonds, internuclear vectors with respect to $\mathbf{B_0}$, the nuclei in a powder sample have a distribution of Larmor frequencies and the resonance offset field. Thus, a ramp CP is used to increase the efficiency of the magnetization transfer.

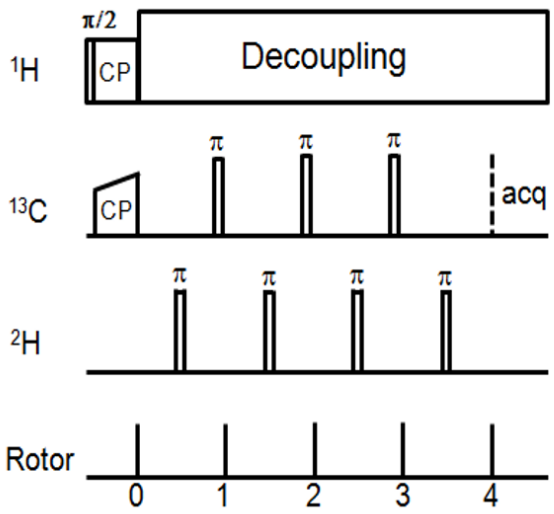

Figure 1.19 ^{13}C-^{2}H REDOR NMR pulse sequence. The columns represent the π/2 or π pulses. CP = cross polarization that transfers ^{1}H transverse magnetization to ^{13}C and can enhance the ^{13}C signal. The CP is followed by a ^{13}C-^{2}H dipolar evolution for a period of time which is called dephasing time (τ). Adjacent ^{13}C π pulses are separated by one rotor period as are adjacent ^{2}H π pulses. ^{13}C is the detecting channel.[79]

After the ^{1}H-^{13}C CP, followed by a dephasing period (τ) during which a series of ^{13}C and ^{2}H π pulses and ^{1}H decoupling field were applied. For each τ, two separate spectra are collected and they are denoted as S_0 and S_1. In the S_0 experiment, only ^{13}C π pulses are applied at the end of each rotor period except at the end of τ. In the S_1 experiment, ^{13}C π pulses are applied at the end of each rotor period while ^{2}H π pulses are applied in the middle of each rotor period. Then followed by the ^{13}C signal acquisition.

The ^{13}C-^{2}H dipolar coupling and chemical shift anisotropy are averaged out by MAS. In REDOR experiments, the function of the ^{13}C and ^{2}H π pulses are to flip the spin by 180°. The secular ^{13}C-^{2}H dipolar coupling is discussed in equation 1.3:

$$\widehat{H}_{hetero} = \frac{\mu_0 h \gamma_C \gamma_D}{16\pi^2 r^3}(3cos^2\theta - 1)\left(2\hat{C}_z\hat{D}_z\right)$$

Thus, the ^{13}C and ^{2}H π pulses change the sign of dipolar coupling. In S_0 experiment, the rotor synchronized ^{13}C π pulses are applied at the end of each rotor period and it flips the sign of $\hat{C}_z$. Their effect on the ^{13}C-^{2}H dipolar coupling is shown in Figure 1.20. Additionally, ^{13}C π pulses refocus the ^{13}C isotropic chemical shift. In S_1 experiment, ^{2}H π pulses are applied at the middle of each rotor period and the ^{13}C-^{2}H dipolar coupling is recoupled. For an isolated ^{13}C-^{2}H spin pair, the signal of S_0 and S_1 follow the relationship as:

$$\frac{S_1}{S_0} = \cos\phi \qquad 1.4$$

$$\phi = \frac{N_c T_r D}{2\pi^2}\sqrt{2}\sin 2b\, sina \qquad 1.5$$

Where N_c is number of rotor period, T_r is sample rotation period, D is dipolar coupling, and a and b are the azimuthal and polar angles of the internuclear vector with respect to the spinning axis. For a powder sample, all internuclear orientations need to be considered, and all values of a and b must be summed over. Thus, S_1 is always smaller than S_0.

The dephasing of REDOR is:

$$\frac{\Delta S}{S_0} = \frac{S_0 - S_1}{S_0} \qquad 1.6$$

In this equation, S_0 and S_1 are the respective signal intensity of S_0 and S_1 experiments. The dephasing is a function of dephasing time τ and the dephasing buildup curve is plotting $\Delta S/S_0$ vs τ.

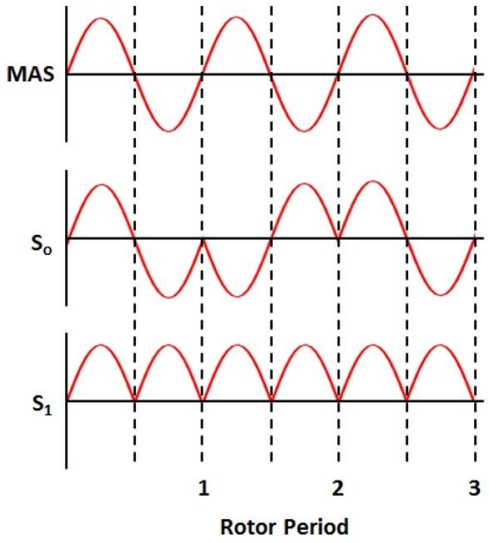

Figure 1.20 Evolution of dipolar coupling as a function of rotor period in REDOR experiments. The dipolar coupling is averaged out over each rotor period by MAS. In S_0 experiment, rotor synchronized ^{13}C π pulses do not interfere with the MAS averaging of the heteronuclear dipolar interaction. In S_1 experiment, rotor-synchronized ^{13}C and ^{2}H π pulses re-introduce the ^{13}C-^{2}H dipolar interaction.

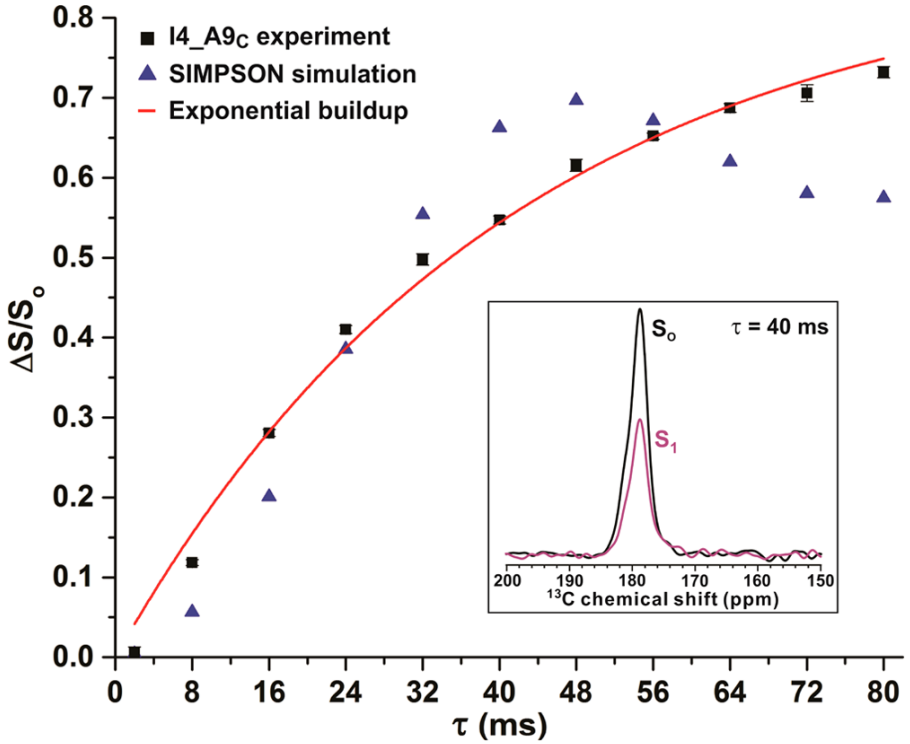

Figure 1.21 ^{13}C-^{2}H REDOR spectra with a 40 ms dephasing time for the I4 peptide as well as $\Delta S/S_0$ vs τ. Black squares are the experimental dephasing of I4 peptide. Blue triangles are the best-fit calculated by the SIMPSON program (without ^{2}H relaxation) with a 22 Hz ^{13}C-^{2}H dipolar coupling. The red line is the best-fit exponential buildup.[81]

The buildup can be fitted with SIMPSON simulation.[85] Exponential fitting can also be used for ^{13}C-^{2}H REDOR.[29, 81] The exponential fitting function is $A(1 - e^{-\gamma\tau})$, in which A and γ are fitting parameters. The buildup rate γ has a relationship with dipolar coupling D. The buildup extent A is correlated to the fraction of ^{13}C nuclei with this coupling. The corresponding value of $1 - A$ is correlated to the fraction of lab ^{13}C nuclei with $D \approx 0$.

Figure 1.21 shows the comparison between these two methods. I4 peptide is used as a standard sample. It has a sequence AEAAAKEAAAKEAAAKAW and it has a has a regular α helical structure. I4 peptide is ^{13}CO label at A9 and $^{2}H_{\alpha}$ label at A8. The isolated ^{13}CO-^{2}H spin pairs all have the same internuclear separation r of 5.0 Å and with a corresponding dipolar coupling D of 37 Hz.[81] The exponential fitting result (red line) has a much smaller deviation from experimental dephasing (black square) compared to sigmoidal shape SIMPSON simulation (blue triangle). Additionally, the maximal values of ~ (2/3) in SIMPSON differs from ~1 in exponential fitting. The best-fit dipolar coupling D of 22 Hz is smaller than the expected d of 37 Hz.

The reason for all these differences are that the nonradiative transitions between the $m = \pm1$ states and the $m = 0$ states of individual ^{2}H nuclei during the dephasing period that are not considered in the SIMPSON calculations. The ^{2}H T_1 relaxation times is $\approx$ 50 ms, which is comparable to τ in REDOR experiments. Thus, there are $m = 0 \leftrightarrow m = \pm1$ ^{2}H transitions during the dephasing time. For each ^{2}H nucleus, it approximately spends two-thirds of the dephasing period at $m = \pm1$ states and approximately one-third at $m = 0$ state. Since there is not buildup when ^{13}CO is at $m = 0$ state, so that the observed buildup rate $\gamma \approx 2D/3$.

The exponential fitting $A \times (1 - e^{-\gamma\tau})$ for the experimental data results in r = 5 Å and γ = 24 Hz $\approx$ (2/3) (37 Hz) and 37 Hz is the dipolar coupling of I4 peptide. This shows that exponential

fitting is an excellent and simple fitting method for ^{13}CO-^{2}H REDOR. D, in units of Hz is given by: $D = \frac{\mu_0 h \gamma_C \gamma_D}{16\pi^3 r^3} = \frac{4642}{r^3}\ Hz$. Thus, the internuclear distance r was calculated as $\sqrt[3]{\frac{4642\ Hz}{\frac{3\gamma}{2}}}$ Å.

1.3.3 Paramagnetic Relaxation Enhancement NMR (PRE)

Paramagnetic Relaxation Enhancement NMR (PRE) is first developed by Solomon in the 1955.[86] Over the past decades, PRE has become a powerful method to provide long-range distance information. In PRE experiments, paramagnetic species such as Mn^{2+}, Gd^{2+}, and nitroxide spin labels have been used in solid-state NMR for studying protein membrane location.[87-90] The PRE arises from dipolar interactions between a nucleus of interest and the unpaired electrons of the paramagnetic center which result in an increase in nuclear relaxation rates. Since the gyromagnetic ratio of electron is three magnitudes greater than most nuclei, this method can be used to get distance information in a range up to 35 Å, which is especially useful for biomacromolecules and biosystems.[87] The first applications of the PRE is to study spin-labeled lysozyme and bovine pancreatic trypsin inhibitor in 1984, and the PRE effects were converted to approximate distance restraints.[91] In addition, PRE is also used to accelerate data acquisition for systems such as biomolecules and human tissue, since PRE increases longitudinal relaxation T_1 rates.[92-95]

Other techniques can also be used to study the protein topology in membrane, but they have some limitations. The Nuclear Overhauser Effect (NOE) is widely used for protein structure determination. However, NOE is based on short-range local interactions, which limits to ~6 Å interproton distance.[61, 96] ^{1}H spin diffusion NMR technique can be used to probe membrane protein topology in membrane.[97-99] This technique detects the rate of ^{1}H magnetization transfer from mobile lipids to the rigid proteins, but not suitable for mobile peptides. Beyond NMR, EPR spectroscopy is also well established for investigating membrane location of peptides.[64, 100]

However, these approaches require the use of bulky spin probes, which may affect the insertion depth or orientation of the peptide in membrane.

To study the membrane protein structure and topology in membrane, the paramagnetic metal ions were added to protein-membrane complex and these ions bind at the membrane surface. The unpaired electron in paramagnetic species enhance the relaxation of the nuclear spins. Let's take Mn^{2+} as an example. The paramagnetic Mn^{2+} contributes to faster dipolar transverse relaxation, T_2, is proposed by Solomon[86] and Bloembergen[101]:

$$\frac{1}{T_2} = W\frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2\frac{\gamma_I^2\mu_{eff}^2\beta^2}{r^6}\left(4\tau_s + \frac{3\tau_s}{1+\omega_I^2\tau_s^2} + \frac{13\tau_s}{1+\omega_e^2\tau_s^2}\right) \quad 1.7$$

Where W is the local concentration of the Mn^{2+} ions, μ_0 is the vacuum permeability, γ_I is the gyromagnetic ratio of nucleus spin I, μ_{eff} is the effective magnetic moment of Mn^{2+} ions, β is the Bohr magneton, r is the average electron-nucleus distance, ω_I and ω_e are the nucleus and election Larmor frequencies. The correlation time τ_s is the inverse sum of the electronic spin-lattice relaxation time T_{1e}; the rotational correlation time of the molecule τ_r; and the residence time of the Mn^{2+} near the nuclear spin τ_m:

$$\frac{1}{\tau_s} = \frac{1}{T_{1e}} + \frac{1}{\tau_r} + \frac{1}{\tau_m} \quad 1.8$$

Based on equation 1.5, $T_2 \propto r^6$, thus the closer distance between the paramagnetic center and the nucleus of interest, the shorter the T_2, resulting in line broadening and peak suppression. By measuring T_2 of the nucleus, the distance information can be obtained.

The quadrupolar echo (quecho) solid-state NMR experiment is one of the most widely used experiments for studying quadrupolar nuclei in solid samples.[102, 103] This technique can be used to measure the T_2 relaxation time and to minimize effects from probe ring-down.[104] The pulse sequence is shown in Figure 1.22. At the beginning of the sequence, a $\pi/2$ pulse with phase x is

applied. After a duration of time τ_1, the second $\pi/2$ pulse with phase y or -y is applied to refocus the magnetization. The echo maximum appears at τ_2 after the second pulse. In actual experiments, the τ_2 is set shorter than the τ_1 to obtain the maximum intensity. To determine the T_2 relaxation time, the decay of the acquired signal was measured as a function of synchronous incrementation of t, and t = [τ_1 + τ_2 + time being shifted]. By plotting the signal intensity I in FID vs t, the T_2 relaxation time can be calculated by:

$$I(t) = I(0) \times exp\left(\frac{t}{T_2}\right) \qquad 1.9$$

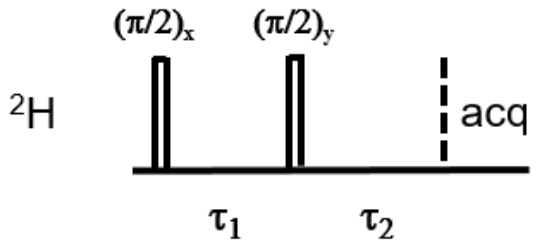

Figure 1.22 Quadrupolar echo (quecho) pulse sequence.

REFERENCES

REFERENCES

[1] Weiss, R. A. (1993) HOW DOES HIV CAUSE AIDS, *Science 260*, 1273-1279.

[2] Briggs, J. A. G., Simon, M. N., Gross, I., Krausslich, H. G., Fuller, S. D., Vogt, V. M., and Johnson, M. C. (2004) The stoichiometry of Gag protein in HIV-1, *Nat. Struct. Mol. Biol. 11*, 672-675.

[3] Levy, J. A. (1993) PATHOGENESIS OF HUMAN-IMMUNODEFICIENCY-VIRUS INFECTION, *Microbiol. Rev. 57*, 183-289.

[4] McCune, J. M., Rabin, L. B., Feinberg, M. B., Lieberman, M., Kosek, J. C., Reyes, G. R., and Weissman, I. L. (1988) ENDOPROTEOLYTIC CLEAVAGE OF GP160 IS REQUIRED FOR THE ACTIVATION OF HUMAN IMMUNODEFICIENCY VIRUS, *Cell 53*, 55-67.

[5] Gelderblom, H. R., Ozel, M., Hausmann, E. H. S., Winkel, T., Pauli, G., and Koch, M. A. (1988) FINE-STRUCTURE OF HUMAN IMMUNODEFICIENCY VIRUS (HIV), IMMUNOLOCALIZATION OF STRUCTURAL PROTEINS AND VIRUS-CELL RELATION, *Micron and Microscopica Acta 19*, 41-60.

[6] Gelderblom, H. R., Ozel, M., and Pauli, G. (1989) MORPHOGENESIS AND MORPHOLOGY OF HIV - STRUCTURE-FUNCTION RELATIONS, *Arch. Virol. 106*, 1-13.

[7] Laskey, S. B., and Siliciano, R. F. (2014) A mechanistic theory to explain the efficacy of antiretroviral therapy, *Nat. Rev. Microbiol. 12*, 772-780.

[8] Grewe, C., Beck, A., and Gelderblom, H. R. (1990) HIV - EARLY VIRUS-CELL INTERACTIONS, *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 3*, 965-974.

[9] Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V., and Melikyan, G. B. (2009) HIV Enters Cells via Endocytosis and Dynamin-Dependent Fusion with Endosomes, *Cell 137*, 433-444.

[10] Gottlieb, M. S., Schroff, R., Schanker, H. M., Weisman, J. D., Fan, P. T., Wolf, R. A., and Saxon, A. (1981) PNEUMOCYSTIS-CARINII PNEUMONIA AND MUCOSAL CANDIDIASIS IN PREVIOUSLY HEALTHY HOMOSEXUAL MEN - EVIDENCE OF A NEW ACQUIRED CELLULAR IMMUNODEFICIENCY, *N. Engl. J. Med. 305*, 1425-1431.

[11] Dalgleish, A. G., Beverley, P. C. L., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984) THE CD4 (T4) ANTIGEN IS AN ESSENTIAL COMPONENT OF THE RECEPTOR FOR THE AIDS RETROVIRUS, *Nature 312*, 763-767.

[12] Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody, *Nature 393*, 648-659.

[13] Broder, C. C., and Dimitrov, D. S. (1996) HIV and the 7-transmembrane domain receptors, *Pathobiology 64*, 171-179.

[14] Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) Core structure of gp41 from the HIV envelope glycoprotein, *Cell 89*, 263-273.

[15] Julien, J. P., Cupo, A., Sok, D., Stanfield, R. L., Lyumkis, D., Deller, M. C., Klasse, P. J., Burton, D. R., Sanders, R. W., Moore, J. P., Ward, A. B., and Wilson, I. A. (2013) Crystal Structure of a Soluble Cleaved HIV-1 Envelope Trimer, *Science 342*, 1477-1483.

[16] Frankel, A. D., and Young, J. A. T. (1998) HIV-1: Fifteen proteins and an RNA, *Annu. Rev. Biochem. 67*, 1-25.

[17] Chan, D. C., and Kim, P. S. (1998) HIV entry and its inhibition, *Cell 93*, 681-684.

[18] Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) Atomic structure of the ectodomain from HIV-1 gp41, *Nature 387*, 426-430.

[19] Lyumkis, D., Julien, J. P., de Val, N., Cupo, A., Potter, C. S., Klasse, P. J., Burton, D. R., Sanders, R. W., Moore, J. P., Carragher, B., Wilson, I. A., and Ward, A. B. (2013) Cryo-EM Structure of a Fully Glycosylated Soluble Cleaved HIV-1 Envelope Trimer, *Science 342*, 1484-1490.

[20] Lu, M., Ji, H., and Shen, S. (1999) Subdomain folding and biological activity of the core structure from human immunodeficiency virus type 1 gp41: Implications for viral membrane fusion, *J. Virol. 73*, 4433-4438.

[21] White, J. M., Delos, S. E., Brecher, M., and Schornberg, K. (2008) Structures and mechanisms of viral membrane fusion proteins: Multiple variations on a common theme, *Crit. Rev. Biochem. Mol. Biol. 43*, 189-219.

[22] Tran, E. E. H., Borgnia, M. J., Kuybeda, O., Schauder, D. M., Bartesaghi, A., Frank, G. A., Sapiro, G., Milne, J. L. S., and Subramaniam, S. (2012) Structural Mechanism of Trimeric HIV-1 Envelope Glycoprotein Activation, *PLoS Pathog. 8*, 18.

[23] Gabrys, C. M., Qiang, W., Sun, Y., Xie, L., Schmick, S. D., and Weliky, D. P. (2013) Solid-State Nuclear Magnetic Resonance Measurements of HIV Fusion Peptide (CO)-C-13 to Lipid P-31 Proximities Support Similar Partially Inserted Membrane Locations of the alpha Helical and beta Sheet Peptide Structures, *J. Phys. Chem. A 117*, 9848-9859.

[24] Gabrys, C. M., and Weliky, D. P. (2007) Chemical shift assignment and structural plasticity of a HIV fusion peptide derivative in dodecylphosphocholine micelles, *Biochim. Biophys. Acta-Biomembr. 1768*, 3225-3234.

[25] Jaroniec, C. P., Kaufman, J. D., Stahl, S. J., Viard, M., Blumenthal, R., Wingfield, P. T., and Bax, A. (2005) Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain, *Biochemistry 44*, 16167-16180.

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

[27] Schmick, S. D., and Weliky, D. P. (2010) Major Antiparallel and Minor Parallel beta Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide, *Biochemistry 49*, 10623-10635.

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

[29] Jia, L. H., Liang, S., Sackett, K., Xie, L., Ghosh, U., and Weliky, D. P. (2015) REDOR solid-state NMR as a probe of the membrane locations of membrane-associated peptides and proteins, *J. Magn. Reson. 253*, 154-165.

[30] Sackett, K., Nethercott, M. J., Zheng, Z. X., and Weliky, D. P. (2014) Solid-State NMR Spectroscopy of the HIV gp41 Membrane Fusion Protein Supports Intermolecular Antiparallel 13 Sheet Fusion Peptide Structure in the Final Six-Helix Bundle State, *J. Mol. Biol. 426*, 1077-1094.

[31] Schibli, D. J., Montelaro, R. C., and Vogel, H. J. (2001) The membrane-proximal tryptophan-rich region of the HIV glycoprotein, gp41, forms a well-defined helix in dodecylphosphocholine micelles, *Biochemistry 40*, 9570-9578.

[32] Sun, Z. Y. J., Oh, K. J., Kim, M. Y., Yu, J., Brusic, V., Song, L. K., Qiao, Z. S., Wang, J. H., Wagner, G., and Reinherz, E. L. (2008) HIV-1 broadly neutralizing antibody extracts its epitope from a kinked gp41 ectodomain region on the viral membrane, *Immunity 28*, 52-63.

[33] Shi, W. X., Bohon, J., Han, D. P., Habte, H., Qin, Y. L., Cho, M. W., and Chance, M. R. (2010) Structural Characterization of HIV gp41 with the Membrane-proximal External Region, *J. Biol. Chem. 285*, 24290-24298.

[34] Gangupomu, V. K., and Abrams, C. F. (2010) All-Atom Models of the Membrane-Spanning Domain of HIV-1 gp41 from Metadynamics, *Biophys. J. 99*, 3438-3444.

[35] Apellaniz, B., Rujas, E., Serrano, S., Morante, K., Tsumoto, K., Caaveiro, J. M. M., Jimenez, M. A., and Nieva, J. L. (2015) The Atomic Structure of the HIV-1 gp41 Transmembrane Domain and Its Connection to the Immunogenic Membrane-proximal External Region, *J. Biol. Chem. 290*, 12999-13015.

[36] Kim, J. H., Hartley, T. L., Curran, A. R., and Engelman, D. M. (2009) Molecular dynamics studies of the transmembrane domain of gp41 from HIV-1, *Biochim. Biophys. Acta-Biomembr. 1788*, 1804-1812.

[37] Chernomordik, L. V., and Kozlov, M. M. (2008) Mechanics of membrane fusion, *Nat. Struct. Mol. Biol. 15*, 675-683.

[38] Markosyan, R. M., Cohen, F. S., and Melikyan, G. B. (2003) HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation, *Mol. Biol. Cell 14*, 926-938.

[39] Wild, C. T., Shugars, D. C., Greenwell, T. K., McDanal, C. B., and Matthews, T. J. (1994) PEPTIDES CORRESPONDING TO A PREDICTIVE ALPHA-HELICAL DOMAIN OF HUMAN-IMMUNODEFICIENCY-VIRUS TYPE-1 GP41 ARE POTENT INHIBITORS OF VIRUS-INFECTION, *Proc. Natl. Acad. Sci. U. S. A. 91*, 9770-9774.

[40] Melikyan, G. B., Markosyan, R. M., Hemmati, H., Delmedico, M. K., Lambert, D. M., and Cohen, F. S. (2000) Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion, *J. Cell Biol. 151*, 413-423.

[41] Banerjee, K., and Weliky, D. P. (2014) Folded Monomers and Hexamers of the Ectodomain of the HIV gp41 Membrane Fusion Protein: Potential Roles in Fusion and Synergy Between the Fusion Peptide, Hairpin, and Membrane-Proximal External Region, *Biochemistry 53*, 7184-7198.

[42] Magnus, C., Rusert, P., Bonhoeffer, S., Trkola, A., and Regoes, R. R. (2009) Estimating the Stoichiometry of Human Immunodeficiency Virus Entry, *J. Virol. 83*, 1523-1531.

[43] Yoon, S. W., Webby, R. J., and Webster, R. G. (2014) Evolution and Ecology of Influenza A Viruses, In *Influenza Pathogenesis and Control - Vol I* (Compans, R. W., and Oldstone, M. B. A., Eds.), pp 359-375, Springer Int Publishing Ag, Cham.

[44] Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., and Kawaoka, Y. (1992) EVOLUTION AND ECOLOGY OF INFLUENZA-A VIRUSES, *Microbiol. Rev. 56*, 152-179.

[45] Claas, E. C. J., Sprenger, M. J. W., Kleter, G. E. M., Vanbeek, R., Quint, W. G. V., and Masurel, N. (1992) TYPE-SPECIFIC IDENTIFICATION OF INFLUENZA VIRUS-A, VIRUS-B AND VIRUS-C BY THE POLYMERASE CHAIN-REACTION, *J. Virol. Methods 39*, 1-13.

[46] Harris, A., Cardone, G., Winkler, D. C., Heymann, J. B., Brecher, M., White, J. M., and Steven, A. C. (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography, *Proc. Natl. Acad. Sci. U. S. A. 103*, 19123-19127.

[47] Wilson, I. A., Skehel, J. J., and Wiley, D. C. (1981) STRUCTURE OF THE HEMAGGLUTININ MEMBRANE GLYCOPROTEIN OF INFLUENZA-VIRUS AT 3-A RESOLUTION, *Nature 289*, 366-373.

[48] Gamblin, S. J., and Skehel, J. J. (2010) Influenza Hemagglutinin and Neuraminidase Membrane Glycoproteins, *J. Biol. Chem. 285*, 28403-28409.

[49] Chen, J., Skehel, J. J., and Wiley, D. C. (1999) N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA(2) subunit to form an N cap that terminates the triple-stranded coiled coil, *Proc. Natl. Acad. Sci. U. S. A. 96*, 8967-8972.

[50] Curtis-Fisk, J., Preston, C., Zheng, Z. X., Worden, R. M., and Weliky, D. P. (2007) Solid-state NMR structural measurements on the membrane-associated influenza fusion protein ectodomain, *J. Am. Chem. Soc. 129*, 11320-+.

[51] Schnell, J. R., and Chou, J. J. (2008) Structure and mechanism of the M2 proton channel of influenza A virus, *Nature 451*, 591-U512.

[52] O'Neill, R. E., Talon, J., and Palese, P. (1998) The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins, *Embo J. 17*, 288-296.

[53] Shi, Y., Wu, Y., Zhang, W., Qi, J. X., and Gao, G. F. (2014) Enabling the 'host jump': structural determinants of receptor-binding specificity in influenza A viruses, *Nat. Rev. Microbiol. 12*, 822-831.

[54] Lakadamyali, M., Rust, M. J., and Zhuang, X. W. (2004) Endocytosis of influenza viruses, *Microbes Infect. 6*, 929-936.

[55] Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) STRUCTURE OF INFLUENZA HEMAGGLUTININ AT THE PH OF MEMBRANE-FUSION, *Nature 371*, 37-43.

[56] Wiley, D. C., and Skehel, J. J. (1987) THE STRUCTURE AND FUNCTION OF THE HEMAGGLUTININ MEMBRANE GLYCOPROTEIN OF INFLUENZA-VIRUS, *Annu. Rev. Biochem. 56*, 365-394.

[57] Skehel, J. J., and Wiley, D. C. (2000) Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin, *Annu. Rev. Biochem. 69*, 531-569.

[58] Sriwilaijaroen, N., and Suzuki, Y. (2012) Molecular basis of the structure and function of H1 hemagglutinin of influenza virus, *Proc. Jpn. Acad. Ser. B-Phys. Biol. Sci. 88*, 226-249.

[59] Nobusawa, E., Aoyama, T., Kato, H., Suzuki, Y., Tateno, Y., and Nakajima, K. (1991) COMPARISON OF COMPLETE AMINO-ACID-SEQUENCES AND RECEPTOR-BINDING PROPERTIES AMONG 13 SEROTYPES OF HEMAGGLUTININS OF INFLUENZA A-VIRUSES, *Virology 182*, 475-485.

[60] Gething, M. J., Doms, R. W., York, D., and White, J. (1986) STUDIES ON THE MECHANISM OF MEMBRANE-FUSION - SITE-SPECIFIC MUTAGENESIS OF THE HEMAGGLUTININ OF INFLUENZA-VIRUS, *J. Cell Biol. 102*, 11-23.

[61] Lorieau, J. L., Louis, J. M., and Bax, A. (2010) The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface, *Proc. Natl. Acad. Sci. U. S. A. 107*, 11341-11346.

[62] Ghosh, U., Xie, L., and Weliky, D. P. (2013) Detection of closed influenza virus hemagglutinin fusion peptide structures in membranes by backbone (CO)-C-13-N-15 rotational-echo double-resonance solid-state NMR, *J. Biomol. NMR 55*, 139-146.

[63] Ghosh, U., Xie, L., Jia, L. H., Liang, S., and Weliky, D. P. (2015) Closed and Semiclosed Interhelical Structures in Membrane vs Closed and Open Structures in Detergent for the Influenza Virus Hemagglutinin Fusion Peptide and Correlation of Hydrophobic Surface Area with Fusion Catalysis, *J. Am. Chem. Soc. 137*, 7548-7551.

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

[65] Sun, Y., and Weliky, D. P. (2009) C-13-C-13 Correlation Spectroscopy of Membrane-Associated Influenza Virus Fusion Peptide Strongly Supports a Helix-Turn-Helix Motif and Two Turn Conformations, *J. Am. Chem. Soc. 131*, 13228-13229.

[66] Chen, J., Lee, K. H., Steinhauer, D. A., Stevens, D. J., Skehel, J. J., and Wiley, D. C. (1998) Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation, *Cell 95*, 409-417.

[67] Harrison, S. C. (2008) Viral membrane fusion, *Nat. Struct. Mol. Biol. 15*, 690-698.

[68] Floyd, D. L., Ragains, J. R., Skehel, J. J., Harrison, S. C., and van Oijen, A. M. (2008) Single-particle kinetics of influenza virus membrane fusion, *Proc. Natl. Acad. Sci. U. S. A. 105*, 15382-15387.

[69] Wiley, D. C., Wilson, I. A., and Skehel, J. J. (1981) STRUCTURAL IDENTIFICATION OF THE ANTIBODY-BINDING SITES OF HONG-KONG INFLUENZA HEMAGGLUTININ AND THEIR INVOLVEMENT IN ANTIGENIC VARIATION, *Nature 289*, 373-378.

[70] Carr, C. M., and Kim, P. S. (1993) A SPRING-LOADED MECHANISM FOR THE CONFORMATIONAL CHANGE OF INFLUENZA HEMAGGLUTININ, *Cell 73*, 823-832.

[71] Chen, J., Skehel, J. J., and Wiley, D. C. (1998) A polar octapeptide fused to the N-terminal fusion peptide solubilizes the influenza virus HA(2) subunit ectodomain, *Biochemistry 37*, 13643-13649.

[72] Ratnayake, P. U., Ekanayaka, E. A. P., Komanduru, S. S., and Weliky, D. P. (2016) Full-length trimeric influenza virus hemagglutinin II membrane fusion protein and shorter constructs lacking the fusion peptide or transmembrane domain: Hyperthermostability of the full-length protein and the soluble ectodomain and fusion peptide make significant contributions to fusion of membrane vesicles, *Protein Expr. Purif. 117*, 6-16.

[73] Carr, C. M., Chaudhry, C., and Kim, P. S. (1997) Influenza hemagglutinin is spring-loaded by a metastable native conformation, *Proc. Natl. Acad. Sci. U. S. A. 94*, 14306-14313.

[74] Kemble, G. W., Danieli, T., and White, J. M. (1994) LIPID-ANCHORED INFLUENZA HEMAGGLUTININ PROMOTES HEMIFUSION, NOT COMPLETE FUSION, *Cell 76*, 383-391.

[75] Griffin, R. G. (1998) Dipolar recoupling in MAS spectra of biological solids, *Nat. Struct. Biol. 5*, 508-512.

[76] Marassi, F. M., and Opella, S. J. (1998) NMR structural studies of membrane proteins, *Curr. Opin. Struct. Biol. 8*, 640-648.

[77] Andrew, E. R., Bradbury, A., and Eades, R. G. (1958) NUCLEAR MAGNETIC RESONANCE SPECTRA FROM A CRYSTAL ROTATED AT HIGH SPEED, *Nature 182*, 1659-1659.

[78] Duer, M. J. (2002) *Solid-state NMR spectroscopy principles and applications*, Blackwell Science Ltd.

[79] Gullion, T. (1998) Introduction to rotational-echo, double-resonance NMR, *Concepts Magn. Resonance 10*, 277-289.

[80] Gullion, T., and Schaefer, J. (1989) ROTATIONAL-ECHO DOUBLE-RESONANCE NMR, *J. Magn. Reson. 81*, 196-200.

[81] Xie, L., Ghosh, U., Schmick, S. D., and Weliky, D. P. (2013) Residue-specific membrane location of peptides and proteins using specifically and extensively deuterated lipids and C-13-H-2 rotational-echo double-resonance solid-state NMR, *J. Biomol. NMR 55*, 11-17.

[82] Gullion, T., Kishore, R., and Asakura, T. (2003) Determining dihedral angles and local structure in silk peptide by C-13-H-2 REDOR, *J. Am. Chem. Soc. 125*, 7510-7511.

[83] Schmidt, A., Kowalewski, T., and Schaefer, J. (1993) LOCAL PACKING IN GLASSY POLYCARBONATE BY CARBON DEUTERIUM REDOR NMR, *Macromolecules 26*, 1729-1733.

[84] Mueller, K. T. (1995) ANALYTIC SOLUTIONS FOR THE TIME EVOLUTION OF DIPOLAR-DEPHASING NMR SIGNALS, *J. Magn. Reson. Ser. A 113*, 81-93.

[85] Bak, M., Rasmussen, J. T., and Nielsen, N. C. (2000) SIMPSON: A general simulation program for solid-state NMR spectroscopy, *J. Magn. Reson. 147*, 296-330.

[86] Solomon, I. (1955) RELAXATION PROCESSES IN A SYSTEM OF 2 SPINS, *Physical Review 99*, 559-565.

[87] Clore, G. M., and Iwahara, J. (2009) Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes, *Chem. Rev. 109*, 4108-4139.

[88] Buffy, J. J., Hong, T., Yamaguchi, S., Waring, A. J., Lehrer, R. I., and Hong, M. (2003) Solid-state NMR investigation of the depth of insertion of protegrin-1 in lipid bilayers using paramagnetic Mn2+, *Biophys. J. 85*, 2363-2373.

[89] Chu, S. D., Maltsev, S., Emwas, A. H., and Lorigan, G. A. (2010) Solid-state NMR paramagnetic relaxation enhancement immersion depth studies in phospholipid bilayers, *J. Magn. Reson. 207*, 89-94.

[90] Su, Y., Mani, R., and Hong, M. (2008) Asymmetric insertion of membrane proteins in lipid bilayers by solid-state NMR paramagnetic relaxation enhancement: A cell-penetrating peptide example, *J. Am. Chem. Soc. 130*, 8856-8864.

[91] Schmidt, P. G., and Kuntz, I. D. (1984) DISTANCE MEASUREMENTS IN SPIN-LABELED LYSOZYME, *Biochemistry 23*, 4261-4266.

[92] Cai, S., Seu, C., Kovacs, Z., Sherry, A. D., and Chen, Y. (2006) Sensitivity enhancement of multidimensional NMR experiments by paramagnetic relaxation effects, *J. Am. Chem. Soc. 128*, 13474-13478.

[93] Linser, R., Chevelkov, V., Diehl, A., and Reif, B. (2007) Sensitivity enhancement using paramagnetic relaxation in MAS solid-state NMR of perdeuterated proteins, *J. Magn. Reson. 189*, 209-216.

[94] Caravan, P. (2006) Strategies for increasing the sensitivity of gadolinium based MRI contrast agents, *Chem. Soc. Rev. 35*, 512-523.

[95] Flacke, S., Fischer, S., Scott, M. J., Fuhrhop, R. J., Allen, J. S., McLean, M., Winter, P., Sicard, G. A., Gaffney, P. J., Wickline, S. A., and Lanza, G. M. (2001) Novel MRI contrast agent for molecular imaging of fibrin implications for detecting vulnerable plaques, *Circulation 104*, 1280-1285.

[96] Lorieau, J. L., Louis, J. M., Schwieters, C. D., and Bax, A. (2012) pH-triggered, activated-state conformations of the influenza hemagglutinin fusion peptide revealed by NMR, *Proc. Natl. Acad. Sci. U. S. A. 109*, 19994-19999.

[97] Huster, D., Yao, X. L., and Hong, M. (2002) Membrane protein topology probed by H-1 spin diffusion from lipids using solid-state NMR spectroscopy, *J. Am. Chem. Soc. 124*, 874-883.

[98] Kumashiro, K. K., Schmidt-Rohr, K., Murphy, O. J., Ouellette, K. L., Cramer, W. A., and Thompson, L. K. (1998) A novel tool for probing membrane protein structure: Solid-state NMR with proton spin diffusion and X-nucleus detection, *J. Am. Chem. Soc. 120*, 5043-5051.

[99] Gallagher, G. J., Hong, M., and Thompson, L. K. (2004) Solid-state NMR spin diffusion for measurement of membrane-bound peptide structure: Gramicidin A, *Biochemistry 43*, 7899-7906.

[100] Macosko, J. C., Kim, C. H., and Shin, Y. K. (1997) The membrane topology of the fusion peptide region of influenza hemagglutinin determined by spin-labeling EPR, *J. Mol. Biol. 267*, 1139-1148.

[101] Bloembergen, N. (1957) PROTON RELAXATION TIMES IN PARAMAGNETIC SOLUTIONS, *J. Chem. Phys. 27*, 572-573.

[102] Davis, J. H., Jeffrey, K. R., Bloom, M., Valic, M. I., and Higgs, T. P. (1976) QUADRUPOLAR ECHO DEUTERON MAGNETIC-RESONANCE SPECTROSCOPY IN ORDERED HYDROCARBON CHAINS, *Chem. Phys. Lett. 42*, 390-394.

[103] Weisman, I. D., and Bennett, L. H. (1969) QUADRUPOLAR ECHOES IN SOLIDS, *Physical Review 181*, 1341-&.

[104] Gabrys, C. M., Yang, R., Wasniewski, C. M., Yang, J., Canlas, C. G., Qiang, W., Sun, Y., and Weliky, D. P. (2010) Nuclear magnetic resonance evidence for retention of a lamellar membrane phase with curvature in the presence of large quantities of the HIV fusion peptide, *Biochim. Biophys. Acta-Biomembr. 1798*, 194-201.

Chapter 2

Materials and Methods

2.1 Materials

The DNA plasmids containing HM, HM_TM and FP_HM and FP_HM_TM genes were ordered from GenScript (Piscataway, NJ). The protein expression cell Escherichia coli BL21(DE3) strain was purchased from Novagen (Gibbstown, NJ). Luria-Bertani broth (LB) medium was purchased from Dot Scientific (Burton, MI); isopropyl β-D-thiogalactopyranoside (IPTG) from Goldbio (St. Louis, MO); Cobalt affinity resin from Thermo Scientific (Waltham, MA). Wang resins, Fmoc protected amino acids and t-Boc protected amino acids were purchased from Peptides International (Louisville, KY), Dupont (Wilmington, Delaware) and Sigma Aldrich (St. Louis, MO). 1-^{13}C Gly and 1-^{13}C Ala were purchased from Cambridge Isotope Laboratories (Andover, MA) and were N-Fmoc or N-t-Boc protected in our laboratory following literature procedures.[1-3] 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-[phospho-rac-(1-glycerol)] (sodium salt) (POPG), 1,2-Dipalmitoyl-sn-glycero-3-phosphocholine (DPPG) and cholesterol were purchased from Avanti Polar Lipids (Alabaster, AL). Deuterium labeled fatty acids Hexadecanoic-2,2-d2 Acid; Hexadecanoic-7,7,8,8-d4 Acid and Hexadecanoic-15,15,16,16,16-d5 Acid were purchased from CDN Isotopes (Quebec, Canada) and shipped to Avanti Polar Lipids (Alabaster, AL) to synthesize DPPC-D4, DPPC-D8 and DPPC-D10 lipids (Structure shown in Figure 2.1). Other materials were obtained from Sigma-Aldrich (St. Louis, MO).

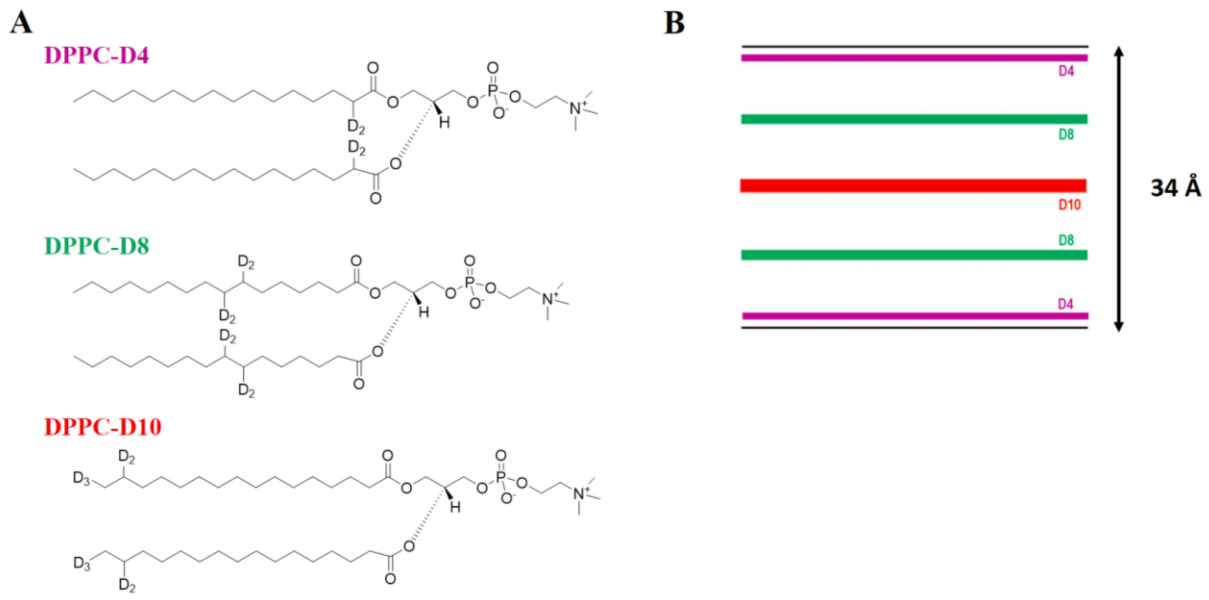

Figure 2.1 (A) Structure of DPPC-D4, DPPC-D8 and DPPC-D10. (B) Approximate membrane locations of the ^{2}H's in the membrane bilayer without protein. The lipid ^{2}H are for the membrane gel-phase.[4]

2.2 Gp41 constructs expression and purification

Figure 2.2 shows the schematic diagram and amino acid sequence of the four gp41 constructs purified in this study. Amino acid sequences of gp41 constructs are for the HXB2 laboratory strain of HIV and use gp160 numbering with color coding of different regions. HM includes residues 535(M535C) – 581 + residue 628 – 683. FP_HM includes residue 512 – 581 + residue 628 - 683. HM_TM includes residue 535(M535C) – 581+ residue 628 – 705. FP_HM_TM includes residue 512 – 581 + residue 628 – 705. All constructs include a non-native sequence SGGRGG in place of residue 582 – 627 and a non-native H_6 affinity tag at their C-terminus that is preceded by a G_6LE or G_8LE spacer. The spacer is needed for exposure of the affinity tag during purification. The M535C mutation in HM and HM_TM is needed for native chemical ligation with FP.

