SOLID STATE NUCLEAR MAGNETIC RESONANCE STUDIES OF STRUCTURES AND MEMBRANE LOCATIONS OF PEPTIDES

By

Li Xie

A DISSERTATION

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

Chemistry - Doctor of Philosophy

2014

ABSTRACT

SOLID STATE NUCLEAR MAGNETIC RESONANCE STUDIES OF STRUCTURES AND MEMBRANE LOCATIONS OF PEPTIDES

By

Li Xie

Solid state nuclear magnetic resonance (SSNMR) can be used to study the structures of molecules such as small peptides and large proteins. Different structures correspond to different chemical environments and thus yield different chemical shifts in NMR spectra. In addition, SSNMR can also be used to probe membrane locations of peptides and proteins and provide insights into their biological functions. This dissertation mainly focuses on the structural and membrane location studies of peptides by rotational-echo double-resonance (REDOR) SSNMR, which is a technique for measuring distances between two coupled hetero nuclei such as ^{13}C and ^{15}N.

Influenza fusion peptide (IFP) is the N-terminal peptide of the HA2 subunit of the influenza hemagglutinin (HA) protein and this peptide plays an important role in the membrane fusion between the virus and the endosome of the host cell. There are 15 different HA subtypes. I studied the structure of membrane-associated H3 subtype IFP (H3_IFP) by ^{13}C-^{15}N REDOR SSNMR. SIMPSON simulations of the data indicate that H3_IFP adopts predominantly closed and semi-closed structures in membranes.

Similar to IFP, human immunodeficiency virus (HIV) fusion peptide (HFP) is the N-terminal peptide of the viral gp41 fusion protein and this peptide plays a key role in the HIV-host cell membrane fusion. HFP adopts major antiparallel and minor parallel β

sheet structures in membranes. Earlier fluorescence spectroscopy and SSNMR studies support a strong positive correlation between the membrane insertion depth and fusogenicity of HFP. However, the ^{19}F labeling in earlier membrane location studies of HFP may perturb the membrane bilayer integrity. Due to this concern, ^{13}C-^{2}H REDOR was developed to detect the residue-specific membrane location of HFP, where one residue of HFP is backbone ^{13}CO labeled and the lipid is either perdeuterated or selectively deuterated in its acyl chains. Since ^{1}H and ^{2}H are chemically equivalent, there is no perturbation on membrane bilayer integrity as well as peptide-lipid interaction regardless of what fraction of deuterated lipid is used. The ^{13}C-^{2}H REDOR pulse sequence was optimized using the setup peptide I4.

The membrane locations of two peptides, HFP and KALP, were studied by ^{13}C-^{2}H REDOR, where KALP is a designed transmembrane α helical peptide. Fitting of the REDOR data revealed that both peptides have multiple locations within the membrane hydrocarbon core. The multiple locations are attributed to the snorkeling of lysine sidechains on both termini for KALP and to the distribution of antiparallel β sheet registries for HFP. HFP has a ~0.7 fraction of deep insertion and a ~0.3 fraction of shallow insertion in the membrane. The predominant deep insertion of HFP may significantly perturb the membrane bilayer structure and lower the activation energy of membrane fusion, which is consistent with the observed positive correlation between the membrane insertion depth and fusion rate for different HFP constructs.

Dedicated to Qingwei Li and Elaine Q. Xie

ACKNOWLEDGMENTS

First of all, I would like to thank my research advisor, Prof. David Weliky, for his tremendous guidance, support, and encouragement throughout my Ph.D. career. He taught me not only NMR principles but also how to think as a real scientist. He has been very patient and I do not remember for how many times that he has spent a couple of hours teaching me in his office over the past few years. He always gave me helpful suggestions and encouragement while I was stuck in NMR troubleshooting. He has been a very supportive advisor and I really appreciate this.

I would also like to thank my committee members, Prof. Daniel Jones, Prof. John McCracken, and Prof. James Geiger, for their valuable inputs and guidance over the past years. Prof. Daniel Jones also gave me very helpful suggestions for job searching, and he has always been generous with his time and insights.

I am grateful for the great help from Dr. Daniel Holmes and Mr. Kermit Johnson at the Max T. Rogers NMR facility center at MSU. Dr. Daniel Holmes has been a very good source for discussing solution state NMR and troubleshooting the instrument.

I am also grateful for the support and help from both the former and current Weliky group members, Dr. Kelly Sackett, Dr. Charles Gabrys, Dr. Matthew Nethercott, Dr. Scott Schmick, Dr. Erica Vogel, Koyeli Banerjee, Punsisi Ratnayake, Ujjayini Ghosh, Shuang Liang, Lihui Jia, and Robert Wolfe. They have been very helpful not only in research but also in many other aspects.

Finally, I would like to thank my wife, Qingwei Li, for her support and unselfish love during my Ph.D. career. It is her support and encouragement that have made my

graduate study much easier than it would have been on my own. I would also like to thank my family and friends for their support and help in many different ways.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

CHCA: α-cyano-4-hydroxycinnamic acid

CHR: C-terminal heptad repeat region

CSA: chemical shift anisotropy

CP: cross polarization

DCM: dichloromethane

DHB: 2,5-dihydroxybenzoic acid

DIEA: N,N-diisopropylethylamine

DMF: N,N-dimethylformamide

DMPC: dimyristoylphosphatidylcholine

DPC: dodecylphosphocholine

DPPC: dipalmitoylphosphatidylcholine

DPPG: dipalmitoylphosphoglycerol

DTPC: 1,2-di-O-tetradecyl-sn-glycero-3-phosphocholine

DTPG: 1,2-di-O-tetradecyl-sn-glycero-3-[phospho-(1'-rac-glycerol)]

EPR: electron paramagnetic resonance

ESI: electrospray ionization

Fmoc: 9-fluorenylmethyloxycarbonyl

FID: free induction decay

FRET: fluorescence resonance energy transfer

HA: hemagglutinin protein

HBTU: 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HFP: HIV fusion peptide

HIV: human immunodeficiency virus

HOBt: 1-hydroxybenzotriazole

HPLC: high pressure liquid chromatography

HSQC: heteronuclear single quantum correlation

IFP: influenza fusion peptide

IR: Infrared spectroscopy

LWHM: line width at half maximum

MALDI: matrix-assisted laser desorption ionization

MES: 2-(N-morpholino)ethanesulfonic acid

MPER: membrane-proximal external region

n.a.: natural abundance

NHR: N-terminal heptad repeat region

NMR: nuclear magnetic resonance

NOE: Nuclear Overhauser effect

PAF: principal axis frame

POPC: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

POPG: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol

PHI: pre-hairpin intermediate

REDOR: rotational-echo double-resonance

r.f. (or RF): radio frequency

SHB: six helix bundle

SPPS: solid phase peptide synthesis

SSNMR: solid state nuclear magnetic resonance

t-Boc: tert-butyloxycarbonyl

TFA: trifluoroacetic acid

TIS: triisopropylsilane

TPPM: two pulse phase modulation

XRD: X ray diffraction

Chapter 1 - Introduction

1.1 NMR background

NMR interactions

For any NMR-active nuclei with nuclear spin quantum number I ≠ 0, they can interact with the electric and magnetic fields and the interactions can be divided into two categories, external and internal. The external interactions are between the nuclear spin and external fields such as the static magnetic field $\mathbf{B_o}$ ($\mathbf{B_o} = B_o\mathbf{z}$) and radio frequency (r.f.) pulse field $\mathbf{B_1}$.[a] The internal interactions are between the nuclear spin and the intrinsic fields such as J-coupling field $\mathbf{B_J}$ and chemical shift field $\mathbf{B_{CS}}$. The total Hamiltonian of the spin is

$$\hat{H}_t = \hat{H}_{ext} + \hat{H}_{int} = (\hat{H}_Z + \hat{H}_{RF}) + (\hat{H}_J + \hat{H}_{CS} + \hat{H}_D + \hat{H}_Q) \quad (1.1)$$

where

$\hat{H}_Z$ is the Hamiltonian of Zeeman interaction between the spin and $\mathbf{B_o}$ field;

$\hat{H}_{RF}$ is the Hamiltonian of the spin interaction with the $\mathbf{B_1}$ field;

$\hat{H}_J$ is the Hamiltonian of J-coupling interaction (indirect nuclear dipole-dipole coupling) between two spins;

---

[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.

$\hat{H}_{CS}$ is the Hamiltonian of the spin interaction with chemical shift field (electronic shielding field) induced by $\mathbf{B_O}$ field;

$\hat{H}_D$ is the Hamiltonian of direct nuclear dipole-dipole coupling between two spins;

$\hat{H}_Q$ is the Hamiltonian of quadrupolar interaction between the electric quadrupole moment of a quadrupolar nucleus (I > 1/2) and the surrounding electric field gradient.

Zeeman interaction

For a nucleus having a spin quantum number I (I ≠ 0), the nuclear spin energy splits into (2I + 1) levels in the static magnetic field $\mathbf{B_O}$. This interaction is referred to as Zeeman interaction and the Hamiltonian is[1]

$$\hat{H}_Z = -\hat{\mu} \cdot \mathbf{B_O} \tag{1.2}$$

where $\hat{\mu}$ is the nuclear magnetic moment operator and is given by

$$\hat{\mu} = \gamma\hbar\hat{I} = \gamma\hbar\,[\mathbf{i}\,\hat{I}_x + \mathbf{j}\,\hat{I}_y + \mathbf{k}\,\hat{I}_z] \tag{1.3}$$

where γ is the nuclear gyromagnetic ratio, $\hbar$ is the Planck constant divided by 2π, $\hat{I}$ is the nuclear spin operator, and **i**, **j**, and **k** are the unit vectors in the x, y, and z axes, respectively.

Incorporation of equation 1.3 into equation 1.2 yields

$$\hat{H}_Z = -\gamma\hbar\hat{I}_z B_O \tag{1.4}$$

The eigenfunctions of $\hat{H}_Z$ are the wavefunctions describing the (2I + 1) states of the spin

in the $\mathbf{B_o}$ field. The eigenvalues of $\hat{H}_z$ are the energies associated with different states of the spin and are obtained through

$$\hat{H}_z \,|I,m\rangle = -\gamma\hbar B_o \hat{I}_z \,|I,m\rangle \qquad (1.5)$$

where $|I,m\rangle$ is an eigenfunction of $\hat{I}_z$ with an eigenvalue of m, i.e.

$$\hat{I}_z \,|I,m\rangle = m\,|I,m\rangle \qquad (1.6)$$

Incorporation of equation 1.6 into equation 1.5 yields

$$\hat{H}_z \,|I,m\rangle = -\gamma\hbar B_o m\,|I,m\rangle = E_{I,m}\,|I,m\rangle \qquad (1.7)$$

where $E_{I,m}$ is the energy associated with the spin state m. For a spin-1/2 nucleus such as ^{13}C, possible values of m are +1/2 and -1/2. For the m = +1/2 state, i.e. α state, the spin energy $E_\alpha = -\frac{1}{2}\gamma\hbar B_o$. For the m = -1/2 state, i.e. β state, the spin energy $E_\beta = \frac{1}{2}\gamma\hbar B_o$. The energy difference between α and β states is

$$\Delta E = E_\beta - E_\alpha = \gamma\hbar B_o \qquad (1.8)$$

The nuclear magnetic moments associated with the α and β states are displayed in Figure 1.1. Each magnetic moment $\boldsymbol{\mu}$ experiences a torque $\boldsymbol{\Gamma} = \boldsymbol{\mu}$ x $\mathbf{B_o}$ from the $\mathbf{B_o}$ field and precesses about $\mathbf{B_o}$ clockwise with an angular frequency of γB_o. This precession is called Larmor precession and the corresponding frequency is called Larmor frequency. At the thermal equilibrium, the magnetization **M**, which is the vector sum of all individual

magnetic moments of all nuclei of a specific type (e.g. ^{13}C), is along the direction of $\mathbf{B_o}$, i.e., z axis.

Since the static magnetic field $\mathbf{B_o}$ is usually orders of magnitude greater than the internal magnetic fields such as dipolar field and chemical shift field, Zeeman interaction is much stronger than the internal interactions. In this case, the internal interactions are treated as perturbations of Zeeman interaction. For any type of internal interaction, its Hamiltonian can be divided into two parts, the secular part which commutes with Zeeman Hamiltonian and the non-secular part which does not. Since only the secular part affects the observable spectrum to the first order, we do not consider the non-secular part, i.e., the non-secular part of an internal Hamiltonian is truncated. This type of treatment is referred to as secular approximation.

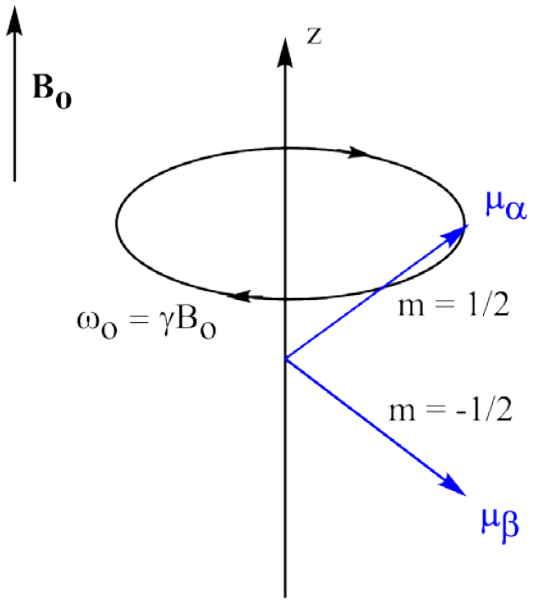

Figure 1.1 The two spin states of a spin-1/2 nucleus in the static magnetic field $\mathbf{B_o}$. The field $\mathbf{B_o}$ is along z axis. The two magnetic moments μ_α and μ_β (blue) are under Larmor precession about the field $\mathbf{B_o}$ with an angular frequency of $\omega_O = \gamma B_O$.

Interaction with B_1 field (r.f. pulses)

In a SSNMR probe, the solenoid coil generates a $\mathbf{B_1}$ field when applying r.f. pulses. Take a 90_X pulse for example, the $\mathbf{B_1}$ field is

$$\mathbf{B_1}(t) = B_1\cos(\omega t)\mathbf{x} \tag{1.9}$$

where $\omega = 2\pi\nu$ and ν is the frequency of the 90_X pulse set by the spectrometer, and **x** is the unit vector along x axis. This oscillating $\mathbf{B_1}$ field can be divided into two components, the resonant part $\mathbf{B_1^{res}}$ and non-resonant part $\mathbf{B_1^{non\text{-}res}}$:

$$\mathbf{B_1^{res}} = \frac{1}{2} B_1[\cos(\omega t)\mathbf{x} - \sin(\omega t)\mathbf{y}] \tag{1.10}$$

$$\mathbf{B_1^{non\text{-}res}} = \frac{1}{2} B_1[\cos(\omega t)\mathbf{x} + \sin(\omega t)\mathbf{y}] \tag{1.11}$$

The $\mathbf{B_1^{res}}$ component rotates clockwise in the xy plane and $\mathbf{B_1^{non\text{-}res}}$ rotates counterclockwise. As previously discussed, the magnetic moment of each spin state is precessing clockwise about the $\mathbf{B_o}$ field. Therefore, only the $\mathbf{B_1^{res}}$ component affects the nuclear spins since it rotates in the same sense as the spin precession (Larmor precession) while the $\mathbf{B_1^{non\text{-}res}}$ component does not.