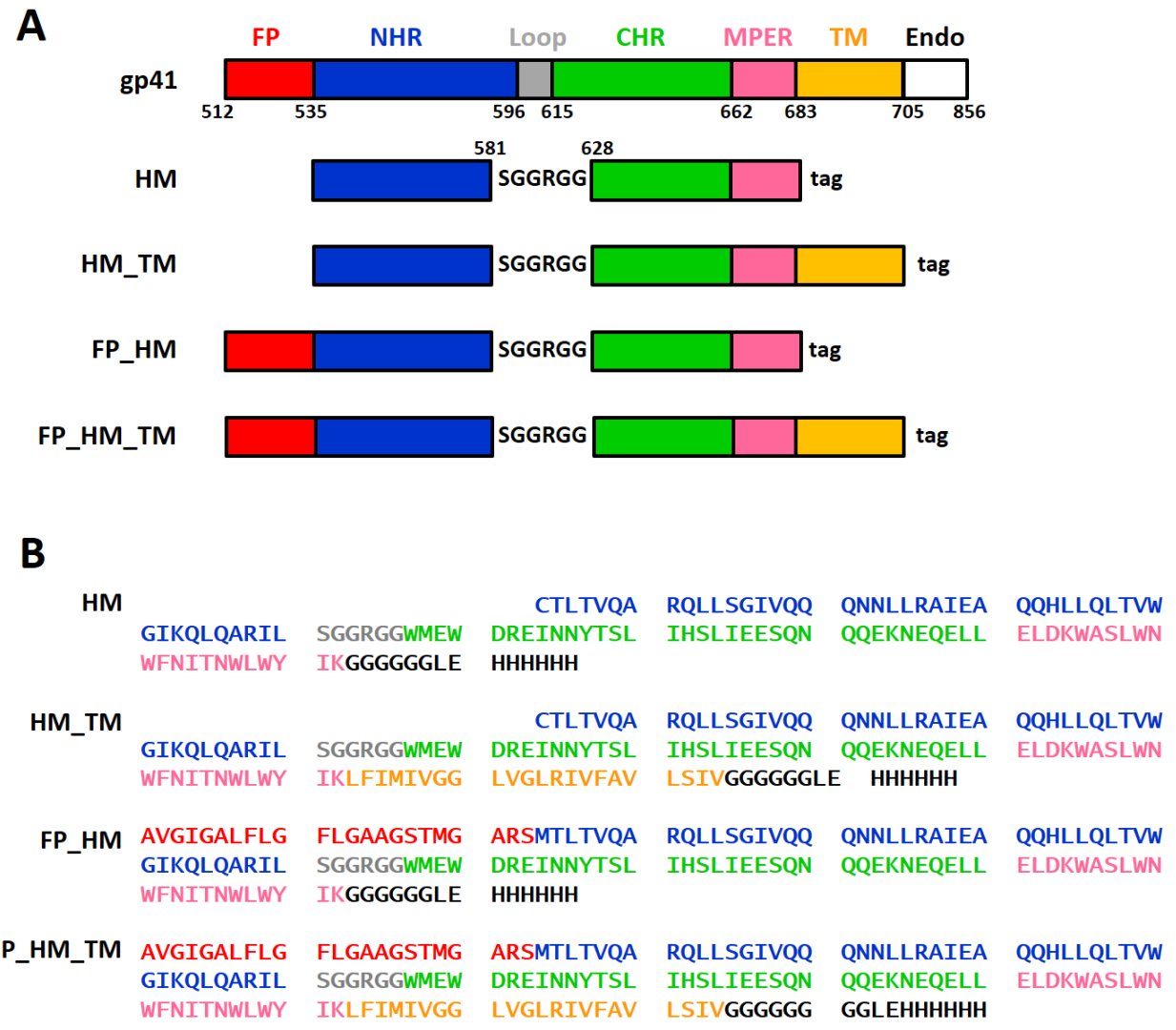

Figure 2.2 (A) Schematic diagrams of full-length HIV gp41 and the four gp41 constructs being studied in this work with domains and corresponding colors: FP ≡ fusion peptide, red; N-helix, blue; Loop, grey; C-helix, green; MPER ≡ membrane-proximal external-region, pink; TM ≡ transmembrane domain, orange; and endo ≡ endodomain, white. The four constructs have non-native SGGRGG replacing native residues 582-627. (B) Amino acid sequences with the same color coding as panel A and the non-native C-terminal G_6H_6 or G_8H_6 in black. The H_6 is for Co^{2+}-affinity chromatography and the G_6/G_8 are necessary spacers for exposure of the H_6 tag.

DNA sequences of these four gp41 constructs are shown in Figure 2.3. All DNA inserts were subcloned into pET-24a(+) vector that contains Lac operon and kanamycin resistance. The plasmid was transformed into *E. coli* BL21(DE3) strain and then grew in 50 mL LB broth at 37 °C overnight. The culture was mixed with 50% glycerol with 1:1 volume ratio and stored at -80 °C as glycerol stock. The protein expression started with adding 50 μL of the bacteria glycerol stock into 50 mL LB medium with Kanamycin 50 mg/L concentration. The 50 mg/L Kanamycin was used to select for bacterial cells that contains the plasmid. After growth at 37 °C and 180 rpm shaking overnight, the 50 mL culture was added to 1 L LB medium and kept growing for ~ 2 hr at the same condition. When the OD_{600} of the culture increased to ~0.8, 2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to start the induction. After 5 hr of induction, the cell suspension was centrifuged at 9000 g, 4 °C for 10 min and the cell pellet was kept in -20 °C freezer.

5 g of wet cell was lysed by tip sonication in 40 mL phosphate-buffered saline (PBS) at pH 7.4. The insoluble fragments were pelleted down by centrifugation (48000 g, 4 °C for 20 min). The pellet contains the desired recombinant protein (RP) since the RP is not soluble in PBS. The PBS wash process was repeated two more times to remove the soluble proteins, molecules and suspended membrane fragments which are only effectively pelleted by >100000 g. The RP pellet was tip sonicated in 40 mL of PBS at pH 7.4 that also contained 8 M Urea, 0.5% sodium dodecyl sulfate (SDS) and 0.8% Sodium lauroyl sarcosinate (Sarkosyl) until fully dissolved. 1 mL of Co^{2+} affinity resin was added to the solution and agitated for 2 hours at room temperature. After 2 hr binding, the unbound proteins were removed by gravity filtration of the suspension. The following steps were different to solubilize gp41 constructs in SDS or dodecylphosphocholine (DPC).

HM

TGTACGCTGACGGTCCAAGCACGTCAGCTGCTGAGCGGCATTGTGCAGCAACAGAACAATCTGCTGCGCGCGATC
GAAGCCCAACAGCATCTGCTGCAGCTGACCGTTTGGGGTATTAAACAACTGCAGGCTCGTATCCTGAGCGGCGGT
CGCGGCGGTTGGATGGAATGGGATCGTGAAATTAACAATTATACGAGCCTGATTCACTCTCTGATCGAAGAAAGT
CAAAACCAACAGGAGAAAAACGAACAGGAACTGCTGGAACTGGACAAATGGGCCTCCCTGTGGAACTGGTTTAAC
ATTACGAACTGGCTGTGGTACATCAAAGGCGGCGGTGGCGGTGGT

HM_TM

TGTACCCTGACGGTCCAAGCTCGCCAACTGCTGAGTGGTATCGTTCAACAACAAAACAATCTGCTGCGTGCTATC
GAAGCCCAACAACATCTGCTGCAGCTGACCGTGTGGGGCATTAAACAGCTGCAGGCCCGTATCCTGAGCGGCGGT
CGCGGCGGTTGGATGGAATGGGATCGTGAAATTAACAATTATACGAGCCTGATTCACTCTCTGATCGAAGAAAGT
CAGAATCAGCAAGAGAAAAACGAACAAGAACTGCTGGAACTGGACAAATGGGCCTCCCTGTGGAACTGGTTTAAT
ATTACCAACTGGCTGTGGTACATCAAACTGTTCATTATGATCGTTGGCGGTCTGGTCGGTCTGCGTATTGTGTTT
GCTGTCCTGAGTATTGTCGGTGGCGGTGGCGGCGGT

FP_HM

GCCGTGGGTATCGGTGCTCTGTTCCTGGGTTTCCTGGGTGCTGCTGGTTCGACGATGGGTGCCCGCTCAATGACG
CTGACGGTCCAAGCACGTCAGCTGCTGAGCGGCATTGTGCAGCAACAGAACAATCTGCTGCGCGCGATCGAAGCC
CAACAGCATCTGCTGCAGCTGACCGTTTGGGGTATTAAACAACTGCAGGCTCGTATCCTGAGCGGCGGTCGCGGC
GGTTGGATGGAATGGGATCGTGAAATTAACAATTATACGAGCCTGATTCACTCTCTGATCGAAGAAAGTCAAAAC
CAACAGGAGAAAAACGAACAGGAACTGCTGGAACTGGACAAATGGGCCTCCCTGTGGAACTGGTTTAACATTACG
AACTGGCTGTGGTACATCAAAGGCGGCGGTGGCGGTGGT

FP_HM_TM

GCGGTTGGCATTGGTGCGCTGTTCCTGGGCTTTCTGGGTGCGGCGGGTAGCACTATGGGTGCGCGTAGCATGACC
CTGACCGTTCAAGCGCGTCAACTGCTGAGCGGTATCGTGCAGCAACAGAACAACCTGCTGCGTGCGATTGAGGCG
CAACAGCACCTGCTGCAGCTGACCGTTTGGGGCATCAAGCAACTGCAGGCGCGTATTCTGAGCGGTGGCCGTGGT
GGCTGGATGGAGTGGGACCGTGAAATCAACAACTACACCAGCCTGATCCACAGCCTGATTGAGGAAAGCCAAAAC
CAACAGGAGAAGAACGAGCAGGAACTGCTGGAACTGGATAAATGGGCGAGCCTGTGGAACTGGTTCAACATCACC
AACTGGCTGTGGTACATCAAACTGTTCATCATGATCGTGGGTGGCCTGGTTGGTCTGCGTATCGTGTTTGCGGTT
CTGAGCATTGTTGGCGGTGGTGGTGGCGGTGGTGGT

Figure 2.3 DNA sequences of gp41 inserts. Each line is 75 nucleotides.

To solubilized proteins in SDS detergent, the Co^{2+} resin was washed with 3×1 mL aliquots of the PBS/urea/SDS/Sarkosyl solution and subsequent gravity filtration to remove unbound protein.[5] Bound protein was eluted from the resin by 4×0.5 mL aliquots of the PBS/urea/SDS/Sarkosyl solution with 250 mM imidazole, and gravity filtration. The eluent fractions were pooled and then mixed with an equal volume of buffer that contained 10 mM Tris-HCl at pH 8.0, 0.17% n-Decyl- β-D-Maltoside (DM), 2 mM EDTA, and 1 M L-arginine, with subsequent agitation overnight at 4 °C. Arginine, urea, and detergents were removed by dialysis against 10 mM Tris at pH 7.4 with 0.2% SDS. If the intermediate mixing step with the arginine solution was skipped, RP precipitated during dialysis, which suggests that RP aggregates are broken up by the arginine solution.

To solubilize proteins in DPC detergent, proteins need to be refolded on column to prevent aggregation. After RP binds on Co^{2+} resin, column was washed by 3×1 mL of 20 mM sodium phosphate at pH 7.4 with 0.25% DPC, with subsequent gravity filtration. RP was then eluted in the pH 7.4 phosphate buffer with 0.25% DPC and 250 mM imidazole. The eluent fractions were pooled and imidazole removed by dialysis against pH 7.4 phosphate buffer with 0.25% DPC, or by dialysis against 20 mM sodium acetate buffer at pH 3.2 with 0.25% DPC. Purified RP concentrations were determined using A280.

2.3 Circular Dichroism (CD)

Spectra were obtained with a Chirascan instrument (Applied Photophysics) at ambient temperature. Proteins with 10-15 μM concentration were added to a 1 mm path length quartz cuvette. 190 nm – 260 nm spectral range was scanned with 0.5 nm steps and 1.5 s per step averaging time. For each protein samples, spectra were averaged and the final spectra were the

(protein + buffer) – (buffer) difference. Thermal stability of folded protein was probed by CD in 5 °C increments over a 25-90 °C range using a J-810 instrument (Jasco) with a water circulation bath. No visible precipitation was observed at any temperature or after cooling down.

2.4 Size-Exclusion Chromatography (SEC)

The chromatography was done using a DuoFlow Pathfinder 20 instrument (Bio-Rad) with a Tricorn Superdex 200 Increase 10/300 GL column (GE Technologies). Flow rate was set to 0.3 mL/min and the detector was set to A280. Proteins were dialyzed against the running buffer, concentrated to ~ 1 mg/mL and centrifuged to remove any precipitate before injected to the instrument. There was no visible precipitate for samples in SDS and a very small precipitate for samples in DPC. A ~100 μL aliquot was injected for each run.

2.5 Protein induced vesicle fusion

The fusion activities of gp41 constructs were assessed by their ability to induce fusion between unilamellar lipid vesicles.[6] Two vesicle compositions were used in this assay: (1) 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG) and Cholesterol (Chol) with 8:2:5 mole ratio; and (2) POPC:Chol with 2:1mole ratio. The Chol mole fraction in both compositions is close to that of the plasma membrane of host cells of HIV. The lipid vesicles with these two compositions are negatively charged and neutral, respectively, so comparison between them allows assessment of the impact of protein/vesicle electrostatics on fusion. To prepare unilamellar lipid vesicles, 1.6 μmol of POPC, 0.4 μmol of POPG and 1 μmol of Chol or 2 μmol of POPC and 1 μmol of Chol were dissolved in 1 mL chloroform and then the solvent was evaporated by dry nitrogen gas and

keep vacuum dry overnight. The dry lipids films were suspended in 2 mL of aqueous buffers (10 mM HEPES, 5 mM MES at pH 7.4 or 10 mM sodium formate at pH 3.2) and a homogeneous suspension was created by 10 freeze/thaw cycles. Unilamellar vesicles were obtained by extrusion of the suspension through a polycarbonate membrane with 100 nm diameter pores for 10 times. For each composition, a set of companions "labeled" vesicles were also prepared that contained additional 2 mol % of the fluorescent lipid NDB-PE and 2 mol % of the quenching lipid Rh-PE. Labeled vesicles were mixed with unlabeled vesicles in 1:9 ratio. Considering a 30% lipid loss during the extrusion, the final concentration of [POPC+POPG] lipids is ~150 μM and chol is ~75 μM. The solution was transferred to a quartz cuvette in a fluorimeter and subjected to constant stirring at 37 °C. Fluorescence was monitored using 467 nm excitation, 530 nm detection, and 1 s time increment. After measurement of the baseline fluorescence F_0, a protein aliquot was added and marked time t = 0. The stock solution contained [protein] ~ 40 μM in 10 mM Tris buffer at pH 7.4 with 150 mM NaCl and 0.2% SDS. The time-dependent fluorescence increase $\Delta F(t) = F(t) - F_0$ is resulted from protein-mediated fusion between a labeled and unlabeled vesicle and causing a longer average distance between fluorophore-quencher lipids. The dead-time in the assay was ~10 s and final asymptotic fluorescence was usually reached by ~600 s. The maximum fluorescence change (ΔF_{max}) was detected after addition of 12 μL 10% Triton X-100 which solubilized the vesicles. The percent fusion was calculated as $[\Delta F(t)/\Delta Fmax] \times 100$. No fusion was detected after the addition of the detergent solution without protein.

2.6 Solid phase peptide synthesis, purification and characterization

The IFP sequence used in this research is: GLFGAIAGFIENGWEGMIDGGGKKKKG. The underlined residues are the first 20 residues of the N-terminal of HA2 subunit, and the

sequence is the H3 subtype of hemagglutinin.[7] The IFP has a non-native C-terminal tag. The four Lysine residues were designed to increase the peptide solubility in water, which makes the purification and NMR sample preparation easier to perform. Four IFP samples were synthesized with different isotopically labeling scheme. They are Leucine 2, Alanine 7 and Glycine 16 residues carbonyl carbon ^{13}C labeled and denoted as IFP_L2C, IFP_A7C and IFP_G16C, respectively.

IFP was manually synthesized by Fmoc solid phase peptide synthesis. The Fmoc peptide synthesis started with Wang resin (polymer-bounded 4-Benzyloxybenzyl alcohol), which has one Fmoc protected Glycine attached to it. The following steps added one amino acid to the resin, and the peptide can be synthesized by repeating these steps:

1. Deprotection: 20% piperidine in DMF was added to the resin to remove the Fmoc group from the existing peptide N-terminal.

2. Amino acid coupling: amino acid with HOBt, HBTU and DIEA in DMF can couple the amino acid to unprotected peptide N-terminal.

3. Capping: 5% acetic anhydride, 2% DIEA and 0.2% HOBt in DMF was used to esterify the N-terminus of the peptide and stop the following side reaction from the unreacted peptide.

Before doing the next step, the solution from the previous step was drained by filtration and the resin was washed by DMF. After the desired sequence is synthesized, the peptide was cleaved from the resin by using 5% water, 2% Tri-isopropyl silane, 2% thioanisole and 2%1, 2-ethane-dithiol in TFA solution and precipitated in cold ether. The peptide was purified using reversed phase HPLC equipped with a semi-preparative C18 column using water - acetonitrile containing 0.1% TFA as mobile phase. TFA helps to maintain the acidic pH (water with 0.1% TFA has a pH ~ 2) and neutralizes the carboxylate group present in the peptide. The purity of the peptide was checked by mass spectrometry and resulted in > 95% peptide purity.

2.7 solid state NMR sample preparation

The lipid composition was DPPC and DPPG with 4:1 ratio. This composition was chosen because: (1) the PC lipids are the major fraction in influenza virus host cell, and (2) the host cell membrane is negatively charged.[8] 50 μmol lipids were dissolved in 2 mL chloroform and methanol solution with a 9:1 volume ratio, and the solvent was removed by dry nitrogen gas flow and vacuum pumping overnight. 3 mL of 10 mM HEPES and 5mM MES buffer at pH 5.0 was used to hydrate the lipid film and followed by 10 times freeze-thaw cycles to make a homogeneous suspension. The lipid-buffer suspension was extruded through a polycarbonate membrane with 100 nm pore size to get large unilamellar vesicles. 1 μmol IFP was dissolved in 10 mL of the HEPES/MES buffer and was added to lipid vesicles drop by drop, then agitate overnight. The peptide-lipid complex was pelleted down by ultra-centrifugation at 100000g for 5 hours. The proteo-liposome complex pellet was lyophilized overnight. Lyophilization helps to reduce sample lost when pack the sample into NMR rotor. The sample was packed in NMR rotor and rehydrate with 10 μL of the HEPES/MES buffer overnight at room temperature.

In PRE experiments, $MnCl_2$ was added in membrane samples as paramagnetic species. The Mn^{2+} was added at the rehydration step followed by 5 freeze-thaw cycles to ensure the accurate amount of Mn^{2+} was added and the Mn^{2+} ions were evenly distributed on both side of the membrane.[9]

2.8 solid state NMR

Rotational Echo Double Resonance (REDOR) solid state NMR and Paramagnetic Relaxation Enhancement (PRE) solid state NMR were used in this study.[10, 11] Spectra were obtained from a 9.4 T Agilent Infinity Plus spectrometer and triple - resonance MAS probe tuned

to ^{1}H, ^{13}C, and ^{2}H frequencies. Figure 2.4A shows the pulse sequence of REDOR and quadrupolar echo (quecho) pulse sequence in PRE experiment.

In REDOR experiment, the rotor was span at 10 kHz and -50 °C, cooling by nitrogen gas flow and the expected sample temperature is ~ -30 °C. The REDOR experiments collects two sets of data: S_0 and S_1. In both experiments: (1) ^{1}H $\pi/2$ pulse; (2) ^{1}H to ^{13}C cross-polarization (CP); (3) ^{1}H decoupling; (4) ^{13}C π pulses at the end of each rotor period; and (5) ^{13}C detection. ^{2}H π pulses are applied in the middle of each rotor period in S_1 experiment but absent in S_0 experiment. The spectra were processed with 100 Hz Gaussian line broadening and referenced to adamantane methylene chemical shift, which is 40.5 ppm. The S_0 and S_1 intensities were calculated with a 3-ppm width of the ^{13}CO peak. The uncertainties were the RMSD's of 6 spectral noise regions with 3-ppm widths. The REDOR dephasing is defined as:

$$\Delta S/S_0 = (S_0 - S_1)/S_0 \quad 2.1$$

The dephasing is a function of dephasing time τ, whose duration is the after ^{1}H to ^{13}C CP and before ^{13}C acquisition. The dephasing is plotted vs τ and the dipolar coupling (D) and internuclear distance (r) between ^{13}C and ^{2}H can be calculated:

$$D(Hz) = \frac{\mu_0 \gamma_{13C} \gamma_{2H} h}{16\pi^2 r^3} = \frac{4642 Å^3 s^{-1}}{r^3} \quad 2.2$$

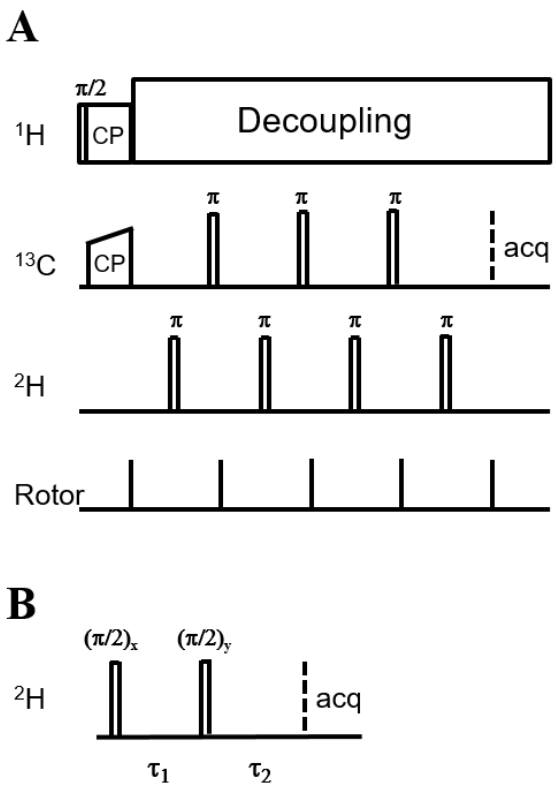

Figure 2.4 (A) ^{13}C - ^{2}H REDOR pulse sequence and (B) quecho pulse sequence.

The PRE experiments were designed for measuring the IFP-bounded membrane lipid acyl chain location by using quadrupolar echo solid-state NMR experiments and the pulsed sequence is shown in Figure 2.4 panel B. ^{2}H spectra were acquired with different τ_1 and τ_2 and a fixed (τ - τ_1) value. Typically, the pulse length was set around 1.5 μs; τ_1, τ_2 were set between 10 and 1000 μs; the recycle delay was set to 1 s. Samples are detected at 25 °C (gel phase lipids) and 50 °C (fluid phase lipids). In data processing, the FID needs to be shifted to make the maximum echo signal at t = 0 and the total relaxation time is denoted as t and t = [τ_1 + τ_2 + time being shifted]. Data processing was done by 1000 Hz exponential line broadening, baseline correction and data shift. ^{2}H T_2 relaxation time is determined by the decay of acquired signal as a function of t. By plotting the signal intensity (I) in FID vs t, the T_2 relaxation time can be calculated by:

$$I(t) = I(0) \times exp\left(\frac{t}{T_2}\right) \quad 2.3$$

The T_2 relaxation rate R_2 can be calculated as:

$$R_2 = \frac{1}{T_2} \quad 2.4$$

Thus, the internuclear distance information between ^2H in deuterated lipid acyl chain and the paramagnetic species Mn^{2+} on membrane surface can be obtained.[12]

REFERENCES

REFERENCES

[1] Chang, C. D., Waki, M., Ahmad, M., Meienhofer, J., Lundell, E. O., and Haug, J. D. (1980) PREPARATION AND PROPERTIES OF N-ALPHA-9-FLUORENYLMETHYLOXYCARBONYLAMINO ACIDS BEARING TERT-BUTYL SIDE-CHAIN PROTECTION, *Int. J. Pept. Protein Res. 15*, 59-66.

[2] Lapatsanis, L., Milias, G., Froussios, K., and Kolovos, M. (1983) SYNTHESIS OF N-2,2,2-(TRICHLOROETHOXYCARBONYL)-L-AMINO ACIDS AND N-(9-FLUORENYLMETHOXYCARBONYL)-L-AMINO ACIDS INVOLVING SUCCINIMIDOXY ANION AS A LEAVING GROUP IN AMINO-ACID PROTECTION, *Synthesis*, 671-673.

[3] Harris, R. B., and Wilson, I. B. (1984) TERT-BUTYL AMINOCARBONATE (TERT-BUTYLOXYCARBONYLOXYAMINE) - A NEW ACYLATING REAGENT FOR AMINES, *Int. J. Pept. Protein Res. 23*, 55-60.

[4] Jia, L. H., Liang, S., Sackett, K., Xie, L., Ghosh, U., and Weliky, D. P. (2015) REDOR solid-state NMR as a probe of the membrane locations of membrane-associated peptides and proteins, *J. Magn. Reson. 253*, 154-165.

[5] Porath, J. (1992) IMMOBILIZED METAL-ION AFFINITY-CHROMATOGRAPHY, *Protein Expr. Purif. 3*, 263-281.

[6] Struck, D. K., Hoekstra, D., and Pagano, R. E. (1981) USE OF RESONANCE ENERGY-TRANSFER TO MONITOR MEMBRANE-FUSION, *Biochemistry 20*, 4093-4099.

[7] Nobusawa, E., Aoyama, T., Kato, H., Suzuki, Y., Tateno, Y., and Nakajima, K. (1991) COMPARISON OF COMPLETE AMINO-ACID-SEQUENCES AND RECEPTOR-BINDING PROPERTIES AMONG 13 SEROTYPES OF HEMAGGLUTININS OF INFLUENZA A-VIRUSES, *Virology 182*, 475-485.

[8] Ghosh, U., Xie, L., Jia, L. H., Liang, S., and Weliky, D. P. (2015) Closed and Semiclosed Interhelical Structures in Membrane vs Closed and Open Structures in Detergent for the Influenza Virus Hemagglutinin Fusion Peptide and Correlation of Hydrophobic Surface Area with Fusion Catalysis, *J. Am. Chem. Soc. 137*, 7548-7551.

[9] Su, Y., Mani, R., and Hong, M. (2008) Asymmetric insertion of membrane proteins in lipid bilayers by solid-state NMR paramagnetic relaxation enhancement: A cell-penetrating peptide example, *J. Am. Chem. Soc. 130*, 8856-8864.

[10] Gullion, T. (1998) Introduction to rotational-echo, double-resonance NMR, *Concepts Magn. Resonance 10*, 277-289.

[11] Clore, G. M., and Iwahara, J. (2009) Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes, *Chem. Rev. 109*, 4108-4139.

[12] Gabrys, C. M., Yang, R., Wasniewski, C. M., Yang, J., Canlas, C. G., Qiang, W., Sun, Y., and Weliky, D. P. (2010) Nuclear magnetic resonance evidence for retention of a lamellar membrane phase with curvature in the presence of large quantities of the HIV fusion peptide, *Biochim. Biophys. Acta-Biomembr. 1798*, 194-201.

Chapter 3

Expression, purification and functional characterization of HIV gp41 ectodomain and transmembrane domain

3.1 Introduction

Human Immunodeficiency Virus (HIV) leads to the Acquired Immunodeficiency Syndrome (AIDS), which is a disease that causes human immune system failure.[1] HIV virus is enveloped by lipid membrane obtained during viral budding from an infected host cell.[2] On HIV envelop, glycoprotein gp120 and gp41 form a "knob" shape complex in a noncovalent manner where three molecules of gp120 forms the "cap" and three molecules of gp41 forms the "stem". The first step of HIV infection cycle is the binding of gp120 onto the host cell CD4 and coreceptors, and this binding can cause conformational change and lead to gp120 displacement.[3-5] gp41 is then exposed and it catalyzes the fusion between viral and host cell membrane. At the end of the fusion, a fusion pore is formed and the viral nucleocapsid is released into the host cell cytoplasm.[6] The RNA is reverse transcribed into DNA catalyzed by reverse transcriptase and then the viral proteins are expressed. Finally new viruses bud from the host cell membrane and start the new infection cycle.[7] This process is detailed discussed in chapter 1 and is also shown in Figure 1.2.[8]

gp41 is the only fusion protein in HIV and it plays a vital role in membrane fusion. Numerous of previous studies illustrate that the gp41 undergoes a huge conformational change before and after fusion. X-ray crystallography and cryo-electron microscopy are used to study the structure of gp41 and big progresses has been made to reveal the structure and its possible fusion mechanism.[6, 9-11]

Crystal and cryo-electron microscopy structure of gp41 at its pre-fusion state included a 50 Å long parallel bundle of three NHR helices in the center, and three CHR 40 Å long helix forming a tripod (Figure 1.5). The monomer structure was NHR-helix/60°-turn/CHR-helix.[9, 11, 12] X-ray crystallography result of gp41 at post-fusion state was a six-helical bundle (SHB) state, in which

three NHR helices formed a parallel trimeric coiled-coil in the center and three CHR helices wrapped antiparallel outside. The monomer structure changed to NHR-helix/180°-turn/CHR-helix (Figure 1.4).[6, 13]

The FP, MPER and TM are the domains that interacting with either host cell or viral membrane and they all play key roles during membrane fusion. However, the structures of these domains were ether truncated or missing in crystallography and cryo-electron microscopy results. The FP alone adopted antiparallel β-sheet structure in physiologically relevant membrane environment.[14] The functional study of FP also suggested its importance in membrane fusion.[15] MPER is the ectodomain region that is closest to the viral membrane with two discrete α helical segments with a central hinge structure.[16, 17] TM is the domain that anchors the gp41 in the viral membrane with a continuous long α-helix with MPER. The structural study results of FP, MPER and TM were shown in Figure 1.7 and 1.8.

The previous functional studies of gp41 also provide information regarding its structure. FP induced vesicle fusion was ~10 times less when Val_2 was mutated to Glu_2, which was consistent with much higher fusion and infection of HIV with wild-type compared to HIV with the V2E mutation. This result supported the catalytic significance of the FP domain in gp41.[18, 19] Solid-phase synthesized FP was monomeric in solution. To obtain FP-dimer or FP-trimer, a tag containing cysteine and lysine at the peptide C-terminal was added. FP-dimer was obtained by cysteine cross-linking between monomers and FP-trimer by synthesizing a peptide bond between the peptide C-terminal carbonyl group in a dimer and a lysine ε-NH_2 group in monomer. Lipid mixing percentage of FP-trimer was 3 folds greater than FP-dimer and 10 folds greater than FP-monomer, which suggested that higher oligomeric constructs induce more vesicle fusion.[20] Vesicle fusion experiments for large gp41 constructs containing N47(L6)C39 (match the SHB crystal

structure in Figure 1.4) and N70(L6)C39 (with appending of FP domain) showed significant lipid mixing for vesicle composition POPC:POPG:Chol (8:2:5) at pH 3.5, but a rapid loss in fusion activity was observed as pH increase, and completely no fusion at pH 5.5. This suggested a strong depend on electrostatic interactions between proteins and vesicles. At low pH, the positively charged protein had electrostatic attraction with negatively charged lipids. As pH increased, the protein charge changed from positive to negative and resulted in repulsion between protein and vesicles, and lead to poor protein binding on vesicles and no fusion.[21] By varying the fraction of negatively charged PG in vesicles, the fusogenicity was also affected. With higher PG fraction, the electrostatic repulsion between vesicles was greater and resulting in lower fusion.[22]

This study is aimed to have an insight of the gp41 structure and function correlation. In this study, I have synthesized the gp41 protein that includes the gp41 construct of residue 512-581+SGGRGG+628-705, and characterized its overall structure, oligomeric state as well as its fusion function. The overall goal is to get a better understanding of the structure and fusion mechanism of gp41 and may help the further vaccine and antiretroviral drug development.

3.2 Result and discussion

3.2.1 Protein solubilization and purification

The constructs and amino acid sequences of full-length HIV gp41 and the four gp41 constructs being studied in this work are shown in Figure 3.1. The constructs had a non-native 6-residue sequence that replaced the SE loop and adjoining terminal regions of the SE N- and C-helices. This modified SE was chosen because of better solubility properties. It also showed that is can retain hyperthermostable helical structure with T_m > 90 °C in HM.[23] The protein expression procedure is discussed in Chapter 2. After expression, the first step in recombinant protein (RP) purification was cell lysis in PBS. The four gp41 constructs are insoluble in PBS, since SDS-PAGE

of the soluble lysate did not show any bands at corresponding molecular weight, as would be expected for membrane proteins. Thus, after centrifugation at 48000 g, 4 °C for 20 min, the RP's were in the pellet. The pellet was resuspended in PBS and centrifugated (2×) to remove non-RP material and this is consistent with visibly decrease in pellet size.

RP expression levels were assessed by static solid-state NMR. The cells were grown in LB broth and then transferred to minimal medium, followed by IPTG induction and adding 1-^{13}C Gly. The procedure was developed from our group's previous work.[24] 50 mL of LB with 0.5 mL of glycerol stock of *E. coli* cells that contained a specific plasmid was incubated at 37 °C and 180 rpm shaking overnight. The cells were harvested by centrifugation at 10,000g at 4°C for 10 minutes and then resuspended into a 250 mL baffled flask containing 50 mL of M9 minimal medium, 100 μL of 1.0 M $MgSO_4$, and 250 μL of 50% glycerol at 180 rpm and 37 °C. 2.0 mM IPTG was added after 1 hr. A dry mixture contained 10 mg of ^{13}CO labeled glycine and 10 mg of each of the other 19 common amino acids were added at the same time. An equivalent mixture was added after one hour of induction. After three hours of expression, the cells were harvested by centrifugation at 10,000 g and 4°C for 10 minutes. The wet cells were resuspended in PBS by tip sonication. The pellet of the lysate after centrifugation was packed into NMR rotor and ^{13}C NMR spectra were obtained (Figure 3.2). The chemical shifts of the peaks in spectra were ~ 175 ppm. By comparing to ^{13}C static NMR of proteins samples, these peaks were assigned to ^{13}CO.[25-27] The similar pellet sizes and the comparable ^{13}CO signal intensities evidenced similar quantities of expressed RP for each construct.

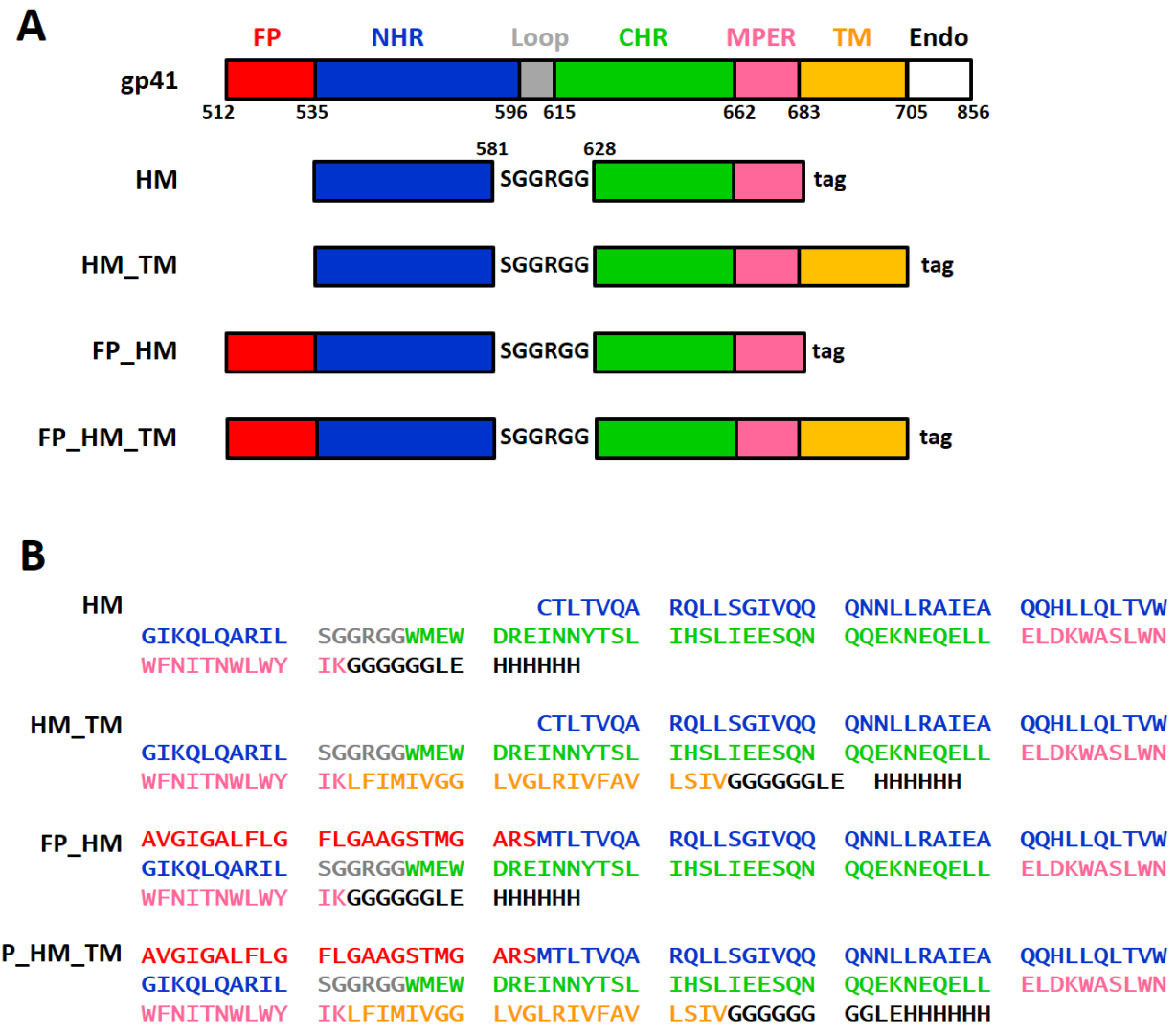

Figure 3.1 (A) Full-length HIV gp41 and the four gp41 constructs. (B) Amino acid sequences with the same color coding as panel A and the non-native C-terminal G_6H_6 or G_8H_6 in black.

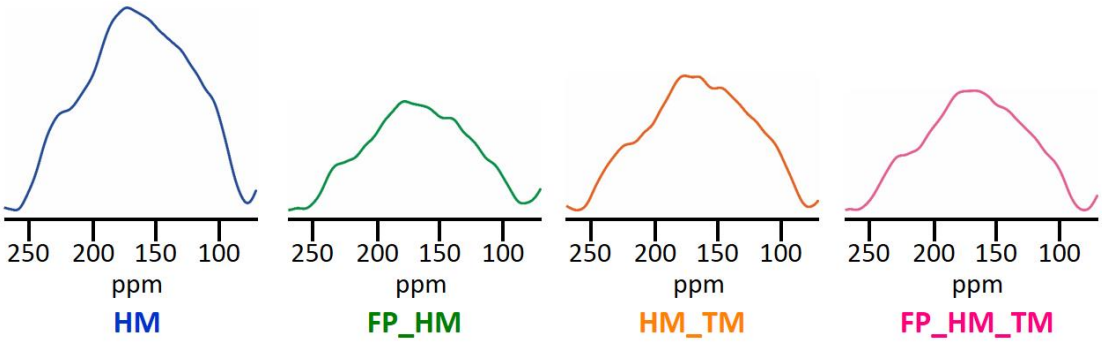

Figure 3.2 Static ^{13}C NMR spectra of lysate pellets labeled with 1-^{13}C Gly. Each spectrum was the sum of 1000 scans.

To find conditions that completely solubilized the RP-rich pellet, four different conditions that contained PBS at pH 7.4 and either: (1) 8M Urea; (2) 6M GuHCl; (3) 8M Urea and 0.8% Sarkosyl; or (4) 8M Urea, 0.5% SDS, and 0.8% Sarkosyl; were tested. The pellets were visually soluble in these buffers after sonication, and the pellet size after centrifugation was similar. Each solution was then subjected to Co^{2+}-affinity chromatography, followed by SDS-PAGE (Figure 3.3 and Figure 3.4). The highest purity and yield were obtained for mixture (4) which was then used for all subsequent protein purification.

All constructs have a non-native C-terminal H_6 affinity tag preceded by either a G_6LE or G_8LE spacer, where the spacer is required for RP binding to the Co^{2+} resin. According to previous crystallography result, the soluble ectodomain (SE, which includes CHR, loop and NHR) by itself likely still adopts helical hairpin structure,[23, 28] so we expect that the spacer provides greater solvent exposure of the H_6 tag.

The longer G_8LE spacer was required for FP_HM_TM binding to the resin. The affinity purification eluent of FP_HM_TM with G_6LE spacer showed no product in SDS-PAGE. The reason might be the H_6 tag buried in protein hydrophobic core, leading to poor affinity column

binding. After adding two other glycine residues to create a G_8LE spacer, a band corresponds to the MW of FP_HM_TM showed on SDS-PAGE, which evidences that a longer spacer increases the exposure of H_6 tag.

To find conditions without urea and Sarkosyl for which solubility was retained for all purified RP's, dialysis was done against a variety of buffers, and soluble RP's were obtained for 10 mM Tris at pH 7.4 and 0.2% SDS. On the other hand, since PC lipids is a significant component of HIV host cell membranes and the DPC has a similar phosphatidylcholine headgroup, the RP's were also dialyzed against buffers with DPC detergent. The RP's precipitated quickly if dialyzed directly against low- or neutral- pH buffers containing 0.25% DPC, which is not consistent with other groups' results.[29, 30] One previous study suggested to refold RP on affinity column.[31] In our observation, all RP's were soluble if the initial exchange into buffer with 0.25% DPC was done with RP bound to Co^{2+}-resin, followed by elution using buffer with DPC and 250 mM imidazole, and then dialysis to remove the imidazole. The final solutions contained 0.25% DPC and either 20 mM sodium acetate buffer at pH 3.2 or 20 mM sodium phosphate buffer at pH 7.4. RP solubility when initial exchange was done with resin-bound RP vs. RP precipitation when initial exchange was done by dialysis of a RP solution suggests that RP aggregates can form quickly, but are not the lowest free-energy state in 0.25% DPC.

Figure 3.4 displays SDS-PAGE of the four RP's after dialysis into 10 mM Tris at pH 7.4 with 0.2% SDS. Bands corresponding to FP_HM, HM_TM, and FP_HM_TM were cut from a gel and subjected to trypsin digestion and mass spectrometry. Peptides were identified from each of the three bands that provided 80, 55, and 75% sequence coverage, respectively (Figure 3.5). In Figure 3.5, the highlighted yellow residues were in a peptide that identified by mass spectrometry.

The highlighted green Met residues had masses consistent with oxidation and the highlighted green Asn and Gln residues had masses consistent with deamination.

Typical purified yields of HM, FP_HM, HM_TM, and FP_HM_TM were respectively ~10, 0.5, 0.5, and 0.3 mg/L bacterial culture. Such significant difference in yields was due to the differences in their binding to the Co^{2+}-resin, since each construct had similar expressed RP quantities. FP_HM_TM required a longer glycine spacer than the other constructs to bind to Co^{2+}-resin. This suggests occlusion of the H_6 tag by FP and TM segments, which is sterically plausible because of SE helical hairpin folding even in the mixtures with denaturant at high concentration.[23]

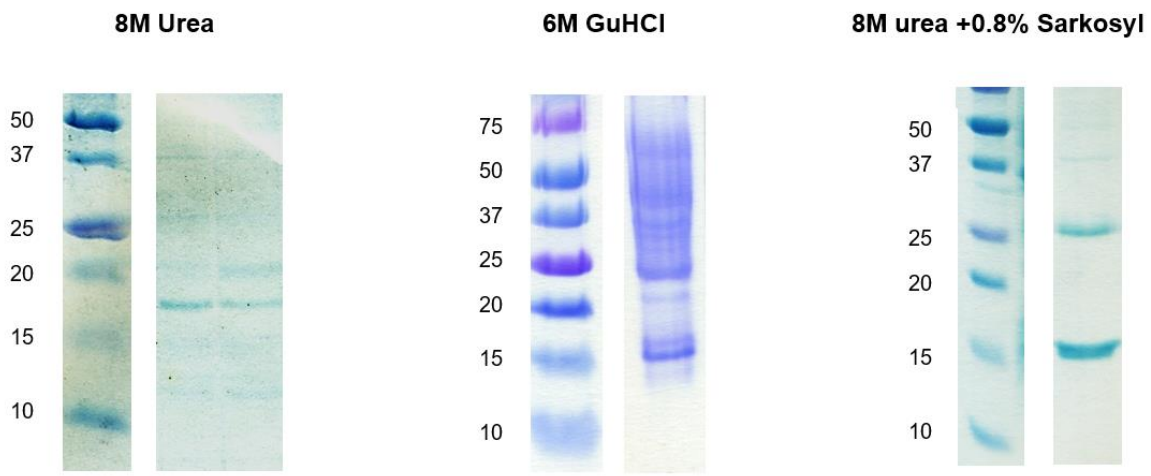

Figure 3.3 SDS-PAGE after Co^{2+}-affinity chromatography of the solubilized pellet enriched in inclusion body protein. The protein is FP_HM and different solubilization conditions for the pellet are noted. Only the MW marker lane and relevant elution lane(s) are displayed.

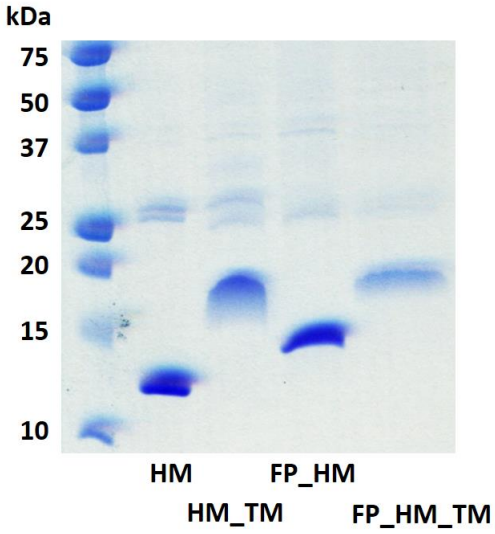

Figure 3.4 SDS-PAGE of the purified HM (MW = 13.7 kDa), HM_TM (MW = 16.7 kDa), FP_HM (MW = 16.5 kDa), and FP_HM_TM (MW = 18.9 kDa) using solubilization condition: PBS at pH 7.4 with 8M Urea, 0.5% SDS, and 0.8% Sarkosyl.

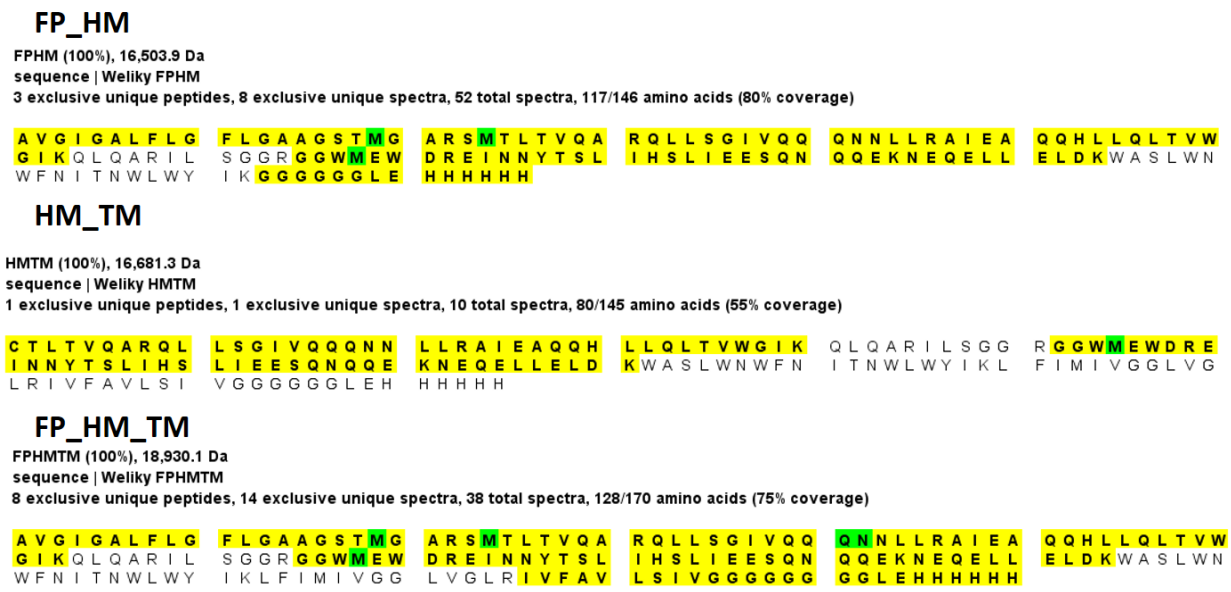

Figure 3.5 Sequence coverage of FP_HM, HM_TM, and FP_HM_TM after trypsin digestion.