In SSNMR, we usually run experiments under magic angle spinning (refer to the "MAS" section for details), where the solenoid coil in the probe has an angle of 54.7^o

(magic angle) with the $\mathbf{B_o}$ field. In this case, for a $\mathbf{B_1}$ field generated along the solenoid coil axis, only a ~0.4 fraction (0.5 x $\sin 54.7^{o}$) of it is utilized in the experiments.

For a magnetization **M** under the $\mathbf{B_1}$ field of a r.f. pulse, it experiences a torque $\Gamma = \mathbf{M} \times \mathbf{B_1}$ and precesses about the $\mathbf{B_1}$ field with an angular frequency of γB_1. Such a precession is called Rabi precession and the frequency is called Rabi frequency. The direction of the torque Γ can be determined using the right-hand rule: first point the four fingers of your right hand along **M**, then curl your fingers toward the direction of $\mathbf{B_1}$, the thumb is now pointing along the torque Γ. For instance, if **M** is along z axis, $\mathbf{B_1}$ is along x axis, then Γ is along y axis. For any r.f. pulse, its flip angle $\theta = \gamma B_1 \tau_p$, where τ_p is the duration of the pulse. $\theta = \pi/2$ for a 90_x pulse and π for a 180_x pulse, where x is the r.f. pulse phase and refers to a $\mathbf{B_1}$ field along the x axis. Figure 1.2 displays the effects of a 90_x pulse and a 180_x pulse on the magnetization **M** which is initially along z axis at the thermal equilibrium. For an x pulse, its Hamiltonian is

$$\hat{H}_{RF} = -\hbar \omega_1 \hat{I}_x = -\gamma \hbar B_1 \hat{I}_x \qquad (1.12)$$

Note that equation 1.12 corresponds to the r.f. pulse Hamiltonian in the rotating frame (refer to the "Rotating frame" section for more details) since the term B_1 is static, as opposed to equation 1.9, which corresponds to the laboratory frame where the $\mathbf{B_1}$ field is oscillating as a function of time.

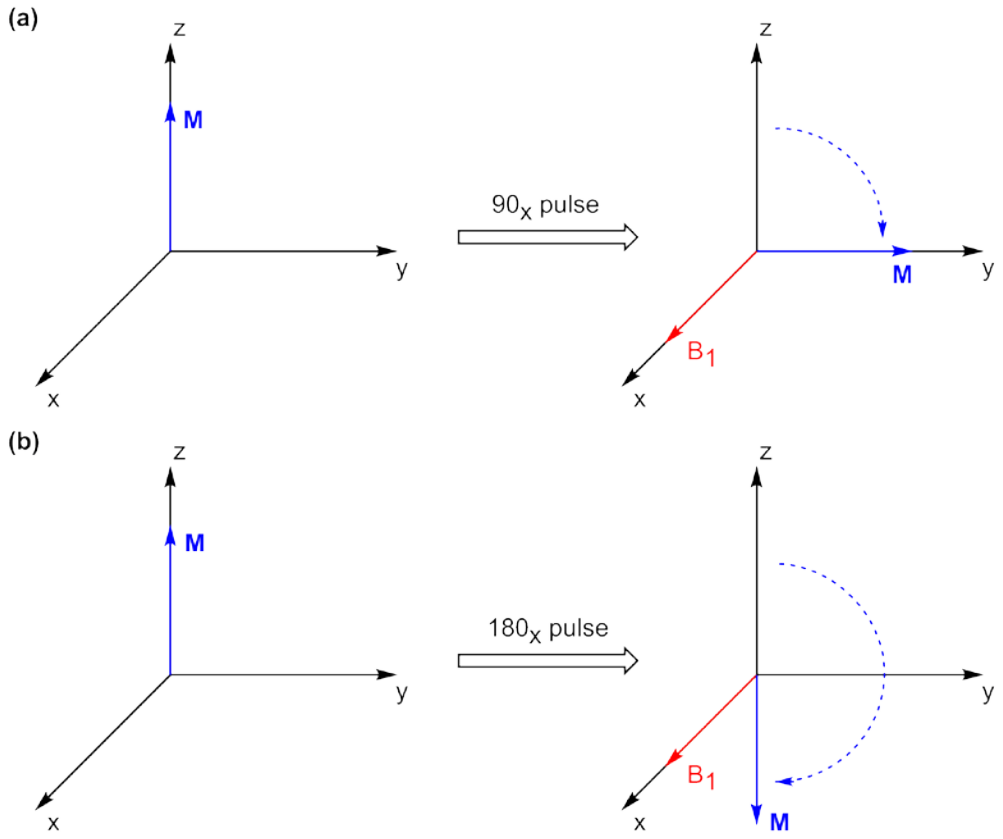

Figure 1.2 Rabi precession of the magnetization **M** (initially along z axis) about the $\mathbf{B_1}$ field of a 90_X pulse (a) and a 180_X pulse (b).

J-coupling interaction

J-coupling, sometimes also known as scalar coupling, is the indirect nuclear dipole-dipole interaction through bonding electrons. There are several mechanisms proposed for J-coupling, of which the most important one is the Fermi contact mechanism[2].

According to this mechanism, both bonding electrons in the X-Y bond (X and Y are both NMR-active, i.e. $I \neq 0$) spend a finite time at the point of nuclei X and Y, in which case the electron and the nucleus are said to be in Fermi contact. In the X-Y bond, it is more stable (i.e. lower energy) when nucleus X (or Y) and the electron in Fermi contact have

antiparallel angular momentum. Since the gyromagnetic ratio is usually positive for a nucleus (e.g. ^{13}C, ^{1}H, etc) and negative for an electron, it means it is more stable when the magnetic moments of the nucleus and electron are parallel. Therefore, when nucleus Y is spin-up, the electron in Fermi contact will be also up. According to the Pauli exclusion principle, the other electron in the X-Y bond must be down since the two electrons occupy the same molecular orbital. In this case, the more favorable spin state of nucleus X will be down. Similarly, when Y is spin-down, the more favorable spin state of X will be up. This tells us the α state of X is more favored (i.e. lower energy) when Y is in β state and less favored (i.e. higher energy) when Y is in α state due to the X-Y J-coupling interaction. Similarly, the β state of X is more favored when Y is in α state and less favored when Y is in β state. The nuclear and electron spins in the X-Y bond and the spin energy diagram of X with and without the X-Y J-coupling interaction are displayed in Figure 1.3a. If Y is a spin 1/2 nucleus, the X-Y J-coupling will result in two X peaks in the spectrum with a spacing equal to the J-coupling constant, J_{XY} (Figure 1.3b).

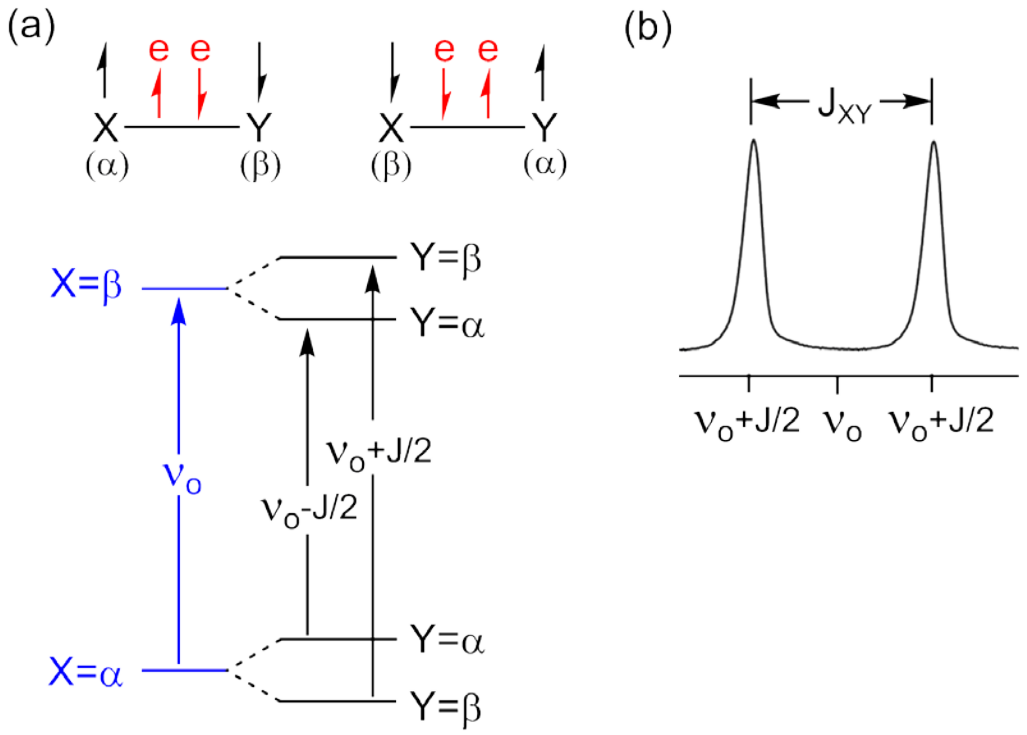

Figure 1.3 J-coupling mechanism. (a) Nuclear and electron spins in the X-Y bond and the spin energy diagram of nucleus X with (black) and without (blue) the X-Y J-coupling interaction. (b) Peak splitting of nucleus X in the spectrum due to the J-coupling between X and Y, where Y is a spin 1/2 nucleus.

The Hamiltonian of J-coupling between two spins I_j and I_k is

$$\hat{H}_J = 2\pi \mathbf{J}_{jk} \hat{\mathbf{I}}_j \cdot \hat{\mathbf{I}}_k \qquad (1.13)$$

where $\hat{\mathbf{I}}_j$ and $\hat{\mathbf{I}}_k$ are the nuclear spin operators and $\mathbf{J}_{jk}$ is the J-coupling tensor. $\mathbf{J}_{jk}$ is a 3 x 3 matrix depending on the molecular orientation but it reduces to a number in isotropic liquids. That is why sometimes J-coupling is also called scalar coupling.

J-coupling is the most important spin interaction for solution NMR but not considered in most cases for SSNMR since it is usually the weakest spin interaction existing in solids and is sometimes eliminated by decoupling pulses.

Isotropic and anisotropic chemical shift interactions

In a molecule, the electrons are either surrounding the nuclei or located at the chemical bonds. For a molecule in the presence of the external magnetic field $\mathbf{B_0}$, the electrons which are moving about a nucleus or within a chemical bond will also move under the Lorentz force from $\mathbf{B_0}$. This type of electron motion generates a magnetic field called shielding field or chemical shift field. Since different molecules in a sample have different orientations with respect to the $\mathbf{B_0}$ field, the directions of the induced electric current and shielding field will also vary from molecule to molecule. Although the shielding field could be in different directions relative to $\mathbf{B_0}$, only the component along $\mathbf{B_0}$ is relevant according to the secular approximation. In most cases, the shielding field decreases the magnitude of $\mathbf{B_0}$ experienced by a nucleus and thus moves the chemical shift of the nucleus to a higher value. However, in a paramagnetic species which contains unpaired electron(s), the shielding field may increase (add to) the magnitude of $\mathbf{B_0}$ experienced by a nucleus. The Hamiltonian of shielding interaction (chemical shift interaction) acting on a spin I is

$$\hat{H}_{CS} = -\gamma\hbar B_0 \boldsymbol{\sigma} \cdot \hat{\mathbf{I}} \qquad (1.14)$$

where $\hat{\mathbf{I}}$ is the nuclear spin operator and $\boldsymbol{\sigma}$ is the shielding tensor which is a second-rank tensor and can be represented by a 3 x 3 matrix. To understand $\boldsymbol{\sigma}$, we can take a look at the electron distribution around a nucleus in a molecule. In general, the electron distribution is not spherically symmetric. In this case, the size of the shielding depends

on the molecular orientation with respect to the external magnetic field $\mathbf{B_o}$. The shielding tensor $\boldsymbol{\sigma}$ is then used to describe how the size of shielding varies with the molecular orientation. There is an axis frame associated with the shielding tensor $\boldsymbol{\sigma}$, called principal axis frame (PAF), in which there are three principal axes, x^{PAF}, y^{PAF}, and z^{PAF} (Figure 1.4). The principal values associated with the three axes in PAF are σ_{xx}, σ_{yy}, and σ_{zz} , respectively. The orientation of PAF is determined by the electronic structure of the part containing the nucleus of interest (e.g. ^{13}C in the ^{13}CO group) in a molecule and is fixed with respect to the molecule.

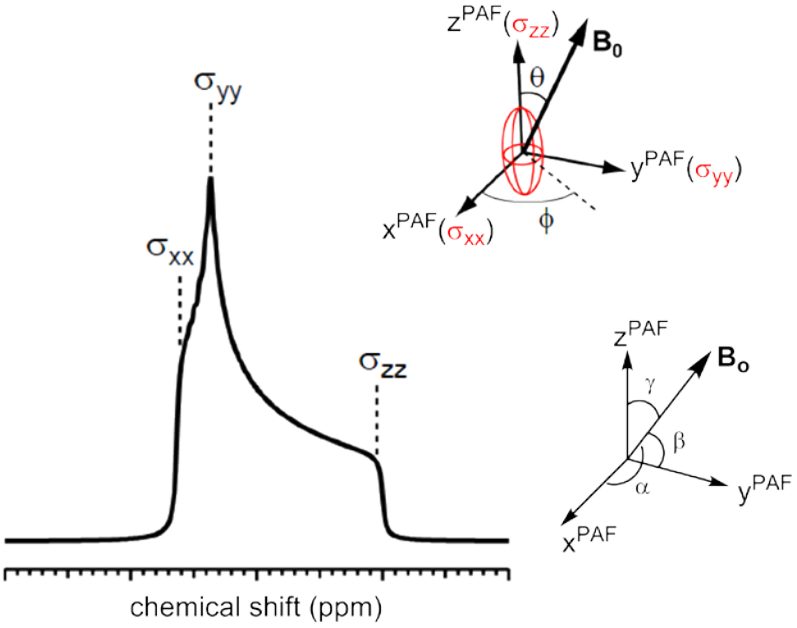

Figure 1.4 Principal axis frame (PAF) and the shielding tensor (red). The two polar angles defining the orientation of $\mathbf{B_o}$ in PAF are θ and ϕ, where θ is the angle between $\mathbf{B_o}$ and the principal axis z^{PAF} and ϕ is the angle between the projection of $\mathbf{B_o}$ in the $(xy)^{PAF}$ plane and the principal axis x^{PAF}. The three principal values associated with the three axes in PAF are σ_{xx}, σ_{yy}, and σ_{zz} , which also correspond to the three principal chemical shifts in the CSA (chemical shift anisotropy) powder pattern. The PAF frame at the lower right corner displays the angles α, β, and γ used in equation 1.15.

In general, the chemical shift of a nucleus depends on the orientation with respect to the external magnetic field $\mathbf{B_0}$. To better understand this, let us take ^{13}CO for example. The backbone ^{13}CO chemical shift of a peptide or protein depends on the orientation of its shielding tensor $\boldsymbol{\sigma}$ relative to $\mathbf{B_0}$. The three PAF axes of $\boldsymbol{\sigma}$ are displayed in Figure 1.5a. Assuming the angles between $\mathbf{B_0}$ and the three axes x^{PAF}, y^{PAF}, and z^{PAF} are respectively α, β, and γ (refer to the PAF frame at the lower right corner of Figure 1.4), the ^{13}CO chemical shift is then defined as:

$$\sigma = \sigma_{xx}\cos^2\alpha + \sigma_{yy}\cos^2\beta + \sigma_{zz}\cos^2\gamma \qquad (1.15)$$

where σ_{xx}, σ_{yy}, and σ_{zz} are the three principal values associated with PAF and are also the three principal chemical shifts displayed in Figure 1.5b. For solids under MAS (refer to the "MAS" section for details) or solutions with rapid molecular tumbling, the isotropic chemical shift results and is given by

$$\sigma_{iso} = \frac{1}{3}(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}) \qquad (1.16)$$

This is because the time-average values $\langle\cos^2\alpha\rangle = \langle\cos^2\beta\rangle = \langle\cos^2\gamma\rangle = \frac{1}{3}$ under MAS or rapid molecular tumbling. Without MAS or rapid molecular tumbling, the backbone ^{13}CO of a peptide or protein may yield a CSA powder pattern similar to the one in Figure 1.5b (the red spectrum).

The chemical shift field (shielding field) can be divided into two components, isotropic field $\mathbf{B}_{CS,iso}$ and anisotropic field $\mathbf{B}_{CS,aniso}$. The total chemical shift Hamiltonian is

$$\hat{H}_{cs} = \{-\sigma_{iso}\gamma\hbar B_o + \frac{1}{2}\delta_{cs}\hbar\ [3\cos^2\theta - 1 - \eta_{cs}\sin^2\theta\cos(2\phi)]\}\hat{I}_z \qquad (1.17)$$

where σ_{iso} is the isotropic chemical shift, $\delta_{cs} = -\gamma B_o(\sigma_{zz} - \sigma_{iso})$, $\eta_{cs} = (\sigma_{yy} - \sigma_{xx})/(\sigma_{zz} - \sigma_{iso})$ and is called asymmetry parameter, θ and ϕ are the polar angles of $\mathbf{B_o}$ in the PAF frame (refer to Figure 1.4)[3]. The first term in equation 1.17 corresponds to isotropic chemical shift interaction whereas the second term corresponds to anisotropic chemical shift interaction. When the nucleus is at a site of axial symmetry, the shielding tensor has a PAF in which the z^{PAF} axis coincides with the symmetry axis and thus $\sigma_{xx} = \sigma_{yy} \neq \sigma_{zz}$. In this case, the asymmetry parameter $\eta_{cs} = 0$.

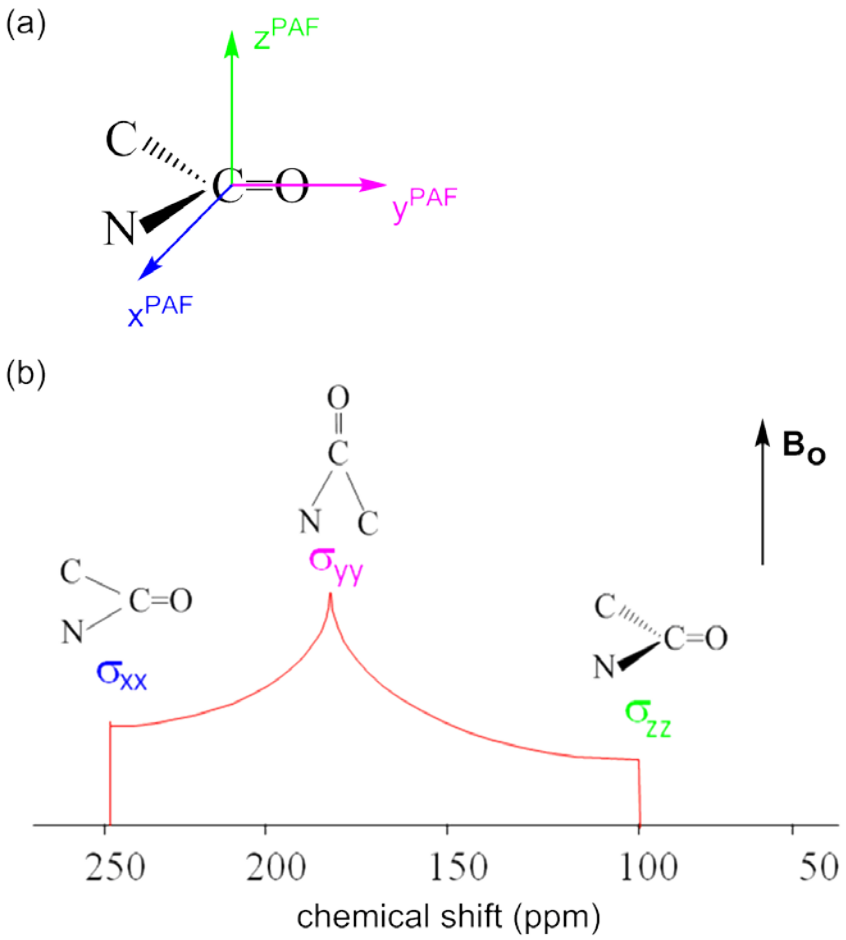

Figure 1.5 (a) The PAF frame of backbone ^{13}CO in a peptide or protein. The z^{PAF} axis is perpendicular to the C—CO—N plane. The x^{PAF} and y^{PAF} axes are both in the C—CO—N plane with x^{PAF} perpendicular to C=O and y^{PAF} parallel to C=O. (b) ^{13}CO CSA powder pattern. The three chemical shifts are σ_{xx} = 247 ppm with $\mathbf{B_0}$ parallel to x^{PAF}, σ_{yy} = 176 ppm with $\mathbf{B_0}$ parallel to y^{PAF}, and σ_{zz} = 99 ppm with $\mathbf{B_0}$ parallel to z^{PAF}.

Dipolar coupling interaction

When two nuclear spins are close to each other in space, the magnetic moment of one spin interacts with the magnetic field generated by the other spin, and vice versa. This type of interaction between two spins is called dipolar coupling and the field is called dipolar field. The dipolar field has a distance dependence of $1/r^3$ and an

orientation dependence of $(3\cos^2\theta\text{-}1)$ for the secular component, where r is the internuclear distance between the two spins and θ is the angle between $\mathbf{B_o}$ and the internuclear vector (Figure 1.6).

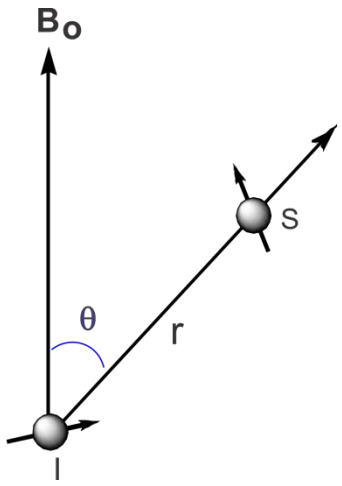

Figure 1.6 Definition of the angle θ and distance r for the dipolar interaction between nucleus I and S. $\mathbf{B_o}$ is the external magnetic field which is along the laboratory frame z axis.

There are two types of dipolar interaction: homonuclear and heteronuclear dipolar coupling. The Hamiltonian of homonuclear dipolar coupling between two spins I_1 and I_2 of the same isotope is

$$\hat{H}_{II} = -\frac{\mu_o}{4\pi}\hbar\frac{\gamma^2}{r^3}\frac{1}{2}(3\cos^2\theta - 1)(3\hat{I}_{1z}\hat{I}_{2z} - \hat{\mathbf{I}}_1\cdot\hat{\mathbf{I}}_2) \qquad (1.18)$$

where $\mu_o = 4\pi \times 10^{-7}$ H/m and is the permeability of free space, θ and r are defined in Figure 1.6, and $\hat{\mathbf{I}}_1$ and $\hat{\mathbf{I}}_2$ are two vector operators of spin 1 and 2, respectively. The Hamiltonian of heteronuclear dipolar coupling between two spins I and S of different isotopes is

$$\hat{H}_{IS} = -\frac{\mu_o}{4\pi}\hbar\frac{\gamma_I\gamma_S}{r^3}\frac{1}{2}(3\cos^2\theta - 1)2\hat{I}_z\hat{S}_z \qquad (1.19)$$

where the term $\frac{\mu_o}{4\pi}\hbar\frac{\gamma_I\gamma_S}{r^3}$ is called dipolar coupling constant (in units of rad/s). For the dipolar coupling constant between I and S in units of Hz, it is defined as

$$d = (\frac{\mu_o}{4\pi}\hbar\frac{\gamma_I\gamma_S}{r^3})/2\pi = \frac{\mu_o h\gamma_I\gamma_S}{16\pi^3 r^3} \qquad (1.20)$$

Using equation 1.20, we can derive that ^{13}C-^{15}N dipolar coupling in Hz equals $3066/r^3$ and ^{13}C-^{2}H dipolar coupling in Hz equals $4642/r^3$, where r is the distance in Å.

Notice that the spin part of equation 1.18 differs from that of equation 1.19, to understand this, we can first take a look at the interaction between Zeeman Hamiltonian and the homonuclear dipolar Hamiltonian. The Zeeman Hamiltonian for two same spins I_1 and I_2 is $\hat{H}_z = -\gamma\hbar Bo(\hat{I}_{1z} + \hat{I}_{2z})$. We can do the calculation to see whether or not the operator $\hat{\mathbf{I}}_1 \cdot \hat{\mathbf{I}}_2$ commutes with $\hat{H}_z$:

$$[\hat{H}_z, \hat{\mathbf{I}}_1 \cdot \hat{\mathbf{I}}_2] = -\gamma\hbar B_0[(\hat{I}_{1z} + \hat{I}_{2z}),(\hat{I}_{1x}\hat{I}_{2x} + \hat{I}_{1y}\hat{I}_{2y} + \hat{I}_{1z}\hat{I}_{2z})]$$

$$= -\gamma\hbar B_0\{\hat{I}_{1z}(\hat{I}_{1x}\hat{I}_{2x} + \hat{I}_{1y}\hat{I}_{2y} + \hat{I}_{1z}\hat{I}_{2z}) + \hat{I}_{2z}(\hat{I}_{1x}\hat{I}_{2x} + \hat{I}_{1y}\hat{I}_{2y} + \hat{I}_{1z}\hat{I}_{2z})\}$$

$$= -\gamma\hbar B_0\{[\hat{I}_{1z}, \hat{I}_{1x}]\hat{I}_{2x} + [\hat{I}_{1z}, \hat{I}_{1y}]\hat{I}_{2y} + [\hat{I}_{1z}, \hat{I}_{1z}]\hat{I}_{2z}$$

$$+[\hat{I}_{2z}, \hat{I}_{2x}]\hat{I}_{1x} + [\hat{I}_{2z}, \hat{I}_{2y}]\hat{I}_{1y} + [\hat{I}_{2z}, \hat{I}_{2z}]\hat{I}_{1z}\}$$

$$= -\gamma\hbar B_0\{i\hbar\hat{I}_{1y}\hat{I}_{2x} - i\hbar\hat{I}_{1x}\hat{I}_{2y} + 0 + i\hbar\hat{I}_{2y}\hat{I}_{1x} - i\hbar\hat{I}_{2x}\hat{I}_{1y} + 0\}$$

$= 0$ (1.21)

Therefore, the operator $\hat{\mathbf{I}}_1 \cdot \hat{\mathbf{I}}_2$ commutes with $\hat{H}_z$ for homonuclear dipolar coupling. However, for heteronuclear dipolar coupling, since the two spins I and S have different gyromagnetic ratios, the γ in equation 1.21 can no longer be put right in the front of the equation. As a result, the component $(\hat{I}_{1x}\hat{I}_{2x} + \hat{I}_{1y}\hat{I}_{2y})$ of the operator $\hat{\mathbf{I}}_1 \cdot \hat{\mathbf{I}}_2$ no longer commutes with Zeeman Hamiltonian of the two spins. According to the secular approximation discussed earlier, $\hat{I}_{1x}\hat{I}_{2x} + \hat{I}_{1y}\hat{I}_{2y}$ is the non-secular part of the heteronuclear dipolar Hamiltonian and is thus truncated. That is why the spin part of homonuclear dipolar Hamiltonian is $3\hat{I}_{1z}\hat{I}_{2z} - \hat{I}_1\hat{I}_2$ (equation 1.18) while for heteronuclear dipolar Hamiltonian it is $2\hat{I}_z\hat{S}_z$ (equation 1.19). The heteronuclear dipolar coupling will be discussed in more details in the "REDOR" section.

Quadrupolar interaction

If a nucleus has a spin quantum number greater than 1/2, it is referred to as a quadrupolar nucleus, e.g. ^{2}H (I = 1) and ^{23}Na (I = 3/2). The charge distribution of a quadrupolar nucleus is displayed in Figure 1.7 and results in a non-zero electric quadrupole moment eQ, where e is the charge carried by a proton and Q is specific to the nuclear isotope. The electric quadrupole moment interacts with the electric field gradient (resulting from the non-uniform electron density distribution) at the location of the nucleus and this interaction is known as quadrupolar coupling. The magnitude of quadrupolar coupling depends on the magnitudes of the electric quadrupole moment of the nucleus and the electric field gradient. For a quadrupolar nucleus, the quadrupolar

interaction affects the nuclear spin energy levels as other previously discussed magnetic interactions do.

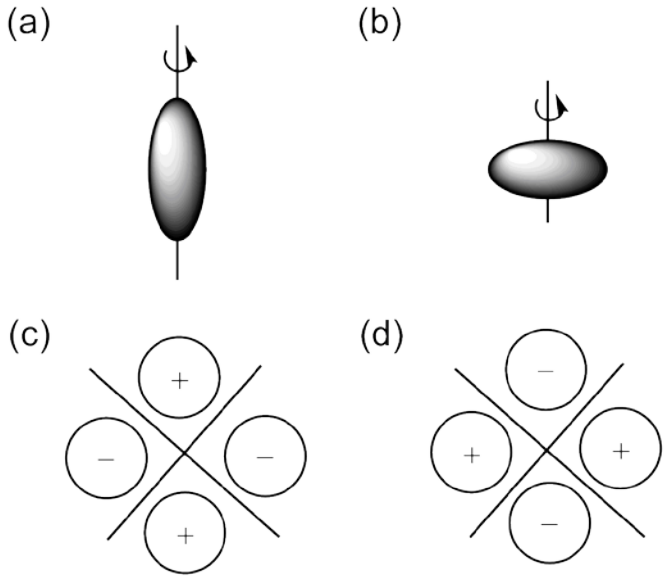

Figure 1.7 Electric properties of a quadrupolar nucleus. (a) prolate charge distribution. (b) oblate charge distribution. (c) electric quadrupole moment of (a) with a positive sign (e.g. it is 2.8 x 10^{-31} m^2 for ^{2}H). (d) electric quadrupole moment of (b) with a negative sign (e.g. it is -2.6 x 10^{-30} m^2 for ^{17}O)[4].