3.2.2 Influence of FP and TM on hyperthermal α helical hairpin structure

Figure 3.6 shows the CD spectra at ambient temperature of the four gp41 constructs in: (A) 0.2% SDS at pH 7.4; (C) 0.25% DPC at pH 7.4; and (D) 0.25% DPC at pH 4.0. The protein concentration and all spectra for a single buffer were acquired during the same day to obtain the most quantitative comparison between constructs. Panel B displays θ_{222} vs. temperature plots derived from CD spectra in panel A (0.2% SDS at pH 7.4). Spectra at 25, 60, and 90 °C are presented in Figure 3.7. The temperature-series of spectra for a single construct were all obtained during the same day. The panel A vs. B differences between the ambient-temperature θ_{222} values for the same construct are in part due to use of different CD instruments. There is the same ordering of θ_{222} values with construct in both panels.

There are several trends among the ambient-temperature CD spectra in Figure 3.6 panel A, C and D. All spectra have the characteristic shape of α helical secondary structure with minima near 208 and 222 nm. This helicity is consistent with the helical hairpin structure of gp41 SE. The

| θ | values for a single construct in DPC are similar at low and neutral pH, which supports a pH-independent hairpin structure. The $|\theta_{DPC}|/|\theta_{SDS}| \approx 3/2$, with some variation both among constructs and with wavelength. This ratio suggests higher helicity in DPC vs. SDS.

The HM_TM construct has the largest | θ | values in all three buffer conditions, with $|\theta_{222}| \approx 28{,}000$ (degrees-cm^2/dmole-residue) in DPC that correlates with 85% average helicity. This is equal to the average helicity calculated using a model of 100% helicity of 123/125 of the native residues and 0% helicity for 2 native residues as well as 20 non-native residues, i.e. SGGRGG loop and C-terminal G_6LEH_6 (Table 3.1). Nearly complete native helicity is consistent with a fully-folded protein containing SE helical hairpin and TM helix structural elements. Similar analysis of $|\theta_{222}|$ of HM yields ~76% experimental helicity and ~10 non-helical native residues. These residues are most likely at the N- and C-termini of the hairpin, based on reasoning that the TM domain stabilizes hairpin helical structure.

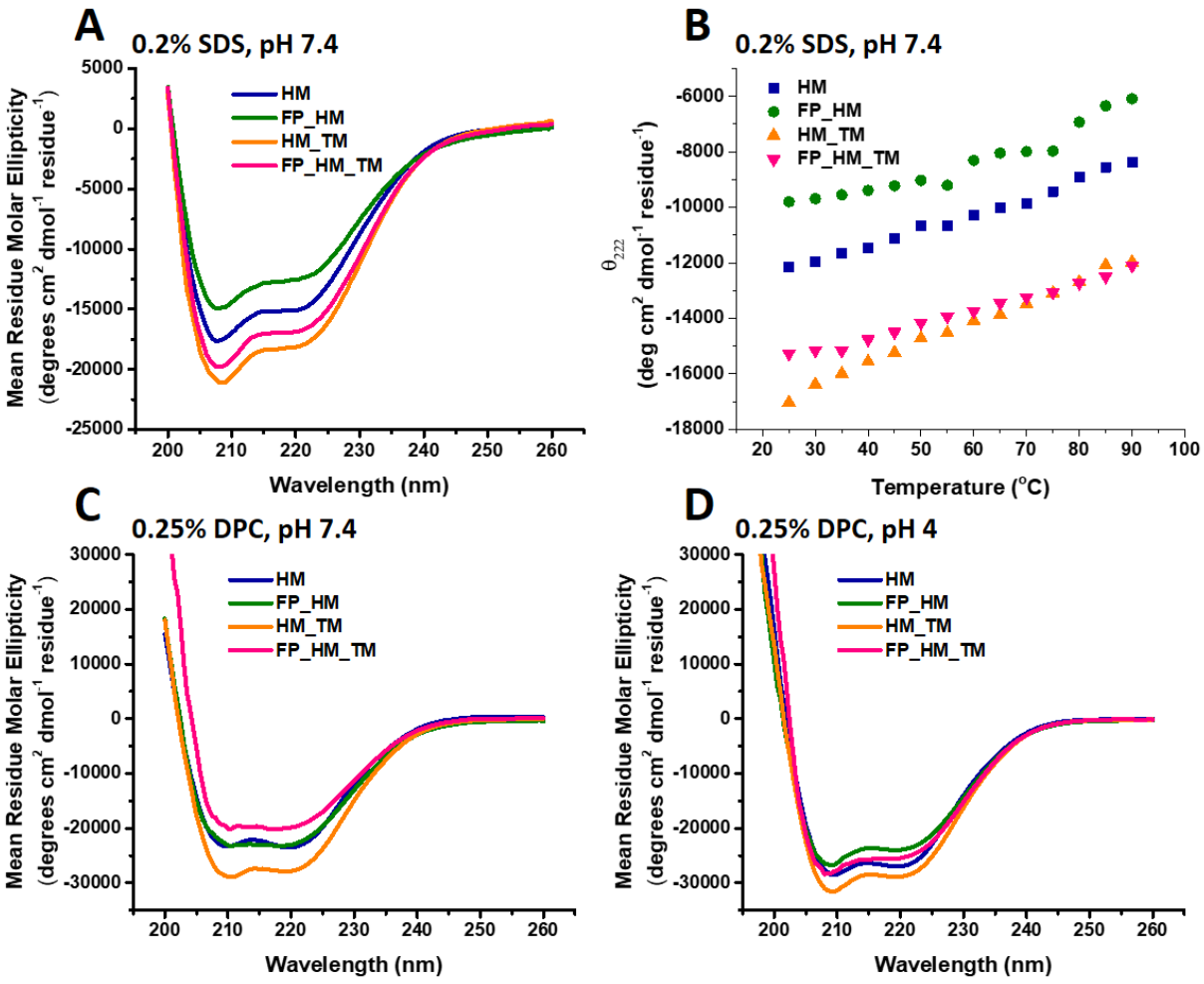

Figure 3.6 Circular dichroism spectra of samples containing ~10 µM protein concentration in different buffer + detergent solutions: (A) 10 mM Tris at pH 7.4 and 0.2% SDS; (C) 20 mM phosphate at pH 7.4 and 0.25% DPC; and (D) 20 mM acetate at pH 4.0 and 0.25% DPC. The spectra in panels A, C, and D spectra were obtained at ambient temperature. (B) the θ_{222} values derived from spectra in 0.2% SDS at pH 7.4 and temperatures between 25 and 90 °C. The panel A vs. B differences between the ambient-temperature θ_{222} values of the same construct may reflect measurement uncertainties in protein concentrations and use of different CD instruments.

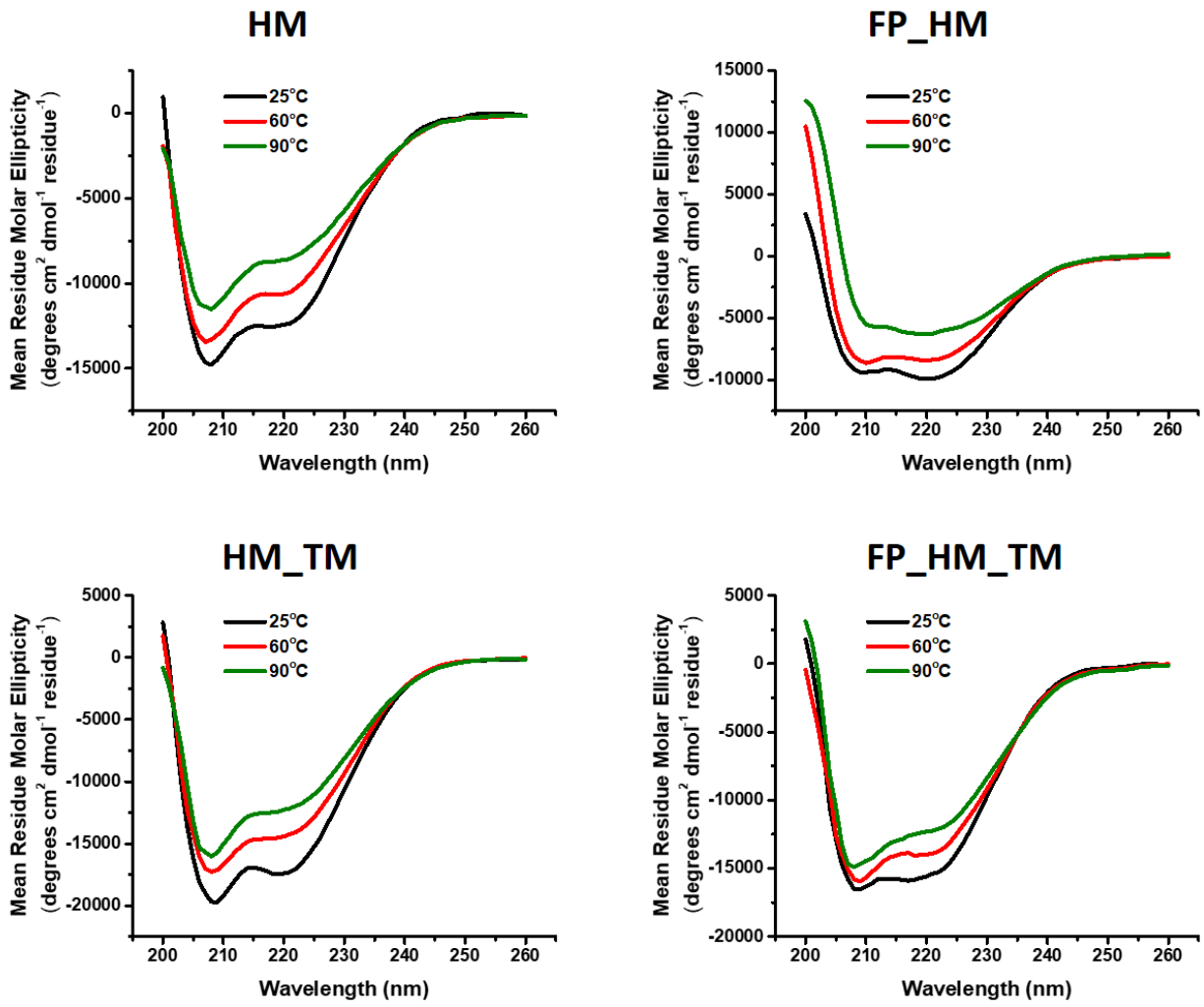

Figure 3.7 Circular dichroism spectra of four gp41 constructs with ~10 µM protein concentration in 10 mM Tris at pH 7.4 and 0.2% SDS at 25, 60 and 90 °C. The panels are noted corresponds to each construct.

Table 3.1 Analysis of CD spectra in 0.25% DPC at pH 4

Construct	Average fractional helicity [a]	Number non-helical residues [b]
HM	0.76	10
HM_TM	0.85	2
FP_HM	0.73	20
FP_HM_TM	0.73	24

[a] Calculated using 100% helicity as $|\theta_{222}|$ = 33000 degrees-cm^2/dmole-residue

[b] Calculated using $N_{non\text{-}helical} = N_{tot} - N_{non\text{-}native} - f_{helix}/N_{tot}$

The CD result showed a general trend that addition of the FP segment decreases average helicity in both SDS and DPC, and this correlates with similar loss of helicity deduced from comparative CD spectra of the shorter "HP" and "FP_HP" constructs.[28] HP and HM have identical N-terminal sequences including SGGRGG non-native loop, but HP lacks the 17 native C-terminal residues of HM, as well as the non-native C-terminal G_6LEH_6. The average helicity of FP_HM and FP_HM_TM in DPC at pH 4 is ~73%, which corresponds to non-helical structure for ~20 (FP_HM) and ~24 (FP_HM_TM) native residues. Most of these residues are likely within the FP, based on ~15-25 non-helical residues when the 23-residue N-terminal FP segment is appended to either HP, HM, or HM_TM. Even the β strand FP is consistent with previous CD result, it is not consistent with earlier NMR chemical shifts.[18, 28, 29] We don't understand the discrepancy between

the CD and NMR data, but note that in membrane, the FP often adopts intermolecular antiparallel β sheet structure.[14]

Figure 3.6 panel B shows only moderate decreases in $|\theta_{222}|$ with temperature in the 25-90 °C range, which supports $T_m \geq 90$ °C for the four constructs. HM and FP_HM exhibit similar $\Delta(|\theta_{222}|) \approx 4000$ degrees-cm^2/dmole-residue over the 25-90 °C range, which is consistent with assignment of the helical structure to the HM region. In some contrast, HM_TM exhibits $\Delta(|\theta_{222}|) \approx 5000$, and FP_HM_TM exhibits $\Delta(|\theta_{222}|) \approx 3000$ degrees-cm^2/dmole-residue, so adding both the FP and TM domains stabilizes the helical hairpin structure of HM. Close FP/TM contact has not been detected by NMR, so increased thermostability may be due to location of the FP and TM segments in the same micelle, so opening of the hairpin requires micelle deformation to maintain detergent contact with both FP and TM.[30, 32, 33] There isn't this penalty if the construct contains only the FP or the TM. Common features of most studies are the very high helicity of the SE and $T_m > 100$ °C, particularly if the construct contains longer N- and C- helix segments with complementarity between their hydrophobic surfaces. These features are exhibited with or without the native loop, for solutions without detergent at low pH, and for solutions with detergent at low and neutral pH.[23, 34, 35] Similar result from our group showing moderate CD changes of the shorter HP and FP_HP constructs, and the combined data with current result support hyperthermostable hairpin SE structure for all constructs.[23] By combining all data, the results support greater thermostability of FP_HM_TM vs. more truncated constructs. The same thermostability correlation was also observed for the influenza virus HA2 construct, whose sequence is non-homologous with gp41, but performs a similar fusion function.[36]

3.2.3 Oligomeric states vary in different detergent

Figure 3.8 displays SEC of the gp41 constructs in: (A) SDS at pH 7.4; (B) DPC at pH 7.4; and (C) DPC at pH 4.0. These detergents were chosen in part because all proteins are thermostable, as assessed by their CD spectra. The running [protein] ≈ 10 μM is comparable to the CD experiments. All buffers also contained 150 mM NaCl to inhibit protein binding on column. The 0.2% SDS is ~5×CMC and the 0.25% DPC is ~8×CMC.[37-39] There is <10% statistical probability of protein occupation of a micelle, so oligomerization is likely due to protein/protein interaction rather than crowding. The protein and detergent form complex and they migrate together on column. The MW determined in SEC is for the complexes of protein + detergent ($MW_{Prot+Det}$).

Table 3.2 shows the MW of SEC peaks and their peak assignment. SEC in 0.2% SDS at pH 7.4 (Figure 3.8 A) mainly has two oligomeric species with $MW_{Prot+Det}$ ≈ 80 – 115 and 190 kDa. A dominant peak is observed for FP_HM at 82 kDa and for FP_HM_TM at 115 kDa. The 78 kDa peak is also observed for HM and the 115 kDa peak for HM_TM, and are accompanied by a peak of approximately equal intensity at 190 kDa for both constructs. Assignments include: 80 and 115 kDa peaks – protein trimers, and 190 kDa peak – protein hexamers. These peak assignments are made based on the known trimer hairpin structure of the SE domain at mM protein concentration, and also considered earlier SEC that supported predominant hexamer species in 6M GuHCl without detergent.[23]

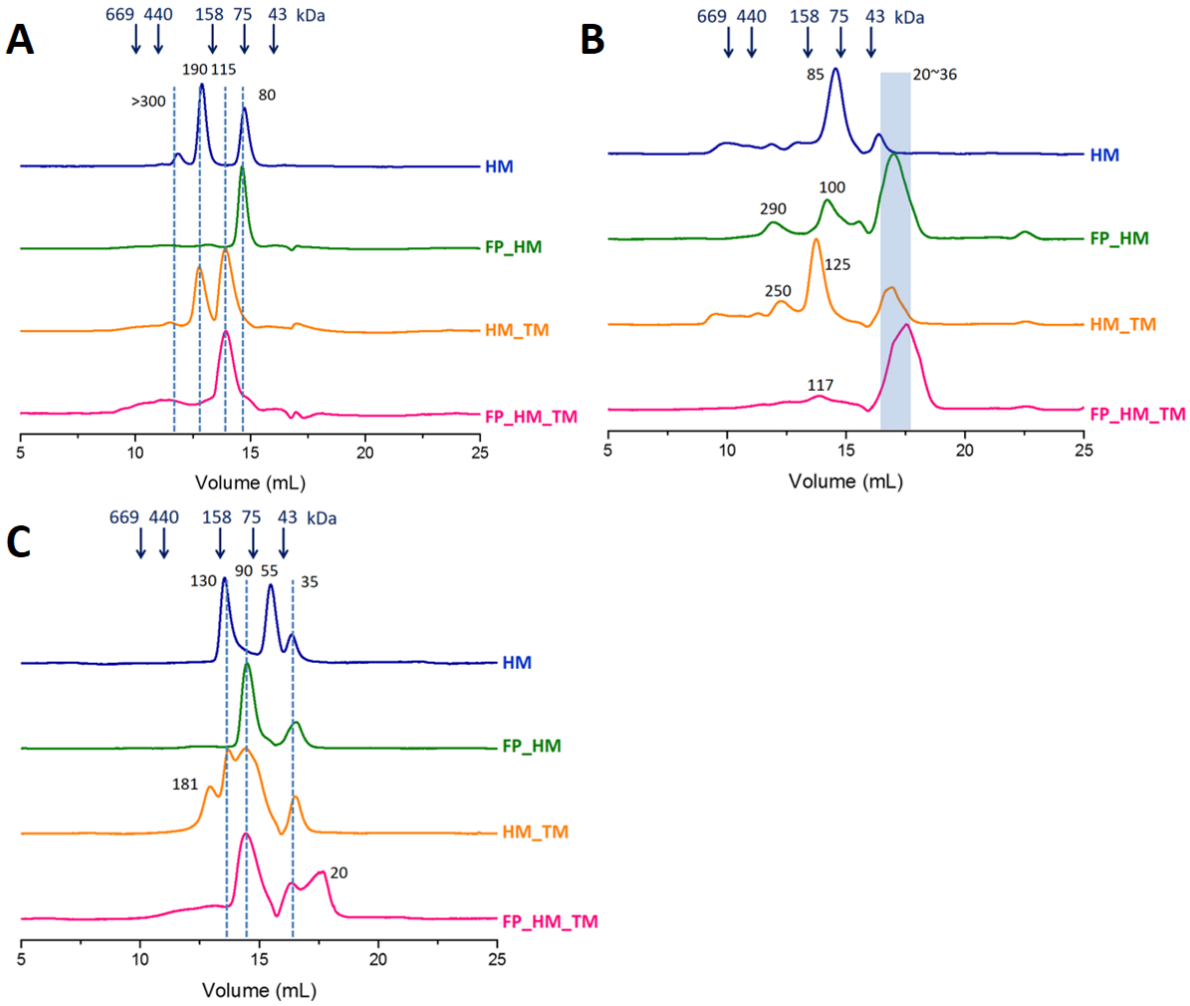

Figure 3.8 SEC of gp41 constructs under the following conditions: (A) 10 mM Tris at pH 7.4, 150 mM NaCl, 0.2% SDS, and ambient temperature; (B) 20 mM phosphate at pH 7.4, 150 mM NaCl, and 0.25% DPC at 4 °C; and (C) 20 mM acetate at pH 4.0, 150 mM NaCl, and 0.25% DPC at 4 °C. SEC was obtained with a Superdex 200-increase column, 1 mg/mL protein loading with ~10-fold dilution in the column, and A280 detection. The arrows in the plots are at the elution volumes of the MW standards, and some of the peaks are identified with dashed lines and with MW's calculated from interpolation between MW standards.

Table 3.2 Analysis of SEC traces [a]

Condition	Construct	MW (kDa)	Oligomer	Protein (kDa)	Detergent (kDa)
SDS at pH 7.4	HM	78	trimer	41	37
	FP_HM [b]	82		49	33
	HM_TM	115		50	65
	FP_HM_TM [b]	115		57	58
	HM	186	hexamer	82	104
	HM_TM	197		100	97
	HM	300	dodecamer	164	136
DPC at pH 7.4	HM	36	monomer	14	22
	FP_HM	31		17	14
	HM_TM	27		17	10
	FP_HM_TM [b]	21		19	3
	HM [b]	85	trimer	41	44
	FP_HM	100		49	51
	HM_TM	125		50	75
DPC at pH 4.0	HM	36	monomer	14	22

Table 3.2 (cont'd)

	FP_HM	33		17	16
	HM_TM	34		17	17
	FP_HM_TM	20		19	1
	FP_HM_TM	36		19	17
	FP_HM [b]	88		49	39
	HM_TM	90	trimer	50	40
	FP_HM_TM	90		57	33
	HM_TM	180	hexamer	100	80

[a] The MW of HM, FP_HM, HM_TM and FP_HM_TM are 13.7, 16.5, 16.7 and 18.9 kDa, respectively

[b] Dominant peak

The assignments are also based on reasonable calculated protein and SDS contributions to the SEC peak masses. The 78 kDa peak is proposed to include HM trimer (41 kDa) and SDS (37 kDa), and the 82 kDa peak includes FP_HM trimer (49 kDa) and SDS (33 kDa). A SDS micelle without protein has a MW of 23 kDa, and additional detergent mass is reasonably needed to solvate protein.[40] The 115 kDa peak includes either HM_TM (50 kDa) or FP_HM_TM (57 kDa) trimer, and SDS (≈ 60 kDa). Relative to the 80 kDa peaks, the additional ~25 kDa SDS mass in the 115

kDa peaks is consistent with solvation of the TM segment by SDS. There is also correlation with the SDS-PAGE in Figure 3.4. The gel shows ~5 kDa increase in monomer MW with inclusion of TM vs. much smaller MW change for inclusion of FP. Hexamer assignment of the 190 kDa peak correlates with HM (82 kDa) or HM_TM (100 kDa), and SDS (≈ 100 kDa), and dodecamer assignment of the 300 kDa peak correlates with HM (164 kDa) and SDS (136 kDa).

SEC traces in DPC (Figure 3.8 B and C) exhibit some peaks with MW's comparable to those in SDS, and assignment is done by similar reasoning. For SEC in 0.25% DPC at pH 7.4, HM, FP_HM and HM_TM peaks at 85 ~ 125 kDa assigned to trimers: HM (41 kDa) + DPC (44 kDa); FP_HM (49 kDa) + DPC (51 kDa); and HM_TM (50 kDa) + DPC (75 kDa). The DPC micelle without protein is 25 kDa.[41] The trimer assignment is supported by previous sedimentation velocity analysis of the related $gp41_{512-711}$ construct in DPC at pH 7.0. This construct includes the full ectodomain and TM. The major species was a trimer (65 kDa) + DPC (62 kDa).[31] For SEC in 0.25% DPC at pH 4.0, the major peaks of FP_HM, HM_TM, and FP_HM_TM are at 90 kDa and assigned as trimer, with ≈ 50 kDa protein and ≈ 40 kDa DPC.

All DPC traces have peaks at lower MW's that are assigned to monomer species. There is often a defined peak at 35 kDa with calculated monomer protein (15 kDa) + DPC (20 kDa) contributions, and sometimes a broader peak at lower MW's. The monomer peaks are dominant for FP_HM and FP_HM_TM at pH 7.4. The SEC monomer assignment is supported by earlier sedimentation velocity data in DPC at pH 7 for several constructs like $gp41_{529-656}$ and $gp41_{546-683}$ that lack the FP and TM domains. Sedimentation coefficients evidenced monomer protein (10 kDa) + DPC (15 kDa).[30]

For DPC, there are high-MW peaks at pH 7.4 that are weak and broad, while there are defined peaks at pH 4.0 including 55 kDa for HM, 130 kDa for HM and HM_TM, and 180 kDa

for HM_TM. The latter peak is assigned to HM_TM hexamers (100 kDa) and DPC (80 kDa), like the corresponding 190 kDa peak of HM_TM in SDS at pH 7.4. I don't have assignments for the 55 kDa and 130 kDa peaks.

Additional SEC traces are provided in Figure 3.9. Panel A displays traces that support reproducibility for replicate samples in 0.2% SDS at pH 7.4. Panel B displays FP_HM and HM_TM traces reproducibility in 0.25% DPC at pH 7.4. The dash lines in these two panels show the alignment of the peaks and support high reproducibility of SEC experiments. Panel C is SEC traces of HM in 0.25% DPC at pH 7.4 with the presents and absent of DTT reducing agent in the injection sample and running buffer. These traces support the lack of cysteine cross-linking in these proteins and little dependence on reducing agent.

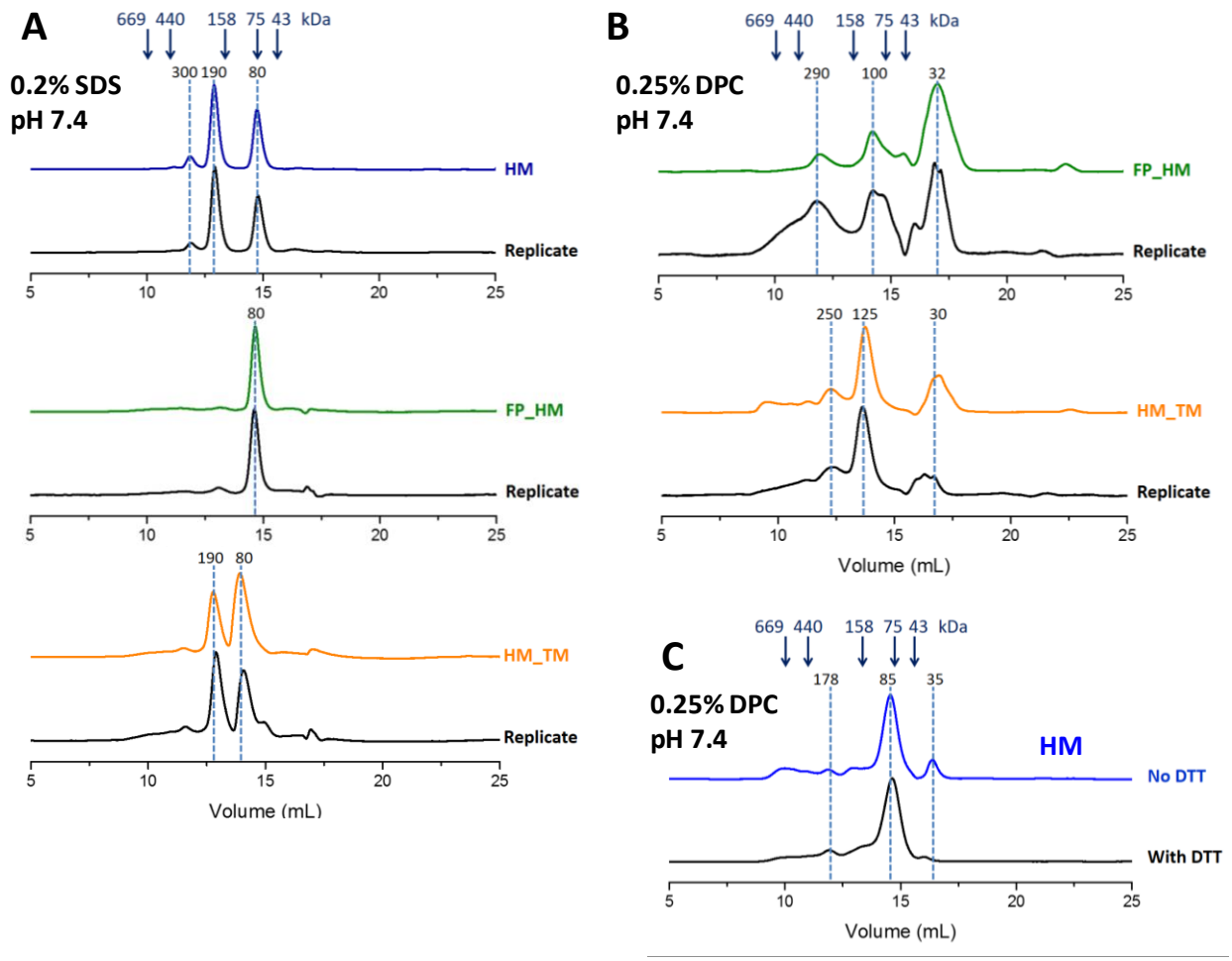

Figure 3.9 SEC traces of replicate samples in: (A) 10 mM Tris buffer at pH 7.4 with 0.2% SDS and 150 mM NaCl; and (B, C) 20 mM phosphate buffer at pH 7.4 with 0.25% DPC and 150 mM NaCl. The (C) replicates differ in the presence vs. absence of 2 mM DTT reducing agent. Constructs and their replicate traces are labeled and dashed lines were used to show peak alignment.

3.2.4 Comparison of gp41 oligomeric states in SEC and AUC

SEC peaks of the present study support the gp41 SE with either monomer, trimer, hexamer, or larger oligomeric states. The current result is compared to SEC data from earlier studies as well as equilibrium and sedimentation velocity analytical ultracentrifugation (AUC) data.[23, 29, 31, 42] In these studies, the constructs all have T_m > 100 °C, which matches the stability of full-length SE. The typical [total protein] ≈ 5 mM for all these SEC and AUC studies, which makes it more reasonable to compare oligomeric interpretations.

Current SEC data supports predominant trimer fraction for all four constructs in SDS at neutral pH, with some hexamer fraction for HM and HM_TM, and no monomer fraction for any construct. There is a significant monomer fraction for all constructs in DPC at both low and neutral pH, as well as trimer and larger oligomer fractions. These data suggest that monomer charges are stabilized by the zwitterionic charges of DPC but not by anionic SDS. The shorter construct HP showed predominant monomer faction at low pH without detergent, whereas large aggregates are predominant at neutral pH.[23] These different oligomeric states correlate with different magnitudes of calculated protein charge, ~ +10e at low pH vs. –1e at neutral pH, with corresponding significant difference in inter-monomer electrostatic repulsion. We don't observe greater monomer fraction in DPC at low vs. neutral pH which supports the significance of electrostatic interactions between the protein and the DPC headgroups with pH-independent zwitterionic charges.

An earlier study used SEC and AUC techniques to study the gp41 512-705 construct in DPC at low pH.[29] The SEC exhibited monomer:trimer ratio ≈ 3:1, and the sedimentation equilibrium AUC data were interpreted to support monomer ↔ trimer $K_a \approx 10^{12}$ M^{-2} which would correspond to a monomer:trimer ratio ≈ 1:4. This contradictory was not explained in this literation. Our SEC of the closely-related FP_HM_TM in DPC at low pH exhibits monomer:trimer ratio ≈

1:1 which is intermediate between the two previous results. For the 538-705 SE+MPER+TM construct, and the 538-665 SE-only construct, sedimentation equilibrium data in DPC at low pH are interpreted to support $K_a \approx 10^{10}$ M^{-2}.[33] Combined consideration of all K_a's suggests that FP and TM segments stabilize the trimer vs. monomer, even though they don't directly interact. There is some similarity with our observation that FP_HM_TM is the most thermostable of the four fully-trimeric constructs in SDS at neutral pH. For 538-665 at low pH without detergent, sedimentation equilibrium K_a's of 5×10^{11} M^{-2} and 7×10^{12} M^{-2} have been independently reported, and likely reflect the magnitude of uncertainty for this technique.[33, 43] The general consensus from the literature is that monomer is stabilized by low pH and also stabilized by DPC detergent. Monomer units may also be stabilized by membrane containing phosphatidylcholine lipids.

Sedimentation velocity AUC has been applied to a construct very similar to FP_HM_TM in DPC.[31] Although K_a's are not derived from sedimentation velocity, the data exhibited a sedimentation coefficient that increased abruptly at pH 5, and was interpreted to support monomer below pH 5 and trimer above pH 5. We did not observe this behavior in SEC and do not understand the discrepancy between methods. The AUC result correlates with the reduction in calculated protein charge from +9e to +3e between pH 4 and 5.

The constructs of the present study were designed based on SE structures to have N- and C-helix segment lengths that yielded T_m > 100 °C, like the full-length protein. There are earlier studies on constructs with similar or shorter helix lengths, with some surprising and sometimes conflicting results about monomer vs. trimer fractions in DPC. For example, the longer 528-579/SGGRGG/628-683 construct at 400 μM concentration and low pH exhibits a tumbling time derived from NMR that is consistent with predominant monomer.[30] This corresponds to $Ka \leq 10^6$ M^{-2} which is >100× smaller than the K_a determined by sedimentation equilibrium for the shorter

538-579/SGGRGG/628-665 construct.[33, 43] At neutral pH, a dramatically larger Ka $\geq 4\times10^{10}$ M^{-2} has been reported for the longer construct based on sedimentation velocity data, whereas much smaller Ka $\leq 2\times10^{8}$ M^{-2} were reported for either 546-579/SGGRGG/628-683 or 528-579/SGGRGG/628-655, which are constructs with respectively a shorter N- or C-helix.[31, 42]

3.2.5 Hairpin protein-induced vesicle fusion at both physiologic and low pH

Figure 3.10 shows the time-courses of vesicle fusion induced by the four different gp41 proteins for vesicles composed of either POPC:POPG:Chol (8:2:5) or POPC:Chol (2:1), at either pH 3.2 or pH 7.4. The data were all acquired on the same day with the same protein stocks. There is ±2% typical variation in long-time fusion extent for replicate assays. POPC and Chol are included to represent some of the physicochemical characteristics of the host cell membrane. For this membrane, PC is a common lipid headgroup, and Chol is ~0.3 mole fraction of total lipid. POPG has a calculated charge of –1 at both pH's, and is included to represent the ~0.15 mole fraction of anionic lipids that are primarily located in the inner leaflet of the host membrane.[44, 45] Neutral POPC:Chol (2:1) may be more representative of the outer leaflet initially encountered by gp41 during HIV/cell fusion.

Vesicle fusion was probed by inter-vesicle lipid mixing detected after addition of 40 μM protein solubilized in 10 mM Tris buffer at pH 7.4, with 0.2% SDS and 150 mM NaCl. This buffer was also used for SEC in Figure 3.8 A, and these SEC data support initial protein trimers for all constructs, and additional hexamer populations for HM and HM_TM. No fusion is observed for buffer without protein, so fusion is due to protein/vesicle rather than SDS/vesicle interaction.

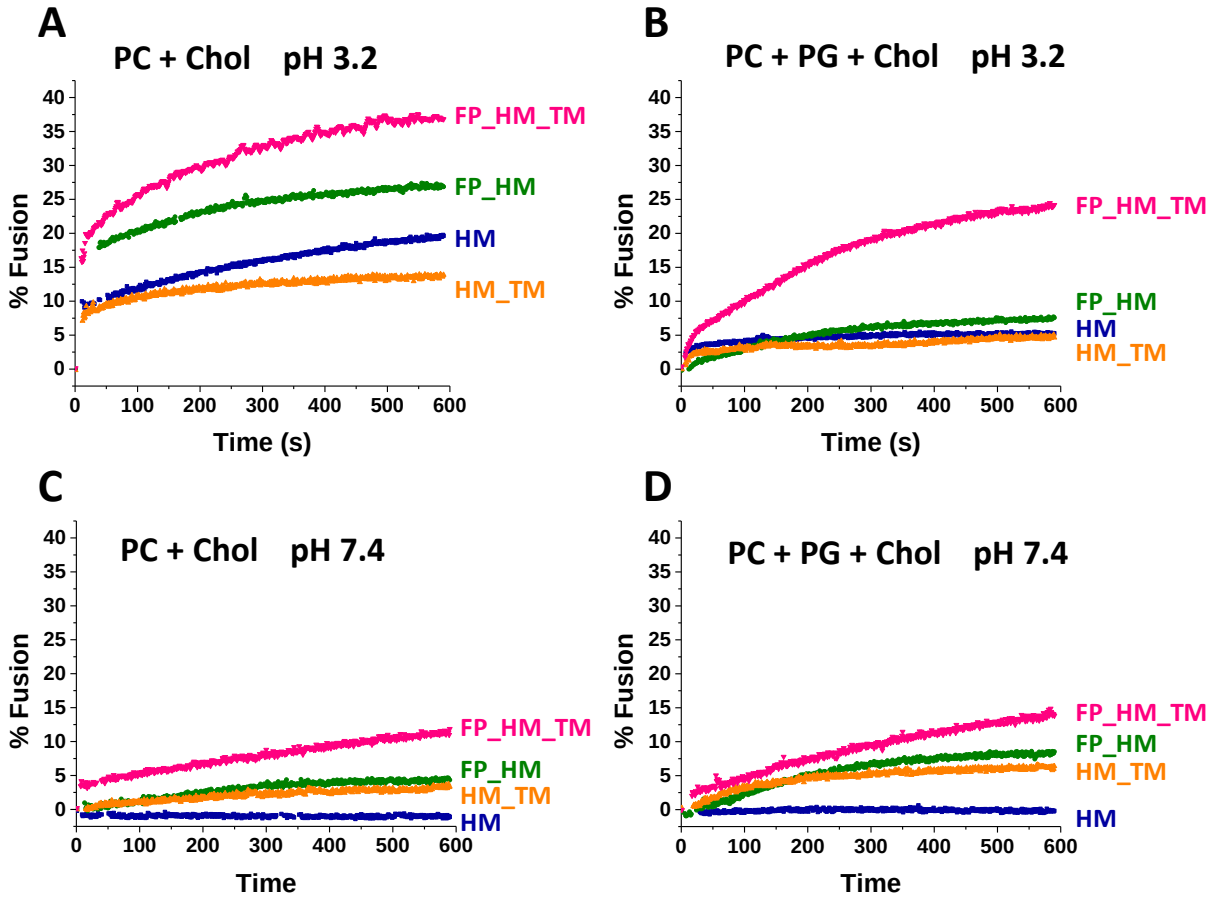

Figure 3.10 Vesicle fusion assays of gp41 proteins. Fusion was initiated by addition of an aliquot of protein stock solution at 0 s, and subsequent fusion was monitored by increased fluorescence associated with inter-vesicle lipid mixing. The stock contained 40 μM protein in buffer at pH 7.4 with 0.2% SDS, and the protein + vesicle mixture contained [protein] = 0.5 μM, [POPC+POPG] = 150 μM, and vesicle molar compositions and pH's: (A) POPC:Chol = 2:1 at pH 3.2; (B) POPC:POPG:Chol = 8:2:5 at pH 3.2; (C) POPC:Chol = 2:1 at pH 7.4; and (D) POPC:POPG:Chol = 8:2:5 at pH 7.4. Fusion wasn't observed for any vesicle composition after addition of an aliquot of buffer + 0.2% SDS without protein. The assay dead time is ~5 s.

Figure 3.10 panel C and D showed the gp41-induced vesicle fusion at pH 7.4, and the fusion of neutral vesicles at neutral pH is the condition similar HIV/cell fusion. There was a study from Chang's group showing the $gp41_{512\text{-}665}$, which is similar to FP-HM, had a fusion activity that was comparable to our data with neutral vesicles and at neutral pH.[46] However, there were two unusual result in that work: (1) the fusion activity of $gp41_{512\text{-}665}$ did not increase as the protein concentration increased; (2) the solubility was very high as ~1 mM in water.[47] From our groups previous experiments, the fusion activity always increases as the concentration of protein increases, and the solubility of gp41 constructs in water is < 1μM. For most of the previous studies, fusion required negatively-charged vesicles.[21, 22] These previous results were interpreted to support attractive electrostatics as a requirement for protein/vesicle binding. For the present study, HM didn't induce fusion at neutral pH, which is consistent with these previous results, but larger constructs induced appreciable fusion.[23] FP_HM_TM exhibited the greatest fusion extent, while FP_HM and HM_TM had smaller extents that were similar to one another. These data support a positive correlation between protein hydrophobicity and vesicle fusion. Fusion was moderately-higher for anionic vs. neutral vesicles, which evidences that electrostatic repulsion between anionic vesicles does not affect fusion, and is also consistent with little bulk electrostatic interaction between a vesicle and the nearly neutral protein, whose calculated charge is ~ –1 at pH 7.4.

The pH 3.2 data provides a comparison with earlier studies on shorter gp41 constructs at low pH. One common feature of all four conditions of Figure 3.10 is highest fusion extent for FP_HM_TM, which supports the hydrophobicity/fusion correlation. Fusion extent at low pH was generally comparable or greater than at neutral pH. This correlates with different magnitudes of protein charge at low vs. neutral pH, +12 vs. –1. Interestingly, the greater extent at low pH is likely

not due to direct protein/vesicle attraction, because extent was also greater for neutral vs. anionic vesicles. This result contrasts with the reverse trend at neutral pH.

3.2.6 FP_HM_TM structural model

Figure 3.11 displays a medium-resolution structural model for FP_HM_TM which incorporates the Table 3.1 analysis of the circular dichroism spectra as well as previous results.[6, 13, 48] This model likely reflects the final gp41 state during fusion, based on T_m > 90 °C (Figure 3.6 B). A monomer is displayed for clarity but the model should also be valid for a trimer bundle or a hexamer (dimer of trimer bundles). The HM region is primarily hairpin structure that contains helices over residues 536-581 and 628-675. This structure is supported by very high helicity of HM and HM_TM, and by a previous crystal structure.[6, 13] The C-terminal MPER and TM regions are also highly helical, based on the HM_TM CD spectrum, and on structures in detergent of peptides corresponding to MPER and/or TM sequences.[48] These structures are typically continuous helices, but there are likely breaks in helical structure for the larger protein because of the topological constraints of membrane interface and traversal locations for the MPER and TM domains, respectively.

The FP region is represented as extended and β strand structure, based both on reduced average helicity for constructs that include the FP, and on earlier NMR and infrared data that evidence FP antiparallel β sheet structure in membrane.[14, 49-51] Such structure is reasonable for a hexamer for which the strands from the two trimers are interleaved. NMR data also support a distribution of N-terminal antiparallel β sheet registries, with significant populations of registries like 512→527/527→512 for adjacent strands.[14] FP insertion in a single leaflet is evidenced by

NMR contacts between multiple FP residues and lipid tails.[49] There isn't close contact between the FP and the MPER or between the FP and the TM, as evidenced by previous NMR studies.[30, 32]

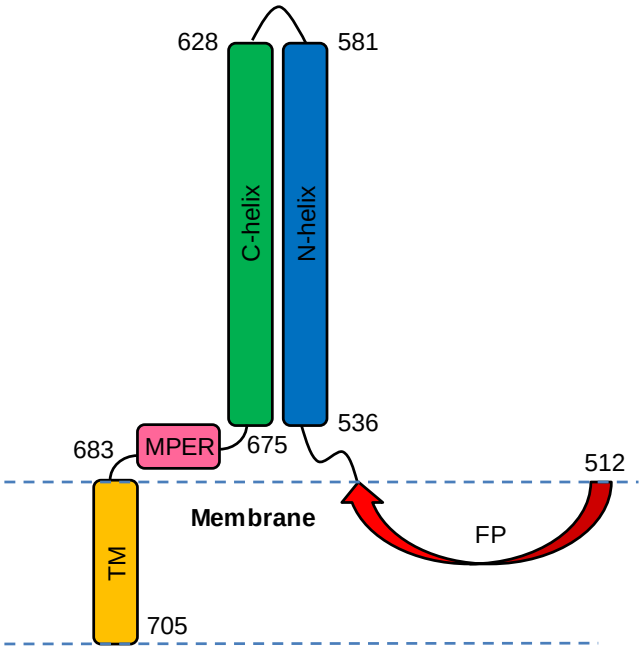

Figure 3.11 Structural model of FP_HM_TM based on circular dichroism spectra of the four constructs as well as other data.[6, 14, 30, 48] A monomer is shown for clarity but the model should be valid for trimers and hexamers. Approximate residue numbers are displayed.

3.2.7 Relationship between vesicle fusion and HIV/cell fusion

The vesicle fusion data of the present study provides information about the relative membrane perturbations by different constructs in defined oligomeric states. This may describe potential contributions of these states to viral gp41-induced fusion. In the pre-fusion complex of gp41 with gp120, electron densities are interpreted to support an interior bundle of trimeric N-helices, and C-helices separated from the bundle.[9] The trimer of hairpins structure has been observed for gp41 without gp120, and large ectodomain constructs exhibit a hyperthermostable monomer at ~5 μM concentration with helices indistinguishable from the monomer units of the trimer. This underlies the hairpin monomer and trimer structures in Figure 3.11 and 3.12.