The quadrupolar Hamiltonian is

$$\hat{H}_Q = \frac{eQeq}{2I(2I-1)\hbar}\frac{1}{2}[3\cos^2\theta - 1 - \eta_Q\sin^2\theta\cos(2\phi)]\frac{1}{2}[3\hat{I}_z^2 - I(I+1)] \qquad (1.22)$$

where e is the elementary charge, Q is the magnitude of the electric quadrupole moment, q is a value associated with the electric field gradient tensor, I is the nuclear spin quantum number, θ and ϕ are the polar angles of the $\mathbf{B_0}$ field in the PAF, η_Q is the asymmetry parameter of the electric field gradient tensor, and $\hat{I}_z$ is the z-component of the nuclear spin operator. Since the electric field gradient tensor has an orientation dependence of $[3\cos^2\theta - 1 - \eta_Q\sin^2\theta\cos(2\phi)]$, the quadrupolar Hamiltonian also has this

dependence. The term $\frac{eQeq}{\hbar}$ in equation 1.22 is called quadrupolar coupling constant (QCC) and is in units of rad/s. As for the QCC in units of Hz, it is defined as

$$\chi = \frac{eQeq}{h} \tag{1.23}$$

For ^2H (I = 1) in aliphatic C-^2H bonds, χ is ~170 kHz[3]. To better understand the quadrupolar interaction, we will take ^2H in aliphatic C-^2H bonds as an example and discuss in some detail. The ^2H quadrupolar energy is given by

$$E_Q = \frac{eQeq}{2I(2I-1)}\frac{1}{2}[3\cos^2\theta - 1 - \eta_Q\sin^2\theta\cos(2\phi)]\frac{1}{2}[3m^2 - I(I+1)]$$

$$= \frac{\pi}{4}\chi\hbar\,[3\cos^2\theta - 1 - \eta_Q\sin^2\theta\cos(2\phi)](3m^2 - 2) \tag{1.24}$$

where χ is ^2H QCC and m has three possible values of -1, 0, and 1 which correspond to the three different spin states of ^2H. For a fixed orientation (i.e. fixed θ and ϕ), different spin states (m = ±1 vs m = 0) have different nuclear charge distributions and thus different electric quadrupole moments, which cause different quadrupolar interactions and energies. For ^2H in aliphatic C-^2H bonds, $\eta_Q \approx 0$. In this case, equation 1.24 can be simplified to

$$E_Q = \frac{\pi}{4}\chi\hbar\,(3\cos^2\theta - 1)(3m^2 - 2) \tag{1.25}$$

where θ is one of the two polar angles of $\mathbf{B_0}$ in the PAF and refers to the angle between the C-^{2}H bond and $\mathbf{B_0}$. Next we will discuss the orientation dependence of the quadrupolar energy E_Q and how it affects the ^{2}H resonance frequency.

(a) When $\theta = 0^o$, $3\cos^2\theta - 1 = 2$, $E_Q = \frac{\pi}{2}\chi\hbar(3m^2 - 2)$. For m = -1, 0, and 1, E_Q is respectively $E_{-1} = \frac{\pi}{2}\chi\hbar$, $E_0 = -\pi\chi\hbar$, and $E_1 = \frac{\pi}{2}\chi\hbar$. There are two transitions for ^{2}H, one is m = 1 → m = 0 and the other is m = 0 → m = -1. Assuming the Zeeman energy for m = -1, 0, and 1 is +E', 0, and -E', respectively, in this case, the energy difference for the m = 1 → m = 0 and m = 0 → m = -1 transitions is $\Delta E_{1\to0} = E' - \frac{3\pi}{2}\chi\hbar$ and $\Delta E_{0\to-1} = E' + \frac{3\pi}{2}\chi\hbar$, respectively. Therefore, the resonance frequency for the m = 1 → m = 0 and m = 0 → m = -1 transitions is $\nu_{1\to0} = \nu' - \frac{3}{4}\chi$ and $\nu_{0\to-1} = \nu' + \frac{3}{4}\chi$, respectively, where $\nu' = E'/h$ is the ^{2}H Larmor frequency without considering the quadrupolar interaction. If we set the ^{2}H transmitter frequency at ν', we will observe two signals in the ^{2}H spectrum, one at $-\frac{3}{4}\chi$ (in Hz, not in ppm) resulting from the m = 1 → m = 0 transition and the other at $\frac{3}{4}\chi$ resulting from the m = 0 → m = -1 transition (Figure 1.8a).

(b) When $\theta = 54.7^o$, $3\cos^2\theta - 1 = 0$, $E_Q = 0$ for all m values. In this case, the m = 1 → m = 0 and m = 0 → m = -1 transitions yield a signal at the same frequency in the ^2H spectrum (Figure 1.8b).

(c) When $\theta = 90^o$, $3\cos^2\theta - 1 = -1$, $E_Q = -\frac{\pi}{4}\chi\hbar\,(3m^2 - 2)$. For m = -1, 0, and 1, E_Q is respectively $E_{-1} = -\frac{\pi}{4}\chi\hbar$, $E_0 = \frac{\pi}{2}\pi\chi\hbar$, and $E_1 = -\frac{\pi}{4}\chi\hbar$. The energy difference for the m = 1 → m = 0 and m = 0 → m = -1 transitions is respectively $\Delta E_{1\to0} = E' + \frac{3\pi}{4}\chi\hbar$ and $\Delta E_{0\to-1} = E' - \frac{3\pi}{4}\chi\hbar$. Therefore, the resonance frequency for the m = 1 → m = 0 and m = 0 → m = -1 transitions is $\nu_{1\to0} = \nu' + \frac{3}{8}\chi$ and $\nu_{0\to-1} = \nu' - \frac{3}{8}\chi$, respectively. For this orientation, we will observe two signals in the ^2H spectrum with one at $\frac{3}{8}\chi$ resulting from the m = 1 → m = 0 transition and the other at $-\frac{3}{8}\chi$ resulting from the m = 0 → m = -1 transition (Figure 1.8c).

When we consider all the possible θ values, i.e. all the possible C-^2H bond orientations with respect to $\mathbf{B_0}$, we will observe a ^2H quadrupolar powder pattern, which is also referred to as Pake doublet, as displayed in Figure 1.7d, where the black and blue component spectra correspond to the m = 0 → m = -1 and m = 1 → m = 0 transitions, respectively.

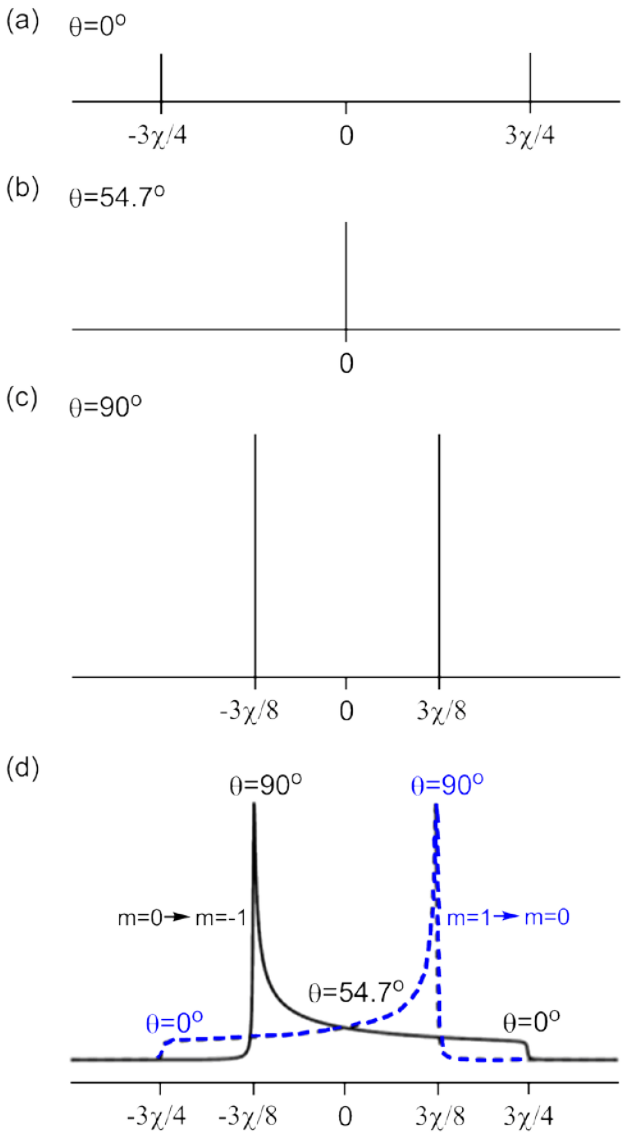

Figure 1.8 Orientation dependence of the ^{2}H spectra for aliphatic C-^{2}H bonds. (a) $\theta = 0^{o}$. (b) $\theta = 54.7^{o}$. (c) $\theta = 90^{o}$. (d) ^{2}H Pake doublet for all possible θ values, where θ refers to the angle between the C-^{2}H bond and $\mathbf{B_o}$ field.

Rotating frame

In the laboratory frame, the $\mathbf{B_1}$ field of a pulse rotates with an angular velocity of ω in the xy plane, where ω is the transmitter frequency of the pulse. We call it transmitter frequency for the reason that a radiofrequency transmitter is used to

generate the pulse. However, in the rotating frame, the xy coordinate system rotates with an angular velocity of ω_r, where ω_r is the so-called rotating frame frequency. When we set the transmitter frequency the same as the rotating frame frequency (as is usually the case), the $\mathbf{B_1}$ field of a pulse appears to be static, i.e. the time dependence of the field is removed. This makes it much simpler to work out the effect of the $\mathbf{B_1}$ field on a nuclear magnetization. In the rotating frame, the apparent precession rate of a magnetization is $(\omega_o - \omega_r)$, where $\omega_o = \gamma B_o$ is the Larmor frequency. This difference frequency, $\omega_o - \omega_r$, is often referred to as the resonance offset frequency $\omega_{r.o.}$, i.e.

$$\omega_{r.o.} = \omega_o - \omega_r \qquad (1.26)$$

According to the relationship $\omega = \gamma B$, the resonance offset field is given by

$$\mathbf{B}_{r.o.} = (\omega_{r.o.}/\gamma)\mathbf{z} \qquad (1.27)$$

where the field $\mathbf{B}_{r.o.}$ is along the z axis and refers to the secular component of the apparent magnetic field experienced by a nucleus in the rotating frame. If we set the transmitter frequency ω close to the Larmor frequency ω_o and set ω the same as the rotating frame frequency ω_r, the resonance offset frequency $\omega_{r.o.}$ will be close to zero and thus the field $\mathbf{B}_{r.o.}$ will be very small. As the magnitude of $\mathbf{B}_{r.o.}$ is small relative to that of the r.f. pulse field $\mathbf{B_1}$, the r.f. pulses can work efficiently on the nuclear magnetization and cause rapid spin transitions. However, for quadrupolar nuclei which have a broad Larmor frequency distribution, no matter where we set the transmitter

frequency, there are large resonance offsets for some of the nuclei and thus $B_{r.o.}/B_1$ may not be close to zero. For such nuclei, the spin transitions caused by the r.f. pulses may be slow. Since the r.f. pulse duration is usually very short (in the order of μs), the nuclei with large resonance offsets may not be affected by the pulses. To solve or minimize this issue, a strong pulse field $\mathbf{B_1}$ is usually applied to compensate the large resonance offsets of quadrupolar nuclei. In the next paragraph, we will discuss in some detail about how can a spectrometer detect an NMR signal (the xy component of a magnetization, i.e. the transverse magnetization) with a resonance offset frequency $\omega_{r.o.}$ in the rotating frame while the transverse magnetization actually precesses with a Larmor frequency ω_o.

In the NMR spectrometer, a mixer is used to mix two input signals together. One is the NMR signal (i.e. transverse magnetization of the observed nuclei) from the probe and the other is a signal generated by the local oscillator. Suppose the NMR signal from the probe is $A\cos\omega_o t$ where A is the overall signal intensity and ω_o is the Larmor frequency. The local oscillator generates two types of signals which are 90^o out of phase: $\cos\omega_r t$ and $-\sin\omega_r t$, where ω_r is the receiver frequency and is usually the same as the transmitter frequency[5]. When the local oscillator generates the signal $\cos\omega_r t$, the mixer multiplies together the NMR signal and the local oscillator signal and yields:

$$A\cos\omega_o t \times \cos\omega_r t = \frac{1}{2}A[\cos(\omega_o + \omega_r)t + \cos(\omega_o - \omega_r)t] \qquad (1.28)$$

After the mixing process, a low-pass filter is used to filter out the term $\cos(\omega_o + \omega_r)t$ and thus only the term $\cos(\omega_o - \omega_r)t$ survives. In this case, the output signal of the mixer is cosine modulated and has a frequency of $\omega_o - \omega_r$ which is exactly the resonance offset frequency $\omega_{r.o.}$ in the rotating frame. This cosine modulated signal is equivalent to the x component of the magnetization detected in the rotating frame and is referred to as the real part of the FID (free induction decay). Next let us consider the other case in which the local oscillator signal is $-\sin\omega_r t$. The output signal of the mixer is given by

$$A\cos\omega_o t \times -\sin\omega_r t = \frac{1}{2}A[\sin(\omega_o - \omega_r)t - \sin(\omega_o + \omega_r)t] \qquad (1.29)$$

where the term $\sin(\omega_o + \omega_r)t$ is removed by the low-pass filter and only the term $\sin(\omega_o - \omega_r)t$ remains. Now the output signal is sine modulated and also has the resonance offset frequency. This sine modulated signal is equivalent to the y component of the magnetization detected in the rotating frame and is referred to as the imaginary part of the FID. The real and imaginary parts of the FID are detected separately and such a scheme is called quadrature detection, as displayed in Figure 1.9.

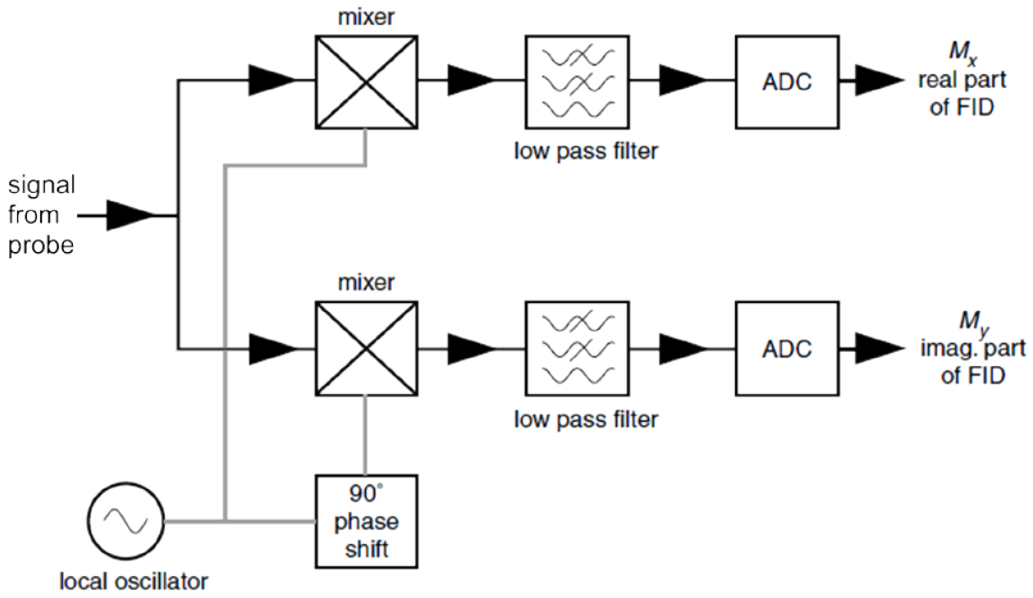

Figure 1.9 Quadrature detection scheme in the rotating frame for all modern NMR spectrometers. The device ADC refers to *analogue to digital converter* and is used to convert the NMR signal from a voltage to a binary number which can be stored in the computer memory[5].

Magic angle spinning (MAS)

In liquids, the rapid molecular tumbling averages out the anisotropic interactions such as CSA, dipolar, and quadrupolar interactions. As a result, sharp peaks with good resolution are usually observed in liquid-state NMR spectra. However, in solids, there is no such rapid molecular tumbling and the anisotropic interactions cause severe line broadening in the static NMR spectra. To solve this problem, magic angle spinning (MAS) was developed[6]. By spinning the sample under MAS, one can average out the anisotropic interactions and thus isotropic chemical shifts result. The geometry of MAS is displayed in Figure 1.10, where the angle between the rotor axis (i.e. sample spinning axis) and the external field $\mathbf{B_0}$ equals the so-called magic angle of 54.7^{o}. To better understand the effect of MAS on anisotropic interactions, let us take the ^{13}C-^{15}N dipolar

coupling as an example. Assume θ is the angle between $\mathbf{B_o}$ and the ^{13}C-^{15}N internuclear vector, α is the angle between $\mathbf{B_o}$ and the sample spinning axis, and β is the angle between the ^{13}C-^{15}N vector and the spinning axis. Under MAS, i.e. $\alpha = 54.7^o$, we can mathematically show that the average of $3\cos^2\theta(t) - 1$ over each rotor period is

$$\langle 3\cos^2\theta(t) - 1 \rangle = \frac{1}{2}(3\cos^2\alpha - 1)(3\cos^2\beta - 1) = 0 \qquad (1.30)$$

Since the heteronuclear dipolar Hamiltonian has an orientation dependence of $3\cos^2\theta - 1$(refer to equation 1.19), the ^{13}C-^{15}N dipolar interaction is averaged out by MAS. Nowadays MAS has been widely used in various SSNMR experiments.

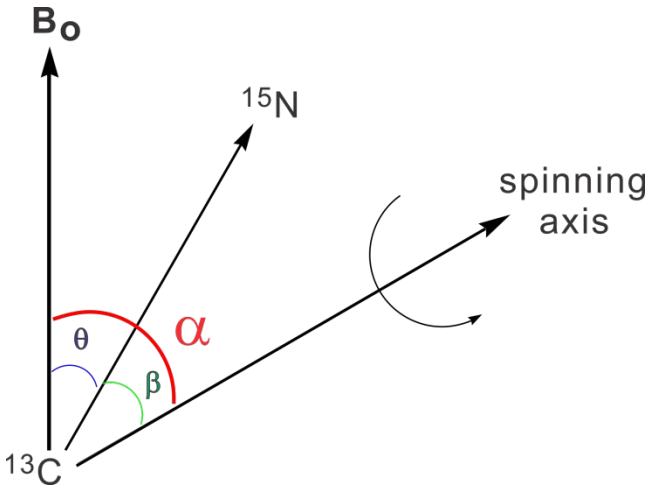

Figure 1.10 The geometry of a ^{13}C-^{15}N internuclear vector in a solid state sample under MAS where α equals 54.7^o (magic angle).

REDOR

Rotational-echo double-resonance (REDOR) is a SSNMR technique originally developed by Gullion and Schaefer[7]. It utilizes rotor-synchronized π pulse trains to

recouple the heteronuclear dipolar interaction under MAS. Since the dipolar coupling d is proportional to $\frac{1}{r^3}$, the distance r between two nuclei can be obtained from REDOR measurement. Next we will take ^{13}C -^{15}N REDOR for example to discuss in some detail about the spin interactions in REDOR under MAS. Figure 1.11 shows the ^{13}C -^{15}N REDOR pulse sequence. In the beginning of the sequence, a ^{1}H 90^{o} pulse is applied to rotate the ^{1}H magnetization from the z axis ($\mathbf{B_o}$ axis) to the xy plane. After that, ^{1}H and ^{13}C cross polarization (CP) pulses are applied simultaneously to transfer magnetization from ^{1}H to ^{13}C via ^{1}H -^{13}C dipolar coupling to enhance the ^{13}C signal intensity. CP occurs in a so-called doubly rotating frame with one in which the ^{1}H $\mathbf{B_1}$ field is static and the other in which the ^{13}C field $\mathbf{B_1}$ is also static. The magnitudes of the ^{1}H and ^{13}C CP pulse fields, $B_{1,H}$ and $B_{1,C}$, must fulfill the Hartmann-Hahn matching condition[8]:

$$\gamma_H B_{1,H} = \gamma_C B_{1,C} \qquad (1.31)$$

This is the condition for ideal δ-function pulses for which no resonance offset field in the rotating frame needs to be considered. However, in practice, we need to take into account the resonance offset field $\mathbf{B_{r.o.}}$ since the r.f. pulses are applied with limited power. As we consider $\mathbf{B_{r.o.}}$, the effective magnetic field $\mathbf{B_{eff}}$ experienced by a nucleus is the vector sum of $\mathbf{B_1}$ and $\mathbf{B_{r.o.}}$ and the matching condition is now given by

$$\gamma_H B_{eff}(^{1}H) = \gamma_C B_{eff}(^{13}C) \qquad (1.32)$$

Note that equation 1.31 and 1.32 describe the Hartmann-Hahn matching condition for the $^{1}H \rightarrow ^{13}C$ CP in a static sample. Under MAS, this matching condition needs to be modified since the sample rotation affects dipolar interactions. As previously discussed, the $^{1}H \rightarrow ^{13}C$ CP is achieved via ^{1}H -^{13}C dipolar coupling. For a peptide with a backbone ^{13}CO label, the closest ^{1}H is ~2 Å apart. Therefore, the largest ^{1}H-^{13}CO dipolar coupling is ~4 kHz. As we spin the sample at a rate of 10 kHz, the ^{1}H-^{13}CO dipolar coupling is supposed to be averaged out by MAS. However, this is not true due to the ^{1}H-^{1}H dipolar coupling. In peptides and other organic compounds, all the protons are dipolar-coupled as a network. There is rapid flip-flop (also known as ^{1}H spin diffusion) between protons via ^{1}H-^{1}H dipolar coupling, that is, when one ^{1}H changes its spin state from β to α, another ^{1}H will change from α to β. The flip-flop rate is roughly equal to the ^{1}H-^{1}H dipolar coupling, which is typically in the 10-50 kHz range[9]. Therefore, under 10 kHz MAS, ^{1}H will change its spin state in a timescale comparable to or even shorter than the rotor period. In this case, the ^{1}H flip-flop will disrupt the coherent averaging of the ^{1}H-^{13}CO dipolar field/coupling by MAS. As a result, the ^{1}H-^{13}CO dipolar coupling in the peptide is not averaged out over each rotor period. This is why you can achieve the $^{1}H \rightarrow ^{13}CO$ CP via ^{1}H-^{13}CO dipolar coupling under 10 kHz

MAS. For the $^{1}H \rightarrow ^{13}C$ CP under MAS, the Hartmann-Hahn matching condition splits into a series of new matching conditions[10]:

$$\gamma_H B_{eff}(^{1}H) = \gamma_C B_{eff}(^{13}C) + n\omega_{mas} \quad (1.33)$$

where ω_{mas} is the angular frequency of MAS and n = ±1, ±2, etc which corresponds to the nth spinning sideband in the ^{13}C spectrum. Recall that the nuclear spin energy $E = -\boldsymbol{\mu} \cdot \mathbf{B}$ where $\boldsymbol{\mu}$ is the nuclear magnetic dipole moment and the energy splitting between the α and β states of a spin 1/2 nucleus is $\Delta E = \gamma \hbar B$. Therefore, for the CP under static condition where $\gamma_H B_{eff}(^{1}H) = \gamma_C B_{eff}(^{13}C)$, the energy splitting of ^{1}H is equal to that of ^{13}C in the rotating frame. As one ^{1}H changes the spin state from α to β, the energy required can be exactly compensated by a transition of ^{13}C from β to α which releases the same amount of energy. In other words, the ^{1}H-^{13}C dipolar interaction causes a redistribution of energy between ^{1}H and ^{13}C spins but the total energy of the ^{1}H and ^{13}C spin system is conserved during the static CP. However, for the CP under MAS, the total spin energy is no longer conserved since $\gamma_H B_{eff}(^{1}H) \neq \gamma_C B_{eff}(^{13}C)$, as shown in equation 1.33.

Due to the different orientations of molecules, chemical bonds, and internuclear vectors relative to $\mathbf{B_0}$, the spin interactions such as shielding and dipolar coupling for a specific nucleus will vary from molecule to molecule. As a result, there is usually a distribution of the Larmor frequency as well as the resonance offset frequency, i.e. there

is a distribution of the resonance offset field $\mathbf{B_{r.o.}}$. Therefore, a ramp is usually applied to either ^{1}H or ^{13}C CP pulse field amplitude to increase the efficiency of the magnetization transfer from ^{1}H to ^{13}C under MAS.

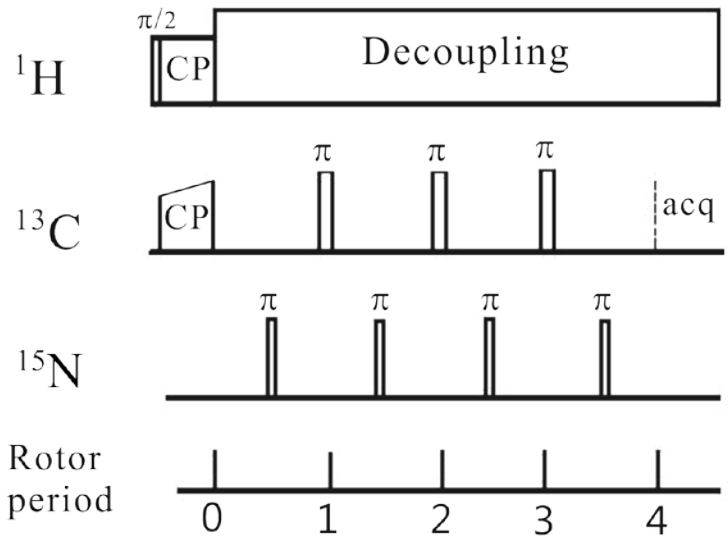

Figure 1.11 ^{13}C -^{15}N REDOR pulse sequence.

After CP, the rotor-synchronized π pulses are applied in two different ways which correspond to the so-called S_0 and S_1 experiments in REDOR. In the S_0 experiment, only ^{13}C π pulses are applied at the end of each rotor period except for the last one. In this case, the ^{13}C -^{15}N dipolar interaction is averaged out over each rotor period by MAS (Figure 1.12). In addition, other anisotropic interactions such as ^{13}C CSA is also eliminated by MAS. The ^{13}C π pulses at the end of each odd-numbered rotor period are applied to refocus the ^{13}C transverse magnetization at the end of each even-numbered

rotor period under the effect of isotropic chemical shift field. In the S_1 experiment, ^{13}C π pulses are applied at the end of each rotor period while ^{15}N π pulses are applied in the middle of each rotor period. In this case, the ^{13}C -^{15}N dipolar coupling is recoupled under MAS (Figure 1.12). Due to the ^{13}C -^{15}N dipolar interaction and the distribution of the dipolar field from ^{15}N spins, the ^{13}C transverse magnetization decays as a function of the dephasing time τ, where τ refers to the time period after CP but before FID acquisition. Therefore, the ^{13}C signal intensity in S_1 experiment is weaker than that in S_0 experiment. Suppose the ^{13}C signal intensity in the S_0 and S_1 experiments is respectively S_0 and S_1, the REDOR dephasing is defined as

$$\Delta S/S_0 = (S_0 - S_1)/S_0 \qquad (1.34)$$

After the dephasing buildup curve of $\Delta S/S_0$ vs τ is achieved, the ^{13}C -^{15}N dipolar coupling can be obtained by comparison with SIMPSON simulation, which is a quantum mechanics-based numerical simulation program for SSNMR[11]. Once the ^{13}C -^{15}N dipolar coupling d is known, one can calculate the ^{13}C -^{15}N internuclear distance r through the equation $d = 3066/r^3$, where d is in Hz and r is in Å.

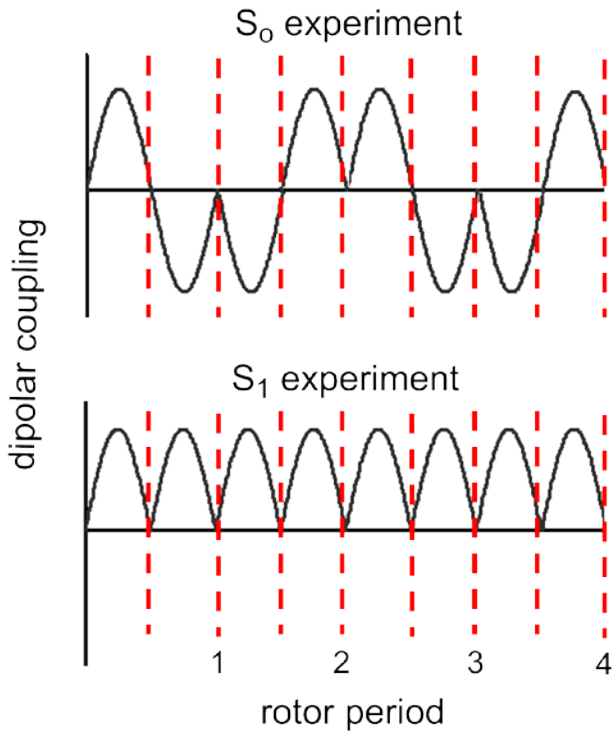

Figure 1.12 Dipolar coupling evolution as a function of rotor period during the dephasing time τ in REDOR. The dipolar interaction is averaged out over each rotor period by MAS in S_0 experiment but not in S_1 experiment.

For REDOR between a spin 1/2 and spin 1/2 pair such as ^{13}C and ^{15}N, there is a so-called universal dephasing curve of $\Delta S/S_0$ vs λ, where λ is the product of the dipolar coupling d and dephasing time τ (Figure 1.13a)[12]. For REDOR between a spin 1/2 and spin 1 pair such as ^{13}C and ^{2}H, there is also a universal dephasing curve of $\Delta S/S_0$ vs λ according to the SIMPSON simulations (Figure 1.13b). However, the maximum dephasing for ^{13}C - ^{2}H REDOR is only ~2/3 without considering T_1 relaxation of ^{2}H. This is because the ^{2}H π pulses only cause spin transitions between m = 1 and m = -1

states whereas the m = 0 state which accounts for 1/3 population of the ^2H spins remains unchanged. In practice, we need to take into account T_1 relaxation of ^2H.