For the present study, stock protein was principally hyperthermostable trimers and the protein:(PC+PG+Chol) ratio was ~1:450. Each vesicle has a diameter d of 100 nm, and the membrane thickness b is ~ 5 nm. Thus, the surface area including the vesicle outer and inner leaflet can be calculated as:

$$A = 4\pi\left[\left(\frac{d}{2}\right)^2 + \left(\frac{d}{2} - b\right)^2\right] = 56834\ nm^2 \quad 3.1$$

The surface area of PC or PG head group is ~ 0.7 nm^2 and Chol is ~ 0.4 nm^2.[52] The PC+PG:Chol ratio is 2:1. The average surface area a of each lipid molecule is:

$$a = \frac{2}{3}\times 0.7\ nm^2 + \frac{1}{3}\times 0.4\ nm^2 = 0.6\ nm^2 \quad 3.2$$

The number of lipid molecules in each vesicle is:

$$N_{tot} = \frac{A}{a} \approx 95{,}000 \quad 3.3$$

This implies ~210 protein molecules or ~70 protein trimers per vesicle when there is quantitative protein binding, and smaller copy number with reduced binding. This is comparable to the ~15

trimers/virion, with significant microscopy and functional evidence that trimers are spatially clustered during fusion.[8, 53]

Earlier studies of vesicle fusion induced by constructs like HM, and FP_HM often showed strong dependences on pH and membrane charge that were interpreted as supporting a large contribution to fusion from protein/vesicle electrostatic attraction. There was similarly much greater leakage of anionic vesicles at low vs. neutral pH.[21-23] For the present study, electrostatic effects were much less pronounced, as evidenced by much smaller dependence of fusion extents on low vs. neutral pH and anionic vs. neutral vesicles (Figure 3.10). Larger constructs induced significant fusion at neutral pH of both neutral and anionic vesicles, which reflects expected physiologic conditions of HIV/host cell fusion. Fusion under physiologic conditions for the present but not earlier studies may be due to inclusion of more hydrophobic segments in the present study, and also stock solutions with predominant trimer in the present study vs. monomer in previous studies.[54] Hydrophobic perturbation of the membrane is increased by both effects, and is magnified in the fusion rate via the Arrhenius Law, assuming that perturbations reduce activation energy by making membranes more like the fusion transition state. The contribution of the hydrophobic effect is evidenced by highest fusion for FP_HM_TM. In addition, the HIV TM sequence is fairly conserved,[55] including the central R696 snorkeling towards membrane surface, which likely contributes membrane perturbation.[56]

Apposition of HIV and host cell membranes is likely aided by conversion from orthogonal N-helix/C-helix geometry in the pre-fusion state to antiparallel geometry in the hairpin state. It is also known that peptides corresponding to N- or C-helix regions inhibit fusion up to the final pore expansion step, which is after inter-membrane lipid mixing and pore formation (Figure 3.12 B).[13, 53, 57] The present study shows that the trimer of hairpins is effective at inducing vesicle fusion

under physiologic conditions, particularly when both the TM and FP segments are included in the construct. The fusion efficiency of the trimer is based on increased local concentration of TM and FP, and correlates with much higher vesicle fusion induced by FP's that are cross-linked at their C-termini, with topology similar to that in the hairpin trimer.[20] The inhibitory peptides would likely not bind to the trimer of hairpins and we propose that they instead bind to monomer hairpins formed after dissociation (Figure 3.12). Such dissociation has been known for twenty years, with additional recent data from our and other groups, and could plausibly occur during the ~1-3 minute HIV/cell fusion time.[8, 23, 30, 32] The bound peptides prevent re-association of the fusion-efficient trimer. For constructs like FP_HM_TM in DPC detergent at low pH, the 3 monomer ↔ trimer $K_a \approx 10^{12}$ M^{-2}, so that $mass_{monomer} \approx mass_{trimer}$ when [total protein] ≈ 1 μM. The membrane K_a hasn't been measured, but there are comparable statistical-average inter-protein molecular separations of ~100 nm for [bulk protein] ≈ 1 μM, and ~50 nm for ~15 protein trimers in a 100 nm diameter virion. Previous work from our and other groups supports the monomer as the hyperthermostable helical folding unit and it is therefore reasonable that the hairpin is the lowest-free-energy structure.[23, 30] Asynchronous folding of monomer protein into hairpin brings the two membranes into apposition and is likely topologically easier than concerted folding of a trimer into a six-helix bundle. This may therefore be evolutionary advantage of retaining some hairpin monomer stability relative to trimer.

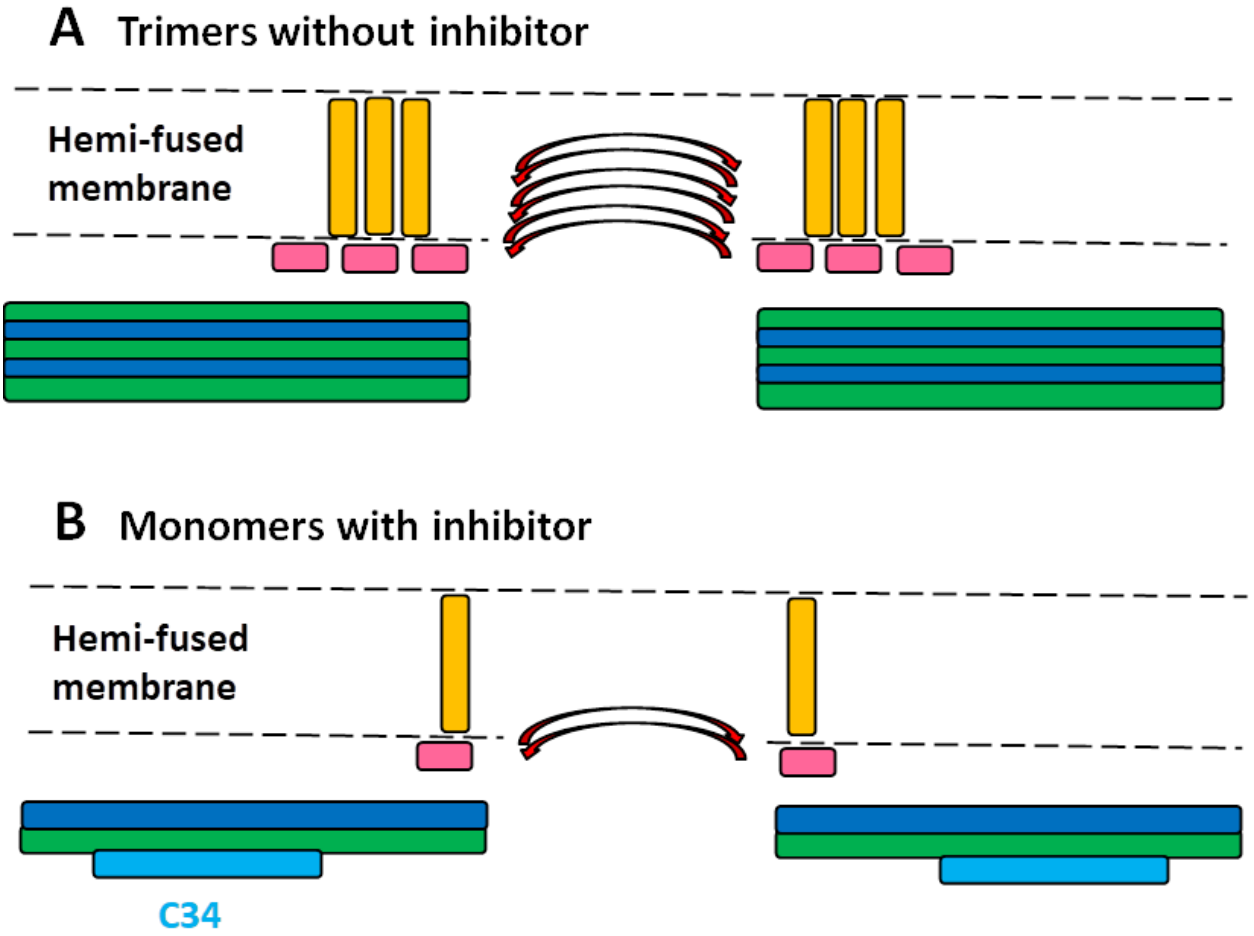

Figure 3.12 Schematic illustrating (A) trimer and (B) monomer respectively favored in the absence and presence of peptide inhibitor. Panel B displays "C34" inhibitor which contains C-helix residues 628-661. The sequence color coding matches Figure 3.1 and 3.11, and loops between structured regions are not displayed for clarity. The FP's from different trimers or monomers adopt antiparallel β sheet structure. Fusion is enhanced in panel A vs. B because of greater clustering of membrane-perturbing protein regions in the trimer vs. monomer. This enhancement exists for the displayed hemifusion state as well as membrane states that precede hemifusion.

3.3 Conclusion

The construct of HIV gp41 membrane fusion protein including the whole ectodomain and transmembrane domain, and shorter constructs have been expressed, purified and stabilized in physiology buffers. The constructs adopt α helical SE and TM, and non-helical FP in SDS and DPC detergents, and they are all hyperthermostable with T_m > 90 °C. The oligomeric states of these proteins vary in different detergent buffer: predominant trimer for all constructs and some hexamer fraction for HM and HM_TM protein in SDS at pH 7.4; and mixtures of monomer, trimer, and higher-order oligomer protein in DPC at pH 4.0 and 7.4. Substantial protein-induced vesicle fusion was observed, including fusion of neutral vesicles at neutral pH, which are the conditions similar HIV/cell fusion. Vesicle fusion by a gp41 ectodomain construct has rarely been observed under these conditions, and is aided by inclusion of both the FP and TM, and by protein which is predominantly trimer rather than monomer. Current data was integrated with existing data, and a structural model and some new interpretations of these data were proposed: (1) gp41 has a monomer ↔ trimer thermodynamic equilibrium; (2) monomer hairpins are formed from trimer dissociation. Asynchronous folding of monomer gp41 into hairpin is likely topologically easier than concerted folding of a trimer into a six-helix bundle; (3) trimer has higher fusion efficiency since higher local concentration of TM and FP and greater membrane perturbation.

REFERENCES

REFERENCES

[1] Weiss, R. A. (1993) HOW DOES HIV CAUSE AIDS, *Science 260*, 1273-1279.

[2] Levy, J. A. (1993) PATHOGENESIS OF HUMAN-IMMUNODEFICIENCY-VIRUS INFECTION, *Microbiol. Rev. 57*, 183-289.

[3] Dalgleish, A. G., Beverley, P. C. L., Clapham, P. R., Crawford, D. H., Greaves, M. F., and Weiss, R. A. (1984) THE CD4 (T4) ANTIGEN IS AN ESSENTIAL COMPONENT OF THE RECEPTOR FOR THE AIDS RETROVIRUS, *Nature 312*, 763-767.

[4] Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J., and Hendrickson, W. A. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody, *Nature 393*, 648-659.

[5] Broder, C. C., and Dimitrov, D. S. (1996) HIV and the 7-transmembrane domain receptors, *Pathobiology 64*, 171-179.

[6] Chan, D. C., Fass, D., Berger, J. M., and Kim, P. S. (1997) Core structure of gp41 from the HIV envelope glycoprotein, *Cell 89*, 263-273.

[7] Frankel, A. D., and Young, J. A. T. (1998) HIV-1: Fifteen proteins and an RNA, *Annu. Rev. Biochem. 67*, 1-25.

[8] Grewe, C., Beck, A., and Gelderblom, H. R. (1990) HIV - EARLY VIRUS-CELL INTERACTIONS, *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 3*, 965-974.

[9] Julien, J. P., Cupo, A., Sok, D., Stanfield, R. L., Lyumkis, D., Deller, M. C., Klasse, P. J., Burton, D. R., Sanders, R. W., Moore, J. P., Ward, A. B., and Wilson, I. A. (2013) Crystal Structure of a Soluble Cleaved HIV-1 Envelope Trimer, *Science 342*, 1477-1483.

[10] Weissenhorn, W., Dessen, A., Harrison, S. C., Skehel, J. J., and Wiley, D. C. (1997) Atomic structure of the ectodomain from HIV-1 gp41, *Nature 387*, 426-430.

[11] Lyumkis, D., Julien, J. P., de Val, N., Cupo, A., Potter, C. S., Klasse, P. J., Burton, D. R., Sanders, R. W., Moore, J. P., Carragher, B., Wilson, I. A., and Ward, A. B. (2013) Cryo-EM Structure of a Fully Glycosylated Soluble Cleaved HIV-1 Envelope Trimer, *Science 342*, 1484-1490.

[12] Tran, E. E. H., Borgnia, M. J., Kuybeda, O., Schauder, D. M., Bartesaghi, A., Frank, G. A., Sapiro, G., Milne, J. L. S., and Subramaniam, S. (2012) Structural Mechanism of Trimeric HIV-1 Envelope Glycoprotein Activation, *PLoS Pathog. 8*, 18.

[13] Lu, M., Ji, H., and Shen, S. (1999) Subdomain folding and biological activity of the core structure from human immunodeficiency virus type 1 gp41: Implications for viral membrane fusion, *J. Virol. 73*, 4433-4438.

[14] Schmick, S. D., and Weliky, D. P. (2010) Major Antiparallel and Minor Parallel beta Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide, *Biochemistry 49*, 10623-10635.

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

[16] Shi, W. X., Bohon, J., Han, D. P., Habte, H., Qin, Y. L., Cho, M. W., and Chance, M. R. (2010) Structural Characterization of HIV gp41 with the Membrane-proximal External Region, *J. Biol. Chem. 285*, 24290-24298.

[17] Sun, Z. Y. J., Oh, K. J., Kim, M. Y., Yu, J., Brusic, V., Song, L. K., Qiao, Z. S., Wang, J. H., Wagner, G., and Reinherz, E. L. (2008) HIV-1 broadly neutralizing antibody extracts its epitope from a kinked gp41 ectodomain region on the viral membrane, *Immunity 28*, 52-63.

[18] Gabrys, C. M., Qiang, W., Sun, Y., Xie, L., Schmick, S. D., and Weliky, D. P. (2013) Solid-State Nuclear Magnetic Resonance Measurements of HIV Fusion Peptide (CO)-C-13 to Lipid P-31 Proximities Support Similar Partially Inserted Membrane Locations of the alpha Helical and beta Sheet Peptide Structures, *J. Phys. Chem. A 117*, 9848-9859.

[19] Freed, E. O., Delwart, E. L., Buchschacher, G. L., and Panganiban, A. T. (1992) A MUTATION IN THE HUMAN-IMMUNODEFICIENCY-VIRUS TYPE-1 TRANSMEMBRANE GLYCOPROTEIN-GP41 DOMINANTLY INTERFERES WITH FUSION AND INFECTIVITY, *Proc. Natl. Acad. Sci. U. S. A. 89*, 70-74.

[20] Qiang, W., and Weliky, D. P. (2009) HIV Fusion Peptide and Its Cross-Linked Oligomers: Efficient Syntheses, Significance of the Trimer in Fusion Activity, Correlation of beta Strand Conformation with Membrane Cholesterol, and Proximity to Lipid Headgroups, *Biochemistry 48*, 289-301.

[21] Sackett, K., TerBush, A., and Weliky, D. P. (2011) HIV gp41 six-helix bundle constructs induce rapid vesicle fusion at pH 3.5 and little fusion at pH 7.0: understanding pH dependence of protein aggregation, membrane binding, and electrostatics, and implications for HIV-host cell fusion, *Eur. Biophys. J. Biophys. Lett. 40*, 489-502.

[22] Ratnayake, P. U., Sackett, K., Nethercott, M. J., and Weliky, D. P. (2015) pH-dependent vesicle fusion induced by the ectodomain of the human immunodeficiency virus membrane fusion protein gp41: Two kinetically distinct processes and fully-membrane-associated gp41 with predominant beta sheet fusion peptide conformation, *Biochim. Biophys. Acta-Biomembr. 1848*, 289-298.

[23] Banerjee, K., and Weliky, D. P. (2014) Folded Monomers and Hexamers of the Ectodomain of the HIV gp41 Membrane Fusion Protein: Potential Roles in Fusion and Synergy Between the Fusion Peptide, Hairpin, and Membrane-Proximal External Region, *Biochemistry 53*, 7184-7198.

[24] Vogel, E. P., and Weliky, D. P. (2013) Quantitation of Recombinant Protein in Whole Cells and Cell Extracts via Solid-State NMR Spectroscopy, *Biochemistry 52*, 4285-4287.

[25] Zhang, H. Y., Neal, S., and Wishart, D. S. (2003) RefDB: A database of uniformly referenced protein chemical shifts, *J. Biomol. NMR 25*, 173-195.

[26] Naito, A., Nagao, T., Norisada, K., Mizuno, T., Tuzi, S., and Saito, H. (2000) Conformation and dynamics of melittin bound to magnetically oriented lipid bilayers by solid-state P-31 and C-13 NMR spectroscopy, *Biophys. J. 78*, 2405-2417.

[27] Costa, P. R., Kocisko, D. A., Sun, B. Q., Lansbury, P. T., and Griffin, R. G. (1997) Determination of peptide amide configuration in a model amyloid fibril by solid-state NMR, *J. Am. Chem. Soc. 119*, 10487-10493.

[28] Sackett, K., Nethercott, M. J., Shai, Y., and Weliky, D. P. (2009) Hairpin Folding of HIV gp41 Abrogates Lipid Mixing Function at Physiologic pH and Inhibits Lipid Mixing by Exposed gp41 Constructs, *Biochemistry 48*, 2714-2722.

[29] Lakomek, N. A., Kaufman, J. D., Stahl, S. J., Louis, J. M., Grishaev, A., Wingfield, P. T., and Bax, A. (2013) Internal Dynamics of the Homotrimeric HIV-1 Viral Coat Protein gp41 on Multiple Time Scales, *Angew. Chem.-Int. Edit. 52*, 3911-3915.

[30] Roche, J., Louis, J. M., Aniana, A., Ghirlando, R., and Bax, A. (2015) Complete dissociation of the HIV-1 gp41 ectodomain and membrane proximal regions upon phospholipid binding, *J. Biomol. NMR 61*, 235-248.

[31] Dai, Z., Tao, Y. S., Liu, N. N., Brenowitz, M. D., Girvin, M. E., and Lai, J. R. (2015) Conditional Trimerization and Lytic Activity of HIV-1 gp41 Variants Containing the Membrane-Associated Segments, *Biochemistry 54*, 1589-1599.

[32] Roche, J., Louis, J. M., Grishaev, A., Ying, J. F., and Bax, A. (2014) Dissociation of the trimeric gp41 ectodomain at the lipid-water interface suggests an active role in HIV-1 Env-mediated membrane fusion, *Proc. Natl. Acad. Sci. U. S. A. 111*, 3425-3430.

[33] Lakomek, N. A., Kaufman, J. D., Stah, S. J., and Wingfield, P. T. (2014) HIV-1 Envelope Protein gp41: An NMR Study of Dodecyl Phosphocholine Embedded gp41 Reveals a Dynamic Prefusion Intermediate Conformation, *Structure 22*, 1311-1321.

[34] Weissenhorn, W., Wharton, S. A., Calder, L. J., Earl, P. L., Moss, B., Aliprandis, E., Skehel, J. J., and Wiley, D. C. (1996) The ectodomain of HIV-1 env subunit gp41 forms a soluble, alpha-helical, rod-like oligomer in the absence of gp120 and the N-terminal fusion peptide, *Embo J. 15*, 1507-1514.

[35] Buzon, V., Natrajan, G., Schibli, D., Campelo, F., Kozlov, M. M., and Weissenhorn, W. (2010) Crystal Structure of HIV-1 gp41 Including Both Fusion Peptide and Membrane Proximal External Regions, *PLoS Pathog. 6*, 7.

[36] Ratnayake, P. U., Ekanayaka, E. A. P., Komanduru, S. S., and Weliky, D. P. (2016) Full-length trimeric influenza virus hemagglutinin II membrane fusion protein and shorter constructs lacking the fusion peptide or transmembrane domain: Hyperthermostability of the full-length protein and the soluble ectodomain and fusion peptide make significant contributions to fusion of membrane vesicles, *Protein Expr. Purif. 117*, 6-16.

[37] Corrin, M. L., and Harkins, W. D. (1947) THE EFFECT OF SALTS ON THE CRITICAL CONCENTRATION FOR THE FORMATION OF MICELLES IN COLLOIDAL ELECTROLYTES, *J. Am. Chem. Soc. 69*, 683-688.

[38] Thongngam, M., and McClements, D. J. (2005) Influence of pH, ionic strength, and temperature on self-association and interactions of sodium dodecyl sulfate in the absence and presence of chitosan, *Langmuir 21*, 79-86.

[39] Palladino, P., Rossi, F., and Ragone, R. (2010) Effective Critical Micellar Concentration of a Zwitterionic Detergent: A Fluorimetric Study on n-Dodecyl Phosphocholine, *J. Fluoresc. 20*, 191-196.

[40] Hayashi, S., and Ikeda, S. (1980) MICELLE SIZE AND SHAPE OF SODIUM DODECYL-SULFATE IN CONCENTRATED NACL SOLUTIONS, *J. Phys. Chem. 84*, 744-751.

[41] le Maire, M., Champeil, P., and Moller, J. V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents, *Biochim. Biophys. Acta-Biomembr. 1508*, 86-111.

[42] Louis, J. M., Baber, J. L., Ghirlando, R., Aniana, A., Bax, A., and Roche, J. (2016) Insights into the Conformation of the Membrane Proximal Regions Critical to the Trimerization of the HIV-1 gp41 Ectodomain Bound to Dodecyl Phosphocholine Micelles, *PLoS One 11*, 21.

[43] Wingfield, P. T., Stahl, S. J., Kaufman, J., Zlotnick, A., Hyde, C. C., Gronenborn, A. M., and Clore, G. M. (1997) The extracellular domain of immunodeficiency virus gp41 protein: Expression in Escherichia coli, purification, and crystallization, *Protein Sci. 6*, 1653-1660.

[44] Brugger, B., Glass, B., Haberkant, P., Leibrecht, I., Wieland, F. T., and Krausslich, H. G. (2006) The HIV lipidome: A raft with an unusual composition, *Proc. Natl. Acad. Sci. U. S. A. 103*, 2641-2646.

[45] Lorizate, M., Sachsenheimer, T., Glass, B., Habermann, A., Gerl, M. J., Krausslich, H. G., and Brugger, B. (2013) Comparative lipidomics analysis of HIV-1 particles and their producer cell membrane in different cell lines, *Cell Microbiol. 15*, 292-304.

[46] Cheng, S. F., Chien, M. P., Lin, C. H., Chang, C. C., Lin, C. H., Liu, Y. T., and Chang, D. K. (2010) The fusion peptide domain is the primary membrane-inserted region and enhances membrane interaction of the ectodomain of HIV-1 gp41, *Mol. Membr. Biol. 27*, 31-44.

[47] Lin, C. H., Chang, C. C., Cheng, S. F., and Chang, D. K. (2008) The application of perfluorooctanoate to investigate trimerization of the human immunodeficiency virus-1 gp41 ectodomain by electrophoresis, *Electrophoresis 29*, 3175-3182.

[48] Apellaniz, B., Rujas, E., Serrano, S., Morante, K., Tsumoto, K., Caaveiro, J. M. M., Jimenez, M. A., and Nieva, J. L. (2015) The Atomic Structure of the HIV-1 gp41 Transmembrane Domain and Its Connection to the Immunogenic Membrane-proximal External Region, *J. Biol. Chem. 290*, 12999-13015.

[49] Jia, L. H., Liang, S., Sackett, K., Xie, L., Ghosh, U., and Weliky, D. P. (2015) REDOR solid-state NMR as a probe of the membrane locations of membrane-associated peptides and proteins, *J. Magn. Reson. 253*, 154-165.

[50] Sackett, K., Nethercott, M. J., Zheng, Z. X., and Weliky, D. P. (2014) Solid-State NMR Spectroscopy of the HIV gp41 Membrane Fusion Protein Supports Intermolecular Antiparallel 13 Sheet Fusion Peptide Structure in the Final Six-Helix Bundle State, *J. Mol. Biol. 426*, 1077-1094.

[51] Martin, I., Schaal, H., Scheid, A., and Ruysschaert, J. M. (1996) Lipid membrane fusion induced by the human immunodeficiency virus type 1 gp41 N-terminal extremity is determined by its orientation in the lipid bilayer, *J. Virol. 70*, 298-304.

[52] Sardan, M., Kilinc, M., Genc, R., Tekinay, A. B., and Guler, M. O. (2013) Cell penetrating peptide amphiphile integrated liposomal systems for enhanced delivery of anticancer drugs to tumor cells, *Faraday Discuss. 166*, 269-283.

[53] Markosyan, R. M., Cohen, F. S., and Melikyan, G. B. (2003) HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation, *Mol. Biol. Cell 14*, 926-938.

[54] Lev, N., Fridmann-Sirkis, Y., Blank, L., Bitler, A., Epand, R. F., Epand, R. M., and Shai, Y. (2009) Conformational Stability and Membrane Interaction of the Full-Length Ectodomain of HIV-1 gp41: Implication for Mode of Action, *Biochemistry 48*, 3166-3175.

[55] Ramakrishnan, R., Mehta, R., Sundaravaradan, V., Davis, T., and Ahmad, N. (2006) Characterization of HIV-I envelope gp4I genetic diversity and functional domains following perinatal transmission, *Retrovirology 3*, 18.

[56] Gangupomu, V. K., and Abrams, C. F. (2010) All-Atom Models of the Membrane-Spanning Domain of HIV-1 gp41 from Metadynamics, *Biophys. J. 99*, 3438-3444.

[57] Wild, C. T., Shugars, D. C., Greenwell, T. K., McDanal, C. B., and Matthews, T. J. (1994) PEPTIDES CORRESPONDING TO A PREDICTIVE ALPHA-HELICAL DOMAIN OF HUMAN-IMMUNODEFICIENCY-VIRUS TYPE-1 GP41 ARE POTENT INHIBITORS OF VIRUS-INFECTION, *Proc. Natl. Acad. Sci. U. S. A. 91*, 9770-9774.

Chapter 4

IFP membrane location studied by ^{13}C-^{2}H REDOR NMR

4.1 Introduction

Influenza virus infects respiratory epithelial cells and causes contagious respiratory illness. Influenza virus is an enveloped virus which means the virion is covered by lipid bilayer. HA is a viral integral membrane protein and has two disulfide bounded subunits: HA1 and HA2.[1-3] Influenza virus infection is a multi-step process and starts with HA1 binds to the sialic acid on host cell membrane. Virion is then endocytosed. The low pH in the endosome triggers the conformational change of HA and HA2 catalyzes the fusion between viral membrane and endosomal membrane, allowing the viral genetic material to get into the host cell.[4, 5] The ~20 N-terminus residues of HA2 is called fusion peptide (IFP) and it is highly conserved.[6] Functional studies of IFP shows that it plays a vital role in membrane fusion, even a single mutation can eliminates its fusion activity.[7-9] IFP is also believed to be the only segment in HA that binds to the endosomal membrane.[10] However, there is no crystal structure of IFP domain in HA, neither evidence of its fusion mechanism. Polypeptide IFP has been extensively studied as fusion model, and the previous research revealed the structure of IFP in detergent micelles and lipid vesicles.[11-16] Lipid membrane is more physiology relevant environment compared to detergent micelles, thus this study is focusing on the IFP-membrane interaction. Ghosh from Weliky's group found out that IFP always adopts two different N-helix/turn/C-helix structures in membrane, denoted as "semi-closed" and "closed" structures.[15] The details have been discussed in chapter 1.

Several models have been proposed either by experiment results or molecular dynamics (MD) simulations to settle the question of IFP membrane location. EPR experiments were done by Tamm's group. The result showed that H3_20 IFP is immerged in membrane as an inverted V-shaped, where the N- and C-terminus deeply inserted into hydrophobic core while the turn was close to the aqueous surface. H3_20 IFP was inserted deeper at pH 5 compare to at pH 7.4 (Figure

4.1 A). The result also suggests that a greater perturbation of lipid bilayer at pH 5 and facilitate the fusion.[11] H1 and H3 membrane associated IFP structure was studied by solid-state NMR, exhibiting amphipathic property: one side of IFP is hydrophobic and the other side is hydrophilic (Figure 4.2 B). It was proposed that IFP hydrophobic sidechains inserted into lipid hydrophobic core and the hydrophilic side facing aqueous layer.[14, 15]

Molecular dynamic simulations are used to understand the details in fusion for both peptide and lipid membrane. In 2013, Kasson's group performed MD simulations of IFP in lipid bilayers. They observed a strong relationship between lipid tail protrusion and the ability of fusion activity, and the tail protrusion is important to the transition state for fusion. It was also predicted the kinked helix structure is more fusogenic than straight or hairpin-like structures (Figure 4.1 C).[17] This is consistent with the solid-state NMR results from Weliky's group, which show a higher fusion activity at low pH when the semi-closed structure population is higher compared to closed structure (Figure 4.2).[15] People also suggested that the lipid head group intrusion contributes to the perturbation of membrane and lipid mixing.[18]

Lazaridis's gourp performed MD simulations of the "closed" H1_23 IFP in lipid membrane and they found out that the IFP deeply inserted in one membrane leaflet (Figure 4.1 D). Even N-terminal helix is buried more deeply into the hydrophobic membrane interior than the C-terminal helix, the N-terminus is solvent accessible. Several water molecules penetrate the membrane interface and interact with it. The pH also influences the orientation of the IFP. On average, the IFP plane is more tilted in the membrane at neutral pH than lower pH, because of the protonation of acidic residues E11 and D19 at low pH allows the C-terminal helix to insert deeper, lowering the rotation angle.[19]

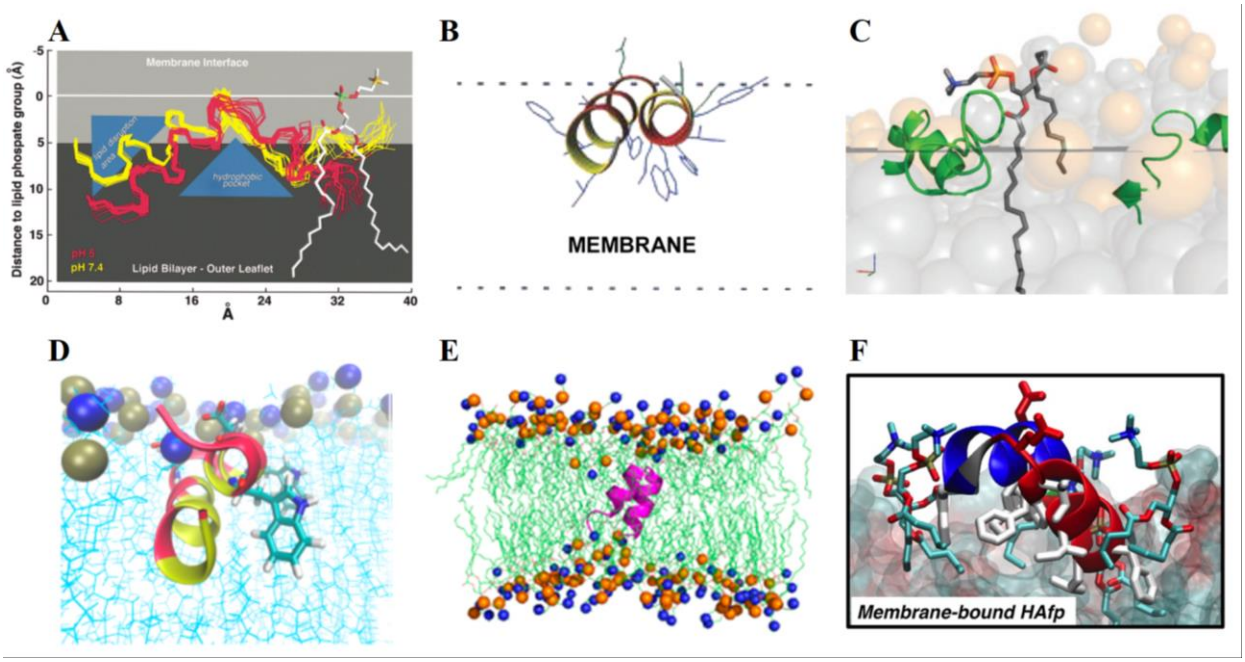

Figure 4.1 (A) The 20 lowest energy conformers of the H3_20 IFP in lipid bilayers by the EPR data at pH 5 (red) and 7.4 (yellow). A phospholipid is shown for reference and nitrogen, oxygen and phosphorus atoms are colored in yellow, red and green, respectively. The white line represents the level of the lipid phosphate groups. The hydrophobic hydrocarbon is in dark gray and interface is in light gray.[11] (B) Membrane locations of closed structure H1_23 IFP. Dashed lines are the hydrocarbon core.[15] (C) Lipid tail protrusion induced by IFP. IFP, lipid tails and phosphates are in green, gray and orange, respectively. A thin gray plane shows the average phosphorus position in the upper leaflet. One lipid is shown in sticks, with an acyl tail protruding into the polar layer.[17] (D) IFP deeply inserted in one membrane leaflet at low pH. Hydrophobic residues are highlighted in yellow and hydrophilic residues in red. Phosphorus and nitrogen atoms are in tan and blue, respectively.[19] (E) Membrane-spanning conformation model of IFP in membrane at low pH. IFP, phosphate group, amine group and acyl chain are in purple, orange, blue and green, respectively.[20] (F) IFP binds on membrane surface at neutral pH. N-terminal helix, C-terminal helix and hydrophobic side chains are in red, blue and white, respectively.[21]

Victor et al. discovered a different model in their MD simulation of IFP in membrane. These simulations revealed that the peptide became deeply inserted into the membrane and adopted a perpendicular or tilted orientation relative to the membrane plane, extending from one leaflet to the other and contacting the lipid headgroups on both sides of the membrane. This model is also called membrane-spanning conformation, which had not been predicted by previous MD simulation studies. The peptide insertion had a strong effect on the membrane, lowering the bilayer thickness, disordering nearby lipids, and promoting lipid tail protrusion (Figure 4.1 E).[20]

Another model of membrane surface was proposed recently by Tajkhorshid's group. The MD simulation result showed that hydrophobic residues Leu-2, Phe-3, Ile-6, Phe-9, and Ile-10 in N-terminal helix are buried in membrane hydrophobic core, while in C-terminal helix, amphipathic side chain Trp-14 and Met-17 are buried, whereas Trp-21 is mostly located at the headgroup region. N-terminal helix inserts much deeper into the membrane than C-terminal helix, suggesting that the N-terminal helix was responsible for hydrophobic anchoring of the peptide into the membrane. C-terminal helix, in contrast, is found to establish major amphipathic interactions at the interfacial region thereby contributing to binding strength of IFP (Figure 4.1 F).[21]

This chapter presents an investigation of the IFP membrane location studied by ^{13}C-^{2}H REDOR solid-state NMR. It is vital to detailed characterize the IFP topology in membrane since IFP is the only segment in HA that interact with target membrane and it is high relevance to the mechanism of viral entry into the host cell. The overall goal is to get a better understanding of how IFP promotes membrane fusion and provide more information for understanding fusion mechanism.

To measure the IFP membrane location, distance was measured between carbonyl carbons in peptide backbone and lipid acyl chain using ^{13}C-^{2}H REDOR. H3_20 IFP was synthesized

manually by Fmoc solid-phase peptide synthesis as described in chapter 2. The IFP constructs were ^{13}C site-specifically labeled at either Leu_2, Ala_7 or Gly_16 residues. The lipid membrane used in this study had a composition of DPPC:DPPG 4:1 ratio as a mimic of host cell membrane. The DPPC lipid was ^{2}H labeled at different acyl chain positions and denoted as DPPC_D4, DPPC_D8 and DPPC_D10 (structure in Figure 2.1). This labeling scheme was chosen based on previous peptide membrane location studies.[22]

4.2 Result and discussion

4.2.1 ^{13}CO-^{2}H REDOR spectra, buildups, and fittings

The H3_20 IFP amino acid sequence is GLFGAIAGFIENGWEGMIDGGGKKKKG where the underlined part is the original sequence and the C-terminal tag is added intended to increase the solubility of IFP. The labeled amino acids are shown in Figure 4.2 and the deuterium atoms in lipid membrane are shown in Figure 2.1.

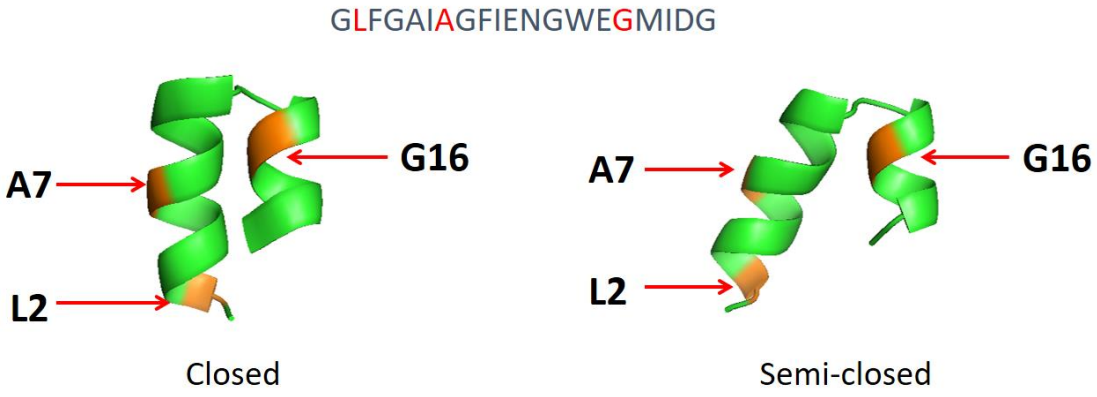

Figure 4.2 H2_20 IFP sequence, closed structure model and semi-closed model. Labeled amino acids are in red in sequence and orange in the cartoon model.

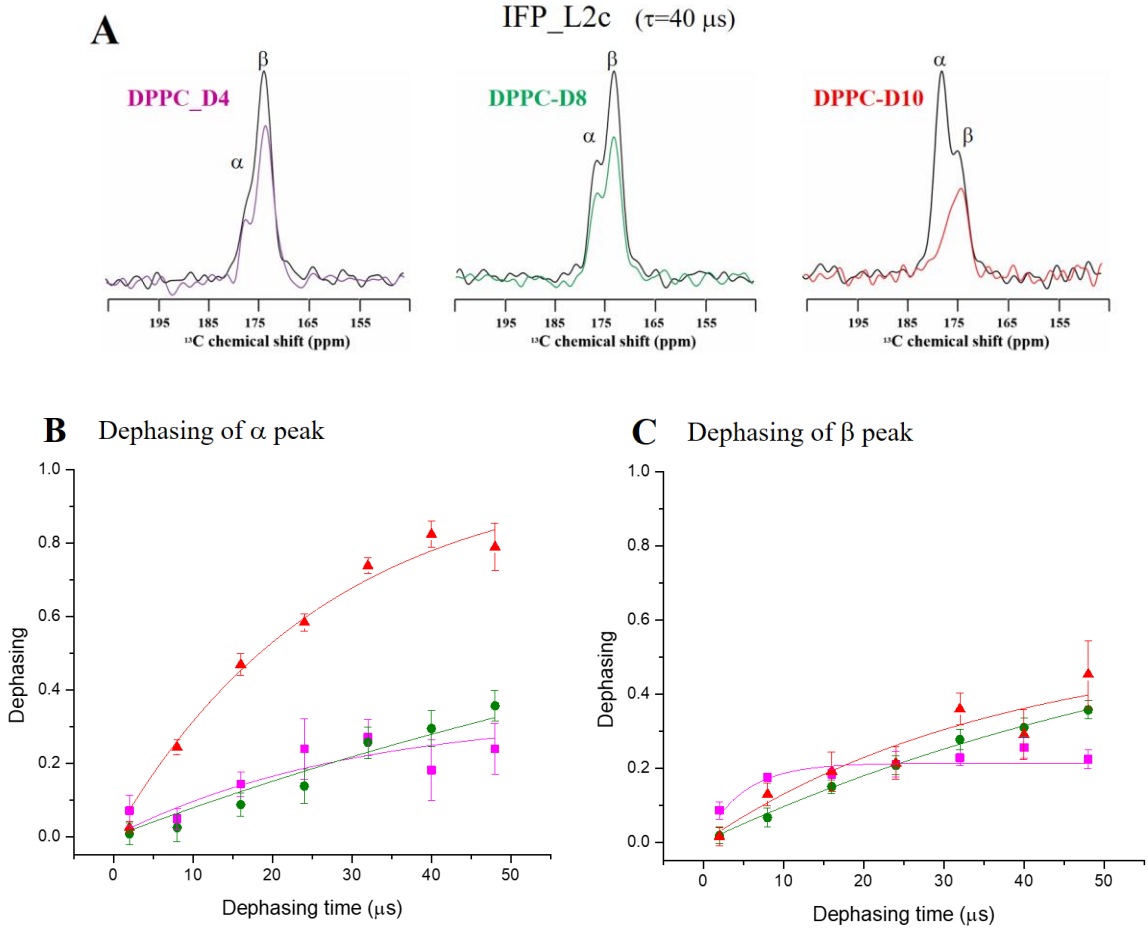

Figure 4.3 $^{13}CO–^{2}H$ REDOR data of samples that contain either IFP_L2c in membrane with DPPC:DPPG 4:1 ratio. Samples were prepared with DPPC_D4, DPPC_D8 and DPPC_D10, and the corresponding data are displayed in purple, green and red, respectively. (A) S_0 (black) and S_1 (colored) REDOR spectra for τ = 40 ms. (B) and (C) Plots of ($\Delta S/S_0$) vs τ are displayed for the α helical and β strand peaks. The solid lines are best-fits to $A(1 - e^{-\gamma\tau})$ and the fitting parameters are in Table 4.1.

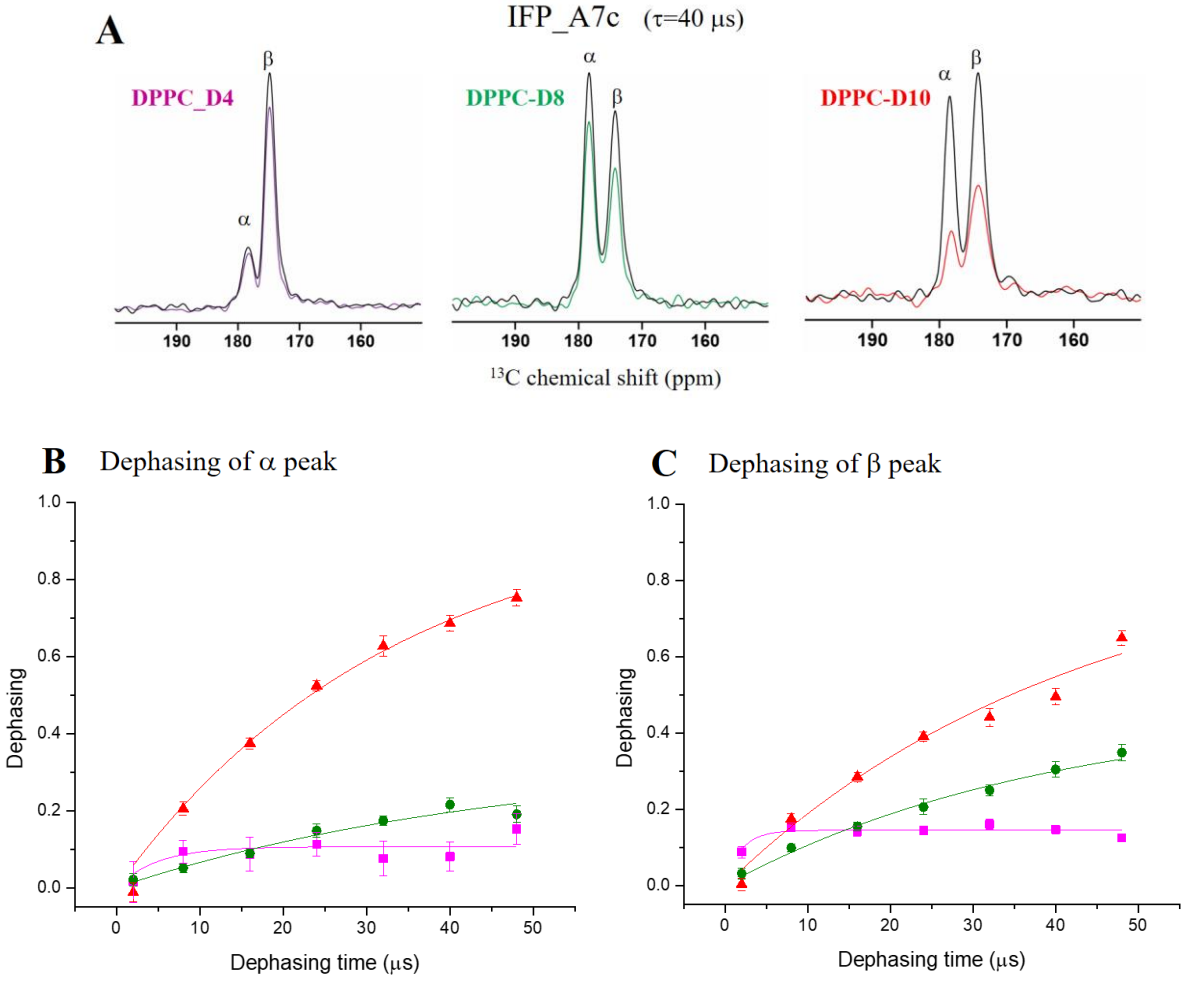

Figure 4.4 $^{13}CO–^{2}H$ REDOR data of samples that contain either IFP_A7c in membrane with DPPC:DPPG 4:1 ratio. Samples were prepared with DPPC_D4, DPPC_D8 and DPPC_D10, and the corresponding data are displayed in purple, green and red, respectively. (A) S_0 (black) and S_1 (colored) REDOR spectra for τ = 40 ms. (B) and (C) Plots of ($\Delta S/S_0$) vs τ are displayed for the α helical and β strand peaks. The solid lines are best-fits to $A(1 - e^{-\gamma\tau})$ and the fitting parameters are in Table 4.1.

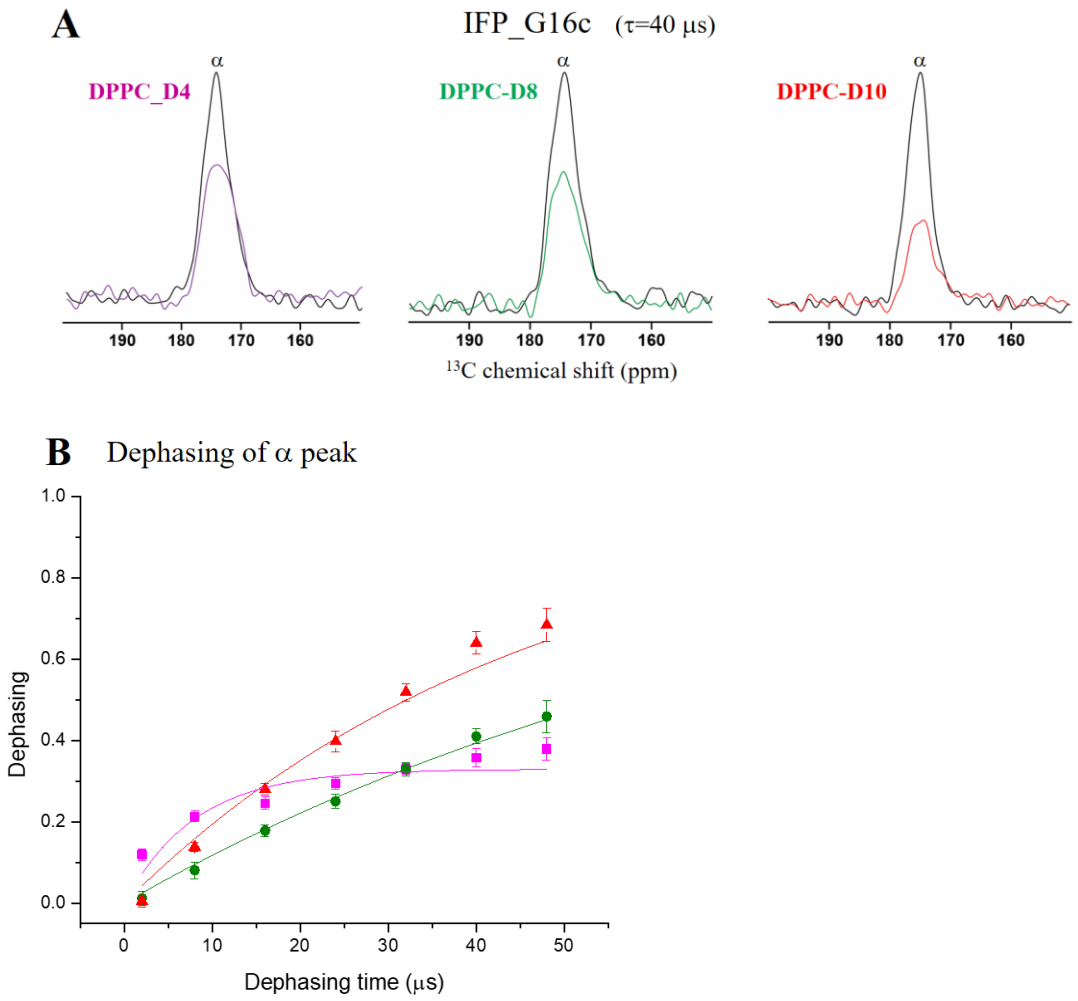

Figure 4.5 $^{13}CO–^{2}H$ REDOR data of samples that contain either IFP_G16c in membrane with DPPC:DPPG 4:1 ratio. Samples were prepared with DPPC_D4, DPPC_D8 and DPPC_D10, and the corresponding data are displayed in purple, green and red, respectively. (A) S_0 (black) and S_1 (colored) REDOR spectra for τ = 40 ms. (B) Plot of ($\Delta S/S_0$) vs τ are displayed for the α helical peak. The solid lines are best-fits to $A(1 - e^{-\gamma\tau})$ and the fitting parameters are in Table 4.1.