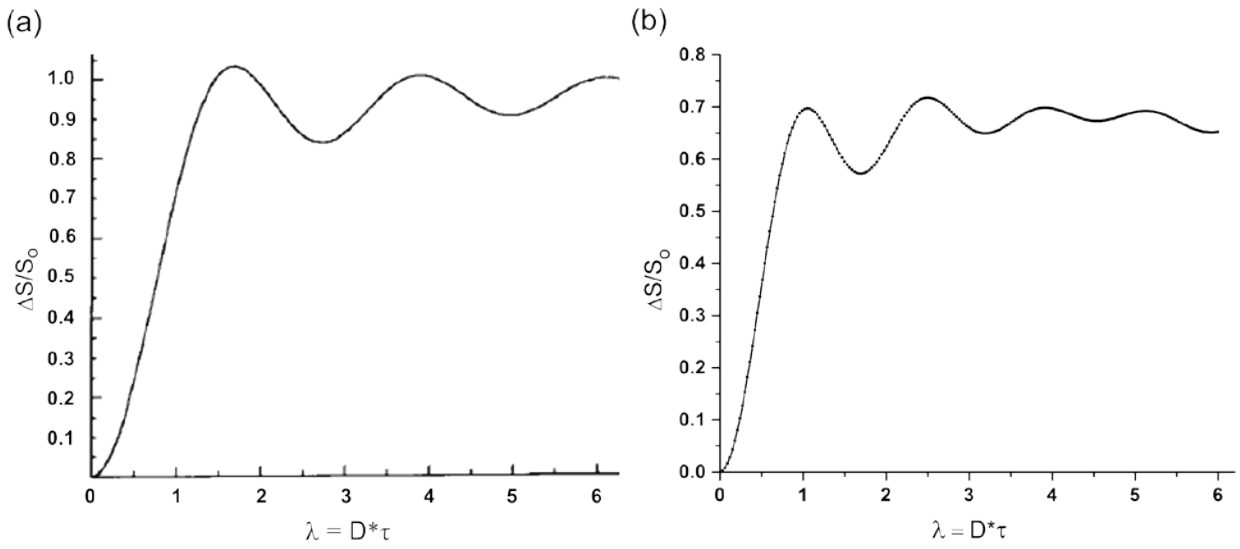

Figure 1.13 REDOR universal dephasing curve of $\Delta S/S_0$ vs λ for a spin 1/2 - spin 1/2 pair (panel a) and for a spin 1/2 - spin 1 pair (panel b), respectively. Panel a is edited from reference [12] and and panel b is obtained from SIMPSON simulations.

1.2 Influenza and HIV fusion peptides

Influenza fusion peptide (IFP)

Influenza virus is a type of retrovirus (i.e. RNA virus) enveloped by a lipid membrane and it infects the respiratory epithelial cell. There are several different subtypes of influenza virus based on two proteins on the surface of the virus. The two viral proteins are hemagglutinin (HA) and neuraminidase (NA). There are 18 HA subtypes and 11 NA subtypes and many different combinations of HA and NA proteins are possible[13]. However, only some of the subtypes (e.g. H1N1, H3N2) are in general circulation among people.

During the infection process, the influenza HA protein plays an important role, where the HA protein is organized as a homotrimer in the viral membrane and consists of two subunits, HA1 (binding subunit) and HA2 (fusion subunit). HA1 stays completely outside the virus. HA2 has a ~185 residue N-terminal ectodomain outside the virus, a ~25 residue transmembrane domain, and a ~10 residue C-terminal endodomain (i.e. cytoplasmic domain) inside the virus. There is a single disulfide bond between cysteine 14 in HA1 and cysteine 137 in HA2[14].

The infection (Figure 1.14) starts with the binding of HA1 to the sialic acids which are attached to the glycoproteins or glycolipids on the host cell surface, followed by the endocytosis of the virus into the host cell. After that, protons are pumped into the newly-formed endosome via the cell physiological processes to lower the endosomal pH to ~5. The low pH then triggers a dramatic conformational change of the HA protein[15], which is cleaved to yield HA1 and HA2. The two subunits remain linked via a single disulfide bond. During the conformational change, the ~20 residue fusion peptide (IFP) at the N-terminus of HA2 is exposed. The IFP plays a vital role in the fusion between the viral and the host cell endosomal membranes, as evidenced by the disruption of the fusion activity of HA2 with point mutations in the IFP region[16]. Substitution of glutamic acid for glycine residue right at the amino-terminus of HA2 even completely abolishes its fusion activity. Earlier studies show that the IFP and transmembrane domain are the only two regions of HA2 which are deeply inserted in the membrane after the fusion between the viral and target membranes[17].

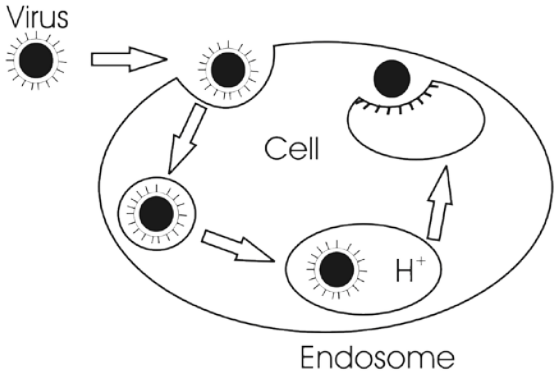

Figure 1.14 Infection model of influenza virus. The virus recognizes the host cell (respiratory epithelial cell) via the viral HA1 binding protein and then enters the cell by endocytosis. The cell physiology pumps protons into the newly-formed endosome and lowers the pH to ~5 to trigger the fusion between the viral and endosomal membranes. The viral nucleocapsid is then released to the cell cytoplasm[18].

More background and ^{13}C-^{15}N REDOR studies of the structures of membrane-associated IFP will be discussed in Chapter 2.

HIV fusion peptide (HFP)

Similar to the influenza virus, human immunodeficiency virus (HIV) is another type of retrovirus and is also enveloped by a lipid membrane. A mature HIV virus has a diameter of 110 to 128 nm[19]. HIV targets human T-cells through the viral envelope glycoprotein (Env). The Env protein is also known as gp160 and consists of two subunits, gp120 and gp41, which are non-covalently associated with each other. Earlier electron microscopy tomography studies have shown that the wild-type Env protein is trimeric and each HIV virion has ~14 Env trimers[20, 21]. When HIV infects a host cell (T cell), the viral gp120 binding protein first recognizes the host cell through the CD4 receptor on the cell surface. After that, gp120 leaves gp41 and binds to both the CD4 receptor and CXCR4 (or CCR5) coreceptor[22]. The exposed gp41 fusion protein binds

to the host cell membrane and initiates the HIV-host cell membrane fusion. The HIV infection process on T cells is displayed in Figure 1.15[23].

During the HIV-host cell membrane fusion process, the HIV gp41 protein undergoes three states sequentially: (i) pre-fusion native state, (ii) pre-hairpin intermediate (PHI) state, and (iii) final hairpin state[24]. These three different states are displayed in Figure 1.16a. The sequence of gp41 fusion protein consists of 345 residues and can be divided into the N-terminal fusion peptide (FP), N-terminal heptad repeat region (NHR), loop region, C-terminal heptad repeat region (CHR), membrane-proximal external region (MPER), transmembrane domain, and the endodomain (Figure 1.16b). All the segments outside the viral membrane form the so-called ectodomain of gp41. The heptad unit in NHR and CHR regions contains seven residues a, b, c, d, e, f, and g, where residues a and d are hydrophobic and residues b, c, e, f, and g are hydrophilic[25]. In the hairpin state, three NHR helices and three CHR helices form a hairpin-like six helix bundle (SHB).

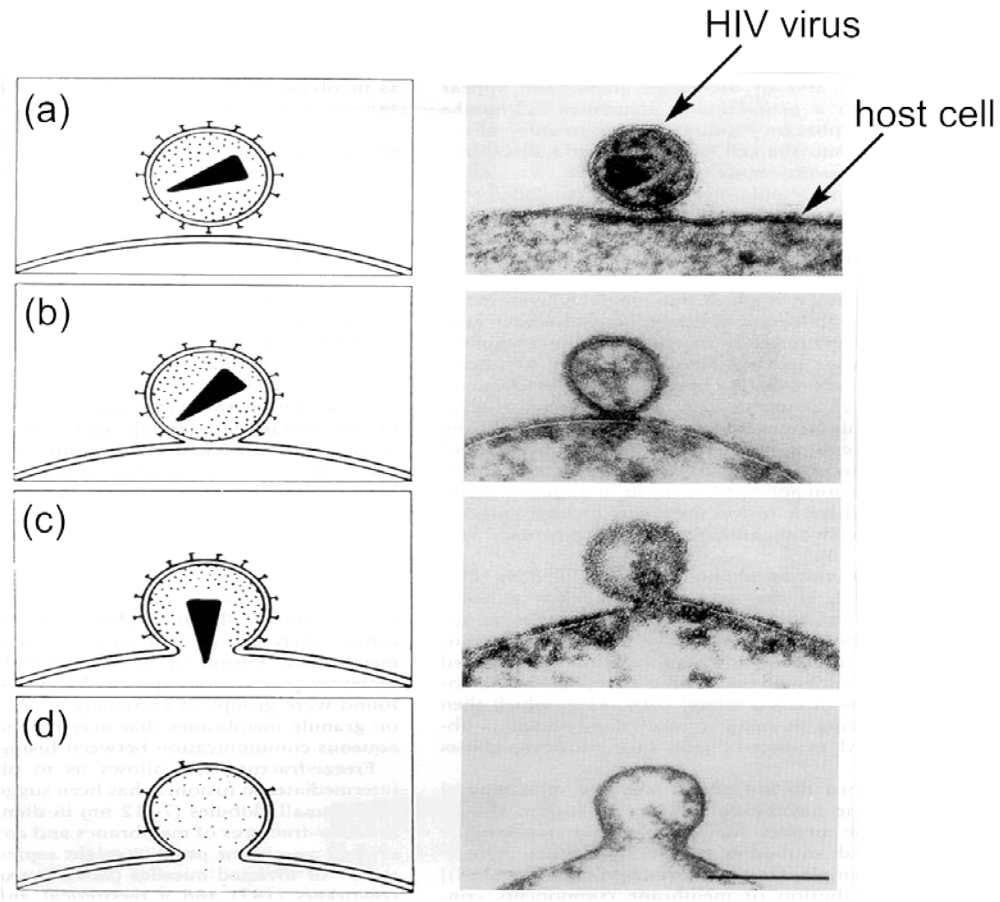

Figure 1.15 (Left) HIV infection model. (Right) HIV infection process illustrated by electron microscopy. (a) HIV-host cell membrane binding. (b) hemi-fusion. (c) complete fusion with the formation of fusion pore. (d) entry of the viral nucleocapsid into the host cell[23].

The crystal structure of HIV gp120 binding protein has been elucidated by X-ray diffraction (XRD)[26]. However, there has been no crystal structure for the full-length gp41 fusion protein. Earlier X-ray crystallography studies have determined the crystal structures of the hairpin (i.e. SHB) and the gp41 ectodomain lacking the fusion peptide (Figure 1.17)[25, 27]. For the hairpin structure (SHB), an essential function of it is to stabilize the membrane fusion pore and prevent it from being collapsed or closed during the HIV infection process. Previous fluorescence spectroscopy studies show that the addition of gp41-derived peptides which inhibit the SHB formation (e.g. C34 and N36)

can cause the cold-arrested fusion pore to quickly and irreversibly close, indicating that the SHB formation is not complete by the time when a fusion pore has formed[28]. Once the SHB formation is complete, this structure stabilizes the membrane fusion pore and ensures its growth.

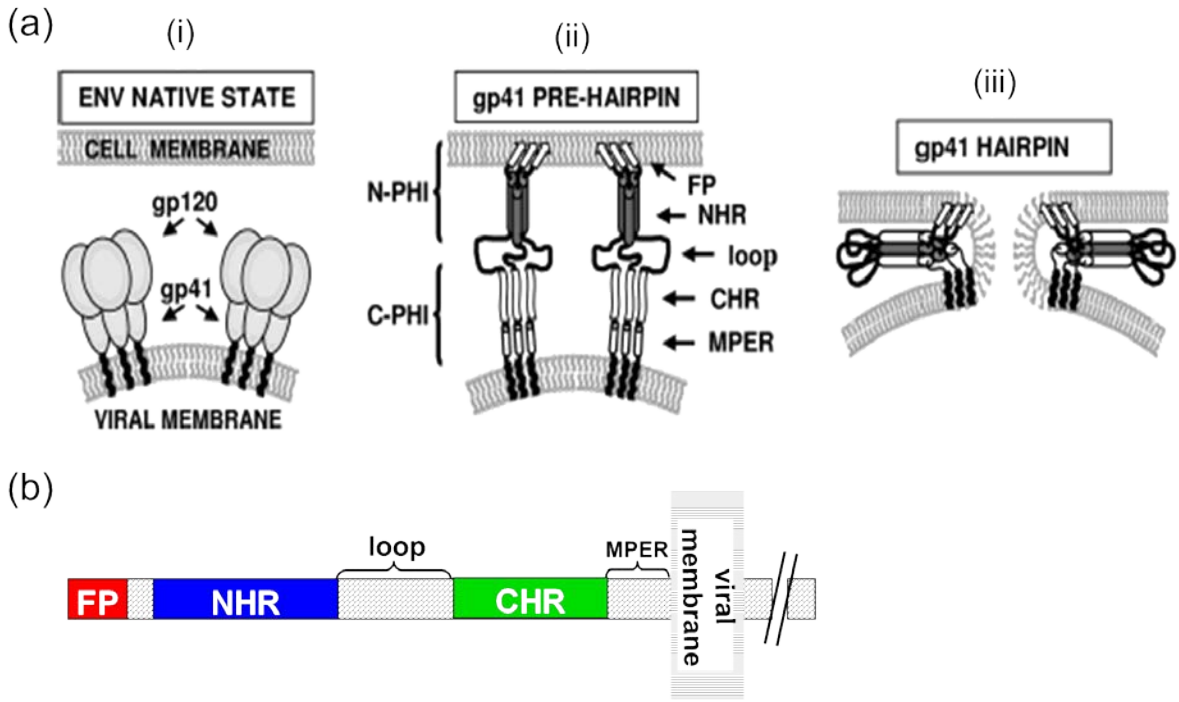

Figure 1.16 (a) HIV-host cell membrane fusion model: (*i*) native-state trimers of gp120 and gp41 prior to fusion, (*ii*) pre-hairpin intermediate (PHI), and (*iii*) final hairpin state[24]. (b) Different segments of the HIV gp41 fusion protein.

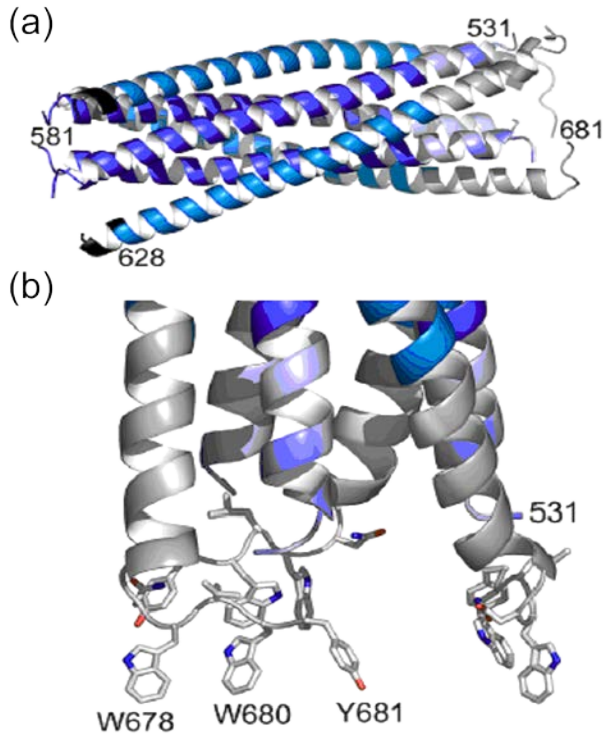

Figure 1.17 Crystal structure of the HIV gp41 ectodomain lacking fusion peptide ($gp41_{528-683}$, residues are numbered according to their positions in the gp160 complex). (*a*) α helical structure of the gp41 ectodomain, including polar region(blue, 528-540), NHR(blue, 541-581), loop region(blue,582-627), CHR(blue,628-666), and MPER (grey,667-683). The structure is 88 Å long. (*b*) Close-up of the MPER region showing the exposure of aromatic side chains of Trp 678, Trp 680 and Tyr 681 towards the membrane, which indicates the potential membrane insertion of MPER residues adjacent to the transmembrane domain of gp41[27].

Earlier solution NMR studies of micelle-associated HIV fusion peptide (HFP), which is the N-terminal ~20-30 residues of the gp41 fusion protein, have shown that HFP adopts a predominantly α helical structure in detergents such as DPC and SDS (Figure 1.18)[29]. For HFP in lipid membranes, previous SSNMR and IR studies have both shown that the membrane-associated HFP majorly adopts an oligomeric anti-parallel β sheet structure[30, 31].

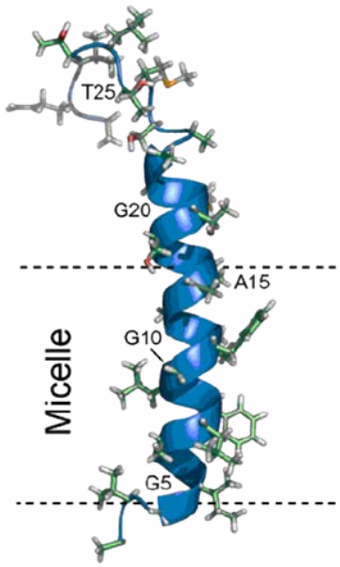

Figure 1.18 Ribbon representation of the HFP30, which is the N-terminal 30 residues of the HIV gp41 fusion protein, in an SDS micelle[29].

Membrane binding and lipid mixing assays indicate that HFP binds to lipid membranes and induces lipid mixing (i.e. vesicle fusion)[32]. On the other hand, lipid mixing assays of HFP23 (the N-terminal 23 residues of gp41), gp41e (the ectodomain of gp41 excluding HFP23), and gp41e-HFP (the ectodomain of gp41 including HFP23) indicate that both HFP23 and gp41e-HFP have a much higher fusion activity than gp41e[33]. All the above data imply that HFP may play an important role in the HIV-host cell membrane fusion during the infection process.

1.3 Membrane insertions of peptides

Earlier studies of membrane insertion depths of IFP and HFP

The insertion depth of protein and peptide residues in lipid membranes is an important feature of membrane-bound proteins and peptides. For instance, the membrane insertion depth of viral fusion peptides such as IFP and HFP is crucial for their membrane fusion activities[34-36]. Earlier electron paramagnetic resonance (EPR)

studies of IFP reveal that the residues in the N- and C-terminal regions penetrate the membrane 3~6 Å more deeply at the fusogenic pH 5.0 than at the non-fusogenic pH 7.4[34]. For HFP, previous lipid mixing assays and REDOR SSNMR studies together have shown a strong positive correlation between the membrane insertion depth and fusion activity of the peptide[35, 37, 38].

Techniques for measuring membrane insertion depths of peptides

A variety of techniques are available to probe the residue-specific locations of proteins and peptides in membranes. In the next paragraphs, we will discuss some of these techniques in some detail.

(a) Fluorescence resonance energy transfer (FRET)

This technique is based on the energy transfer from the fluorophore such as tryptophan to the quencher such as acrylamide or bromolipid. The fluorescence quenching effect has a fluorophore-quencher distance dependence and can be described by the Stern-Volmer equation:

$$F_0/F = 1 + K_{SV}[Q] \qquad (1.35)$$

where F_0 and F represent the fluorescence intensities in the absence and presence of the quencher Q, respectively, K_{SV} is the Stern-Volmer quenching constant which is proportional to $1/r^6$ and r is the fluorophore-quencher distance, and [Q] is the concentration of the quencher Q. The membrane insertion depth of HFP has been studied by the FRET technique[39, 40]. The residue Phe8 of HFP was mutated to Trp (i.e. F8W mutant) to detect the membrane location of the Trp indole group. The

quencher was either acrylamide in aqueous solution or bromine atom attached to the lipid acyl chain. Comparison of the Trp fluorescence intensities with and without the LUVs (large unilamellar vesicles) in the aqueous solution containing the quencher acrylamide showed that the Trp residue was embedded in the membrane interior. A more precise membrane location of the Trp residue was probed by using the bromolipid which was either 6,7 Bromo-PC (1-palmitoyl-2-stearoyl (6-7 dibromo)-sn-glycero-3-phosphocholine) or 11,12 Bromo-PC (1-palmitoyl-2-stearoyl (6-7 dibromo)-sn-glycero-3-phosphocholine). The quenching data indicated that the Trp residue was located closer to the C6-C7 than to the C11-C12 position in the membrane hydrocarbon core.

(b) Paramagnetic relaxation enhancement (PRE) SSNMR

This technique involves the use of paramagnetic ions such as Dy^{3+} and Mn^{2+}[41, 42]. The unpaired electron(s) in a paramagnetic ion can significantly enhance the T_1 and T_2 relaxation rates of nuclei, which is known as the PRE effect. Due to the faster T_2 relaxation, the signal of the observed nuclei decays faster in the presence of paramagnetic ions. The PRE effect has a distance dependence of $1/r^6$, where r is the distance between the unpaired electron and the nuclear spin. This distance dependence can be described by the Solomon-Bloembergen equation. Below is a modified version of the equation used in biomolecular NMR[43].

$$r = [\frac{K}{R_2}(4\tau_c + \frac{3\tau_c}{1+\omega_h^2\tau_c^2})]^{1/6} \qquad (1.36)$$

where r is the distance between the unpaired electron and the nuclear spin, R_2 is the nuclear T_2 relaxation rate solely from the PRE contribution (i.e. the intrinsic T_2 relaxation rate is not included in equation 1.36), τ_c is the correlation time for the electron-nuclear interaction, ω_h is the nuclear Larmor frequency, and K is a constant given by

$$K = S(S+1)\gamma^2 g^2 \beta^2/15 \qquad (1.37)$$

where S is the nuclear spin quantum number, γ is the nuclear gyromagnetic ratio, g is the electronic g factor, and β is the Bohr magneton. To calculate the distance r, an approximation could be made that τ_c is equal to the correlation time of the protein or peptide[43].

Since the aqueous paramagnetic ions are either in the bulk solution or at the water-membrane interface, one can probe the location of a protein or peptide with respect to the membrane surface.

(c) Spin diffusion

X-nucleus (^{13}C, ^{15}N, etc) detected, ^1H spin diffusion is another recently developed SSNMR technique for membrane location measurement of peptides and proteins[44-46]. The experiment proceeds in a time sequence as (1) select the ^1H magnetization from water or the mobile lipid chain-end methyl group using a T_2 filter; (2) allow a mixing time t_m for the ^1H spin diffusion from water or lipids to a rigid protein or

peptide; (3) apply $^{1}H \rightarrow {}^{13}C$ or $^{1}H \rightarrow {}^{15}N$ cross polarization (CP) where the protein or peptide is ^{13}C or ^{15}N labeled; and (4) detect the ^{13}C or ^{15}N signal. This technique can semiquantitatively measure the insertion depth of the protein or peptide relative to the membrane surface or bilayer center. The pulse sequence of the X-nucleus detected, ^{1}H spin diffusion experiment is displayed in Figure 1.19[44].

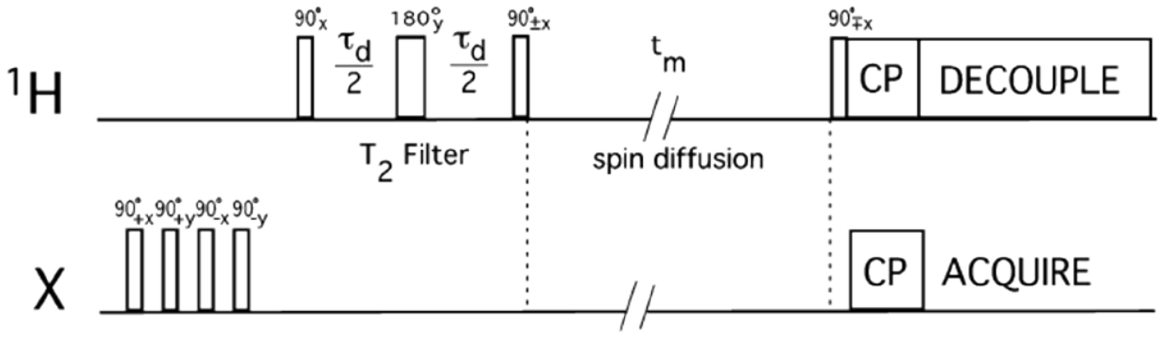

Figure 1.19 X-nucleus detected, ^{1}H spin diffusion pulse sequence[44]. The sequence starts with an initial saturation (destroy) of any X magnetization by several 90^{o} pulses, followed by a dipolar dephasing period. After the ^{1}H excitation, a T_2 filter is applied during the period τ_d to select the ^{1}H magnetization which has a long T_2. The 180^{o} pulse in the middle of the period τ_d is applied to refocus chemical shift and field inhomogeneity. The 90^{o} pulse at the end of τ_d stores the remaining ^{1}H magnetization with a long T_2 on the z axis, followed by a spin diffusion period with a variable t_m (e.g. 0-100 ms). During the mixing time t_m, the existing ^{1}H z-magnetization is transferred to nearby protons via ^{1}H-^{1}H dipolar coupling. In the end, a $^{1}H \rightarrow X$ CP is applied and then the X signal is detected with ^{1}H decoupling.

In the X-nucleus (^{13}C, ^{15}N, etc) detected, ^{1}H spin diffusion experiments, the signal intensity of X is detected as a function of the mixing time t_m. The X signal intensity at each t_m is then normalized to that at the final mixing time and plotted as a

function of $(t_m)^{1/2}$. Finally, the best-fit simulation of the buildup curve of the normalized X signal intensity versus $(t_m)^{1/2}$ yields a semiquantitative ^{1}H-^{1}H distance, where the former ^{1}H refers to those which are either directly bonded to or right next to (~2 Å) the X-nucleus in the protein or peptide and which are the major ^{1}H source for the CP while the latter ^{1}H refers to those from water or the mobile lipid chain-end methyl group which have a long T_2 time and thus their magnetization remains after the T_2 filter in the pulse sequence.

(d) REDOR SSNMR

As previously discussed in the "REDOR" section, this technique is a very useful tool for measuring heteronuclear dipolar couplings from which the internuclear distances can be calculated. There are several different types of REDOR such as ^{13}C-^{31}P, ^{13}C-^{19}F, and ^{13}C-^{2}H REDOR which can be used to probe the membrane locations (insertion depths) of peptides and proteins[35, 47, 48]. For instance, ^{13}C-^{31}P REDOR can probe the membrane insertion depth of the ^{13}CO-labeled residue of a peptide relative to the membrane surface since the membrane headgroup region contains ^{31}P. For a peptide-membrane sample in which the peptide contains one ^{13}CO-labeled residue and the membrane contains ~10% (mole fraction) 1-palmitoyl-2-[16-fluoropalmitoyl]-phosphatidylcholine (16-^{19}F-DPPC) lipid, ^{13}C-^{19}F REDOR can be used to probe the membrane insertion depth of the ^{13}CO-labeled residue with respect to the

membrane bilayer center. Three HFP constructs, which are respectively the V2E mutated HFP monomer (HFPmn_V2E), wild-type HFP monomer (HFPmn), and wild-type HFP trimer (HFPtr), have been studied by ^{13}C-^{31}P and ^{13}C-^{19}F REDOR and the data show that the membrane insertion depths of the three HFP constructs have an increasing order of HFPmn_V2E < HFPmn < HFPtr (Figure 1.20)[35].

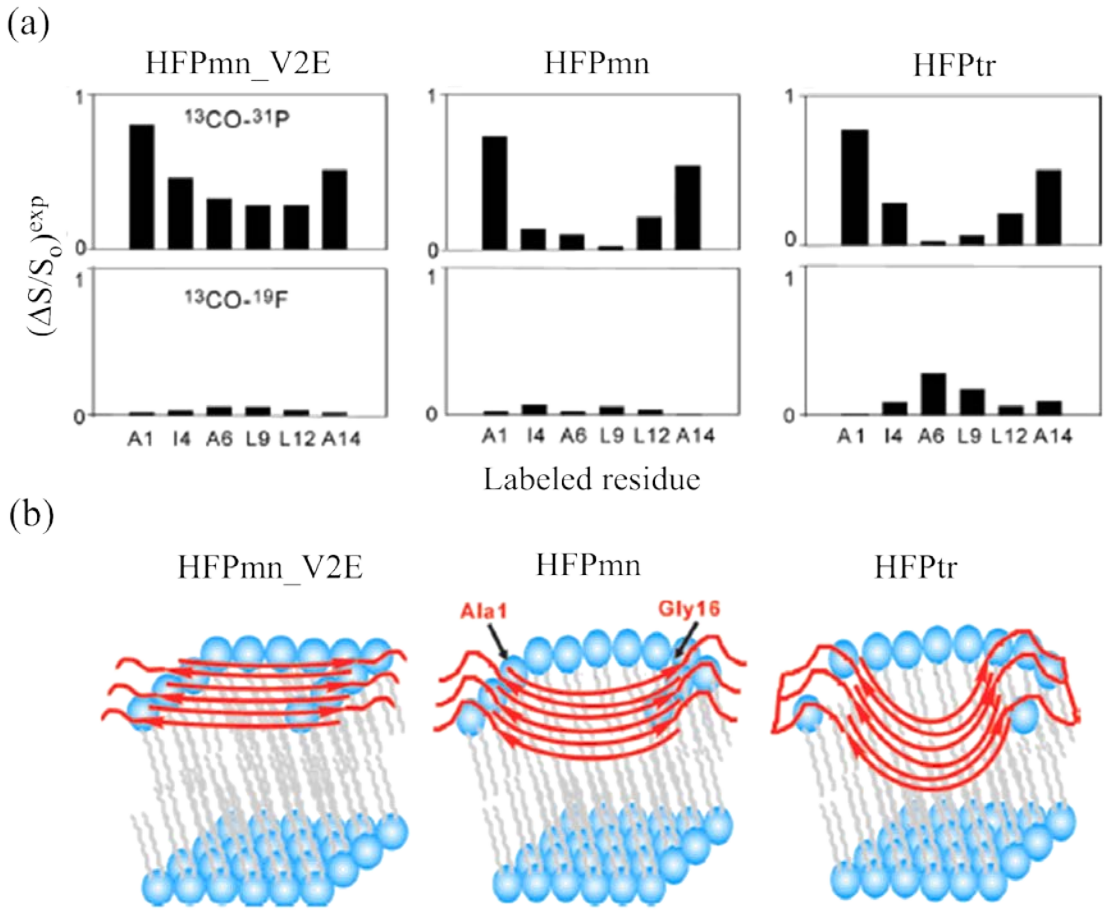

Figure 1.20 (a) ^{13}C-^{31}P and ^{13}C-^{19}F REDOR experimental dephasing $(\Delta S/S_o)^{exp}$ as a function of the ^{13}CO-labeled residue. (b) Membrane insertion models of the anti-parallel β sheet HFPmn_V2E, HFPmn, and HFPtr, respectively[35].