Table 4.1 Best-fit exponential buildup parameters for ^{13}CO–^{2}H REDOR of IFP in DPPC and DPPG [a] [b]

Peptide	Peak	Membrane	A	γ (Hz)	r (Å)
L2c	α	DPPC_D4	0.36 (15)	34 (24)	4 (2)
L2c [c]	α	DPPC_D10	1.00 (15)	38 (10)	4.3 (5)
L2c	β	DPPC_D8	0.72 (10)	14 (3)	6.0 (5)
L2c [c]	β	DPPC_D10	0.55 (23)	27 (18)	5 (2)
A7c	α	DPPC_D8	0.34 (10)	21 (9)	5 (1)
A7c [c]	α	DPPC_D10	0.99 (10)	30 (5)	4.7 (3)
A7c	β	DPPC_D8	0.47 (6)	25 (5)	5.0 (4)
A7c	β	DPPC_D10	0.90 (19)	23 (7)	5.1 (7)
G16c	α	DPPC_D8	1.00 (27)	12 (4)	6.4 (9)
G16c [c]	α	DPPC_D10	1.00 (28)	22 (8)	5.2 (8)

[a] Samples are prepared by aqueous binding method with 1 μmol IFP, 40 μmol deuterated DPPC and 10 μmol DPPG

[b] Data that with big error and unreasonable fitting is not shown

[c] Samples with substantial dephasing

The spectra of the IFP L2c, A7c and G16c and their fitting are shown in Table 4.1, Figure 4.3, Figure 4.4 and Figure 4.5 respectively. The spectra were obtained at ~-30 °C and the $\Delta S/S_0$ dephasing buildup was fitted to $A(1 - e^{-\gamma\tau})$, in which A and γ are fitting parameters. A represents the fraction of protein with dipolar coupling $D \approx 3\gamma/2$ which is based on equal time spent in the

three ^{2}H m states during τ because of T_1 relaxation. (1 - A) is the fraction protein with D ≈ 0. D, in units of Hz is given by: $D = \frac{\mu_0 h \gamma_C \gamma_D}{16\pi^3 r^3}$. Thus, the ^{13}C – ^{2}H dipolar coupling in Hz is $D = \frac{4642}{r^3}\ Hz$. The internuclear distance r was calculated as $\sqrt[3]{\frac{4642\ Hz}{\frac{3\gamma}{2}}}$ Å. The ^{13}CO chemical shifts are correlated to the peptide backbone secondary structures and thus the secondary structure can be determined by comparing the chemical shifts in spectra to literature.[23, 24]

A9_^{13}CO, A8_^{2}H$_\alpha$ labeled I4 peptide was used as a standard compound, the IFP-membrane interaction has some differences. (1) The membrane environment is locally non-crystalline so that the distance between the ^{13}CO in IFP and a particular ^{2}H in membrane may vary among peptide molecules even if all the IFP have the same membrane location. (2) Each IFP is surrounded by several lipid molecules, which means the $\Delta S/S_0$ is a sum of multi-spin geometry. Since $D \propto r^{-3}$ for each spin pair, the $\Delta S/S_0$ buildup is dominated by the D associated with the closest ^{2}H. Overall, these considerations for peptide ^{13}CO–lipid ^{2}H REDOR imply that fitting parameters will be semi-quantitatively rather than quantitatively related to membrane location.[22, 25]

The L2c and A7c have two peaks in DPPC-D4, DPPC-D8 and DPPC-D10 membrane, however, G16c only exhibits one peak. The L2c, A7c and G16c spectra have peaks with respective 177, 178 and 175 ppm chemical shifts and are assigned to α helical structure of L2, A7 and G16 residues. The peaks at 174 and 175 ppm in L2c and A7c spectra are assigned to β sheet structure.[24] Although all samples were prepared using the same protocol, there was some variation in the α:β population ratio. L2 α helical peak population in three lipids are D10 > D8 > D4. A7 α helical peak population are D8 > D10 > D4.

As for G16, there is only one broader peak at 175 ppm and this peak is assigned to α structure based on two reasons: (1) The chemical shift of β sheet Gly should have a chemical shift

of ~170 ppm[25] and there is no peak or shoulder correspond to this region, and (2) the chemical shifts of lipid carbonyl carbon natural abundance is also at ~ 175 ppm, and the broader peak may resulted from the overlapping of ^{13}CO in IFP and natural abundance in lipids. 1,2-di-O-tetradecyl-sn-glycero-3-phosphocholine (DTPC) and 1,2-di-O-tetradecyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DTPG) are lipids that do not have carbonyl group. By comparing the full width at half maximum (FWHM) of G16c S_0 in DPPC:DPPG membrane to DTPC:DTPG membrane, the natural abundance contribution can be obtained (shown in Table 4.2). IFP in DTPC:DTPG has 4.6 ppm FWHM at 2 ms and 4.3 ppm FWHM at 40 ms. Even though IFP in DPPC:DPPG has an average 4.9 ppm FWHM at both dephasing time, which indicates that lipid carbonyl carbon natural abundance indeed contributes to broader lines, the L2 and A7 still showed a much narrower FWHM of ~ 2 ppm. Thus, lipid natural abundance is not the main reason of broad G16 peak. The reason might be because the Gly has less steric restriction than other residues.

Previous studies of IFP in detergent micelle and lipid membrane both showed that the IFP adopts a N-helix (residues 1–11)/turn/C-helix (residues 13–19) structure. The two helices are in close contact and this structure is denoted as "closed" structure. In lipid membrane, IFP also adopts a "semi-closed" structure, which is also a helix/turn/helix structure but with a wider interhelical angle.[12, 15] There is not much information about the β structure. It might be an β sheet structure at residue 1 – 11 region and α helix at residue 13 – 19.

Table 4.2 FWHM of membrane associated IFP [a]

Labeling	2 ms	40 ms
G16c – D4	5.5	5.2
G16c – D8	5.2	5.2
G16c – D10	4.2	4.2
G16c – F9n [b]	4.6	4.3

[a] IFP with DPPC:DPPG 4:1 at pH 5. S_0 spectra of 2 ms or 40 ms dephasing time were processed with 20 Hz line broadening, baseline DC offset correction and baseline polynomial correction of the order 5. The FWHM values are in unit ppm.

[b] IFP with DTPC:DTPG 4:1 at pH 5. The data is obtained from Ujjayini Ghosh's dissertation.

The α peak dephasing buildups for L2c, A7c and G16c in membrane containing DPPC_D10 is comparable large and buildups in DPPC_D8 or DPPC_D4 are smaller. As shown in Table 4.1, the best-fit $A \approx 1$ for all three residues with the DPPC_D10 lipid supports a single membrane location of IFP. Considering the IFP structure, L2 ^{13}CO is on the hydrophobic face and A7, G16 ^{13}CO nuclei are on the hydrophilic faces of the IFP. However, they all showed a similar 4 – 5 Å internuclear distance with the D10 deuterium, which is consistent with van der Waals contact.

For membrane that contains DPPC_D10, there are smaller β than α buildups for L2c and A7c. This suggests there are smaller β than α buildups, which supports an overall shallower location for β IFP. Additionally, the fitting results of A7c in D10 and D8 show $r_{D10} \approx r_{D8}$ and $A_{D10} \approx 2A_{D8}$. This supports two different membrane locations for IFP with a major population with deep

insertion in contact with D10 ^{2}H's and a minor population with shallower insertion in contact with D8 ^{2}H's.[25]

4.2.2 Effects of sample preparation method and lipid charge

The result discussed in 4.2.1 was IFP in negatively charge DPPC and DPPG lipid membrane prepared by aqueous binding (AB) and the protocol is described in Chapter 2. This method was designed as incorporation during viral fusion. Organic co-solubilization (OC) sample preparation method was also used for IFP in neutral DPPC membrane. Table 4.3 and Figure 4.5 displays the τ = 40 ms ^{13}C-^{2}H REDOR spectra and fittings of IFP L2c in 100 % DPPC-D10 and DPPC-D8 lipids. This method is to achieve thermodynamic equilibrium integration of the two components. The details process is as below.

50 μmole of DPPC-D10 or DPPC-D8 lipids was dissolved in chloroform and then solvent was removed by dry nitrogen gas flow and vacuum pumping overnight. 1 μmol dry IFP_L2c was added to lipid film and then dissolved in a solvent mixture containing 2,2,2-trifluoroethanol, 1,1,1,3,3,3-hexafluoroisopropanol, and chloroform with 2:2:3 volume ratio followed by subsequent solvent removal via nitrogen gas and vacuum pumping. 3 mL of 10 mM HEPES and 5mM MES buffer at pH 5.0 was used to hydrate the film and followed by 10 times freeze-thaw cycles to make a homogeneous suspension. Another 20 mL buffer was added followed by ultra-centrifugation at 100000g for 5 hours at 4 °C to pellet the peptide-lipid complex.

As shown in Figure 4.6 panel A, the chemical shifts of L2c in DPPC-D8 and DPPC-D10 are 174 ppm which is correspond to the β strand structure only. This result is different from the situation where L2c binds to negatively charged membrane (DPPC:DPPG 4:1 ratio) in aqueous phase and the peptide adopts both α helix and β strands structure.

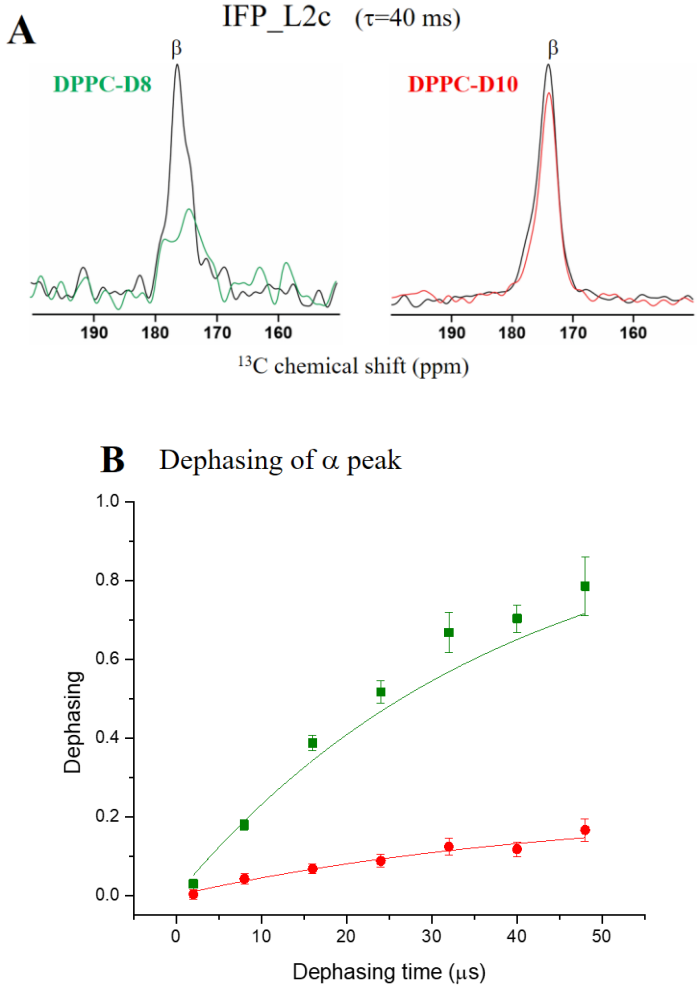

Figure 4.6 ^{13}CO–^{2}H REDOR data of samples that contain either IFP_L2c in membrane with DPPC prepared by organic co-solubilization method. Samples were prepared with DPPC_D8 and DPPC_D10, and the corresponding data are displayed in green and red, respectively. (A) S_0 (black) and S_1 (colored) REDOR spectra for τ = 40 ms. (B) Plot of ($\Delta S/S_0$) vs τ are displayed for the β strand peak. The solid lines are best-fits to $A(1 - e^{-\gamma\tau})$ and the fitting parameters are in Table 4.3.

Table 4.3 Best-fit exponential buildup parameters for $^{13}CO–^{2}H$ REDOR of IFP in DPPC [a]

Peptide	Peak	Membrane	A	γ (Hz)	r (Å)
L2c[b]	β	DPPC_D8	1.00 (29)	26 (11)	5 (1)
L2c	β	DPPC_D10	0.22 (7)	22 (10)	5 (1)

[a] Samples are prepared by organic co-solubilization method with 1 μmol IFP and 50 μmol deuterated DPPC.

[b] Samples with substantial dephasing

The difference in signal to noise ratio of D8 and D10 spectra (Figure 4.6 A right panel) is due to the amount of IFP that bound to the membrane. The D8 sample contained 0.65 μmol peptide, while the D10 sample contained 1.08 μmol peptide. The signal to noise of D8 and D10 samples at 2 μs dephasing time are the same, while the acquisition number of D8 is three times greater than D10. Thus the signal of D10 is 1.7 times greater than D8, which is consistent with the 1.08:0.65 mole ratio.

L2c has a much higher buildup rate in DPPC_D8 lipids than in DPPC_C10 lipids (Figure 4.6 B). The ^{13}CO-^{2}H distances of L2c with D10 and D8 are both ≈ 5 Å, and $A_{D10} \approx 1$ is much greater than $A_{D8} \approx 0.22$ (Table 4.3). This suggest that when IFP and neutral lipid were prepared by organic co-solubilization method, the L2 ^{13}CO is predominantly contacting D8 ^{2}H's. However, when the sample is prepared by aqueous binding method and with negatively charged lipids, $r_{D10} \approx 5$ Å is closer than $r_{D8} \approx 6$ Å, and $A_{D10} > A_{D8}$. Besides, comparing the dephasing of D8, $(\Delta S/S_0)_{AB} = 0.3$ is smaller than $(\Delta S/S_0)_{OC} = 0.7$ at large τ. All the information showed that the IFP L2c has close contact with D8 ^{2}H's in neutral membrane, while close contact with D10 ^{2}H's in negatively charged membrane.

The sample preparation methods and the charge of the lipids can be the reasons that cause such significant difference. Aqueous binding process is a mimic of actual infection, while organic co-solubilization is to achieve thermodynamic equilibrium integration of the two components. Previous study of HIV gp41 fusion peptide (FP) in lipid membrane containing DPPC:DPPG:Cholesterol with 8:2:5 ratio composition using aqueous binding and organic co-solubilization sample preparation methods obtained both β sheet structure spectra and similar extent of buildups.[25] As for the charge effect, in that paper, organic co-solubilization prepared FP with neutral PC only membrane and negative PC:PG membrane were also compared. The two spectra both showed β sheet structure spectra and qualitatively similar buildups. Thus, from the results, the sample preparation and lipid composition did not affect the peptide secondary structure and membrane location of HIV fusion peptide.

However, when IFP was prepared by different sample preparation methods and with different lipid composition, both secondary structure and membrane location were change. When prepared by aqueous binding with negatively charged membrane, IFP adopts both α strand and β sheet structures. The α IFP showed fast dephasing buildup with D10, on the other hand, β IFP has small buildups with all D4, D8 and D10. When prepared by organic co-solubilization with neutral membrane, IFP only adopts β sheet structure and showed high buildup with D8. The reason for different secondary structures and different β structure membrane locations maybe due to the sample preparation methods. The organic co-solubilization method achieves the thermodynamic favorable integration of the IFP and membrane, while the aqueous binding method leads to kinetic favorable conformation.

The electrostatic interaction between positively charged IFP and negatively charged DPPG lipid at pH 5.0 also affects the binding percentage. As mentioned in Chapter 2, the NMR sample

was the centrifugation pellet containing membrane + bound peptide and the unbound IFP exists in the supernatant. After ultra-centrifugation, the supernatant A280 absorbance was measured to quantify the amount of unbounded IFP. For samples prepared by aqueous binding method, no IFP exist in supernatant (A280 = 0), which means 100% binding. But for organic co- solubilization, only 60% to 80% binding percentage was observed.

The DPPC:DPPG 4:1 composition and aqueous binding was chosen for most of the samples based on two reasons: (1) PC:PG 4:1 combination is a better mimic of host cell membrane since the human cell membrane has similar fraction of negatively charged lipids;[26] (2) aqueous binding is similar to the viral infection process, in which the IFP is exposed during conformational change of HA and then binds to the host cell membrane.

4.2.3 Close contact of α helical IFP and ^{2}H in DPPC_D10

IFP with N-helix/turn/C-helix structure appears to have a single membrane location as evidenced by rapid L2c, A7c and G16c buildups in membrane with DPPC_D10 and correlate best-fit A ≈ 1 (Figure 4.3 and Table 4.1). Much smaller buildups were observed in membrane with DPPC_D4 and DPPC_D8. There are close contacts between the lipid acyl chain tail and α helical IFP in both N-helix and C-helix region. This may be relevant for fusion catalysis because of local perturbation of the membrane bilayer with consequent reduced activation energy to the highly perturbed fusion transition state.

By comparing the REDOR result to the current MD simulations (Figure 4.1 C - F), close contacts of all the L2, A7 and G16 residues with lipid acyl chain tail is inconsistent with N-helix deeply inserted while C-helix at membrane surface (Figure 4.1 D) and inconsistent with interfacial location of a IFP (Figure 4.1 F).[19, 21] However, based on the current REDOR results, it is hard to

distinguish whether IFP adopts a membrane-spanning conformation or IFP promotes lipid acyl chain protrusion. The possible IFP membrane location models are shown in Figure 4.7.

Both models may facilitate membrane fusion by perturbing local membrane structure. In membrane spanning model, the IFP inserts deeply in membrane hydrophobic center and break the continuous membrane structure (Figure 4.7 A), and in lipid protrusion model, the lipid acyl chain is induced to protrude and contact with water (Figure 4.7 B). When IFP bonds to host cell membrane, this perturbation may facilitate the lipid acyl chain to interact with viral outer leaflet and thus induce hemifusion.

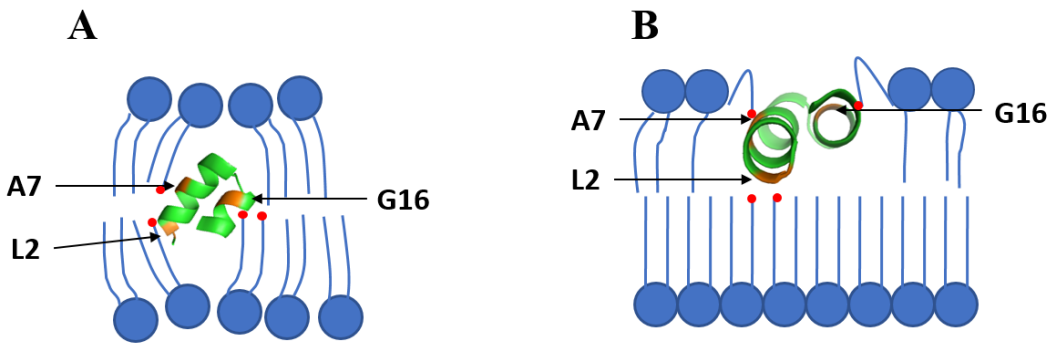

Figure 4.7 Possible models for IFP topology in membrane. The lipid molecules are shown in blue, D10 ^{2}H atoms in lipid acyl chain shown as red dots, IFP in green and labeled residues in orange. (A) Membrane spanning model and (B) lipid tail protrusion model.

4.3 Conclusion

^{13}C-^{2}H REDOR solid-state NMR is used to study the IFP membrane location. In this work, H3_20 IFP was synthesized manually by Fmoc solid-phase peptide synthesis, and ^{13}CO labeled at either Leu_2, Ala_7 or Gly_16 residues. IFP was bonded to acyl chain deuterated lipid vesicle with a composition of DPPC:DPPG 4:1 ratio as a mimic of host cell membrane. The ^{13}CO-^{2}H dipolar coupling was fitted based on REDOR dephasing and internuclear distances were calculated.

The IFP adopts major α helical, minor β strand secondary structure in PC/PG membrane. The α helical IFP's with respectively ^{13}CO labeled Leu-2, Ala-7 and Gly-16 all show close contacts with the lipid acyl chain tail, suggesting IFP has strong interaction with the membrane: it either has a membrane-spanning confirmation, or it promotes lipid trail protrusion. IFP bounded lipid membrane structure studied by paramagnetic relaxation enhancement (PRE) solid-state NMR will be discussed in Chapter 5 and provides more information about the detailed IFP membrane location model.

REFERENCES

REFERENCES

[1] Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., and Kawaoka, Y. (1992) EVOLUTION AND ECOLOGY OF INFLUENZA-A VIRUSES, *Microbiol. Rev. 56*, 152-179.

[2] Wilson, I. A., Skehel, J. J., and Wiley, D. C. (1981) STRUCTURE OF THE HEMAGGLUTININ MEMBRANE GLYCOPROTEIN OF INFLUENZA-VIRUS AT 3-A RESOLUTION, *Nature 289*, 366-373.

[3] Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) STRUCTURE OF INFLUENZA HEMAGGLUTININ AT THE PH OF MEMBRANE-FUSION, *Nature 371*, 37-43.

[4] Kemble, G. W., Danieli, T., and White, J. M. (1994) LIPID-ANCHORED INFLUENZA HEMAGGLUTININ PROMOTES HEMIFUSION, NOT COMPLETE FUSION, *Cell 76*, 383-391.

[5] Skehel, J. J., and Wiley, D. C. (2000) Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin, *Annu. Rev. Biochem. 69*, 531-569.

[6] Nobusawa, E., Aoyama, T., Kato, H., Suzuki, Y., Tateno, Y., and Nakajima, K. (1991) COMPARISON OF COMPLETE AMINO-ACID-SEQUENCES AND RECEPTOR-BINDING PROPERTIES AMONG 13 SEROTYPES OF HEMAGGLUTININS OF INFLUENZA A-VIRUSES, *Virology 182*, 475-485.

[7] Gething, M. J., Doms, R. W., York, D., and White, J. (1986) STUDIES ON THE MECHANISM OF MEMBRANE-FUSION - SITE-SPECIFIC MUTAGENESIS OF THE HEMAGGLUTININ OF INFLUENZA-VIRUS, *J. Cell Biol. 102*, 11-23.

[8] Qiao, H., Armstrong, R. T., Melikyan, G. B., Cohen, F. S., and White, J. M. (1999) A specific point mutant at position 1 of the influenza hemagglutinin fusion peptide displays a hemifusion phenotype, *Mol. Biol. Cell 10*, 2759-2769.

[9] Steinhauer, D. A., Wharton, S. A., Skehel, J. J., and Wiley, D. C. (1995) STUDIES OF THE MEMBRANE-FUSION ACTIVITIES OF FUSION PEPTIDE MUTANTS OF INFLUENZA-VIRUS HEMAGGLUTININ, *J. Virol. 69*, 6643-6651.

[10] Durrer, P., Galli, C., Hoenke, S., Corti, C., Gluck, R., Vorherr, T., and Brunner, J. (1996) H+-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region, *J. Biol. Chem. 271*, 13417-13421.

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

[12] Lorieau, J. L., Louis, J. M., and Bax, A. (2010) The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface, *Proc. Natl. Acad. Sci. U. S. A. 107*, 11341-11346.

[13] Sun, Y., and Weliky, D. P. (2009) C-13-C-13 Correlation Spectroscopy of Membrane-Associated Influenza Virus Fusion Peptide Strongly Supports a Helix-Turn-Helix Motif and Two Turn Conformations, *J. Am. Chem. Soc. 131*, 13228-13229.

[14] Ghosh, U., Xie, L., and Weliky, D. P. (2013) Detection of closed influenza virus hemagglutinin fusion peptide structures in membranes by backbone (CO)-C-13-N-15 rotational-echo double-resonance solid-state NMR, *J. Biomol. NMR 55*, 139-146.

[15] Ghosh, U., Xie, L., Jia, L. H., Liang, S., and Weliky, D. P. (2015) Closed and Semiclosed Interhelical Structures in Membrane vs Closed and Open Structures in Detergent for the Influenza Virus Hemagglutinin Fusion Peptide and Correlation of Hydrophobic Surface Area with Fusion Catalysis, *J. Am. Chem. Soc. 137*, 7548-7551.

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

[17] Larsson, P., and Kasson, P. M. (2013) Lipid Tail Protrusion in Simulations Predicts Fusogenic Activity of Influenza Fusion Peptide Mutants and Conformational Models, *PLoS Comput. Biol. 9*, 9.

[18] Legare, S., and Laggue, P. (2014) The influenza fusion peptide promotes lipid polar head intrusion through hydrogen bonding withphosphates and N-terminal membrane insertion depth, *Proteins 82*, 2118-2127.

[19] Brice, A. R., and Lazaridis, T. (2014) Structure and Dynamics of a Fusion Peptide Helical Hairpin on the Membrane Surface: Comparison of Molecular Simulations and NMR, *J. Phys. Chem. B 118*, 4461-4470.

[20] Victor, B. L., Lousa, D., Antunes, J. M., and Soares, C. M. (2015) Self-Assembly Molecular Dynamics Simulations Shed Light into the Interaction of the Influenza Fusion Peptide with a Membrane Bilayer, *J. Chem Inf. Model. 55*, 795-805.

[21] Baylon, J. L., and Tajkhorshid, E. (2015) Capturing Spontaneous Membrane Insertion of the Influenza Virus Hemagglutinin Fusion Peptide, *J. Phys. Chem. B 119*, 7882-7893.

[22] Xie, L., Ghosh, U., Schmick, S. D., and Weliky, D. P. (2013) Residue-specific membrane location of peptides and proteins using specifically and extensively deuterated lipids and C-13-H-2 rotational-echo double-resonance solid-state NMR, *J. Biomol. NMR 55*, 11-17.

[23] Morcombe, C. R., and Zilm, K. W. (2003) Chemical shift referencing in MAS solid state NMR, *J. Magn. Reson. 162*, 479-486.

[24] Zhang, H. Y., Neal, S., and Wishart, D. S. (2003) RefDB: A database of uniformly referenced protein chemical shifts, *J. Biomol. NMR 25*, 173-195.

[25] Jia, L. H., Liang, S., Sackett, K., Xie, L., Ghosh, U., and Weliky, D. P. (2015) REDOR solid-state NMR as a probe of the membrane locations of membrane-associated peptides and proteins, *J. Magn. Reson. 253*, 154-165.

[26] van Meer, G., Voelker, D. R., and Feigenson, G. W. (2008) Membrane lipids: where they are and how they behave, *Nat. Rev. Mol. Cell Biol. 9*, 112-124.

Chapter 5

IFP effects on membrane studied by ^{2}H paramagnetic relaxation enhancement (PRE) solid-state NMR

5.1 Introduction

In Chapter 4, REDOR NMR results showed that the Leu2, Ala7 and Gly16 residues in IFP have close contact with lipid acyl chain tail. According to the current models, the REDOR data is consistent with two of them: membrane-spanning model and lipid acyl chain protrusion model, and the possible IFP membrane location models are shown in Figure 4.7. To further study the membrane structure perturbed by IFP, Paramagnetic Relaxation Enhancement (PRE) solid-state NMR is used. The information about IFP-membrane interaction may help us to understand the fusion process and mechanism.

Solomon developed the PRE in 1955.[1] PRE experiments utilizes the high gyromagnetic ratio of unpaired electron in paramagnetic species to enhance dipolar interactions between the nuclear of interest and the electron and can be used to get distance information in a range up to 35 Å. Thus, it is a very powerful tool for long-range distance determination, and for data acquisition acceleration, especially suitable for biomolecules and human tissue.[2, 3]

Previous studies used paramagnetic species such as Mn^{2+}, Gd^{2+}, and nitroxide spin in solid state NMR for studying protein membrane location. For example, to study the Protegrin-1 peptide membrane location, ^{13}C site-specifically labeled the peptide was bonded to membrane and the signal attenuation caused by Mn^{2+} was measured. Paramagnetic Mn^{2+} induces distance-dependent line broadening and signal attenuation of the nuclear spin signals by enhancing the spin-spin relaxation rate ($1/T_2$). The Mn^{2+} location in a lipid vesicle system is well defined: it only attaches to membrane surface and does not penetrate the membrane. By comparing the signal decrease in peptide to the signal decrease in lipid carbons, the relative depth of the peptide with respect to the lipid can be extracted.[4] This method can also be used to study the asymmetric insertion of membrane proteins in membrane by only applying Mn^{2+} ions on the outer but not the inner leaflet

of lipid bilayers.[5] Another example is the location of cholesterol in membrane determined by comparing the ^{13}C spin-lattice relaxation rates of phospholipid and cholesterol in the presence of paramagnetic Gd^{3+}.[6]

In this study, PRE was used to illustrate the effect of IFP on membrane structure and also to solve the model of IFP insertion in membrane.

PRE was chosen based on its long distance detection range. Besides, compare to other techniques such as NOE, ^{1}H spin diffusion or EPR, PRE has less limitations. The NOE detection range is ~ 5 Å, which is too short compared to the 34 Å membrane hydrophobic core thickness. To study the peptide insertion in membrane, ^{1}H spin diffusion NMR is another approach. The magnetization transfer rate from lipid methyl protons in the center of the bilayer to the target peptide is measured in this experiment, and this transfer is facilitated by ^{1}H-^{1}H dipolar coupling. There is a significant difference in the magnetization transfer rates between a rigid peptide and the mobile lipids at ambient temperature. The spin diffusion from the lipid methyl protons to a rigid peptide close to the center of the bilayer is rapid. In contrast, for a peptide that bind on membrane surface, the methyl ^{1}H magnetization must first diffuse through the lipid acyl chains before reaching the protein. Since spin diffusion in the lipids is extremely slow because of motion-averaged ^{1}H-^{1}H dipolar couplings, the surface-bound protein receives little ^{1}H magnetization from the methyl protons and showed low transfer rate. If the peptide does not have enough rigidity, the spin diffusion from methyl proton to peptide is also slow, then it is hard to distinguish the diffusion to peptide or lipid since they have closer rate. Thus, to distinguish the peptide location on membrane surface or membrane hydrophobic core, a rigid the peptide is required.[7] EPR experiments required to introduce of bulky spin probes, which may perturb the membrane packing and complicate the data interpretation.[8]

In the experiment, lipid membrane with composition of DPPC:DPPG 4:1 ratio was used, and this fraction of negatively charged lipid was chosen as a mimic the HIV host cell membrane.[9] To study the membrane structure, DPPC-D8 and DPPC-D10 lipids were chosen, as they are deuterated on lipid acyl chain as shown in Figure 2.1. Deuterium atoms in DPPC-D8 and DPPC_D10 are at center of membrane leaflet and center of membrane hydrophobic core, thus their positions in membrane can provide information of the lipid acyl chain conformation. Paramagnetic species, Mn^{2+}, was added to vesicles. Labeling scheme shown in Figure 5.1.

The PRE arises from dipolar interactions between the ^2H in lipid acyl chain and the unpaired electrons in Mn^{2+} and contributes to faster T_2 relaxation which can be described as:

$$\frac{1}{T_2} = R_2 = W\frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2\frac{\gamma_D^2\mu_{eff}^2\beta^2}{r^6}\left(4\tau_s + \frac{3\tau_s}{1+\omega_D^2\tau_s^2} + \frac{13\tau_s}{1+\omega_e^2\tau_s^2}\right) \quad 5.1$$

Where R_2 is the T_2 relaxation rate; W is the local concentration of the Mn^{2+} ions; μ_0 is the vacuum permeability; γ_D is the gyromagnetic ratio of ^2H; μ_{eff} is the effective magnetic moment of Mn^{2+} ions; β is the Bohr magneton; r is the average electron-nucleus distance; ω_D and ω_e are the ^2H and election Larmor frequencies. The correlation time τ_s is the inverse sum of the electronic spin-lattice relaxation time T_{1e}; the rotational correlation time of the molecule τ_r; and the residence time of the Mn^{2+} near ^2H τ_m:

$$\frac{1}{\tau_s} = \frac{1}{T_{1e}} + \frac{1}{\tau_r} + \frac{1}{\tau_m} \quad 5.2$$

Since Mn^{2+} only binds on membrane surface, it can represent the aqueous layer. By measuring ^2H Mn^{2+} induced R_2 increase, the distance of ^2H to membrane surface can be obtained.

Deuterium T_2 relaxation rate was determined by quadrupolar echo (quecho) NMR experiment (pulse sequence shown in Figure 5.2). The details have been discussed in chapter 1

and chapter 2. The signal intensity (I) in FID was plotted vs t, and t = [τ_1 + τ_2 + time being shifted]. The T_2 relaxation time can be fitted by:

$$I(t) = I(0) \times exp\left(\frac{t}{T_2}\right) \quad 5.3$$

Where $I(0)$ and T_2 are fitting parameters. $I(t)$ is the FID intensity at t = [τ_1 + τ_2 + time being shifted]; $I(0)$ is the FID intensity at t=0; and T_2 is the transverse relaxation time or spin-spin relaxation time. The R_2 is calculated by:

$$R_2 = \frac{1}{T_2} \quad 5.4$$

R_2 of (1) pure lipid; (2) pure lipid with Mn^{2+}; (3) peptide-bounded lipid; and (4) peptide-bounded lipid with Mn^{2+} were measured. If smaller R_2 difference is observed between (1) and (2) than (3) and (4), the ^2H atoms in IFP-bounded membrane have a closer distance to membrane surface. If the R_2 difference between (1) (2) and (3) (4) is similar, no lipid protrusion is promoted by IFP. Thus, the distance information between ^2H and membrane surface can be obtained.

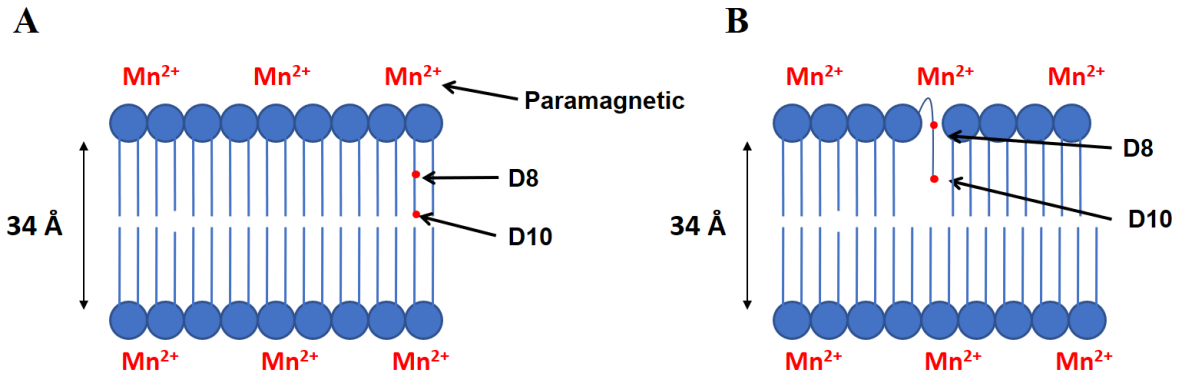

Figure 5.1 Labeling scheme of PRE experiments. Paramagnetic species are Mn^{2+} ions, and they bind to the surface of lipid membrane. Lipid composition used was DPPC:DPPG 4:1 mole ratio. DPPC-D8 and DPPC-D10 were used. The hydrophobic core is 34 Å, which is a comparable length with the PRE detection range. The Mn^{2+} ions are labeled in red, lipid molecules in blue and deuterium atoms shown as red dots. (A) Lipid acyl chain without protrusion. (B) Lipid acyl chain protrudes towards aqueous surface.

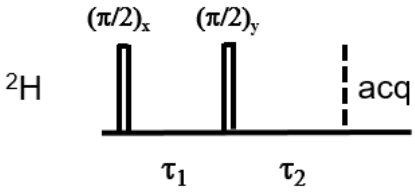

Figure 5.2 Quecho pulse sequence. ^2H spectra were acquired with different τ_1 and τ_2 and a fixed (τ_1 -τ_2) value. Typically, the pulse length was set ≈1.5 μs; τ_1, τ_2 were set between 10 and 1000 μs; the recycle delay was set to 1 s. Samples are detected at 25 °C (gel phase lipids) and 50 °C (fluid phase lipids).

5.2 Result and discussion

5.2.1 Sample preparation methods

The lipid vesicle preparation and peptide aqueous binding methods are discussed in chapter 2. The relaxation rate in PRE experiments depends on the concentration of Mn^{2+} ions. Thus, the quantity of Mn^{2+} that binds on membrane surface is crucial. Two sample preparation methods were tested:

Method 1. According to literature, Mn^{2+} ions all binds to the lipid membrane surface if the mole percentage of Mn^{2+} is smaller than 50%.[4] To prepare lipid vesicle with 20 mole % Mn^{2+}, lipid film containing 20 μmol of DPPC (D8 or D10 deuterated) and 5 μmol DPPG was hydrated with 2 mL 10 mM HEPES and 5mM MES buffer at pH 5.0 containing 5 μmole of Mn^{2+}, followed by 10 freeze-thaw cycle and extrusion. 1 μmol IFP in the same buffer was added to vesicles, then ultracentrifugation at 100000g for 5 hours. Presumably all Mn^{2+} ions should bind to the membrane surface and no Mn^{2+} left in supernatant. The pellet was lyophilized, packed into NMR rotor and rehydrate with 5 μL pH 5 buffer.

Method 2. Lipid film containing 20 μmol of DPPC (D8 or D10 deuterated) and 5 μmol DPPG lipid was hydrated with 2 mL 10 mM HEPES and 5mM MES buffer at pH 5.0, followed by 10 freeze-thaw cycle and extrusion. 1 μmol IFP in the same buffer was added to vesicles, then ultracentrifugation at 100000g for 5 hours. The pellet was lyophilized, packed into NMR rotor and rehydrate with 5 μL 10 mM HEPES and 5mM MES buffer that contains 5 μmole of Mn^{2+}. Five freeze-thaw cycles were done to evenly distribute the Mn^{2+} on both side of the membrane.

The key difference in these two methods are when to introduce Mn^{2+} ions. In **Method 1**, Mn^{2+} was added to rehydrate the lipid film, thus, Mn^{2+} binds to membrane surface during

preparation of unilamellar vesicles, while in **Method 2** Mn^{2+} was added at when pack the sample into NMR rotor.

The binding result in **Method 1** was tested. Mn^{2+} in solution can be oxidized by $NaIO_4$, following this equation:

$$2Mn^{2+} + 5IO_4^- + 3H_2O \rightarrow 2MnO_4^- + 5IO_3^- + 6H^+$$

The product permanganate ion has an intense purple color and can be determined by A525. The supernatant after ultracentrifugation was mixed with access amount of $NaIO_4$ and boiled, then the absorbance at 525 nm was measured. By comparing the Mn^{2+} quantity in supernatant to the control experiment (1.25 μmole Mn^{2+} dissolved in the same volume as the supernatant), the mole of Mn^{2+} that binds to membrane can be calculated. Five conditions for DPPC-D8 or DPPC-D10, and with or without IFP were tested. The results showed Mn^{2+} mole percentage to lipid molecules are respectively 16%, 15%, 9%, 12% and 3%, not as expected 20%. The uncertainty of Mn^{2+} binding causes unpredictable change in T_2 relaxation time measurements.

Mn^{2+} was added at the last step in **Method 2**, thus the Mn^{2+} quantity is consistent from sample to sample. In **Method 2**, IFP solution was added to unilamellar vesicles with 100 nm diameter. IFP bond to vesicles and due to the fusion activity of IFP, the vesicles may have various diameter. Five cycles of freeze-thaw were applied to reach Mn^{2+} distributions equilibrium on both side of the membrane.[10, 11] It was also reported that repeating freeze-thaw cycles can form unilamellar vesicles with diameter < 200 nm.[12] Thus, the freeze-thaw effect on membrane structure should be neglectable.

5.2.2 ^{2}H FID, spectra and T_2 relaxation time of pure lipid membrane

To obtain molecular level view of how IFP modulates lipid organization, the pure lipid structure and T_2 relaxation rate was studied. Solid-state ^{2}H NMR FID and spectra of DPPC-D10:DPPG or DPPC-D8:DPPG (D10 and D8 as abbreviation respectively) with 4:1 ratio at 25 or 50 °C were obtained and analyzed. Figure 5.3 – Figure 5.6 show the FID, stacked FID and stacked spectra with synchronize increasing τ_1 and τ_2 values. As shown in these figures, both the FID intensity and spectra integration decrease as τ_1 and τ_2 increase, and the decrease in signal follows exponential decay.

The ^{2}H NMR spectra D10 or D8 at 25 or 50 °C are shown in Figure 5.7. The peak splitting analysis is in Table 5.1. The ^{2}H spectra at 50 °C in Figure 5.7 showed the sharp cut off signal indicate that the sample is in the liquid phase, which is consistent with the 41 °C theoretical phase transition temperature of pure DPPC and DPPG lipids. The big differences in spectra line shape at 25 and 50 °C also illustrate a membrane phase change.

The spectrum of D10 at 50 °C had the smallest peak splitting of 2.6 kHz, which is from the terminal methyl group of the lipid acyl chain. The broader 9.9 kHz splitting came from the methylene groups. The 25.9 kHz splitting in D8 was greater than the 9.9 kHz in D10, since when the methylene groups that are further from acyl chain terminal, a larger splitting occurs. At lower temperature, both D10 and D8 are in gel phase and had larger peak splitting compare to liquid phase. The methylene peak -CD_2 is not resolved at 25 °C (Table 5.1).

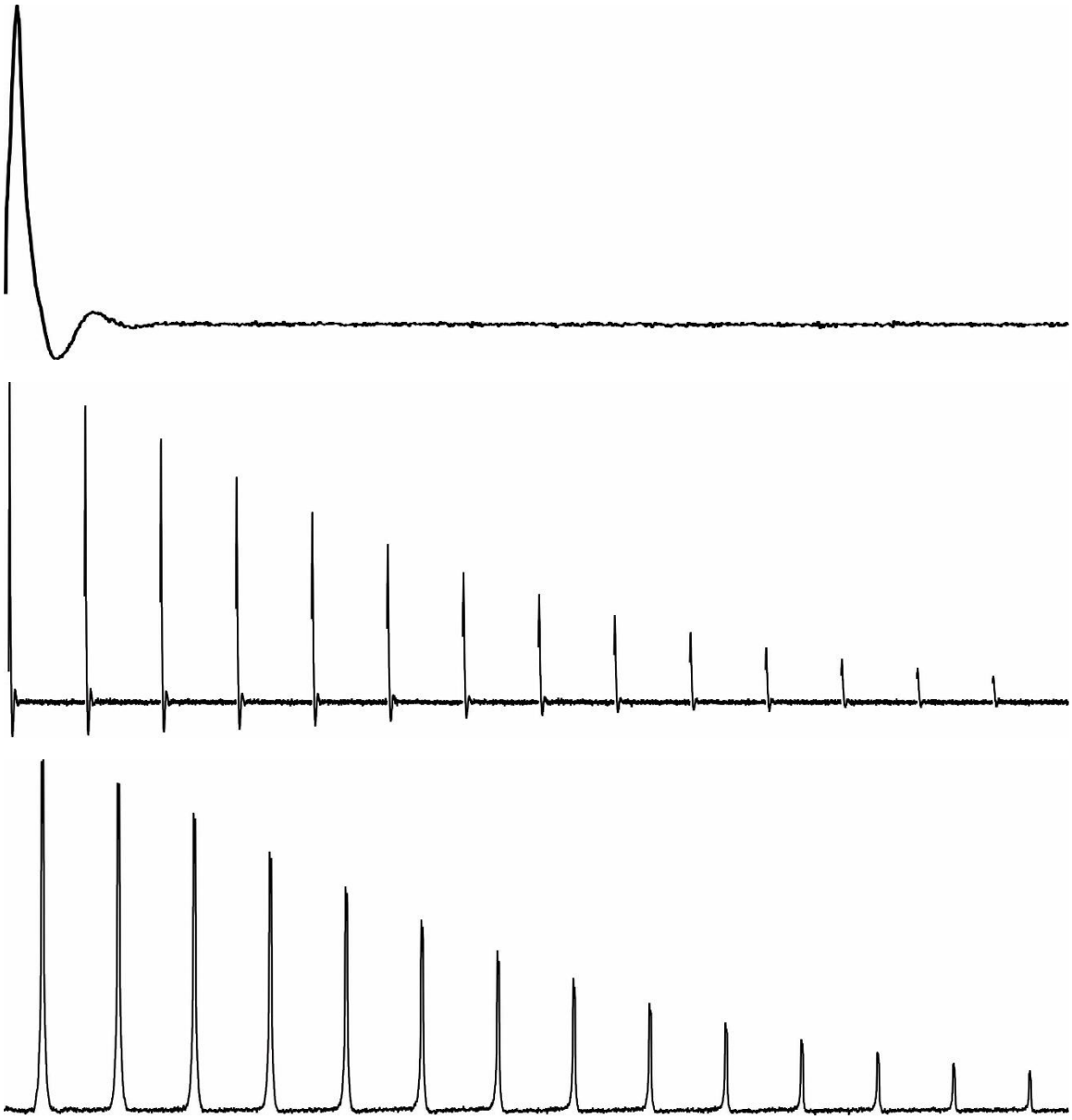

Figure 5.3 D10 at 25 °C. Top: FID. Middle: FID stacked plots with increasing τ_1 and τ_2. Bottom: spectra stacked plots. All spectra were acquired with 5000 scans, processed with 300 Hz exponential line broadening, data shift = -11, and baseline correction.

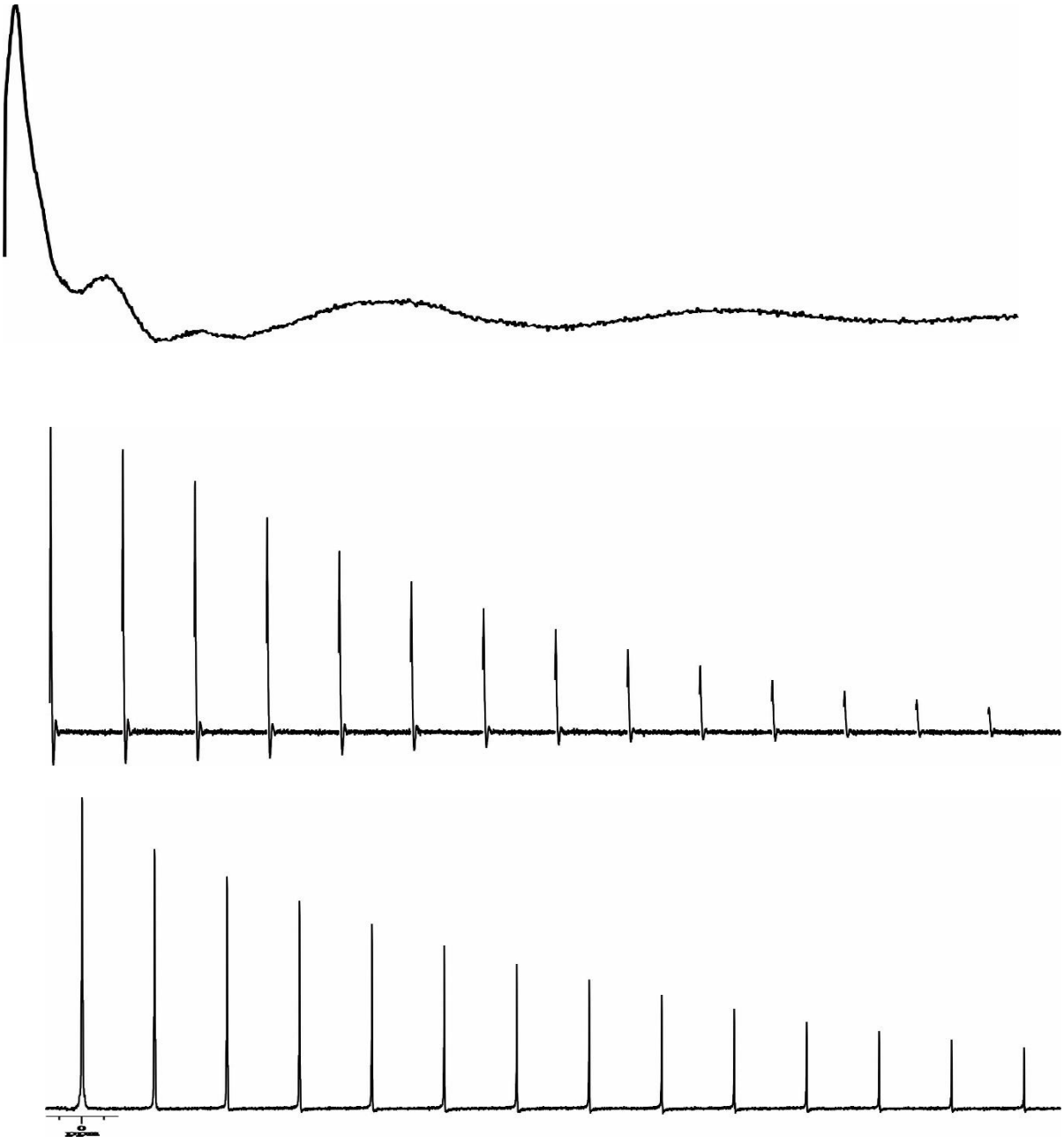

Figure 5.4 D10 at 50 °C. Top: FID. Middle: FID stacked plots with increasing τ_1 and τ_2. Bottom: spectra stacked plots. All spectra were acquired with 5000 scans, processed with 300 Hz exponential line broadening, data shift = -11, and baseline correction.

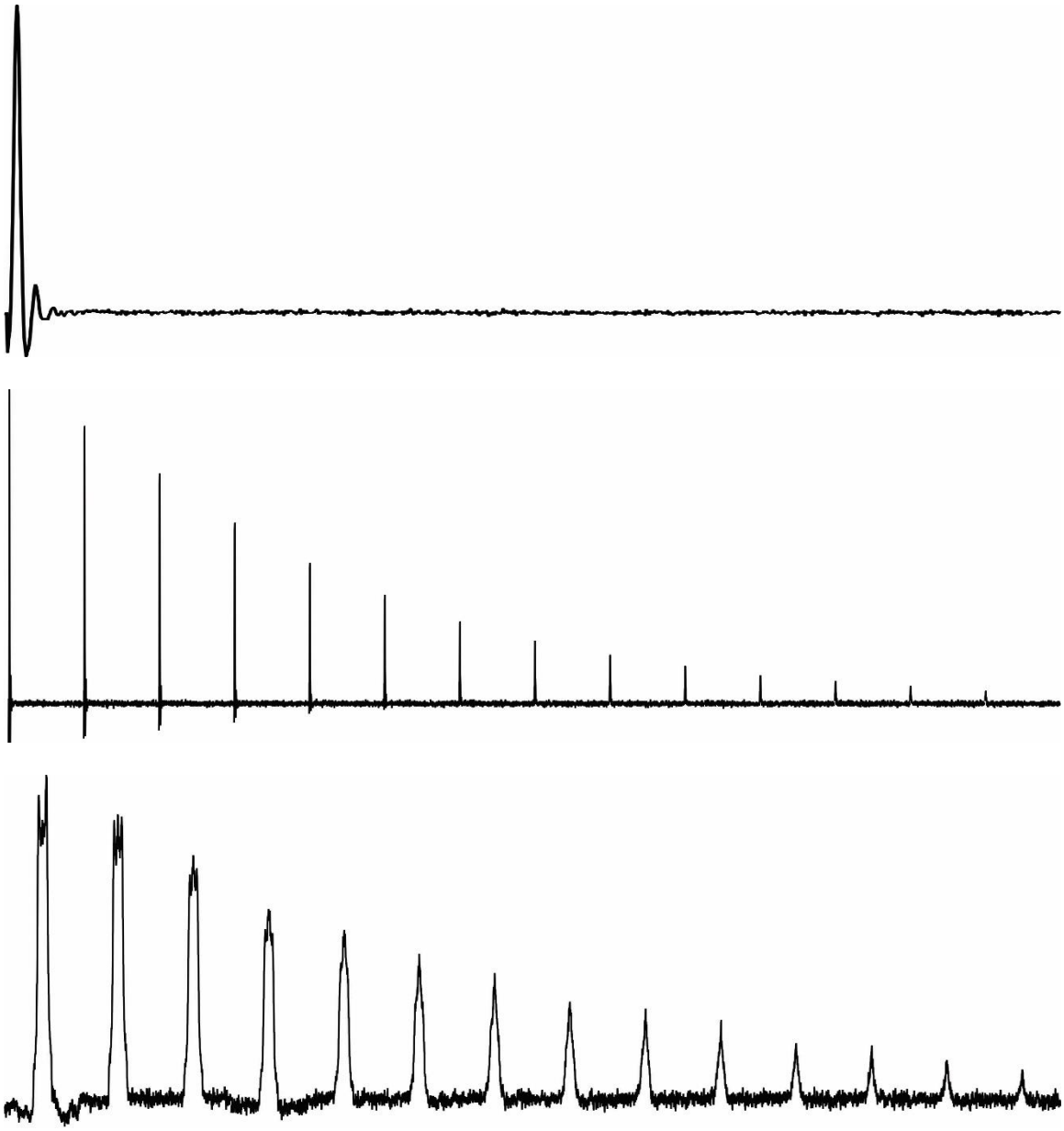

Figure 5.5 D8 at 25 °C. Top: FID. Middle: FID stacked plots with increasing τ_1 and τ_2. Bottom: spectra stacked plots. All spectra were acquired with 5000 scans, processed with 1000 Hz exponential line broadening, data shift = -11, and baseline correction.

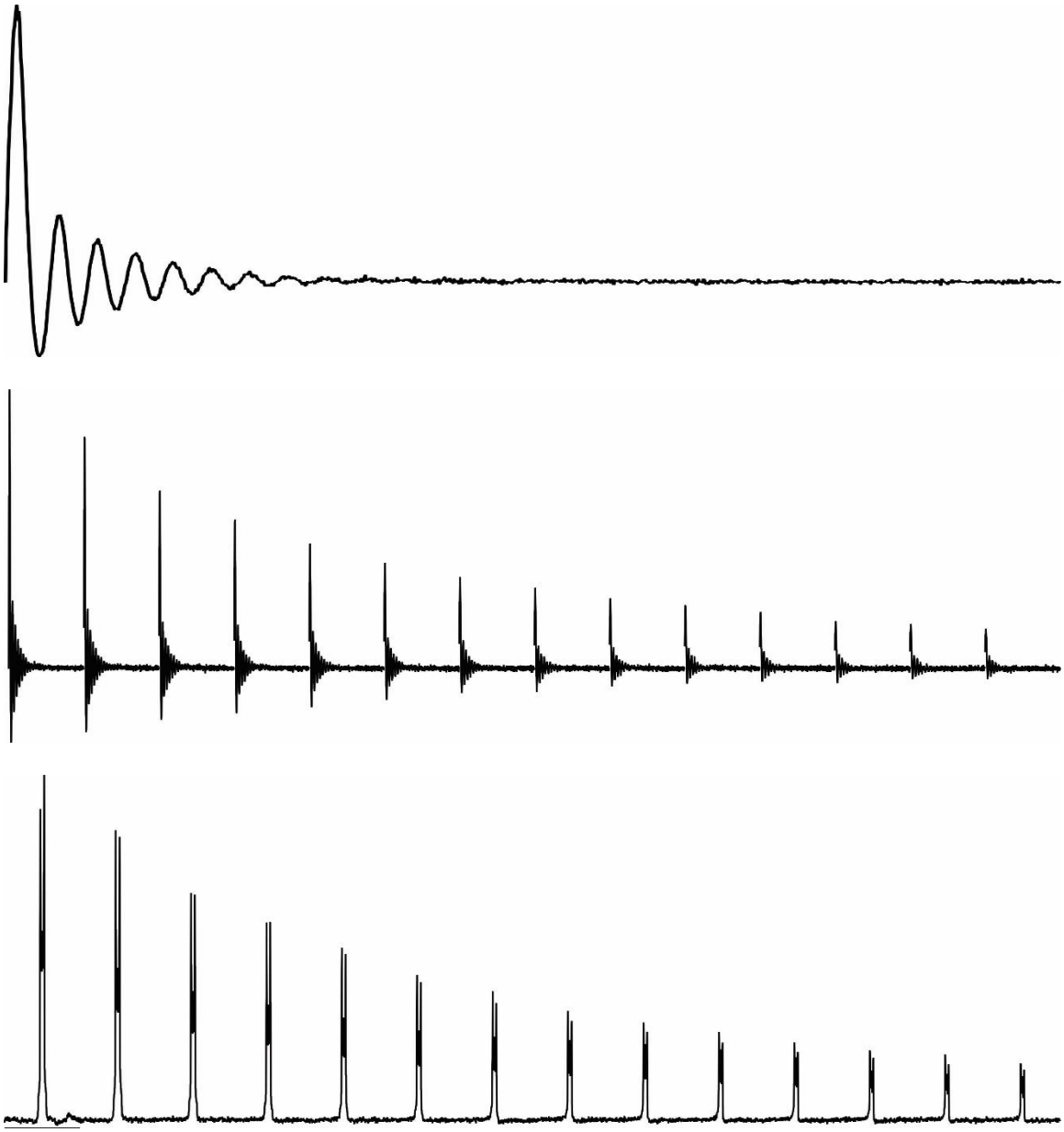

Figure 5.6 D8 at 50 °C. Top: FID. Middle: FID stacked plots with increasing τ_1 and τ_2. Bottom: spectra stacked plots. All spectra were acquired with 5000 scans, processed with 1000 Hz exponential line broadening, data shift = -11, and baseline correction.

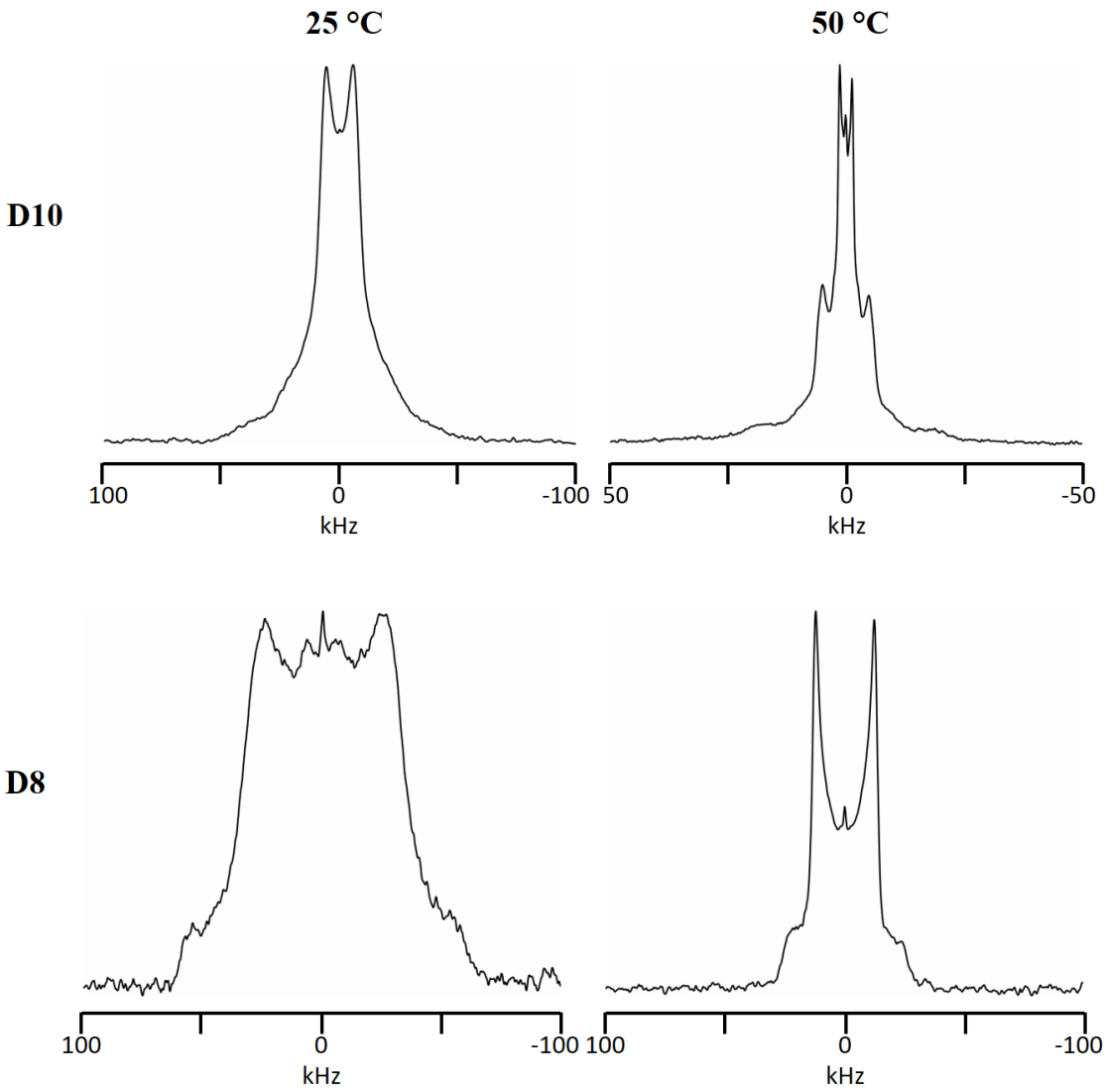

Figure 5.7 ^{2}H NMR spectra D10 or D8 at 25 or 50 °C. Spectra were acquired with 5000 scans, data shift = -11, and baseline correction. D10 spectra were processed with 300 Hz exponential line broadening, and D8 with 1000 Hz exponential line broadening.

Table 5.1 ^{2}H NMR spectra peak splitting (kHz)

	25 °C	50 °C
D10 [a]	11.8	2.6 / 9.9
D8	50.6	25.9

[a] Methylene peak -CD_2 of D10 at 25 °C was not determined because of the lack of the well resolved -CD_2 peaks.

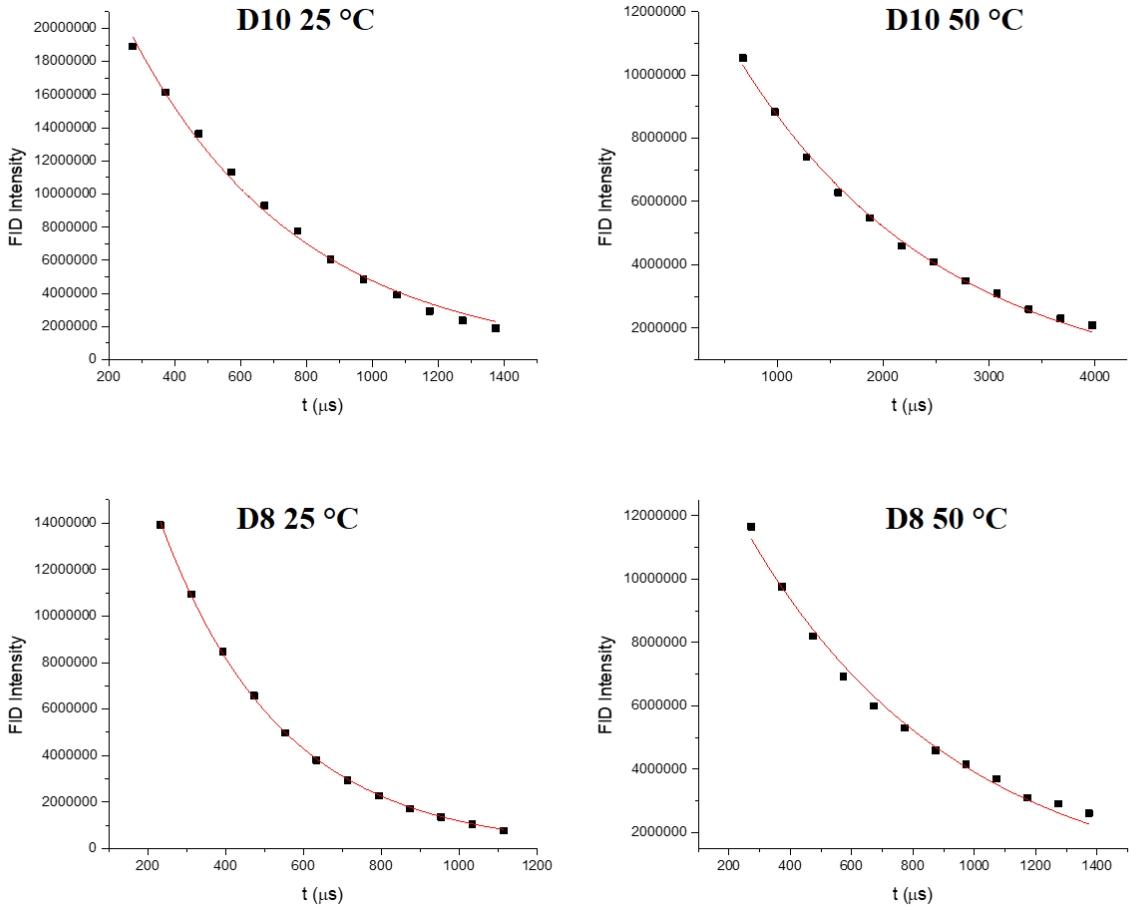

Figure 5.8 Quecho experimental (black squares) and best fit (red lines) plots of D10 and D8 at 25 and 50 °C.

Table 5.2 Best-fit ^2H T_2 (μs) of D10 and D8 lipids measured by quecho experiment [a]

	25 °C	50 °C
D10	516 (14)	1937 (36)
D8	310 (2)	688 (21)

[a] The T_2 relaxation time values are in unit μs. The uncertainties are shown in parenthesis

The T_2 relaxation rate was fitted by equation 5.3.

$$I(t) = I(0) \times exp\left(\frac{t}{T_2}\right) \quad 5.3$$

Where $I(0)$ and T_2 are fitting parameters. $I(t)$ is the FID intensity at t = [τ_1 + τ_2 + time being shifted]; time being shifted ≈ 22 μs in these experiments; $I(0)$ is the FID intensity at t=0; and T_2 is the transverse relaxation time or spin-spin relaxation time.[13]

Figure 5.8 is shown as the FID intensity decay and their best fitting. The T_2 values from exponential fitting are shown in Table 5.2. From 25 to 50 °C, the T_2 relaxation time of D10 increased from 516 to 1937 μs, which is almost four folds longer. As for D8, the T_2 relaxation time increased by two folds, from 310 to 688 μs. Even the increase differs in these two samples, but it can be concluded that the T_2 has a positive as correlation with temperature. Additionally, D8 had a T_2 of 310 μs at 25 °C, which is shorter compare to the 516 μs T_2 of D10 at the same temperature. The trend is also true at 50 °C, T_2 of D8 was 688 μs, shorter than D10's 1937 μs.

5.2.3 The effect of Mn^{2+} concentration on lipid T_2

As discussed in equation 5.1:

$$\frac{1}{T_2} = R_2 = W\frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2\frac{\gamma_D^2\mu_{eff}^2\beta^2}{r^6}\left(4\tau_s + \frac{3\tau_s}{1+\omega_D^2\tau_s^2} + \frac{13\tau_s}{1+\omega_e^2\tau_s^2}\right) \quad 5.1$$

The concentration of Mn^{2+} (W) affects the T_2 and R_2. High concentration of Mn^{2+} may also affect the membrane structure. Therefore, in this section, the concentration effect is discussed. The stacked D10 ^2H spectra with different concentration of Mn^{2+} were shown in Figure 5.9. The left panel shows the spectra at 25 °C and the right panel shows the spectra at 50 °C. From top to bottom are D10 with 20%, 5%, 1%, 0.2% and pure lipids. The peak splitting in each spectrum was shown in Table 5.3.

At 25 °C, the peak splitting of the methyl groups fell in 11.8 – 12 kHz range, not effected by the increasing concentration of Mn^{2+}. By looking at the spectra, the line shapes of various concentration of Mn^{2+} were overall the same. However, at 50 °C, when lipid in fluid phase, the line shape was greatly influenced by Mn^{2+}. The linewidths were narrower with smaller Mn^{2+} concentration and spectrum was better resolved. This is also evidenced by a decreasing peak splitting of both methyl and methylene groups. When Mn^{2+} concentration decrease from 5% to 0%, the peak splitting of methyl group slightly decreased from 2.9 to 2.5 kHz, and 11 to 10 kHz for methylene group. However, with 20% Mn^{2+}, the shape as well as peak width did not follow the trend. The peak was much broader but the peak splitting was smaller. The reason for contradictory result may due to the high concentration of Mn^{2+} on membrane surface changed the membrane structure.

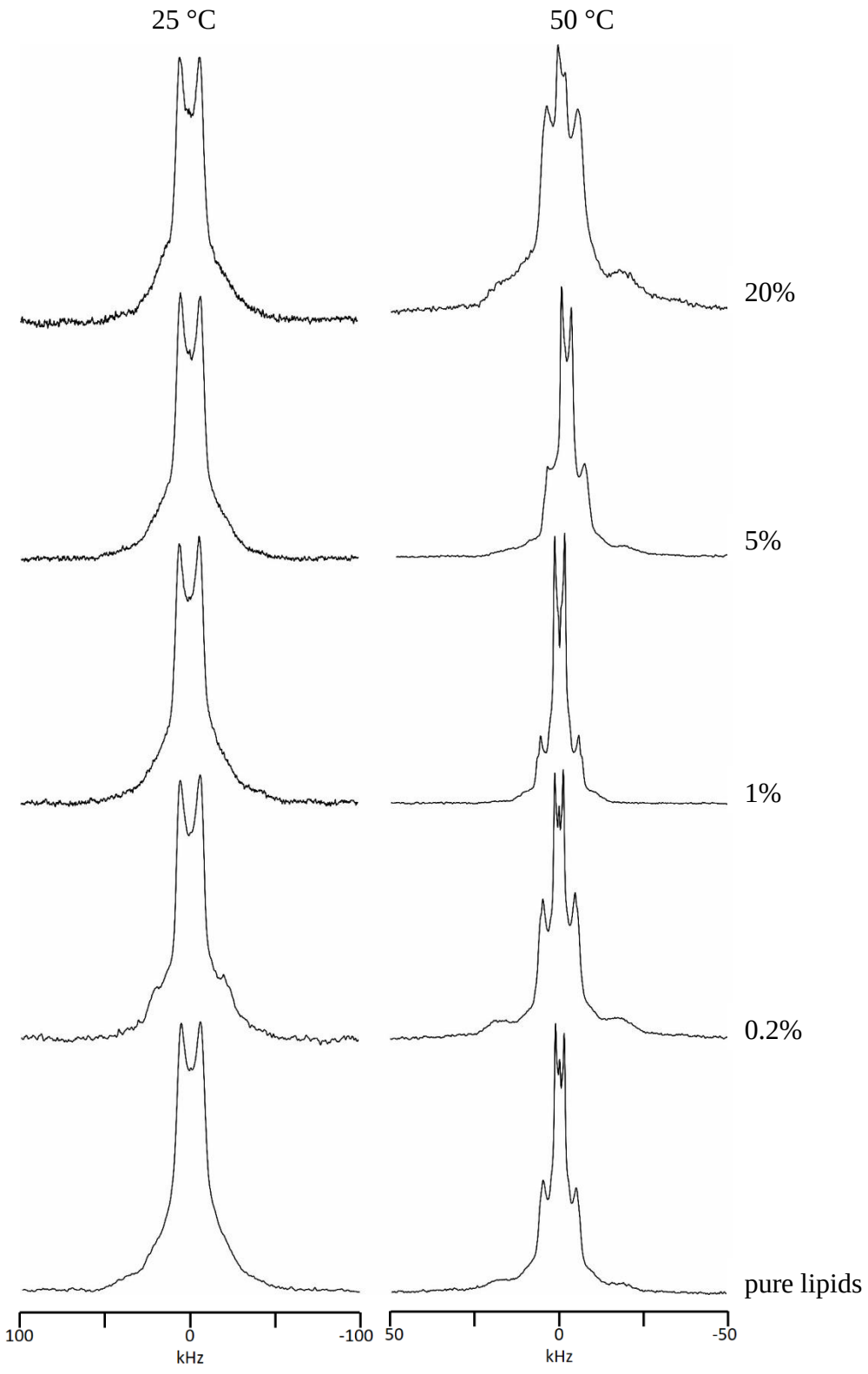

Figure 5.9 ^{2}H NMR spectra of D10 with different mole percentage of Mn^{2+}.

Table 5.3 D10 ^{2}H NMR spectra peak splitting affected by Mn^{2+} concentration [a]

	25 °C		50 °C	
	Methyl -CD_3	Methylene -CD_2 [b]	Methyl -CD_3	Methylene -CD_2
D10_20% Mn^{2+}	11.9	-	2.1	9.2
D10_5% Mn^{2+}	12.0	-	2.8	11.0
D10_1% Mn^{2+}	11.8	-	2.9	11.3
D10_0.2% Mn^{2+}	11.9	39.4	2.5	9.5
D10	11.8	-	2.6	9.8

[a] The numbers represents the peak splitting in unit kHz.

[b] Some methylene peak splitting at 25 °C was not determined because of the lack of the well resolved -CD_2 peaks.

Table 5.4 Best-fit ^{2}H T_2 (μs) of D10 with various Mn^{2+} concentration [a]

	25 °C	50 °C
D10_20%	477 (10)	245 (5)
D10_5%	471 (2)	593 (18)
D10_1%	458 (9)	1035 (13)
D10_0.2%	520 (10)	1018 (9)
D10	516 (14)	1937 (36)

[a] The T_2 relaxation time values are in unit μs. The uncertainties are shown in parenthesis.

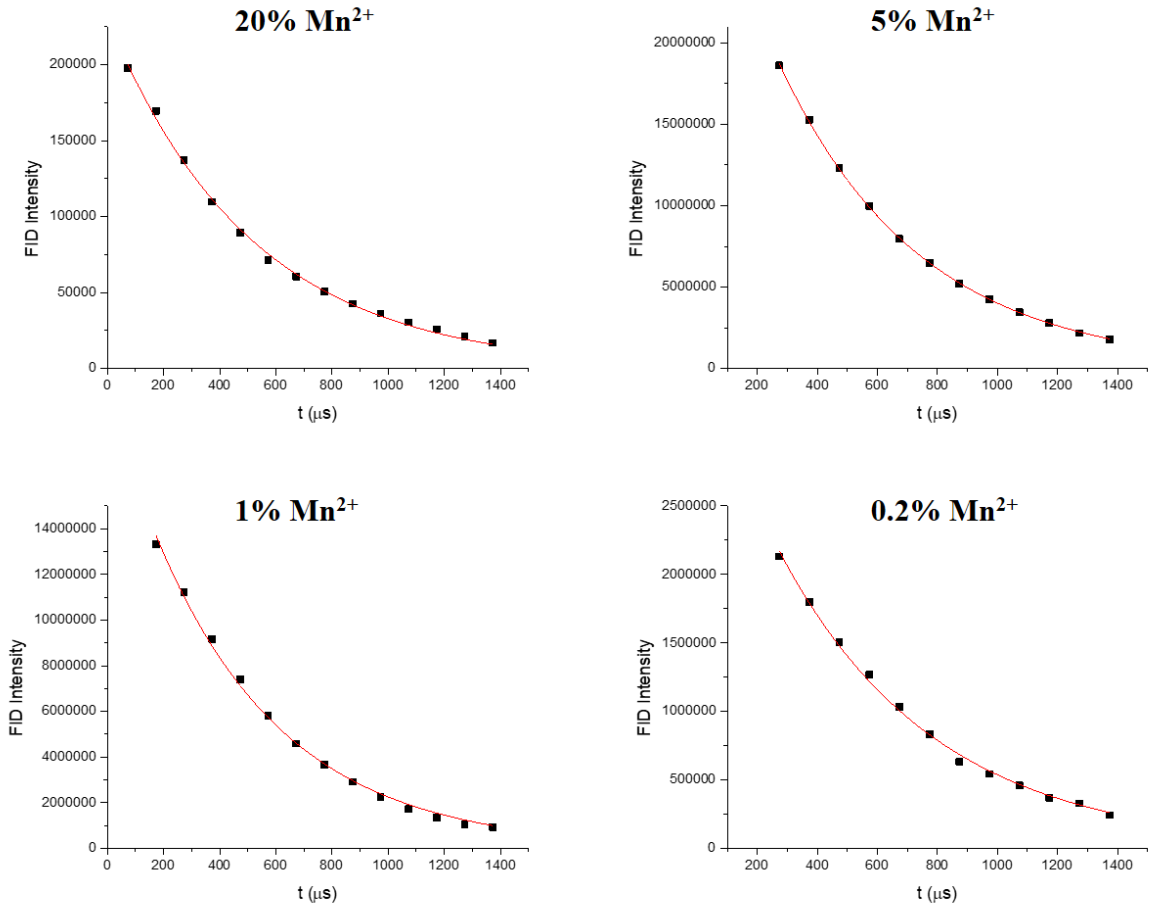

Figure 5.10 Quecho experimental (black squares) and best fit (red lines) plots of D10 with various concentration of Mn^{2+} at 25 °C.

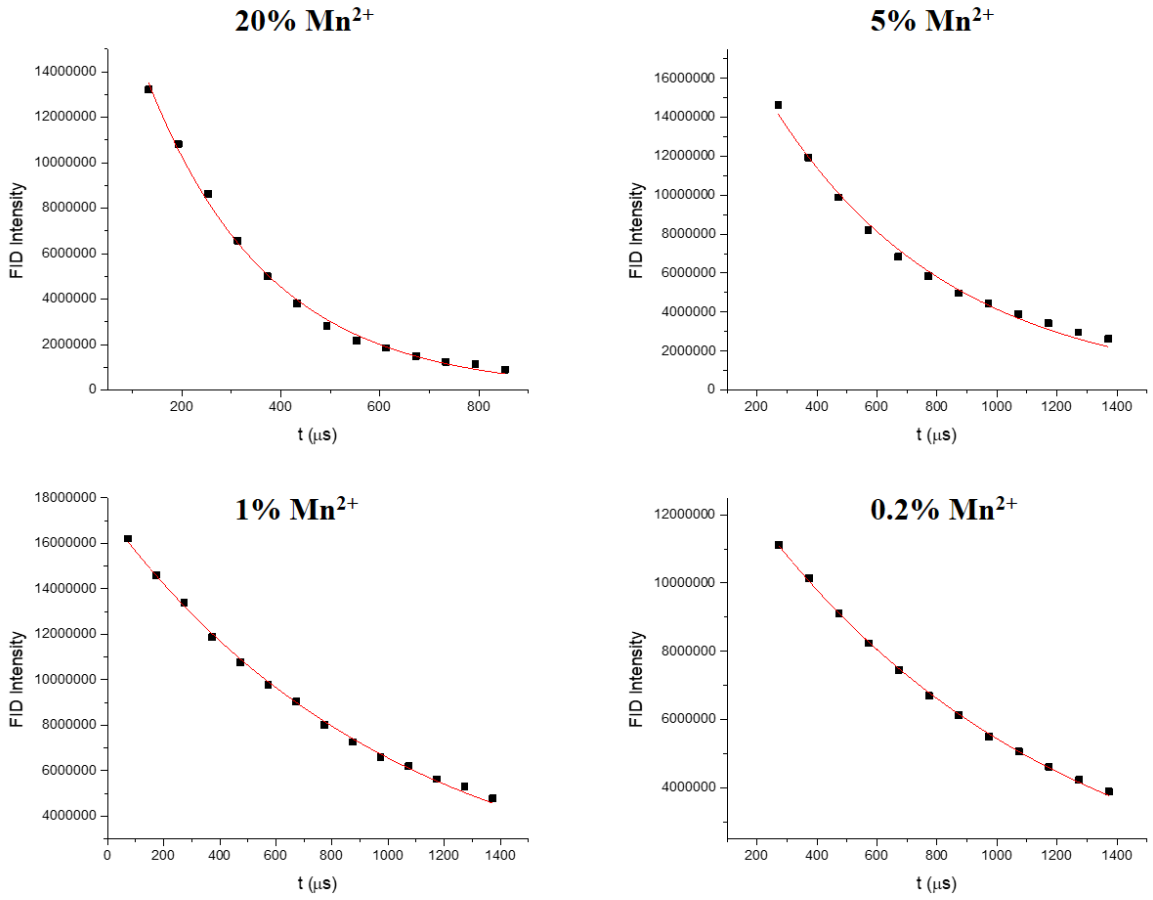

Figure 5.11 Quecho experimental (black squares) and best fit (red lines) plots of D10 with various concentration of Mn^{2+} at 50 °C.

The fitting of T_2 relaxation time of D10 lipid with various Mn^{2+} concentration is shown in Figure 5.10, 5.11 and Table 5.4. With 0 – 20% Mn^{2+}, the T_2 relaxation time at gel phase was in 470 – 520 μs range, which is consistent with the line shape and peak splitting similarities at 25 °C. At fluid phase, D10 with 20% Mn^{2+} had the shortest T_2 = 245 μs. With decrease in Mn^{2+} concentration from 5%, 1%, 0.2% to 0%, the T_2 values increased to 593, 1035, 1018 and 1937 μs respectively. This is due to with higher concentration Mn^{2+}, the average distance between unpaired electron in Mn^{2+} and ^2H is shorter.

IFP bounded membrane samples were prepared with 20% and 5% Mn^{2+} and will be shown in the following sections. However, from the previous data, the big differences in line shape and peak splitting of D10 with 20% Mn^{2+} illustrate that high concentration of Mn^{2+} may affect the overall membrane structure. Therefore, the R_2 difference between IFP bounded membrane with and without 20% Mn^{2+} may not only due to the addition of IFP, since Mn^{2+} also changes the membrane. The 1% and 0.2% Mn^{2+} D10 samples were prepared most recently and the IFP bounded membrane samples correspond to these two Mn^{2+} concentration were not prepared.

5.2.4 FID, spectra and T_2 of IFP-bounded membrane with 5% Mn^{2+}

Figure 5.12 and 5.13 are shown as examples of the FID, stacked FID and stacked spectra of IFP-bounded membrane at gel of liquid phase. The spectra of pure D10, D10 with 5% Mn^{2+}, IFP-bounded D10, and IFP-bounded D10 Mn^{2+} (denoted as D10, D10_Mn, IFP_D10, and FIP_D10_Mn) and pure D8, D8 with 5% Mn^{2+}, IFP-bounded D8, and IFP-bounded D8 Mn^{2+} (denoted as D8, D8_Mn, IFP_D8, and FIP_D8_Mn) at 25 or 50 °C are shown in Figure 5.14 and 5.15. The lipid composition were DPPC:DPPG with 4:1 mol ratio. All the samples were prepared using buffer containing 10 mM HEPES and 5mM MES at pH 5.0. The IFP-bounded lipid samples

contained 20 μmol DPPC-D10 or DPPC-D8, 5 μmol DPPG and 1 μmol IFP and were prepared by aqueous binding method (As discussed in Chapter 2). The Mn^{2+} containing samples are prepared separately from Mn^{2+} free samples. The sample preparation method is the same for samples with or without Mn^{2+}. The difference is the rehydration step. For example, D10 is rehydrated with pH 5 buffer and D10_Mn is rehydrated with pH 5 buffer and additional 5% mol Mn^{2+}.

Figure 5.12 and 5.13 show the IFP_D10 sample at 25 and 50 °C. The top panels in these two figures are the FID at gel and fluid phase, which showed high similarity to the FID of pure D10 at these temperatures (Figure 5.3 and 5.4). The middle and bottom panels show the stacked FID intensity as well as stack spectra, and they both decrease as τ_1 and τ_2 increase, and the decrease in signal follows exponential decay. The FID, stacked FID and stacked spectra of D10_Mn, IFP_D10_Mn, D8_Mn, IFP_D8, and IFP_D8_Mn all showed similar decay in signal (not shown in this dissertation).

The spectra of all samples at both 25 and 50 °C are shown in Figure 5.14 and 5.15. The peak splitting and fitted T_2 are shown in Table 5.5 – 5.8. For D10 at gel phase and for D8 at liquid phase, adding IFP or Mn^{2+} or both does not change the line shape in the spectra. However, for D10 at liquid phase, adding both IFP and Mn^{2+} lead to a less resolved spectrum and the methylene peak splitting cannot be obtained. As for D8 at 25 °C, addition of IFP or Mn^{2+} or both result in narrower lines. Besides, with same quantity of lipid, same number of scans and same date processing, IFP and Mn^{2+} cause the D8 spectra signal to noise decrease. Even there were slightly differences in these ^2H NMR spectra, we still can conclude that addition of IFP or Mn^{2+} to DPPC+DPPG lipid does not change the lamellar membrane phase.

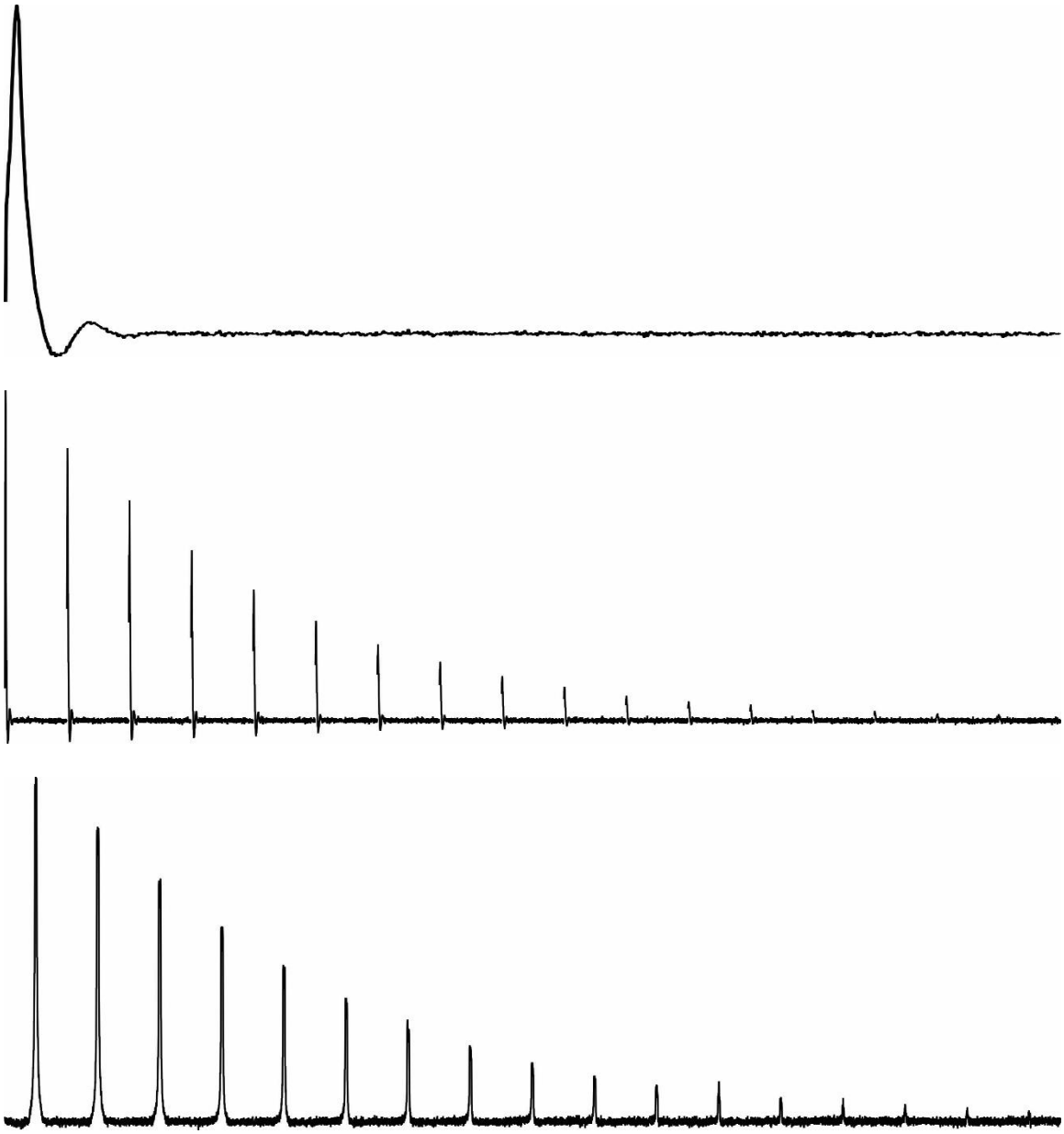

Figure 5.12 IFP_D10 at 25 °C. Top: FID. Middle: FID stacked plots with increasing τ_1 and τ_2. Bottom: spectra stacked plots. All spectra were acquired with 5000 scans, processed with 300 Hz exponential line broadening, data shift = -11, and baseline correction.

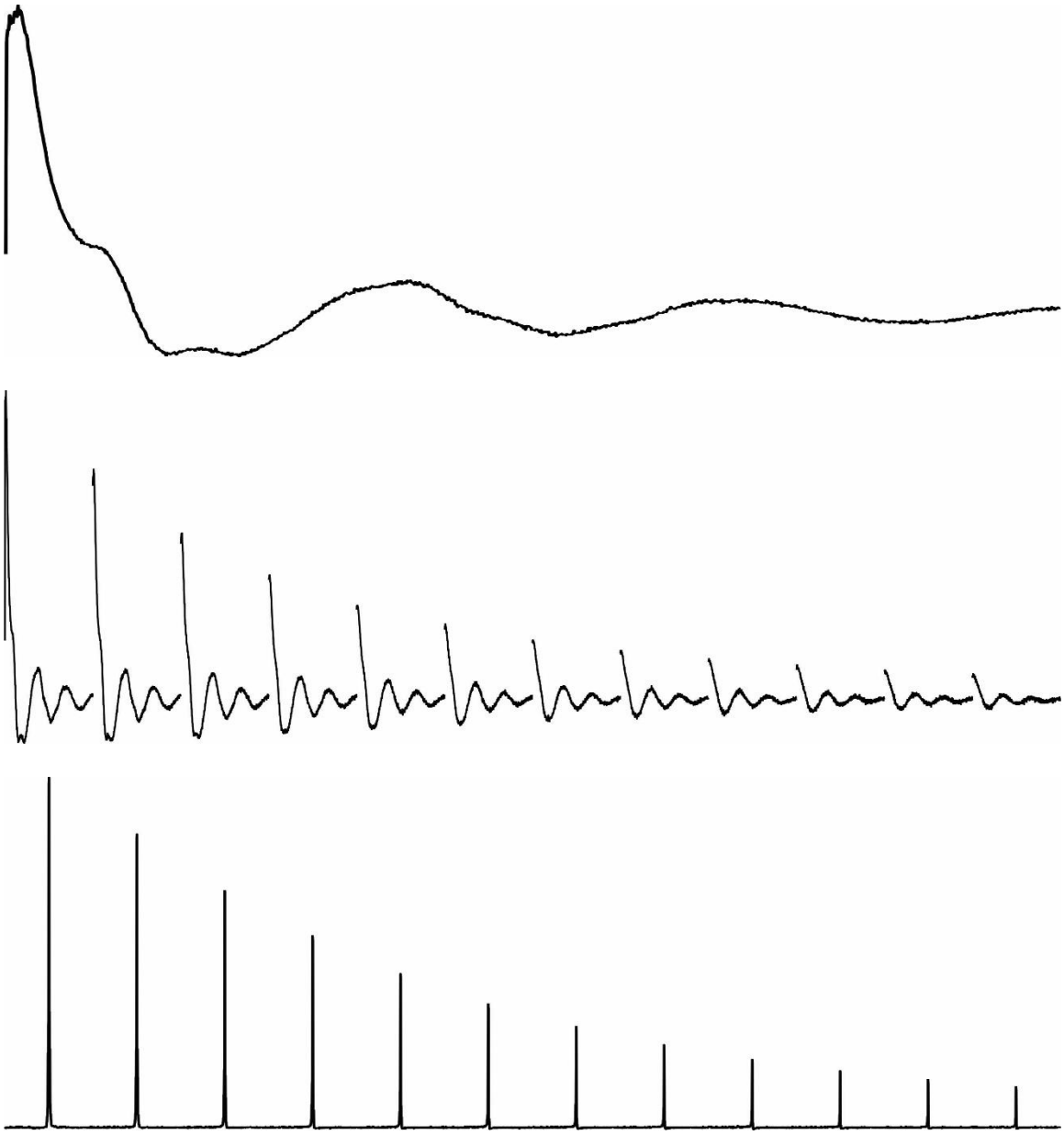

Figure 5.13 IFP_D10 at 50 °C. Top: FID. Middle: FID stacked plots with increasing τ_1 and τ_2. Bottom: spectra stacked plots. All spectra were acquired with 5000 scans, processed with 300 Hz exponential line broadening, data shift = -11, and baseline correction.

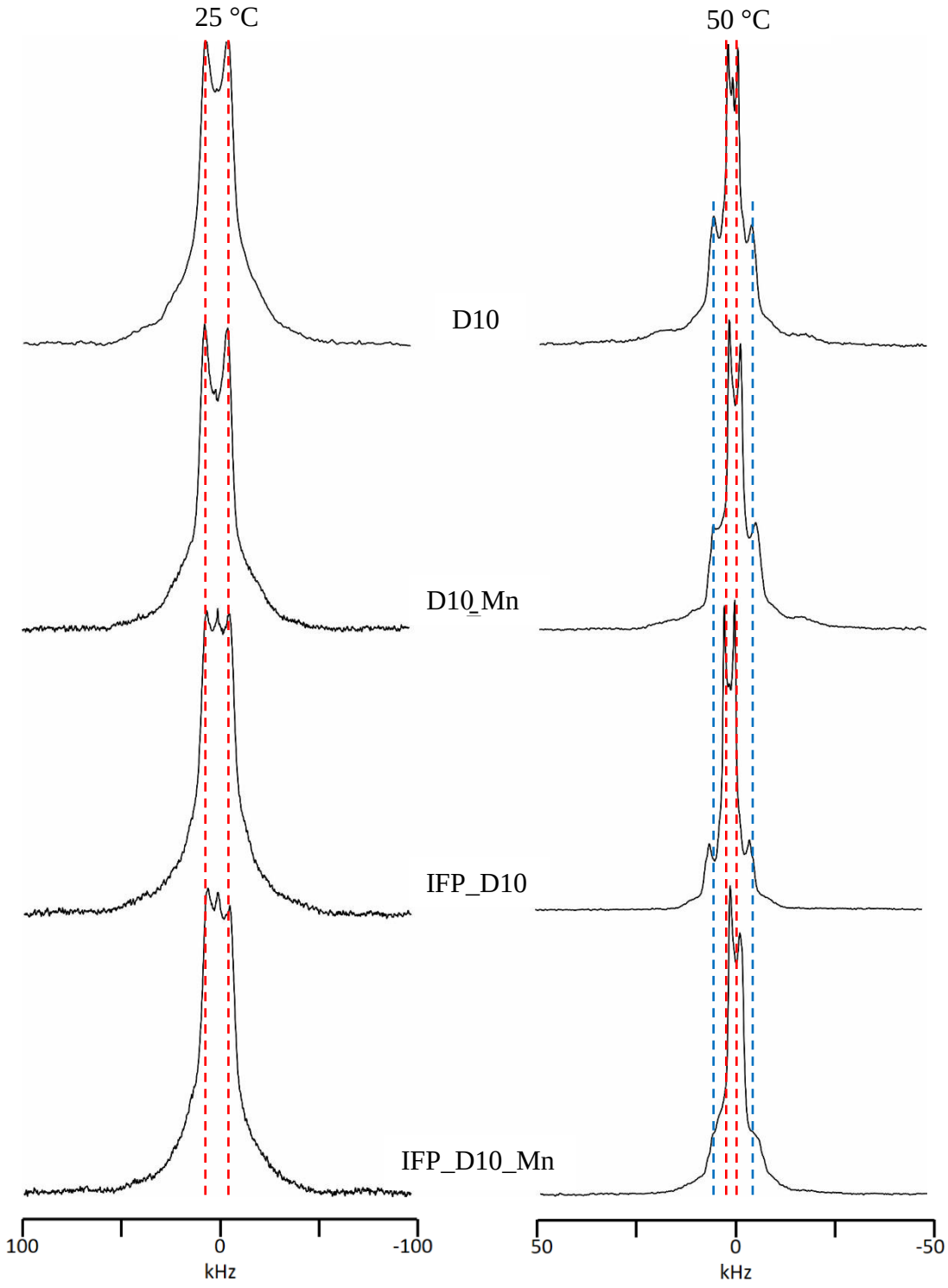

Figure 5.14 ^{2}H NMR spectra of D10 samples at 25 and 50 °C with 5% Mn^{2+}.

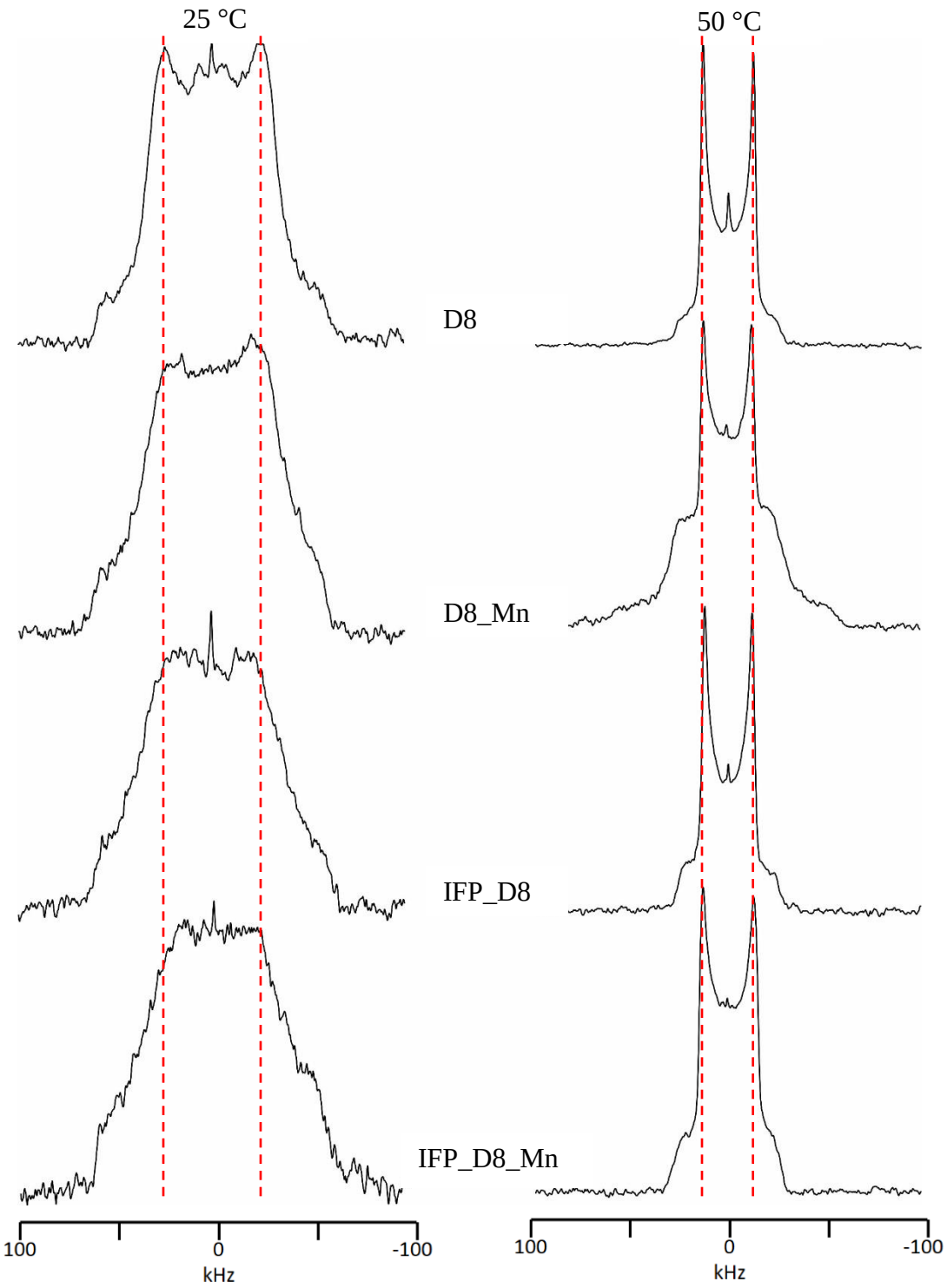

Figure 5.15 ^{2}H NMR spectra of D8 samples at 25 and 50 °C with 5% Mn^{2+}.

Table 5.5 IFP-bounded D10 membrane ^{2}H NMR spectra peak splitting with 5% Mn^{2+} [a]

	25 °C	50 °C	
	Methyl -CD_3	Methyl -CD_3	Methylene -CD_2
D10	11.8	2.6	9.8
D10_Mn	12.0	2.8	11.0
IFP_D10	11.1	2.7	10.4
IFP_D10_Mn [b]	10.8	2.6	-

[a] The numbers represents the peak splitting in unit kHz.

[b] The methylene peak splitting was not determined because of the lack of the well resolved -CD_2 peaks.

Table 5.6 IFP-bounded D8 membrane ^{2}H NMR spectra peak splitting with 5% Mn^{2+} [a]

	25 °C	50 °C
	Methylene -CD_2	Methylene -CD_2
D8	48.1	26.1
D8_Mn	36.3	24.6
IFP_D8	40.4	24.2
IFP_D8_Mn	35.3	25.5

[a] The numbers represents the peak splitting in unit kHz.

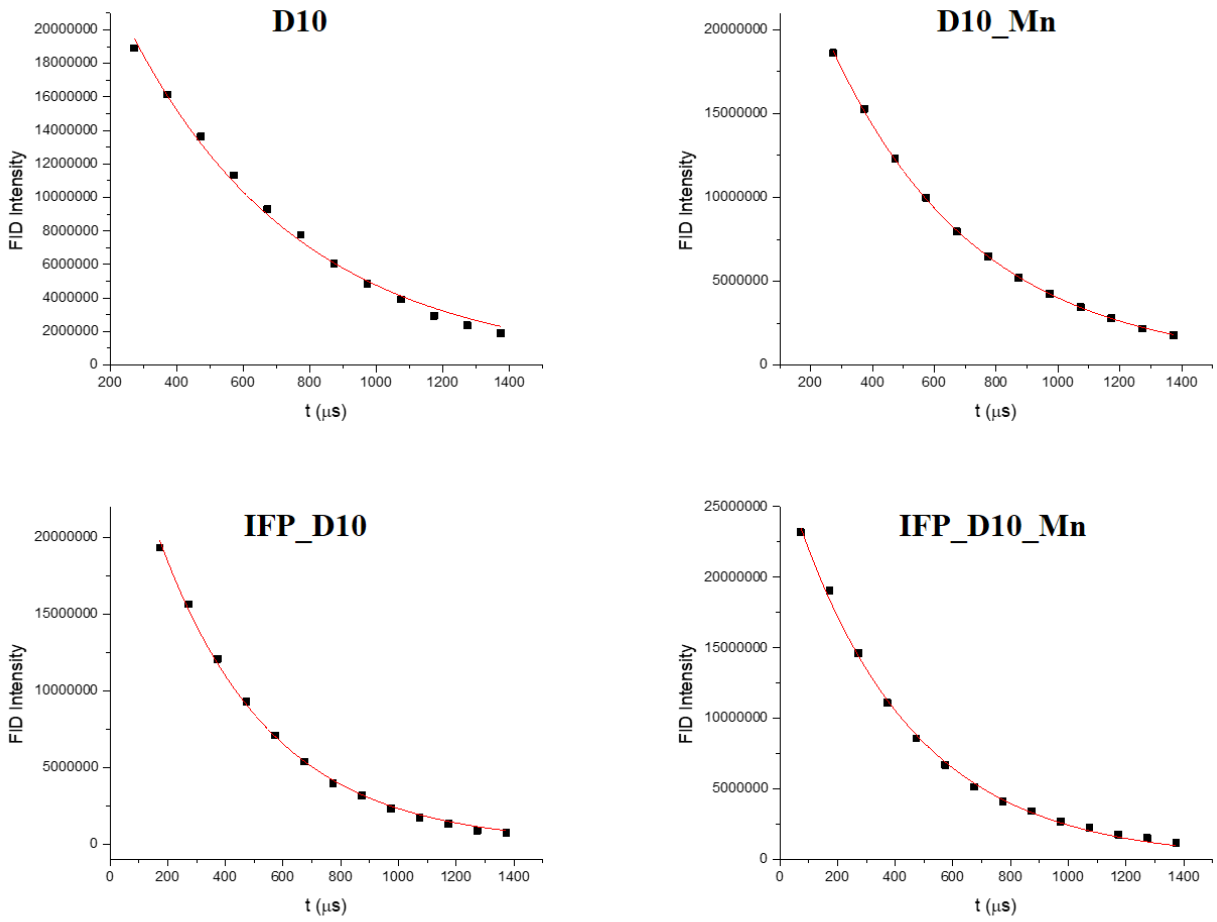

Figure 5.16 Quecho experimental (black squares) and best fit (red lines) plots of D10, D10_Mn, IFP_D10 and IFP_D10_Mn with 5% Mn^{2+} at 25 °C.

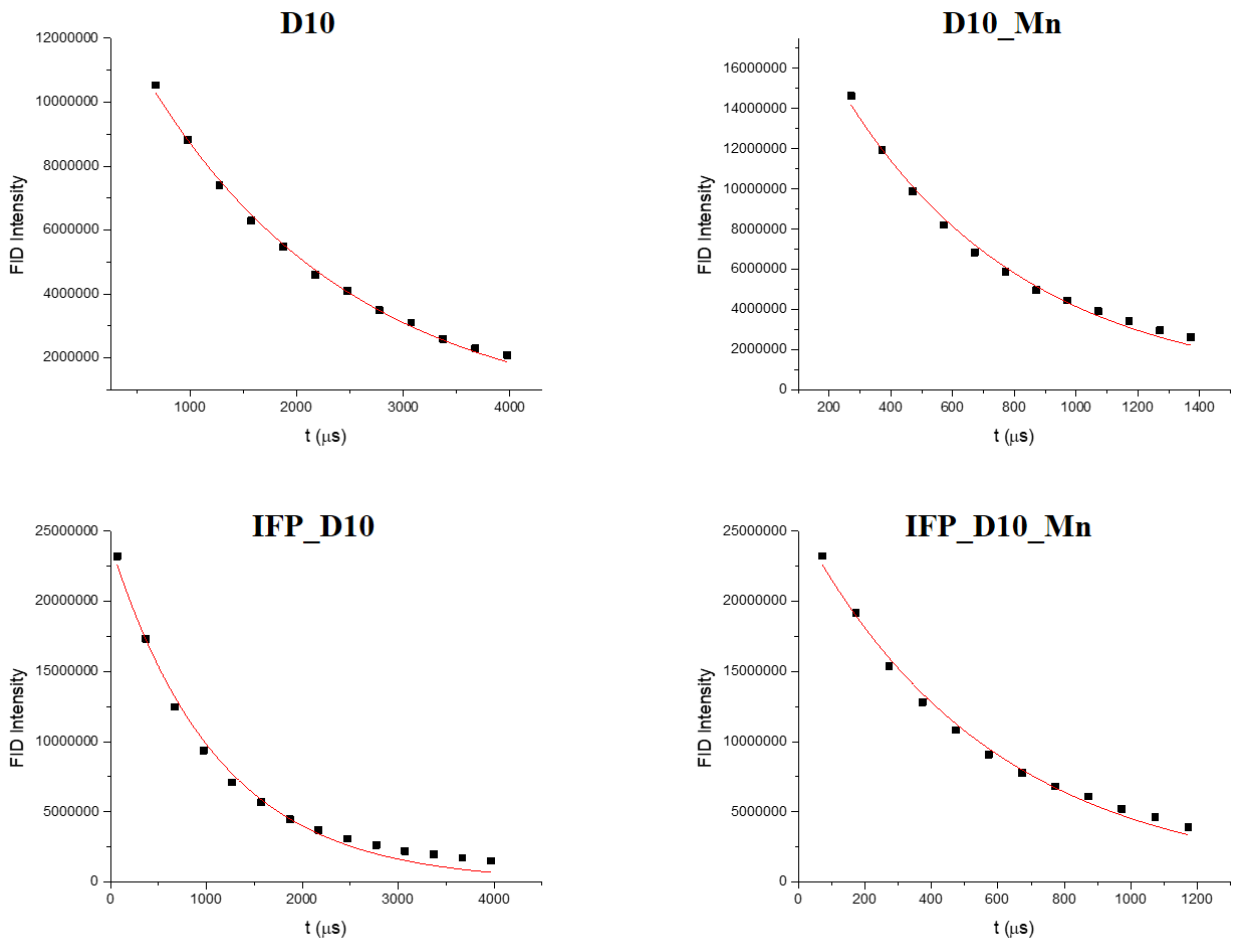

Figure 5.17 Quecho experimental (black squares) and best fit (red lines) plots of D10, D10_Mn, IFP_D10 and IFP_D10_Mn with 5% Mn^{2+} at 50 °C.

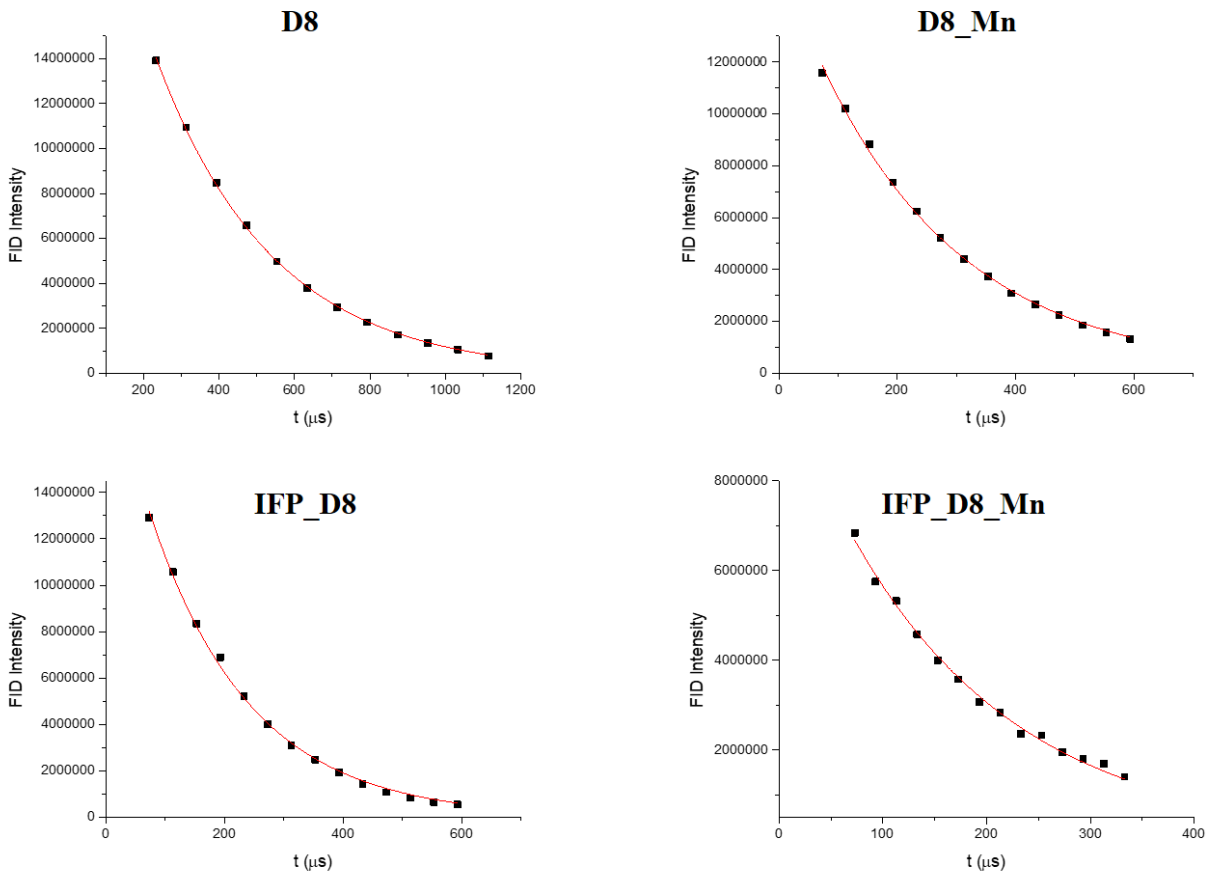

Figure 5.18 Quecho experimental (black squares) and best fit (red lines) plots of D8, D8_Mn, IFP_D8 and IFP_D8_Mn with 5% Mn^{2+} at 25 °C.

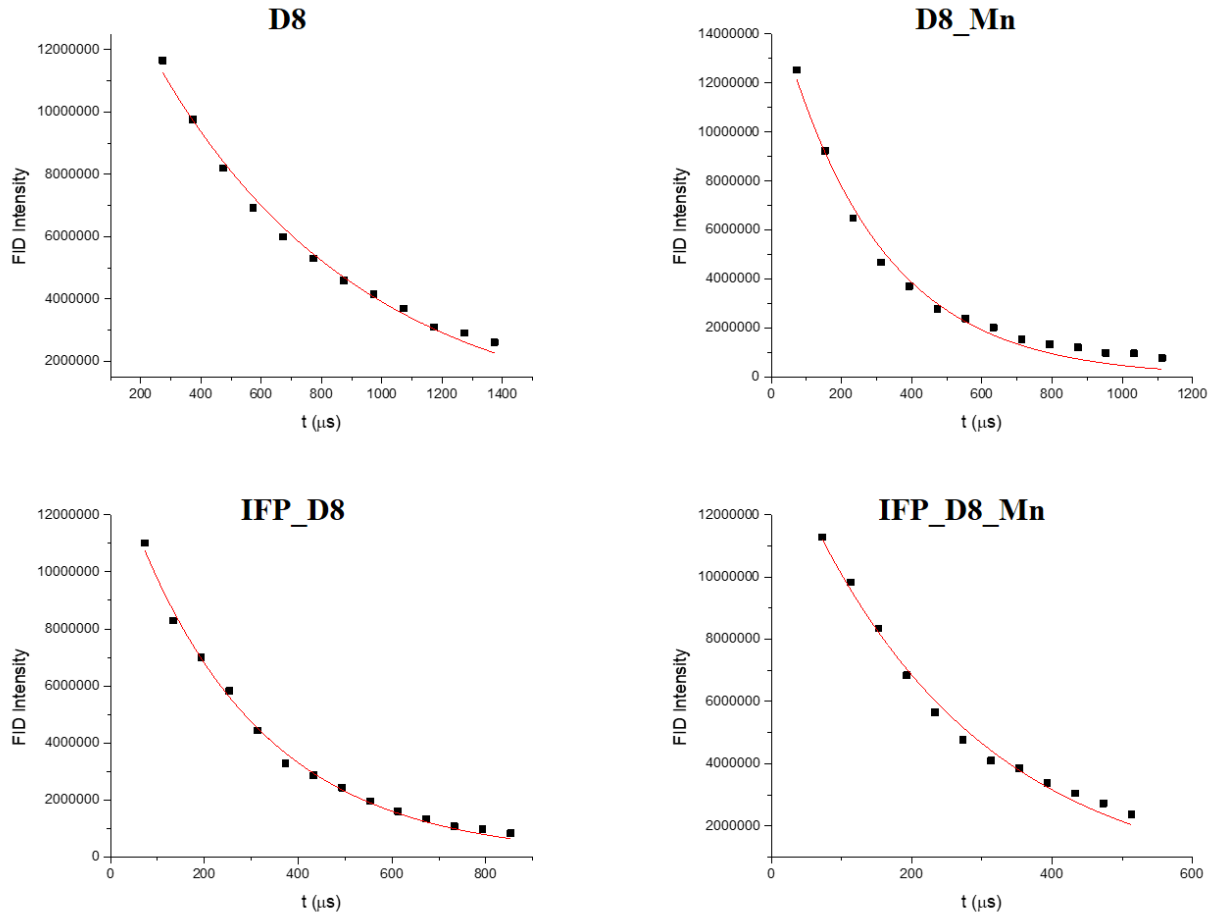

Figure 5.19 Quecho experimental (black squares) and best fit (red lines) plots of D8, D8_Mn, IFP_D8 and IFP_D8_Mn with 5% Mn^{2+} at 50 °C.

Table 5.7 Best-fit ^2H T_2 (μs) of IFP-bounded membrane with 5% Mn^{2+}

	25 °C	50 °C
D10	516 (14)	1937 (36)
D10_Mn	471 (2)	593 (18)
IFP_D10	385 (7)	1116 (42)
IFP_D10_Mn	408 (6)	578 (17)
D8	310 (2)	688 (21)
D8_Mn	243 (4)	259 (9)
IFP_D8	169 (3)	278 (6)
IFP_D8_Mn	163 (3)	244 (5)

Table 5.8 Best-fit ^2H T_2 relaxation rate (R_2, kHz) of IFP-bounded membrane with 5% Mn^{2+}

	25 °C	50 °C
D10	1.94 (5)	0.52 (1)
D10_Mn	2.12 (1)	1.69 (5)
IFP_D10	2.60 (5)	0.90 (4)
IFP_D10_Mn	2.45 (4)	1.73 (5)
D8	3.23 (2)	1.45 (5)
D8_Mn	4.12 (7)	3.86 (14)
IFP_D8	5.92 (11)	3.60 (8)
IFP_D8_Mn	6.13 (12)	4.10 (9)

The D10 peak splitting were presented in Table 5.5. At 25 °C, the peak splitting of pure D10 was 11.8 kHz, which is close to the 12.0 kHz peak splitting of D10_Mn. As discussed in section 5.2.3, concentration of Mn^{2+} did not affect the D10 peak splitting at gel phase. However, IFP_D10 had a 11.1 kHz peak splitting, which is ≈ 1 kHz narrower than D10. The IFP_D10_Mn also showed this narrowing. As for 50 °C, there were slightly changings in methyl peak splitting, but Mn^{2+} and IFP increased the methylene peak splitting of D10_Mn and IFP_D10 by 1.2 and 0.6 kHz, respectively. The trend of IFP_D10_Mn was not clear due to the methylene peak was not well resolved.

Mn^{2+} ions and IFP decreased the peak splitting of D8 both 25 and 50 °C (Table 5.6). The addition of Mn^{2+} to D8 membrane decrease the peak splitting from 48.1 to 36.3 kHz at 25 °C, and from 26.1 to 24.6 at 50 °C. The IFP also reduced the splitting by 7.7 and 1.9 kHz at 25 and 50 °C respectively. It can be concluded that IFP changes the order of the lipid acyl chain.

Figure 5.16 – 5.19 and Table 5.7 – 5.8 showed the best-fit plots, T_2 and R_2 values for the eight samples at different temperature. For pure D10 lipid at 25 °C, it had a T_2 relaxation rate (R_2) of 1.94 kHz, which is close to the R_2 of D10_Mn. This suggests that the addition of Mn^{2+} does not affect the D10 ^2H's T_2 relaxation rate. As for IFP bounded membrane, the IFP_D10 and IFP_D10_Mn also showed close R_2 values. Therefore, both IFP and Mn^{2+} does not change the D10 ^2H's membrane location. The D8 had a R_2 of ~ 3.2 kHz and D8_Mn had a R_2 of ~ 4.1 kHz, which shows that the addition of Mn^{2+} leads to a 0.7 kHz R_2 increase. This increase is due to the spin-nuclear coupling between paramagnetic election in Mn^{2+} and the D8 ^2H's. IFP_D8 and IFP_D8_Mn both had a R_2 of ~ 6 kHz. By comparing the D8 data, IFP induced less R_2 increase, which indicates the average distance between IFP bounded D8 ^2H's and Mn^{2+} is longer compare

to IFP free D8. Thus, the results at gel phase showed IFP does not decrease the average distance between D10 and D8 ^{2}H's and aqueous layer and are inconsistent with lipid protrusion model.

At fluid phase, the effect of Mn^{2+} on D8 and D10 also has similar trends. D10 has a R2 of 0.52 kHz and the addition of Mn^{2+} increase the R_2 by 1.2 kHz. Since at liquid phase, the lipid acyl chain is not packed as ordered at gel phase and has more motion, there is a greater fraction of D10 ^{2}H's locate closer the membrane surface compare to gel phase. Thus, the R_2 increases when Mn^{2+} is introduced, while D10 and D10_Mn have similar R_2 value at gel phase. IFP bounded membrane with Mn^{2+} showed a 0.8 kHz increase in R_2 compare to IFP_D10, and this increase is smaller than the 1.2 kHz of D10 and D10_Mn. On the other hand, Mn^{2+} increase the R_2 of pure D8 lipid by 2.4 kHz, from 2.1 to 5.6 kHz, and 0.5 kHz for IFP-bounded D8. The D8 also showed a smaller increase of IFP_D8 than D8. Since the IFP-bounded D10 and D8 did not have a greater R_2 increase with and without Mn^{2+} compared to pure membrane, the IFP does not decrease the average distance between D10 or D8 deuterium atoms and the aqueous layer. Thus, it can be concluded that the IFP does not induce lipid tail protrusion at both gel phase and liquid phase.

As for the reproducibility, the pure D8 membrane at 25 °C T_2 was measured with different τ_1 and τ_2 values, and the fitting was shown in Figure 5.20. The result R_2 were 3.23 and 3.53 kHz, respectively, with a ~ 0.3 kHz difference.

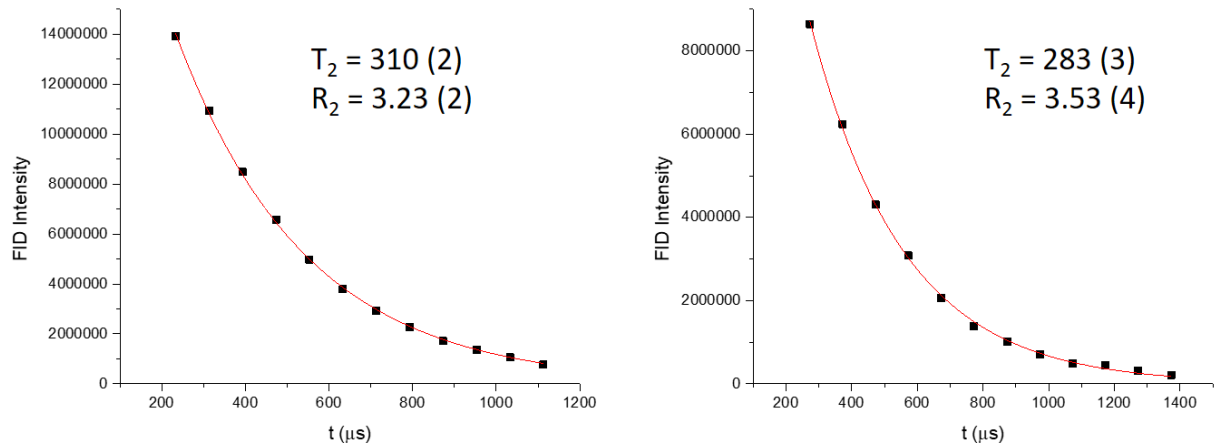

Figure 5.20 Reproducibility of D8 at 25°C. The quecho experimental (black squares) and best fit (red lines) plots as well as the fitted T_2 and R_2 values are shown.

5.2.5 T_2 of IFP-bounded membrane with 20% Mn^{2+}

The spectra of pure D10, D10 with 20% Mn^{2+}, IFP-bounded D10, and IFP-bounded D10 Mn^{2+} (denoted as D10, D10_Mn', IFP_D10, and FIP_D10_Mn') at 25 or 50 °C are shown in Figure 5.21, and the fitting of D10_Mn' and FIP_D10_Mn' at both temperatures are shown in Figure 5.22. The peak splitting, fitted T_2 and R_2 are shown in Table 5.9 – 5.11.

At gel phase, the addition 20% Mn^{2+} did not change the line shape or peak splitting of the spectra, which is consistent with the 5% Mn^{2+} results. However, at liquid phase, the 20% Mn^{2+} greatly affected the spectra with a smaller peak splitting of both methyl and methylene groups. With both 20% Mn^{2+} and IFP, the doublet was even not well resolved.

The D10 at 25 °C showed a R_2 of D10 ≈ R_2 of D10_Mn' and R_2 of IFP_D10 ≈ R_2 of IFP_D10_Mn'. At 50 °C the addition of Mn^{2+} increase the R_2 by ~3.6 kHz in D10 and ~2.6 kHz in IFP bounded D10. This trend is consistent with the 5% Mn^{2+} data and both show no IFP induced lipid tail protrusion at gel phase or liquid phase.

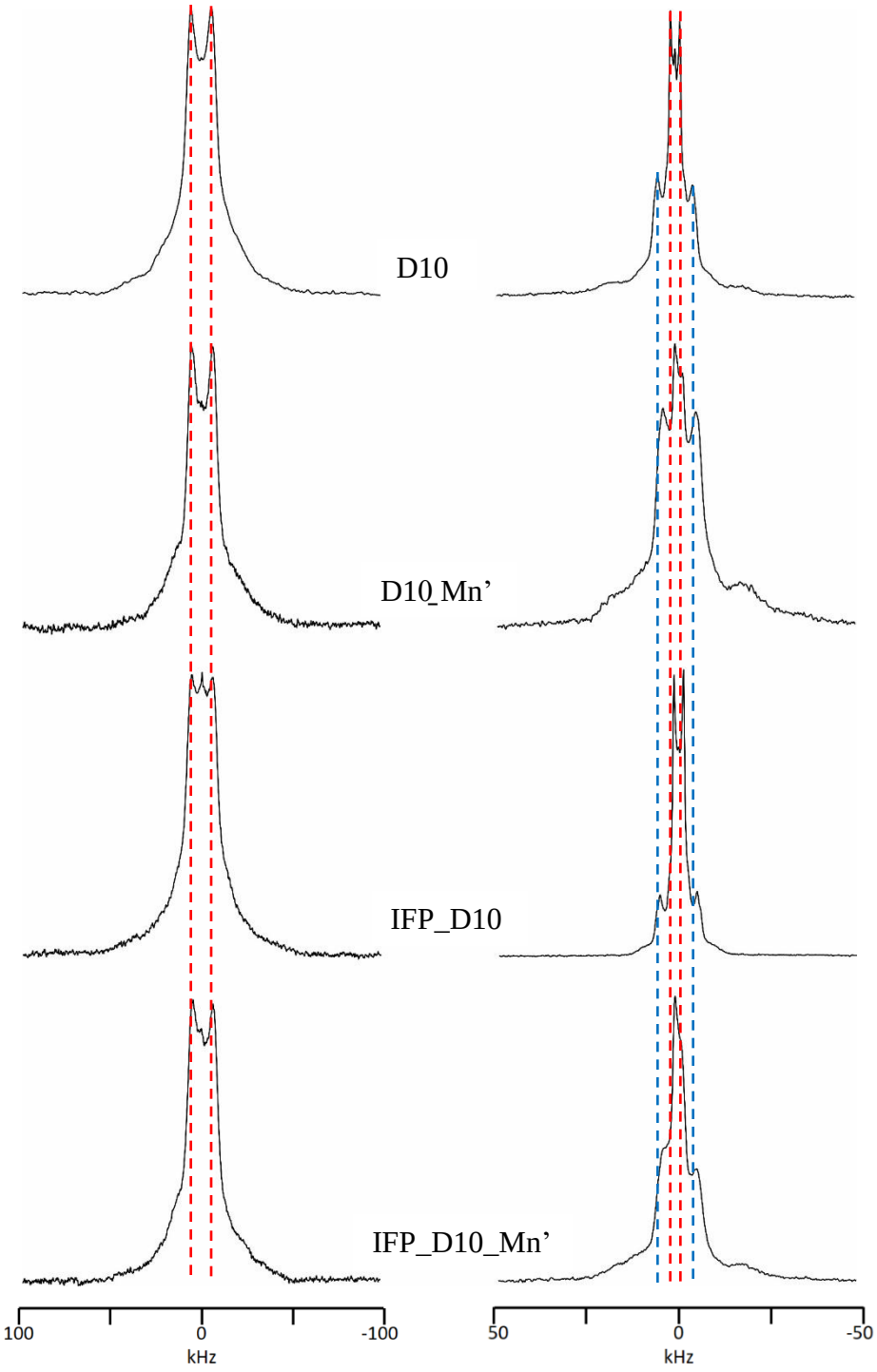

Figure 5.21 ^{2}H NMR spectra of D10 samples at 25 and 50 °C with 20% Mn^{2+}.

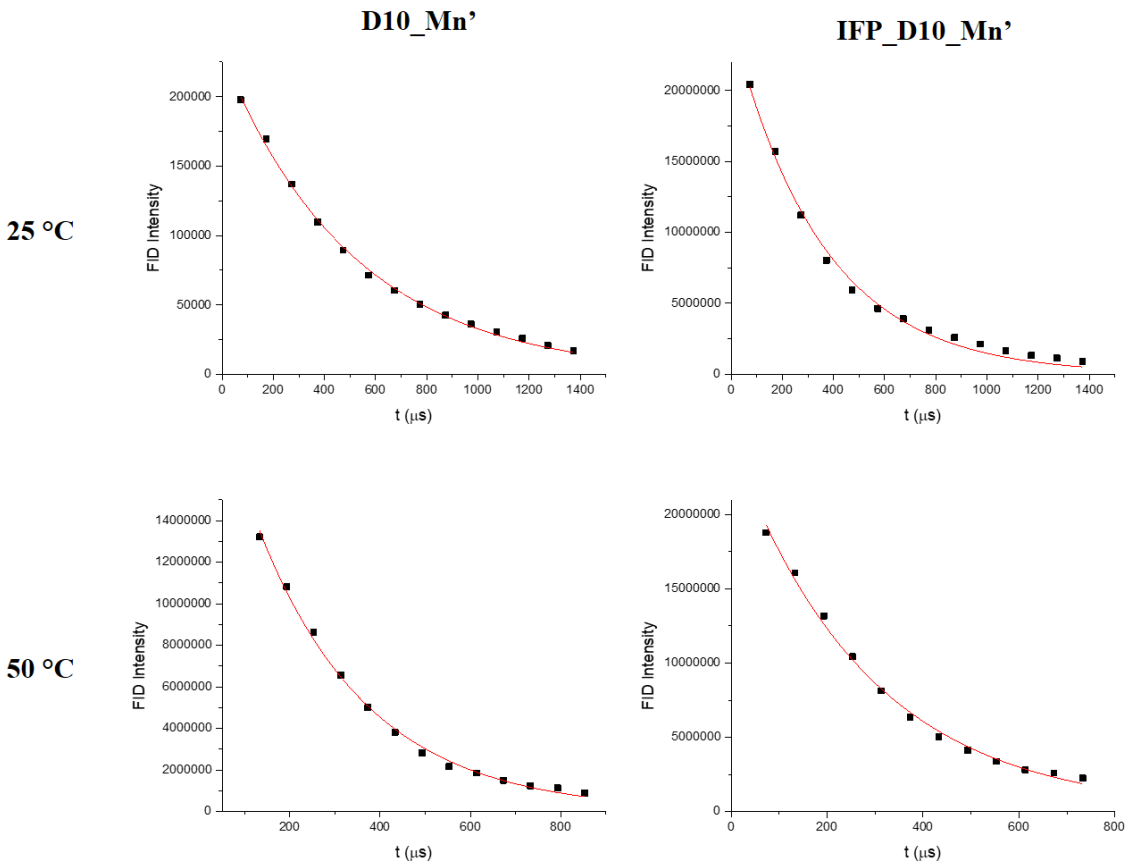

Figure 5.22 Quecho experimental (black squares) and best fit (red lines) plots of D10_Mn' and IFP_D10_Mn' with 20% Mn^{2+} at 25 °C and 50 °C.

Table 5.9 IFP-bounded D10 membrane ^2H NMR spectra peak splitting (kHz) with 20% Mn^{2+}

	25 °C	50 °C	
	Methyl -CD_3	Methyl -CD_3	Methylene -CD_2
D10	11.8	2.6	9.8
D10_Mn'	11.9	2.1	9.2
IFP_D10	11.1	2.7	10.4
IFP_D10_Mn'	11.0	-	-

Table 5.10 Best-fit ^2H T_2 (μs) of IFP-bounded membrane with 20% Mn^{2+}

	25 °C	50 °C
D10	516 (14)	1937 (36)
D10_Mn'	477 (10)	245 (5)
IFP_D10	385 (7)	1116 (42)
IFP_D10_Mn'	355 (11)	283 (7)

Table 5.11 Best-fit ^2H T_2 relaxation rate (R_2, in unit kHz) of IFP-bounded membrane with 20% Mn^{2+}

	25 °C	50 °C
D10	1.94 (5)	0.52 (1)
D10_Mn'	2.10 (4)	4.08 (9)
IFP_D10	2.60 (5)	0.90 (4)
IFP_D10_Mn'	2.82 (9)	3.53 (9)

5.2.6 IFP bounded membrane structure

The T_2 relaxation rate of the D10 and D8 samples at both gel phase and liquid phase showed that IFP does not promote lipid protrusion. This is concluded from the evidence that the Mn^{2+} induced R_2 increase is the same for IFP free D10 and IFP bounded D10, and the Mn^{2+} induced R_2 increase is the smaller for IFP bounded D8 than IFP free D8 at gel phase. At liquid phase, both IFP bounded D8 or D10 showed smaller Mn^{2+} induced R_2 increase than membrane only.

In Chapter 4, the IFP topology in membrane at gel phase was studied by REDOR solid-state NMR and the result showed close contacts between the lipid acyl chain terminus with Gly_2, Ala_7 and Gly_16 residues at gel phase, which is consistent with either the membrane spanning model or the lipid protrusion model (Figure 4.1). The membrane structure studied by PRE solid-state NMR experiments showed the IFP does not promote lipid protrusion at both gel phase and liquid phase. Thus, the only possible model of IFP is the membrane spanning conformation. A model of IFP topology in membrane is proposed by integrating the REDOR and PRE results, and is shown in Figure 5.23. The red dots on the lipid acyl chain represents the positions of D10 and D8 ^2H's.

To my knowledge, the IFP membrane spanning conformation is only proposed by two other groups and evidenced by simulation results done by Victor et al. and Worch et al.[14, 15] In Victor and Worch's simulation, the membrane and IFP were randomly distributed in a box of water and allowed to spontaneously assemble. Even though this simulation process was not a mimic of the actual infection, it was an efficient and unbiased way of determining the insertion mode of membrane-interacting IFP. The results showed that the dominant IPF conformation is a deeply buried helical hairpin, and it is the most probable and maybe the lowest free energy configuration. There is one discrepancy in these two studies of the orientation of IFP. Victor et al. showed a 47°

angle between peptide N-terminal helix axis and the membrane plane normal, while Worch et al. observed a parallel orientation.[14, 15]

There have been several studies for the IFP orientation in membrane from both our group as well as other groups. Wasniewski from our group used solid-state NMR to study the ^{15}N labeled Ala7 and Phe3 IFP in oriented membrane. The NMR ^{15}N peak was sharp when the membrane normal is parallel to the external magnetic field, and poor signal was obtained when the membrane normal is perpendicular to the external magnetic field. His results suggested that the IFP adopts an orientation approximately perpendicular to membrane normal.[16] In Yan Sun's dissertation, she also presented the study of the IFP N-terminal helix orientation in bicelle with DTPC:DMPC:DHPC 48:1:15 ratio by solid-state NMR. She observed that the IFP N-helix has a ~45° relative to membrane bicelle normal and this angle is independent of sample pH. This result is consistent with the ATR-FTIR experimental data showing this angle to be 45° and 41° from Luneberg et al. and Wu et al., respectively.[17, 18] Thus, in Figure 5.23, the IFP is shown with a ~45° tilt angle with respect to membrane normal, and this conformation is based on the MD simulation, solid-state NMR and ATR-FTIR results discussed above.[14, 17, 18]

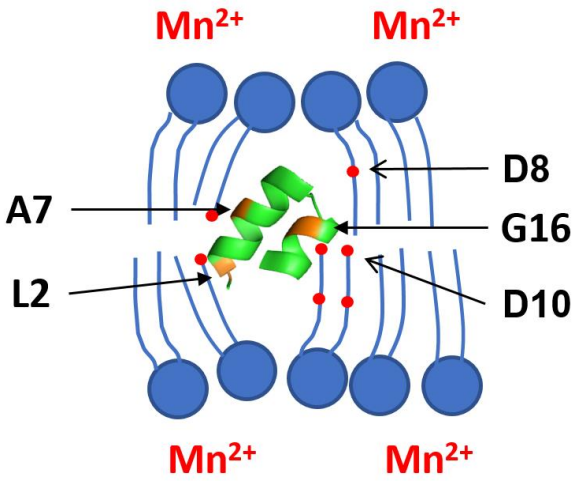

Figure 5.23 Model for IFP-bounded membrane. The lipid molecules are shown as blue sphere with two lines, ^{2}H atoms in lipid acyl chain shown as red dots, IFP in green, labeled amino acids in orange and Mn^{2+} binds on membrane surface.

5.3 Conclusion

Secondary structure and conformation of IFP have been intensively studied in both DPC micelle and lipid membrane.[8, 19-22] It has come to an agreement that IFP adopts a helix-turn-helix structure in membrane. In the recent study from our group, the IFP structure is independent of pH, HA subtype and sequence length, and there are always two distinct conformation: closed and semi-closed.

However, there has been arguments about the IFP membrane location. In some of the previous studies, the IFP is reported to be parallel to the membrane plane, and located near the lipid-water interface.[22-25] In this study, the REDOR and PRE solid-state NMR experiments were designed to testify this hypothesis. The REDOR experiments discussed in Chapter 4 showed close contacts between the lipid acyl chain tail with Gly_2, Ala_7 and Gly_16 residues. In this chapter, PRE solid-state NMR was used to study the IFP bounded membrane structure.

Sample preparation methods, effect of Mn^{2+} concentration on T_2 relaxation time and Mn^{2+} induced R_2 of IFP bounded membrane were studied and discussed. Based on both 5% and 20% Mn^{2+} PRE results, it can be concluded that IFP does not promote lipid protrusion at both gel phase and liquid phase, which is evidenced by that the R_2 difference with and without Mn^{2+} is smaller for IFP free membrane than IFP bounded membrane, meaning IFP does not induced smaller the average distance between lipid acyl chain and aqueous layer. By integrating these results, a IFP-membrane interaction model is proposed, in which the IFP deeply inserted in membrane hydrophobic core and the N-terminal helix has a ~45° with respect to membrane normal.

REFERENCES

REFERENCES

[1] Solomon, I. (1955) RELAXATION PROCESSES IN A SYSTEM OF 2 SPINS, *Physical Review 99*, 559-565.

[2] Clore, G. M., and Iwahara, J. (2009) Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes, *Chem. Rev. 109*, 4108-4139.

[3] Cai, S., Seu, C., Kovacs, Z., Sherry, A. D., and Chen, Y. (2006) Sensitivity enhancement of multidimensional NMR experiments by paramagnetic relaxation effects, *J. Am. Chem. Soc. 128*, 13474-13478.

[4] Buffy, J. J., Hong, T., Yamaguchi, S., Waring, A. J., Lehrer, R. I., and Hong, M. (2003) Solid-state NMR investigation of the depth of insertion of protegrin-1 in lipid bilayers using paramagnetic Mn2+, *Biophys. J. 85*, 2363-2373.

[5] Su, Y., Mani, R., and Hong, M. (2008) Asymmetric insertion of membrane proteins in lipid bilayers by solid-state NMR paramagnetic relaxation enhancement: A cell-penetrating peptide example, *J. Am. Chem. Soc. 130*, 8856-8864.

[6] Villalain, J. (1996) Location of cholesterol in model membranes by magic-angle-sample-spinning NMR, *Eur. J. Biochem. 241*, 586-593.

[7] Huster, D., Yao, X. L., and Hong, M. (2002) Membrane protein topology probed by H-1 spin diffusion from lipids using solid-state NMR spectroscopy, *J. Am. Chem. Soc. 124*, 874-883.

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

[9] van Meer, G., Voelker, D. R., and Feigenson, G. W. (2008) Membrane lipids: where they are and how they behave, *Nat. Rev. Mol. Cell Biol. 9*, 112-124.

[10] Mayer, L. D., Hope, M. J., Cullis, P. R., and Janoff, A. S. (1985) SOLUTE DISTRIBUTIONS AND TRAPPING EFFICIENCIES OBSERVED IN FREEZE-THAWED MULTILAMELLAR VESICLES, *Biochimica Et Biophysica Acta 817*, 193-196.

[11] Hope, M. J., Bally, M. B., Mayer, L. D., Janoff, A. S., and Cullis, P. R. (1986) GENERATION OF MULTILAMELLAR AND UNILAMELLAR PHOSPHOLIPID-VESICLES, *Chem. Phys. Lipids 40*, 89-107.

[12] Traikia, M., Warschawski, D. E., Recouvreur, M., Cartaud, J., and Devaux, P. F. (2000) Formation of unilamellar vesicles by repetitive freeze-thaw cycles: characterization by

electron microscopy and P-31-nuclear magnetic resonance, *Eur. Biophys. J. Biophys. Lett. 29*, 184-195.

[13] Gabrys, C. M., Yang, R., Wasniewski, C. M., Yang, J., Canlas, C. G., Qiang, W., Sun, Y., and Weliky, D. P. (2010) Nuclear magnetic resonance evidence for retention of a lamellar membrane phase with curvature in the presence of large quantities of the HIV fusion peptide, *Biochim. Biophys. Acta-Biomembr. 1798*, 194-201.

[14] Victor, B. L., Lousa, D., Antunes, J. M., and Soares, C. M. (2015) Self-Assembly Molecular Dynamics Simulations Shed Light into the Interaction of the Influenza Fusion Peptide with a Membrane Bilayer, *J. Chem Inf. Model. 55*, 795-805.

[15] Worch, R., Krupa, J., Filipek, A., Szymaniec, A., and Setny, P. (2017) Three conserved C-terminal residues of influenza fusion peptide alter its behavior at the membrane interface, *Biochim. Biophys. Acta-Gen. Subj. 1861*, 97-105.

[16] Wasniewski, C. M., Parkanzky, P. D., Bodner, M. L., and Weliky, D. P. (2004) Solid-state nuclear magnetic resonance studies of HIV and influenza fusion peptide orientations in membrane bilayers using stacked glass plate samples, *Chem. Phys. Lipids 132*, 89-100.

[17] Wu, C. W., Cheng, S. F., Huang, W. N., Trivedi, V. D., Veeramuthu, B., Kantchev, A. B., Wu, W. G., and Chang, D. K. (2003) Effects of alterations of the amino-terminal glycine of influenza hemagglutinin fusion peptide on its structure, organization and membrane interactions, *Biochim. Biophys. Acta-Biomembr. 1612*, 41-51.

[18] Luneberg, J., Martin, I., Nussler, F., Ruysschaert, J. M., and Herrmann, A. (1995) STRUCTURE AND TOPOLOGY OF THE INFLUENZA-VIRUS FUSION PEPTIDE IN LIPID BILAYERS, *J. Biol. Chem. 270*, 27606-27614.

[19] Lorieau, J. L., Louis, J. M., and Bax, A. (2010) The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface, *Proc. Natl. Acad. Sci. U. S. A. 107*, 11341-11346.

[20] Lorieau, J. L., Louis, J. M., Schwieters, C. D., and Bax, A. (2012) pH-triggered, activated-state conformations of the influenza hemagglutinin fusion peptide revealed by NMR, *Proc. Natl. Acad. Sci. U. S. A. 109*, 19994-19999.

[21] Ghosh, U., Xie, L., and Weliky, D. P. (2013) Detection of closed influenza virus hemagglutinin fusion peptide structures in membranes by backbone (CO)-C-13-N-15 rotational-echo double-resonance solid-state NMR, *J. Biomol. NMR 55*, 139-146.

[22] Ghosh, U., Xie, L., Jia, L. H., Liang, S., and Weliky, D. P. (2015) Closed and Semiclosed Interhelical Structures in Membrane vs Closed and Open Structures in Detergent for the Influenza Virus Hemagglutinin Fusion Peptide and Correlation of Hydrophobic Surface Area with Fusion Catalysis, *J. Am. Chem. Soc. 137*, 7548-7551.

[23] Larsson, P., and Kasson, P. M. (2013) Lipid Tail Protrusion in Simulations Predicts Fusogenic Activity of Influenza Fusion Peptide Mutants and Conformational Models, *PLoS Comput. Biol. 9*, 9.

[24] Brice, A. R., and Lazaridis, T. (2014) Structure and Dynamics of a Fusion Peptide Helical Hairpin on the Membrane Surface: Comparison of Molecular Simulations and NMR, *J. Phys. Chem. B 118*, 4461-4470.

[25] Baylon, J. L., and Tajkhorshid, E. (2015) Capturing Spontaneous Membrane Insertion of the Influenza Virus Hemagglutinin Fusion Peptide, *J. Phys. Chem. B 119*, 7882-7893.

Chapter 6

Summary and future directions

My research during the past five years has been focused on HIV gp41 and influenza hemagglutinin (HA). These two proteins are responsible for the initial step of virus infection by catalyzing the fusion between viral and host cell membrane. HIV infection happens on the surface of immune cells. After the HIV glycoprotein gp120 binds to the host cell, gp41 is exposed and fusion between host cell membrane and viral membrane starts. The fusion peptide (FP) and transmembrane (TM) domains are the segments that binds to membrane and are critical for fusion.[1] As for influenza virus, the virion is endocytosed by host cell and the pH drop inside endosome triggers a conformational change in HA. The ~25 N-terminal residues of HA subunit 2 is fusion peptide domain (IFP). IFP binds to endosomal membrane and is necessary for fusion.[2] However, the structure of gp41 and HA as well as the infection mechanism are still not fully understood. The overall goal is to understand the gp41 and IFP structure and their interaction with membrane

Chapter 3 mainly discussed the structure and function of HIV gp41 fusion protein. A gp41 construct including the whole ectodomain and transmembrane domain, and three other shorter constructs were expressed, purified and stabilized in physiology buffers containing SDS or DPC detergents at neutral or low pH's. The constructs adopt α helical SE and TM, and non-helical FP in both detergents, and they are all hyperthermostable with T_m > 90 °C. The oligomeric states of these proteins vary in different detergent buffer: predominant trimer for all constructs and some hexamer fraction for HM and HM_TM protein in SDS at pH 7.4; and mixtures of monomer, trimer, and higher-order oligomer protein in DPC at pH 4.0 and 7.4. Substantial protein-induced vesicle fusion was observed, including fusion at the condition similar to HIV/cell fusion. The fusion activity is aided by the membrane associate FP and TM domains, and by protein which is predominantly trimer rather than monomer.

We proposed that there is no contact between FP and TM domains in detergent solutions according to current CD data. This can be testified by further hydrogen-deuterium exchange experiments.[3] To perform the experiment, the gp41 proteins is mixed with D_2O and incubate for a certain time, then the H-D exchange is quenched by lowering the pH to ~2 and decrease temperature to 0 °C. The deuterated protein is digested by pepsin and the digested peptides are injected into mass spectrometry. The molecular weight of the peptides can be analyzed. If the FP and TM formed a complex in the final SHB structure, the exchange rate of deuterons present in the FP region of FP_HM_TM can be different from the exchange rate of deuterons present in the FP region of FP_HM.

The FP and TM interaction in membrane also helps to understand the conformation of gp41 at the post-fusion state. Future studied can be done to determine the proximity of FP_HM_TM incorporated in membrane using ^{13}C-^{15}N REDOR solid-state NMR. The FP_HM_TM can be obtained by native chemical ligation of ^{13}C labeled FP peptide and ^{15}N labeled HM_TM. The FP peptide can be synthesized by solid-phase peptide synthesis with site-specific labeling, while the HM_TM can be labeled with ^{15}N Phe residue by bacterial expression.[4] There are three Phe in HM_TM: one of them in MPER and two in TM. The Phe in MPER could be mutated to other residues to avoid the FP and MPER interaction being observed. The result of a big dephasing buildup of the ^{13}C-^{15}N REDOR indicates close contact between FP and TM, and small dephasing buildup indicates no contact. The possible problem lies in the expression of ^{15}N Phe labeled HM_TM and other residues might also get labeled due to *E.coli* scrambling. It is possible that except for Phe, other residues are also labeled. If the yield of ligation experiment is high, it is worth trying to ligate FP and TM to HM.

Previous studies of IFP were done in both detergent and membrane, and it has come to an agreement that IFP adopts a helix-turn-helix structure in membrane.[5, 6] The studies shown in Chapter 4 and 5 are trying to illustrate the IFP membrane location. The ^{13}C-^{2}H REDOR solid-state NMR results showed that the IFP adopts major α helical, minor β strand secondary structure in PC/PG membrane. The Leu-2, Ala-7 and Gly-16 in α helical IFP all show close contacts with the lipid acyl chain tail, suggesting IFP has strong interaction with the membrane. The PRE solid-state NMR is used to study the IFP bounded lipid membrane structure. The T_2 relaxation time and rate were measured for membrane with or without IFP and with or without Mn^{2+}. The results showed not IFP induced lipid acyl chain protrusion at both gel phase and liquid phase, which is evidenced by that the Mn^{2+} induced R_2 increase is smaller or equal for IFP bounded membrane than IFP free membrane. This suggest that the IFP does not induce a smaller the average distance between lipid acyl chain and aqueous layer. By comparing the REDOR and PRE results to the existing IFP topology models, a IFP membrane spanning model was proposed, in which IFP N-terminal helix adopts a 45° angle with respect to membrane normal.

It will be interesting to study the IFP mutant membrane location. The IFP is highly conserved and even a single site mutation might affect the IFP fusion activity. For example, mutation of Gly_1 to Val, Glu, Gln or Lys completely eliminates the fusion.[7] The secondary structure, membrane location and peptide-membrane interaction might be different from the wild type and can provides more information about the possible fusion process. The experiment can be done by the ^{13}C-^{2}H REDOR experiment of site specifically labeled IFP mutant and deuterated lipid membrane. The membrane structure perturbed by the IFP mutant can be studied by PRE with deuterated lipids and Mn^{2+} ions.

It will also be meaningful to study the IFP in a larger construct and how the IFP interacts with membrane. The large construct such as FHA2, which includes the soluble ectodomain SHA2 and fusion peptide, can be obtained by ligate ^{13}C labeled FP peptide and unlabeled SHA2. The IFP interaction with the soluble ectodomain is also very interesting. The ^{13}C labeled FP peptide can be ligated with deuterated SHA2. Compared to isolated IFP, these experiments are better mimic of the actual infection.

APPENDICES

APPENDIX A

Location of NMR data

Figure 4.3

(a) /home/khare0/mb4b/data/Shuang/13C2H/IFP/L2C/031214 (L2c-D4)

(b) /home/khare0/mb4b/data/Shuang/13C2H/IFP/L2C/031814 (L2c-D8)

(c) /home/khare0/mb4b/data/Shuang/13C2H/IFP/L2C/032714 (L2c-D10)

Figure 4.4

(a) /home/khare0/mb4b/data/Shuang/13C2H/IFP/A7C/041614 (A7c-D4)

(b) /home/khare0/mb4b/data/Shuang/13C2H/IFP/A7C/050614 (A7c-D8)

(c) /home/khare0/mb4b/data/Shuang/13C2H/IFP/A7C/052114 (A7c-D10)

Figure 4.5

(a) /home/khare0/mb4b/data/Shuang/13C2H/IFP/G16C/111314 (G16c-D10)

(b) /home/khare0/mb4b/data/Shuang/13C2H/IFP/G16C/112414 (G16c-D8)

(c) /home/khare0/mb4b/data/Shuang/13C2H/IFP/G16C/120214 (G16c-D4)

Figure 4.6

(a) /home/hapi0/mb4b/data/Shuang/13C2H/IFP/IFP_040213 (L2c-D10)

(b) /home/hapi0/mb4b/data/Shuang/13C2H/IFP/IFP_091113 (L2c-D8)

Figure 5.3

/home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10/100316_25_T2

Figure 5.4

/home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10/101116_50_T2

Figure 5.5

/home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8/121216_25_T2

Figure 5.6

/home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8/120316_50_T2

Figure 5.8

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10/100316_25_T2 (D10 25°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10/101116_50_T2 (D10 50°C)

(c) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8/121216_25_T2 (D8 25°C)

(d) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8/120316_50_T2 (D8 50°C)

Figure 5.10

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/101216_25_T2 (20% Mn^{2+} 25°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/103116_25_T2_5% (5% Mn^{2+} 25°C)

(c) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/020117_25_T2_1% (1% Mn^{2+} 25°C)

(d) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/032417_25_T2_0.2% (0.2% Mn^{2+} 25°C)

Figure 5.11

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/102216_50_T2 (20% Mn^{2+} 50°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/110116_50_T2_5% (5% Mn^{2+} 50°C)

(c) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/020217_50_T2_1% (1% Mn^{2+} 50°C)

(d) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/031417_50_T2_0.2% (0.2% Mn^{2+} 50°C)

Figure 5.12

/home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D10/100616_25_T2

Figure 5.13

/home/khare0/mb4b/data/Shuang/quecho/repeat_101016/IFP_D10/100616_50_T2

Figure 5.16

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10/100316_25_T2 (D10 25°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/103116_25_T2 (D10_Mn 25°C)

(c) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D10_Mn/100616_25_T2 (IFP_D10 25°C)

(d) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D10_Mn/112916_25_T2 (IFP_D10_Mn 25°C)

Figure 5.17

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10/101116_50_T2 (D10 50°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/110116_50_T2 (D10_Mn 50°C)

(c) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D10_Mn/101016_50_T2 (IFP_D10 50°C)

(d) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D10_Mn/120116_50_T2 (IFP_D10_Mn 50°C)

Figure 5.18

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8/121216_25_T2 (D8 25°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8_Mn/020717_25_T2 (D8_Mn 25°C)

(c) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D8_Mn/112416_25_T2 (IFP_D8 25°C)

(d) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D8_Mn/121616_25_T2 (IFP_D8_Mn 25°C)

Figure 5.19

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8/120316_50_T2 (D8 50°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8_Mn/020816_50_T2 (D8_Mn 50°C)

(c) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D8_Mn/112616_50_T2 (IFP_D8 50°C)

(d) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D8_Mn/120416_50_T2 (IFP_D8_Mn 50°C)

Figure 5.20

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8/121216_25_T2 (D8 25°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D8/111716_25_T2 (D8 25°C replica)

Figure 5.22

(a) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/101216_25_T2 (D10_Mn' 25°C)

(b) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/D10_Mn/102216_50_T2 (D10_Mn' 50°C)

(c) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D10_Mn/101716_25_T2 (IFP_D10_Mn' 25°C)

(d) /home/khare0/mb4b/data/Shuang/quecho/repeat_100316/IFP_D10_Mn/102016_50_T2 (IFP_D10_Mn' 50°C)

APPENDIX B

Solid-state NMR raw data

Table B1 IFP L7c in DPPC-D4 studied by REDRO raw data (Figure 4.3)

Dephasing time (μs)	S_0	Error of S_0	S_1	error of S_1	$\Delta S/ S_0$	Error of $\Delta S/ S_0$
α						
2	9.1931	0.344	8.5472	0.211	0.07026	0.04168
8	15.7928	0.337	15.0415	0.331	0.04757	0.02919
16	17.1303	0.441	14.6836	0.438	0.14283	0.03377
24	6.4586	0.419	4.9174	0.429	0.23863	0.08278
32	15.4763	0.666	11.2919	0.582	0.27037	0.04899
40	7.0760	0.521	5.7963	0.407	0.18085	0.08334
48	12.7943	0.757	9.7390	0.680	0.23880	0.06966
β						
2	16.07	0.344	14.6996	0.211	0.085276914	0.02358
8	39.8608	0.337	32.8811	0.331	0.175101854	0.01084
16	45.6214	0.441	37.2641	0.438	0.183188153	0.01243
24	15.8156	0.419	12.4953	0.429	0.209938289	0.03426
32	37.3028	0.666	28.8126	0.582	0.227602218	0.02082
40	19.622	0.521	14.6103	0.407	0.255412292	0.02865
48	35.483	0.757	27.5546	0.68	0.223442212	0.02533

Table B2 IFP L7c in DPPC-D8 studied by REDRO raw data (Figure 4.3)

Dephasing time (μs)	S_0	Error of S_0	S_1	error of S_1	$\Delta S/ S_0$	Error of $\Delta S/ S_0$
α						
2	16.1121	0.374	16.0004	0.288	0.0069	0.0292
8	12.4556	0.396	12.1574	0.273	0.0239	0.0380
16	15.8868	0.426	14.5082	0.304	0.0868	0.0311
24	14.1491	0.552	12.2086	0.445	0.1371	0.0461
32	22.4405	0.848	16.6942	0.722	0.2561	0.0427
40	14.8187	0.794	10.4636	0.448	0.2939	0.0484
48	16.2268	0.811	10.4543	0.444	0.3557	0.0423
β						
2	21.078	0.374	20.7092	0.288	0.0175	0.0221
8	17.9311	0.396	16.7442	0.273	0.0662	0.0256
16	25.1182	0.426	21.3658	0.304	0.1494	0.0188
24	24.7114	0.552	19.5980	0.445	0.2069	0.0253
32	34.3695	0.848	24.8544	0.722	0.2768	0.0276
40	26.5748	0.794	18.3733	0.448	0.3086	0.0267
48	27.5736	0.811	17.7389	0.444	0.3567	0.0248

Table B3 IFP L7c in DPPC-D10 studied by REDRO raw data (Figure 4.3)

Dephasing time (μs)	S_0	Error of S_0	S_1	error of S_1	$\Delta S/ S_0$	Error of $\Delta S/ S_0$
α						
2	65.2858	0.695	63.6866	0.775	0.0245	0.0158
8	35.2148	0.439	26.6633	0.630	0.2428	0.0202
16	21.8837	0.507	11.6344	0.598	0.4684	0.0300
24	33.8663	0.849	14.1174	0.704	0.5831	0.0233
32	30.6851	0.934	8.0556	0.616	0.7375	0.0216
40	28.1228	0.836	4.9677	0.975	0.8234	0.0351
48	14.1873	0.679	2.9976	0.907	0.7887	0.0647
β						
2	41.0415	0.695	40.4178	0.775	0.0152	0.0252
8	24.2896	0.439	21.1622	0.630	0.1288	0.0303
16	13.6729	0.507	11.0643	0.598	0.1908	0.0530
24	21.7212	0.849	17.0868	0.704	0.2134	0.0447
32	19.8017	0.934	12.6815	0.616	0.3596	0.0434
40	16.9414	0.836	12.0296	0.975	0.2899	0.0674
48	10.7682	0.679	5.8910	0.907	0.4529	0.0910

Table B4 IFP A7c in DPPC-D4 studied by REDRO raw data (Figure 4.4)

Dephasing time (μs)	S_0	Error of S_0	S_1	error of S_1	$\Delta S/ S_0$	Error of $\Delta S/ S_0$
α						
2	32.4448	1.220	31.9501	1.224	0.01525	0.05286
8	20.7829	0.381	18.8372	0.500	0.09362	0.02924
16	30.6827	0.891	28.0396	1.076	0.08614	0.04398
24	29.6244	0.754	26.2857	0.635	0.11270	0.03114
32	33.4121	0.911	30.8848	1.256	0.07564	0.04526
40	20.4066	0.566	18.7614	0.566	0.08062	0.03768
48	21.8580	0.739	18.5341	0.622	0.15207	0.04039
β						
2	106.8569	1.220	97.4845	1.224	0.087709825	0.01548
8	73.9308	0.381	62.6787	0.500	0.152197731	0.00805
16	98.2642	0.891	84.3455	1.076	0.141645686	0.01343
24	110.911	0.754	94.9251	0.635	0.144132683	0.00816
32	108.6387	0.911	91.2165	1.256	0.160368267	0.01354
40	79.2688	0.566	67.6721	0.566	0.146295894	0.00939
48	96.2936	0.739	84.2791	0.622	0.124769455	0.00932

Table B5 IFP A7c in DPPC-D8 studied by REDRO raw data (Figure 4.4)

Dephasing time (μs)	S_0	Error of S_0	S_1	error of S_1	$\Delta S/ S_0$	Error of $\Delta S/ S_0$
α						
2	49.7243	0.676	48.7030	0.449	0.02054	0.01609
8	83.8645	0.762	79.6153	0.593	0.05067	0.01115
16	87.1904	0.894	79.4545	0.580	0.08872	0.01147
24	36.4263	0.573	31.0494	0.436	0.14761	0.01797
32	41.2316	0.323	34.0730	0.432	0.17362	0.01232
40	33.4359	0.521	26.2444	0.450	0.21508	0.01819
48	37.6289	0.855	30.4702	0.430	0.19024	0.02166
β						
2	55.9898	0.676	54.2555	0.449	0.030975285	0.01418
8	83.9161	0.762	75.6585	0.593	0.098403048	0.01081
16	82.008	0.894	69.3302	0.580	0.154592235	0.01162
24	30.8507	0.573	24.5143	0.436	0.205389181	0.02043
32	35.8539	0.323	26.9150	0.432	0.249314579	0.01382
40	28.0622	0.521	19.5341	0.450	0.303899908	0.02060
48	31.482	0.855	20.5213	0.430	0.348157677	0.02236

Table B6 IFP A7c in DPPC-D10 studied by REDRO raw data (Figure 4.4)

Dephasing time (μs)	S_0	Error of S_0	S_1	error of S_1	$\Delta S/ S_0$	Error of $\Delta S/ S_0$
α						
2	26.6149	0.494	26.9447	0.394	-0.01239	0.02392
8	35.6549	0.405	28.3532	0.515	0.20479	0.01704
16	36.4014	0.444	22.7760	0.447	0.37431	0.01446
24	38.2096	0.591	18.1990	0.440	0.52371	0.01367
32	31.6442	0.700	11.8022	0.786	0.62703	0.02617
40	29.5316	0.930	9.2844	0.536	0.68561	0.02067
48	28.7216	0.585	7.1254	0.597	0.75191	0.02139
β						
2	38.4733	0.494	38.3586	0.394	0.002981288	0.01639
8	46.8079	0.405	38.6348	0.515	0.174609414	0.01312
16	44.6926	0.444	31.9632	0.447	0.284821201	0.01227
24	45.2512	0.591	27.5809	0.440	0.390493512	0.01257
32	37.1126	0.700	20.7705	0.786	0.440338322	0.02366
40	33.4251	0.930	16.8915	0.536	0.494646239	0.02133
48	32.1653	0.585	11.2943	0.597	0.648866947	0.01963

Table B7 IFP G16c in DPPC-D4, DPPC-D8 and DPPC-D10 studied by REDRO raw data (Figure 4.5)

Dephasing time (μs)	S_0	Error of S_0	S_1	error of S_1	$\Delta S/ S_0$	Error of $\Delta S/ S_0$
D4						
2	56.0648	0.745	49.3945	0.509	0.1190	0.015
8	98.5737	1.214	77.5882	0.843	0.2129	0.013
16	122.5577	1.676	92.4948	1.358	0.2453	0.015
24	99.5157	1.134	70.2728	1.115	0.2939	0.014
32	71.7066	0.751	48.0781	0.847	0.3295	0.014
40	81.6746	1.721	52.4972	1.411	0.3572	0.022
48	72.1971	2.262	44.8710	1.380	0.3785	0.027
D8						
2	112.9807	1.277	111.6880	1.554	0.0114	1.554
8	78.7638	1.295	72.4339	1.110	0.0804	1.110
16	134.6262	1.054	110.6980	1.673	0.1777	1.673
24	82.9444	1.573	62.2264	0.835	0.2498	0.835
32	73.8451	1.162	49.5225	0.932	0.3294	0.932
40	78.8261	1.347	46.5444	1.210	0.4095	1.210
48	54.4701	1.357	29.4906	2.056	0.4586	2.056
D10						
2	67.3786	0.544	67.0672	0.768	0.0046	0.0139

Table B7 (cont'd)

8	200.2962	1.524	172.6975	1.769	0.1378	0.0110
16	154.4760	1.964	111.3407	1.682	0.2792	0.0142
24	73.3691	1.227	44.2168	1.721	0.3973	0.0255
32	79.8731	1.820	38.4816	1.493	0.5182	0.0217
40	81.1137	1.483	29.2690	2.181	0.6392	0.0277
48	45.4283	1.254	14.3887	1.796	0.6833	0.0405

Table B8 IFP L2c in pure DPPC-D8 and DPPC-D10 studied by REDRO raw data (Figure 4.6)

Dephasing time (μs)	S_0	Error of S_0	S_1	error of S_1	$\Delta S/ S_0$	Error of $\Delta S/ S_0$
D8						
2	55.3897	0.483	53.8299	0.546	0.02816	0.01300
8	57.2080	0.647	46.9816	0.565	0.17876	0.01356
16	34.3314	0.719	21.0580	0.510	0.38663	0.01964
24	30.6357	0.757	14.8135	0.809	0.51646	0.02898
32	20.1005	1.000	6.6848	0.958	0.66743	0.05045
40	18.1662	0.840	5.3939	0.574	0.70308	0.03445
48	12.0504	1.040	2.5912	0.866	0.78497	0.07422
D10						
2	37.1279	0.360	37.0288	0.264	0.00267	0.01200
8	57.5783	0.598	55.1622	0.438	0.04196	0.01252
16	39.0030	0.364	36.3823	0.331	0.06719	0.01216
24	39.2912	0.522	35.8535	0.442	0.08749	0.01654
32	43.3655	0.779	38.0072	0.653	0.12356	0.02179
40	56.7507	0.812	50.0904	0.756	0.11736	0.01836
48	47.8901	0.821	39.9486	1.178	0.16583	0.02845

Table B9 D10 with various concentration of Mn^{2+} at 25 °C studied by PRE (Figure 5.10)

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
20%				5%			
30	11	73	16741177	30	11	73	25351580
80	61	173	13945282	80	61	173	22215764
130	111	273	10734524	130	111	273	18618496
180	161	373	8184182	180	161	373	15270932
230	211	473	6424096	230	211	473	12326184
280	261	573	5254574	280	261	573	9971276
330	311	673	4295137	330	311	673	7971368
380	361	773	3641079	380	361	773	6485960
430	411	873	2973930	430	411	873	5196364
480	461	973	2472762	480	461	973	4233112
530	511	1073	1995562	530	511	1073	3442028
580	561	1173	1728693	580	561	1173	2793020
630	611	1273	1317678	630	611	1273	2176876
680	661	1373	1163524	680	661	1373	1782552
1%				0.2%			
30	11	73	15349736	30	11	73	2698428
80	61	173	13316156	80	61	173	2380908
130	111	273	11215504	130	111	273	2131156

Table B9 (cont'd)

180	161	373	9158632	180	161	373	1798660
230	211	473	7401352	230	211	473	1507128
280	261	573	5819964	280	261	573	1268288
330	311	673	4591952	330	311	673	1032948
380	361	773	3657716	380	361	773	834864
430	411	873	2909052	430	411	873	632496
480	461	973	2238292	480	461	973	542252
530	511	1073	1720156	530	511	1073	458456
580	561	1173	1337784	580	561	1173	369484
630	611	1273	1045780	630	611	1273	328884
680	661	1373	926568	680	661	1373	241948

Table B10 D10 with various concentration of Mn^{2+} at 50 °C studied by PRE (Figure 5.11)

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
20%				5%			
30	11	73	14691436	30	11	71	20660388
60	41	133	13212680	80	61	171	17831868
90	71	193	10819840	130	111	271	14636092
120	101	253	8627060	180	161	371	11931888
150	131	313	6565180	230	211	471	9883984
180	161	373	4996032	280	261	571	8210824
210	191	433	3810804	330	311	671	6839976
240	221	493	2813868	380	361	771	5855812
270	251	553	2169360	430	411	871	4972756
300	281	613	1846720	480	461	971	4448160
330	311	673	1494368	530	511	1071	3905876
360	341	733	1236584	580	561	1171	3421528
390	371	793	1138232	630	611	1271	2957996
420	401	853	891008	680	661	1371	2635640
1%				0.2%			
30	11	73	16205600	30	11	73	2698428
80	61	173	14598764	80	61	173	2380908
130	111	273	13389080	130	111	273	2131156

Table B10 (cont'd)

180	161	373	11880292	180	161	373	1798660
230	211	473	10761116	230	211	473	1507128
280	261	573	9770600	280	261	573	1268288
330	311	673	9035032	330	311	673	1032948
380	361	773	8015876	380	361	773	834864
430	411	873	7253204	430	411	873	632496
480	461	973	6605856	480	461	973	542252
530	511	1073	6191140	530	511	1073	458456
580	561	1173	5625868	580	561	1173	369484
630	611	1273	5300064	630	611	1273	328884
680	661	1373	4779708	680	661	1373	241948

Table B11 D10 at 25 °C studied by PRE (Figure 5.16)

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D10				D10_Mn			
30	11	73	22939148	30	11	73	25351580
80	61	173	21274456	80	61	173	22215764
130	111	273	18905892	130	111	273	18618496
180	161	373	16142057	180	161	373	15270932
230	211	473	13634104	230	211	473	12326184
280	261	573	11328846	280	261	573	9971276
330	311	673	9301617	330	311	673	7971368
380	361	773	7763542	380	361	773	6485960
430	411	873	6041013	430	411	873	5196364
480	461	973	4843347	480	461	973	4233112
530	511	1073	3912750	530	511	1073	3442028
580	561	1173	2934186	580	561	1173	2793020
630	611	1273	2366845	630	611	1273	2176876
680	661	1373	1896893	680	661	1373	1782552
IFP_D10				IFP_D10_Mn			
30	11	73	23445574	30	11	73	23175468
80	61	173	19328814	80	61	173	19028900
130	111	273	15634267	130	111	273	14593728

Table B11 (cont'd)

180	161	373	12078985	180	161	373	11093240
230	211	473	9278228	230	211	473	8564448
280	261	573	7090318	280	261	573	6690932
330	311	673	5377061	330	311	673	5133752
380	361	773	3980167	380	361	773	4107736
430	411	873	3159236	430	411	873	3417952
480	461	973	2316542	480	461	973	2700456
530	511	1073	1747176	530	511	1073	2219808
580	561	1173	1311151	580	561	1173	1763332
630	611	1273	878248	630	611	1273	1515724
680	661	1373	716006	680	661	1373	1178916

Table B12 D10 at 50 °C studied by PRE (Figure 5.17)

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D10				D10_Mn			
30	11	73	20429904	30	11	71	20660388
180	161	373	13719244	80	61	171	17831868
330	311	673	10529842	130	111	271	14636092
480	461	973	8826575	180	161	371	11931888
630	611	1273	7400605	230	211	471	9883984
780	761	1573	6286838	280	261	571	8210824
930	911	1873	5476900	330	311	671	6839976
1080	1061	2173	4598277	380	361	771	5855812
1230	1211	2473	4091422	430	411	871	4972756
1380	1361	2773	3490750	480	461	971	4448160
1530	1511	3073	3099347	530	511	1071	3905876
1680	1661	3373	2597324	580	561	1171	3421528
1830	1811	3673	2302469	630	611	1271	2957996
1980	1961	3973	2087342	680	661	1371	2635640
IFP_D10				IFP_D10_Mn			
30	11	71	23213666	30	11	73	23224660
180	161	371	17338260	80	61	173	19154584
330	311	671	12472822	130	111	273	15390656

Table B12 (cont'd)

480	461	971	9379849	180	161	373	12821624
630	611	1271	7100031	230	211	473	10820300
780	761	1571	5680466	280	261	573	9088668
930	911	1871	4467065	330	311	673	7790152
1080	1061	2171	3699964	380	361	773	6816512
1230	1211	2471	3080141	430	411	873	6083032
1380	1361	2771	2630065	480	461	973	5225736
1530	1511	3071	2197503	530	511	1073	4604304
1680	1661	3371	1979658	580	561	1173	3897424
1830	1811	3671	1723330	630	611	1273	3619868
1980	1961	3971	1508588	680	661	1373	3472196

Table B13 D8 at 25 °C studied by PRE (Figure 5.18)

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D8				D8_Mn			
30	11	73	18979972	30	11	73	11574004
70	51	153	16777240	50	31	113	10196976
110	91	233	13911604	70	51	153	8835052
150	131	313	10937824	90	71	193	7362160
190	171	393	8483648	110	91	233	6240424
230	211	473	6575180	130	111	273	5226200
270	251	553	4970600	150	131	313	4401616
310	291	633	3794728	170	151	353	3733872
350	331	713	2939836	190	171	393	3079276
390	371	793	2267060	210	191	433	2653260
430	411	873	1710868	230	211	473	2231720
470	451	953	1355856	250	231	513	1849952
510	491	1033	1055080	270	251	553	1571860
550	531	1113	763484	290	271	593	1325920
IFP_D8				IFP_D8_Mn			
30	11	73	12914244	30	11	73	6839532
50	31	113	10577272	40	21	93	5753080
70	51	153	8350124	50	31	113	5323424

Table B13 (cont'd)

90	71	193	6890136	60	41	133	4575368
110	91	233	5205524	70	51	153	4000256
130	111	273	4006824	80	61	173	3575728
150	131	313	3083364	90	71	193	3069020
170	151	353	2477144	100	81	213	2829832
190	171	393	1932796	110	91	233	2362168
210	191	433	1444700	120	101	253	2328236
230	211	473	1081860	130	111	273	1952584
250	231	513	828176	140	121	293	1810428
270	251	553	646252	150	131	313	1692552
290	271	593	547804	160	141	333	1401204

Table B14 D8 at 50 °C studied by PRE (Figure 5.19)

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D8				D8_Mn			
30	11	73	18297716	30	11	73	11284620
80	61	173	15195036	50	31	113	9828736
130	111	273	11656272	70	51	153	8352444
180	161	373	9759664	90	71	193	6843708
230	211	473	8197936	110	91	233	5652852
280	261	573	6931136	130	111	273	4777816
330	311	673	5993344	150	131	313	4096996
380	361	773	5302892	170	151	353	3857288
430	411	873	4594732	190	171	393	3393380
480	461	973	4150888	210	191	433	3050220
530	511	1073	3699508	230	211	473	2728400
580	561	1173	3100776	250	231	513	2376980
630	611	1273	2909236	270	251	553	2285012
680	661	1373	2614820	290	271	593	2203524
IFP_D8				IFP_D8_Mn			
30	11	73	11008496	30	11	73	14011488
60	41	133	8294992	50	31	113	11421860
90	71	193	6992826	70	51	153	10259204

Table B14 (cont'd)

120	101	253	5824356	90	71	193	8121512
150	131	313	4443080	110	91	233	7090256
180	161	373	3288032	130	111	273	5868476
210	191	433	2868336	150	131	313	5026988
240	221	493	2434644	170	151	353	4249668
270	251	553	1960872	190	171	393	3619308
300	281	613	1614376	210	191	433	3141236
330	311	673	1341708	230	211	473	2861672
360	341	733	1080152	250	231	513	2471484
390	371	793	978816	270	251	553	2148172
420	401	853	831404	290	271	593	1961748

Table B15 D8 reproducibility at 25 °C studied by PRE (Figure 5.20)

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D8				D8_replicate			
30	11	73	18297716	30	11	73	14017444
80	61	173	15195036	80	61	173	11658036
130	111	273	11656272	130	111	273	8644612
180	161	373	9759664	180	161	373	6239496
230	211	473	8197936	230	211	473	4303348
280	261	573	6931136	280	261	573	3077584
330	311	673	5993344	330	311	673	2068132
380	361	773	5302892	380	361	773	1385992
430	411	873	4594732	430	411	873	1012780
480	461	973	4150888	480	461	973	709404
530	511	1073	3699508	530	511	1073	493972
580	561	1173	3100776	580	561	1173	446048
630	611	1273	2909236	630	611	1273	315104
680	661	1373	2614820	680	661	1373	215124

Table B16 D10 with 20% Mn^{2+} studied by PRE (Figure 5.22)

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D10_Mn' at 25 °C				IFP_D10_Mn' at 25 °C			
30	11	73	16741177	30	11	73	20422396
80	61	173	13945282	80	61	173	15695484
130	111	273	10734524	130	111	273	11204164
180	161	373	8184182	180	161	373	8035720
230	211	473	6424096	230	211	473	5927740
280	261	573	5254574	280	261	573	4633164
330	311	673	4295137	330	311	673	3916324
380	361	773	3641079	380	361	773	3123200
430	411	873	2973930	430	411	873	2584748
480	461	973	2472762	480	461	973	2130140
530	511	1073	1995562	530	511	1073	1679304
580	561	1173	1728693	580	561	1173	1335496
630	611	1273	1317678	630	611	1273	1149484
680	661	1373	1163524	680	661	1373	896388
D10_Mn' at 50 °C				IFP_D10_Mn' at 50 °C			
30	11	73	14691436	30	11	73	18792088
60	41	133	13212680	60	41	133	16079740
90	71	193	10819840	90	71	193	13179672

Table B16 (cont'd)

120	101	253	8627060	120	101	253	10427644
150	131	313	6565180	150	131	313	8130728
180	161	373	4996032	180	161	373	6341088
210	191	433	3810804	210	191	433	5036468
240	221	493	2813868	240	221	493	4125032
270	251	553	2169360	270	251	553	3381408
300	281	613	1846720	300	281	613	2802932
330	311	673	1494368	330	311	673	2594356
360	341	733	1236584	360	341	733	2240276
390	371	793	1138232	390	371	793	2142636
420	401	853	891008	420	401	853	1885260

APPENDIX C

Additional PRE data

In this section, the PRE data of samples prepared by **Method 1** as described in Chapter 5 and with 20% Mn^{2+} are shown. Since the amount of Mn^{2+} that binds to membrane surface is unknown, the data is not reliable. Besides, the ^{2}H chanel is not on resonance, the data also showed fluctuation in echo intensity. All the samples contains 20 μmole DPPC-D10 or DPPC-D8, 5 μmole DPPG. The IFP bounded membrane samples contains 1 μmole IFP. The 20% Mn^{2+} are introduced before extrusion.

Table C1 D10 samples at 25 °C

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D10				D10_Mn			
100	50	212	2427552	100	50	212	1194804
120	70	252	2012156	120	70	252	979348
140	90	292	1854984	140	90	292	915672
160	110	332	1593588	160	110	332	962132
180	130	372	1370424	180	130	372	901556
200	150	412	1259636	200	150	412	714420
220	170	452	1112872	220	170	452	691116
240	190	492	951628	240	190	492	588460

Table C1 (cont'd)

260	210	532	905464	260	210	532	585264
280	230	572	797932	280	230	572	483384
IFP_D10				IFP_D10_Mn			
100	50	212	580796	100	50	212	1588944
120	70	252	333272	120	70	252	1158172
140	90	292	208024	140	90	292	844656
160	110	332	119904	160	110	332	846028
180	130	372	76492	180	130	372	947076
200	150	412	13092	200	150	412	921324
220	170	452	16912	220	170	452	1008936
240	190	492	1580	240	190	492	1061484
260	210	532	-8500	260	210	532	1054540
280	230	572	-14516	280	230	572	1016268

File locations of data shown in this table are:

(d) /home/khare0/mb4b/data/Shuang/quecho/D10/121514_array_25 (D10)

(e) /home/khare0/mb4b/data/Shuang/quecho/D10_Mn/122214_array_25 (D10_Mn)

(f) /home/khare0/mb4b/data/Shuang/quecho/IFP_D10/123014_array_25 (IFP_D10)

(g) /home/khare0/mb4b/data/Shuang/quecho/IFP_D10_Mn/010615_array_25 (IFP_D10_Mn)

Table C2 D10 samples at 50 °C

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D10				D10_Mn			
100	50	212	1638944	100	50	212	842488
120	70	252	1265572	120	70	252	1043264
140	90	292	977224	140	90	292	970648
160	110	332	787708	160	110	332	879260
180	130	372	661260	180	130	372	892140
200	150	412	527400	200	150	412	747548
220	170	452	415532	220	170	452	639164
240	190	492	369428	240	190	492	590300
260	210	532	338308	260	210	532	445184
280	230	572	235244	280	230	572	424200
IFP_D10				IFP_D10_Mn			
100	50	212	451708	100	50	204	1845600
120	70	252	876512	120	70	244	1590460
140	90	292	898724	140	90	284	1433212
160	110	332	777956	160	110	324	1016560
180	130	372	651280	180	130	364	835572
200	150	412	628744	200	150	404	840264
220	170	452	605104	220	170	444	862324

Table C2 (cont'd)

240	190	492	398048	240	190	484	885156
260	210	532	240316	260	210	524	941796
280	230	572	117136	280	230	564	955076

File locations of data shown in this table are:

(a) /home/khare0/mb4b/data/Shuang/quecho/D10/121614_array_50 (D10)

(b) /home/khare0/mb4b/data/Shuang/quecho/D10_Mn/122714_array_50 (D10_Mn)

(c) /home/khare0/mb4b/data/Shuang/quecho/IFP_D10/123114_array_50 (IFP_D10)

(d) /home/khare0/mb4b/data/Shuang/quecho/IFP_D10_Mn/010715_array_50 (IFP_D10_Mn)

Table C3 D8 samples at 25 °C

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D8				D8_Mn			
100	50	212	3327919	100	50	212	2228917
120	70	252	2961711	120	70	252	1806647
140	90	292	2621768	140	90	292	1458881
160	110	332	2257266	160	110	332	1258533
180	130	372	1924261	180	130	372	992903
200	150	412	1731391	200	150	412	871658
220	170	452	1573543	220	170	452	788846
240	190	492	1312340	240	190	492	699777
260	210	532	1100150	260	210	532	528551
280	230	572	968456	280	230	572	488626
IFP_D8				IFP_D8_Mn			
100	50	212	5243646	100	50	212	3119518
120	70	252	4598274	120	70	252	2432473
140	90	292	4009754	140	90	292	1999605
160	110	332	3486237	160	110	332	1608763
180	130	372	2977790	180	130	372	1350645
200	150	412	2560767	200	150	412	1172473
220	170	452	2196623	220	170	452	1001531

Table C3 (cont'd)

240	190	492	1879294	240	190	492	871345
260	210	532	1582613	260	210	532	683008
280	230	572	1370021	280	230	572	579342
				300	250	612	519509
				320	270	652	448488
				340	290	692	390976

File locations of data shown in this table are:

(a) /home/hapi0/mb4c/data/Shuang/quecho/D8/012015_array_25 (D8)

(b) /home/hapi0/mb4c/data/Shuang/quecho/D8_Mn/012215_array_25 (D8_Mn)

(c) /home/hapi0/mb4c/data/Shuang/quecho/IFP_D8/012515_array_25 (IFP_D8)

(d) /home/hapi0/mb4c/data/Shuang/quecho/IFP_D8_Mn/013015_array_25 (IFP_D8_Mn)

Table C4 D8 samples at 50 °C

τ1	τ2	t	FID intensity	τ1	τ2	t	FID intensity
D8				D8_Mn			
100	50	212	1891933	100	50	212	2591942
120	70	252	1305385	120	70	252	1684197
140	90	292	889396	140	90	292	1121416
160	110	332	644175	160	110	332	729318
180	130	372	502305	180	130	372	577378
200	150	412	349133	200	150	412	421690
220	170	452	274068	220	170	452	268906
240	190	492	257280	240	190	492	285795
260	210	532	147527	260	210	532	265650
280	230	572	175812	280	230	572	201262
IFP_D8				IFP_D8_Mn			
100	50	212	2496467	100	50	212	2600675
120	70	252	1992008	120	70	252	1581667
140	90	292	1563270	140	90	292	1063378
160	110	332	1295604	160	110	332	681049
180	130	372	1080933	180	130	372	491643
200	150	412	927278	200	150	412	347409
220	170	452	792277	220	170	452	250122

Table C4 (cont'd)

240	190	492	652293	240	190	492	243333
260	210	532	573384	260	210	532	238036
280	230	572	504266	280	230	572	153824

File locations of data shown in this table are:

(a) /home/hapi0/mb4c/data/Shuang/quecho/D8/012115_array_50 (D8)

(b) /home/hapi0/mb4c/data/Shuang/quecho/D8_Mn/012315_array_50 (D8_Mn)

(c) /home/hapi0/mb4c/data/Shuang/quecho/IFP_D8/012615_array_50 (IFP_D8)

(d) /home/hapi0/mb4c/data/Shuang/quecho/IFP_D8_Mn/012815_array_50 (IFP_D8_Mn)

REFERENCES

REFERENCES

[1] Levy, J. A. (1993) PATHOGENESIS OF HUMAN-IMMUNODEFICIENCY-VIRUS INFECTION, *Microbiol. Rev. 57*, 183-289.

[2] Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., and Kawaoka, Y. (1992) EVOLUTION AND ECOLOGY OF INFLUENZA-A VIRUSES, *Microbiol. Rev. 56*, 152-179.

[3] Engen, J. R. (2009) Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS, *Anal. Chem. 81*, 7870-7875.

[4] Curtis-Fisk, J., Spencer, R. M., and Weliky, D. P. (2008) Isotopically labeled expression in E-coli, purification, and refolding of the full ectodomain of the influenza virus membrane fusion protein, *Protein Expr. Purif. 61*, 212-219.

[5] Ghosh, U., Xie, L., Jia, L. H., Liang, S., and Weliky, D. P. (2015) Closed and Semiclosed Interhelical Structures in Membrane vs Closed and Open Structures in Detergent for the Influenza Virus Hemagglutinin Fusion Peptide and Correlation of Hydrophobic Surface Area with Fusion Catalysis, *J. Am. Chem. Soc. 137*, 7548-7551.

[6] Lorieau, J. L., Louis, J. M., and Bax, A. (2010) The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface, *Proc. Natl. Acad. Sci. U. S. A. 107*, 11341-11346.

[7] Qiao, H., Armstrong, R. T., Melikyan, G. B., Cohen, F. S., and White, J. M. (1999) A specific point mutant at position 1 of the influenza hemagglutinin fusion peptide displays a hemifusion phenotype, *Mol. Biol. Cell 10*, 2759-2769.