(e) Electron paramagnetic resonance (EPR)

EPR is a commonly used technique for distance measurement which requires site-specific mutagenesis. One example of such mutagenesis is the introduction of a cysteine into a protein or peptide to replace the residue of interest. The thiol group of cysteine can specifically react with a spin label which contains a paramagnetic center (e.g. nitroxyl radical) and thus attach the spin label to the protein or peptide. The distance measurement by EPR is based on the electron-electron spin interaction where one electron is from the spin label and the other could be from metal ions such as Ni^{2+}. EPR has been used to detect the membrane insertion depths of both IFP and HFP[34, 49, 50]. For the membrane-associated IFP, the residues in the N- and C-terminal regions penetrate the membrane 3~6 Å more deeply at the fusogenic pH 5.0 than at the non-fusogenic pH 7.4[34]. For HFP in the membrane containing 80% (mole fraction) POPC (1-palmitoyl-2-oleoyl-sn-gycero-3-phosphocholine) and 20% POPG (1-palmitoyl-2-oleoyl-sn-gycero-3-phosphoglycerol), EPR studies have shown that residues 4, 7, 12, and 15 penetrate into the membrane to a similar depth which is ~8 Å below the level of the lipid phosphate groups at the membrane surface[50].

The membrane location studies of peptides by ^{13}C-^{2}H REDOR SSNMR will be discussed in more detail in chapter 4.

REFERENCES

REFERENCES

1. Duer, M.J., *Solid-State NMR Spectroscopy: Principles and Applications*. 2002.

2. Lambert, J.B. and E.P. Mazzola, *Nuclear Magnetic Resonance Spectroscopy: An Introduction to Principles, Applications, and Experimental Methods*. 2004.

3. Schmidt-Rohr, K. and H.W. Spiess, *Multidimentional Solid-State NMR and Polymers*. 2005.

4. Harris, R.K., *Nuclear Magnetic Resonance Spectroscopy*. 1986.

5. Keeler, J., *Understanding NMR Spectroscopy*. 2002.

6. Andrew, E.R., A. Bradbury, and R.G. Eades, *NUCLEAR MAGNETIC RESONANCE SPECTRA FROM A CRYSTAL ROTATED AT HIGH SPEED.* Nature, 1958. **182**(4650): p. 1659-1659.

7. Gullion, T. and J. Schaefer, *Rotational-Echo Double-Resonance NMR.* Journal of Magnetic Resonance, 1989. **81**(1): p. 196-200.

8. Hartmann, S.R. and E.L. Hahn, *NUCLEAR DOUBLE RESONANCE IN ROTATING FRAME.* Physical Review, 1962. **128**(5): p. 2042-&.

9. Yannoni, C.S., *HIGH-RESOLUTION NMR IN SOLIDS - THE CPMAS EXPERIMENT.* Accounts of Chemical Research, 1982. **15**(7): p. 201-208.

10. Meier, B.H., *CROSS POLARIZATION UNDER FAST MAGIC ANGLE SPINNING - THERMODYNAMICAL CONSIDERATIONS.* Chemical Physics Letters, 1992. **188**(3-4): p. 201-207.

11. Bak, M., J.T. Rasmussen, and N.C. Nielsen, *SIMPSON: A general simulation program for solid-state NMR spectroscopy.* Journal of Magnetic Resonance, 2000. **147**(2): p. 296-330.

12. Gullion, T., *Introduction to rotational-echo, double-resonance NMR.* Concepts in Magnetic Resonance, 1998. **10**(5): p. 277-289.

13. CDC, *Types of influenza viruses.* http://www.cdc.gov/flu/about/viruses/types.htm, 2014.

14. Chen, J., J.J. Skehel, and D.C. Wiley, *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.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(16): p. 8967-8972.

15. Hughson, F.M., *STRUCTURAL CHARACTERIZATION OF VIRAL FUSION PROTEINS.* Current Biology, 1995. **5**(3): p. 265-274.

16. Gething, M.J., et al., *STUDIES ON THE MECHANISM OF MEMBRANE-FUSION - SITE-SPECIFIC MUTAGENESIS OF THE HEMAGGLUTININ OF INFLUENZA-VIRUS.* Journal of Cell Biology, 1986. **102**(1): p. 11-23.

17. Durrer, P., et al., *H+-induced membrane insertion of influenza virus hemagglutinin involves the HA2 amino-terminal fusion peptide but not the coiled coil region.* Journal of Biological Chemistry, 1996. **271**(23): p. 13417-13421.

18. Curtis-Fisk, J., *Structural studies of the Influenza and HIV viral fusion proteins and bacterial inclusion bodies.* Ph.D. Thesis, Michigan State University, 2009.

19. Gentile, M., et al., *DETERMINATION OF THE SIZE OF HIV USING ADENOVIRUS TYPE-2 AS AN INTERNAL LENGTH MARKER.* Journal of Virological Methods, 1994. **48**(1): p. 43-52.

20. Zhu, P., et al., *Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(26): p. 15812-15817.

21. Zhu, P., et al., *Distribution and three-dimensional structure of AIDS virus envelope spikes.* Nature, 2006. **441**(7095): p. 847-852.

22. White, J.M., et al., *Structures and mechanisms of viral membrane fusion proteins: Multiple variations on a common theme.* Critical Reviews in Biochemistry and Molecular Biology, 2008. **43**(3): p. 189-219.

23. Grewe, C., A. Beck, and H.R. Gelderblom, *HIV - EARLY VIRUS-CELL INTERACTIONS.* Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology, 1990. **3**(10): p. 965-974.

24. Sackett, K., et al., *Hairpin Folding of HIV gp41 Abrogates Lipid Mixing Function at Physiologic pH and Inhibits Lipid Mixing by Exposed gp41 Constructs.* Biochemistry, 2009. **48**(12): p. 2714-2722.

25. Chan, D.C., et al., *Core structure of gp41 from the HIV envelope glycoprotein.* Cell, 1997. **89**(2): p. 263-273.

26. Kwong, P.D., et al., *Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody.* Nature, 1998. **393**(6686): p. 648-659.

27. Buzon, V., et al., *Crystal Structure of HIV-1 gp41 Including Both Fusion Peptide and Membrane Proximal External Regions.* Plos Pathogens, 2010. **6**(5).

28. Markosyan, R.M., F.S. Cohen, and G.B. Melikyan, *HIV-1 envelope proteins complete their folding into six-helix bundles immediately after fusion pore formation.* Molecular Biology of the Cell, 2003. **14**(3): p. 926-938.

29. Jaroniec, C.P., et al., *Structure and dynamics of micelle-associated human immunodeficiency virus gp41 fusion domain.* Biochemistry, 2005. **44**(49): p. 16167-16180.

30. Nieva, J.L., S. Nir, and J. Wilschut, *Destabilization and fusion of zwitterionic large unilamellar lipid vesicles induced by a beta-type structure of the HIV-1 fusion peptide.* Journal of Liposome Research, 1998. **8**(2): p. 165-182.

31. Schmick, S.D. and D.P. Weliky, *Major Antiparallel and Minor Parallel beta Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide.* Biochemistry, 2010. **49**(50): p. 10623-10635.

32. Buzon, V., E. Padros, and J. Cladera, *Interaction of fusion peptides from HIV gp41 with membranes: A time-resolved membrane binding, lipid mixing, and structural study.* Biochemistry, 2005. **44**(40): p. 13354-13364.

33. Cheng, S.F., et al., *The fusion peptide domain is the primary membrane-inserted region and enhances membrane interaction of the ectodomain of HIV-1 gp41.* Molecular Membrane Biology, 2010. **27**(1): p. 31-44.

34. Han, X., et al., *Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin.* Nature Structural Biology, 2001. **8**(8): p. 715-720.

35. Qiang, W., Y. Sun, and D.P. Weliky, *A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(36): p. 15314-15319.

36. Tamm, L.K. and X. Han, *Viral fusion peptides: A tool set to disrupt and connect biological membranes.* Bioscience Reports, 2000. **20**(6): p. 501-518.

37. Yang, R., et al., *A trimeric HIV-1 fusion peptide construct which does not self-associate in aqueous solution end which Has 15-fold higher membrane fusion rate.* Journal of the American Chemical Society, 2004. **126**(45): p. 14722-14723.

38. Gabrys, C.M., et al., *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.* Journal of Physical Chemistry A, 2013. **117**(39): p. 9848-9859.

39. Agirre, A., et al., *Interactions of the HIV-1 fusion peptide with large unilamellar vesicles and monolayers. A cryo-TEM and spectroscopic study.* Biochimica Et Biophysica Acta-Biomembranes, 2000. **1467**(1): p. 153-164.

40. Haque, M.E., et al., *Properties and structures of the influenza and HIV fusion peptides on lipid membranes: Implications for a role in fusion.* Biophysical Journal, 2005. **89**(5): p. 3183-3194.

41. Grobner, G., C. Glaubitz, and A. Watts, *Probing membrane surfaces and the location of membrane-embedded peptides by C-13 MAS NMR using lanthanide ions.* Journal of Magnetic Resonance, 1999. **141**(2): p. 335-339.

42. Su, Y., R. Mani, and M. Hong, *Asymmetric insertion of membrane proteins in lipid bilayers by solid-state NMR paramagnetic relaxation enhancement: A cell-penetrating peptide example.* Journal of the American Chemical Society, 2008. **130**(27): p. 8856-8864.

43. Battiste, J.L. and G. Wagner, *Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data.* Biochemistry, 2000. **39**(18): p. 5355-5365.

44. Gallagher, G.J., M. Hong, and L.K. Thompson, *Solid-state NMR spin diffusion for measurement of membrane-bound peptide structure: Gramicidin A.* Biochemistry, 2004. **43**(24): p. 7899-7906.

45. Huster, D., X.L. Yao, and M. Hong, *Membrane protein topology probed by H-1 spin diffusion from lipids using solid-state NMR spectroscopy.* Journal of the American Chemical Society, 2002. **124**(5): p. 874-883.

46. Wang, T., H. Yao, and M. Hong, *Determining the depth of insertion of dynamically invisible membrane peptides by gel-phase H-1 spin diffusion heteronuclear correlation NMR.* Journal of Biomolecular Nmr, 2013. **56**(2): p. 139-148.

47. Xie, L., et al., *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.* Journal of Biomolecular NMR, 2013. **55**(1): p. 11-17.

48. Toke, O., et al., *Secondary structure and lipid contact of a peptide antibiotic in phospholipid bilayers by REDOR.* Biophysical Journal, 2004. **87**(1): p. 662-674.

49. Gordon, L.M., et al., *THE AMINO-TERMINAL PEPTIDE OF HIV-1 GLYCOPROTEIN-41 INTERACTS WITH HUMAN ERYTHROCYTE-MEMBRANES - PEPTIDE CONFORMATION, ORIENTATION AND AGGREGATION.* Biochimica Et Biophysica Acta, 1992. **1139**(4): p. 257-274.

50. Lai, A.L., et al., *Fusion Activity of HIV gp41 Fusion Domain Is Related to Its Secondary Structure and Depth of Membrane Insertion in a Cholesterol-Dependent Fashion.* Journal of Molecular Biology, 2012. **418**(1-2): p. 3-15.

Chapter 2 - Structural studies of Influenza fusion peptide (IFP)

2.1 Background

Influenza virus is a type of retrovirus (i.e. RNA virus) enveloped by a lipid membrane and it infects the respiratory epithelial cell. During the infection process, the influenza fusion peptide (IFP), which is the N-terminal ~20 residues of the HA2 subunit of the hemagglutinin protein, plays a vital role in the fusion between the viral and the host cell endosomal membranes, as evidenced by the disruption of the fusion activity of HA2 with point mutations in the IFP region[1]. Substitution of glutamic acid for glycine residue right at the amino-terminus of HA2 even completely abolishes its fusion activity.

The structure of IFP depends on the sequence and whether it is in detergents or lipid membranes. For instance, the H3_20 peptide, which has the sequence GLFGAIAGFIENGWEGMIDGGCGKKKK with the underlined part representing the wild type 20-residue IFP of serotype H3, adopts open structures at both pH 5.0 and 7.4 in detergents whereas the H1_23 peptide, which has the sequence GLFGAIAGFIEGGWTGMIDGWYGSGKKKKD with the underlined part representing the wild type 23-residue IFP of serotype H1, adopts a closed N-helix/turn/C-helix hairpin structure at pH 7.4 in detergents and a similar structure at pH 4.0 according to the similar HSQC (heteronuclear single quantum correlation) spectra at both pHs (Figure 2.1)[2, 3]. The H3_20 sequence differs from H1_23 in that the 12th and 15th residues are Asn and Glu in H3_20 and Gly and Thr in H1_23, respectively. Besides, H1_23 has three extra C-terminal residues, WYG. Although the H3_20 peptide has open structures in detergents at both pH 5.0 and 7.4, there are some conformational changes between the two pHs according to the study by Han and Tamm[2]. At pH 5.0, H3_20 adopts an

N-terminal α helix from residue Leu2 to Ile10, followed by a turn which redirects the C-terminal region of the peptide and is stabilized by hydrogen bonds from the amide NHs of Glu11 and Asn12 to the carbonyls of Gly8 and Phe9, and a C-terminal 3_{10} helix. At pH 7.4, the N-terminal α helix ranges from residue Leu2 to Phe9, followed by a turn stabilized by hydrogen bonds from the amide NHs of Glu11 and Asn12 to the carbonyls of Gly8 and Phe9. The C-terminal region forms an extended structure. Interestingly, a very recent study by Lorieau and Bax has shown that the H1_23 peptide adopts a helical structure at both N- and C-termini at pH 7.4[4]. This study also reports that the H1_20 peptide is highly dynamic and in equilibrium between a ~11% population of closed structure and a ~89% population of open structure when solubilized by DPC (dodecylphosphocholine) micelles at pH 7.4. Comparison of the major open structure of H1_20 with the predominant closed structure of H1_23 indicates that the three C-terminal residues, WYG, play a key role in stabilizing the closed structure of H1_23 in detergents.

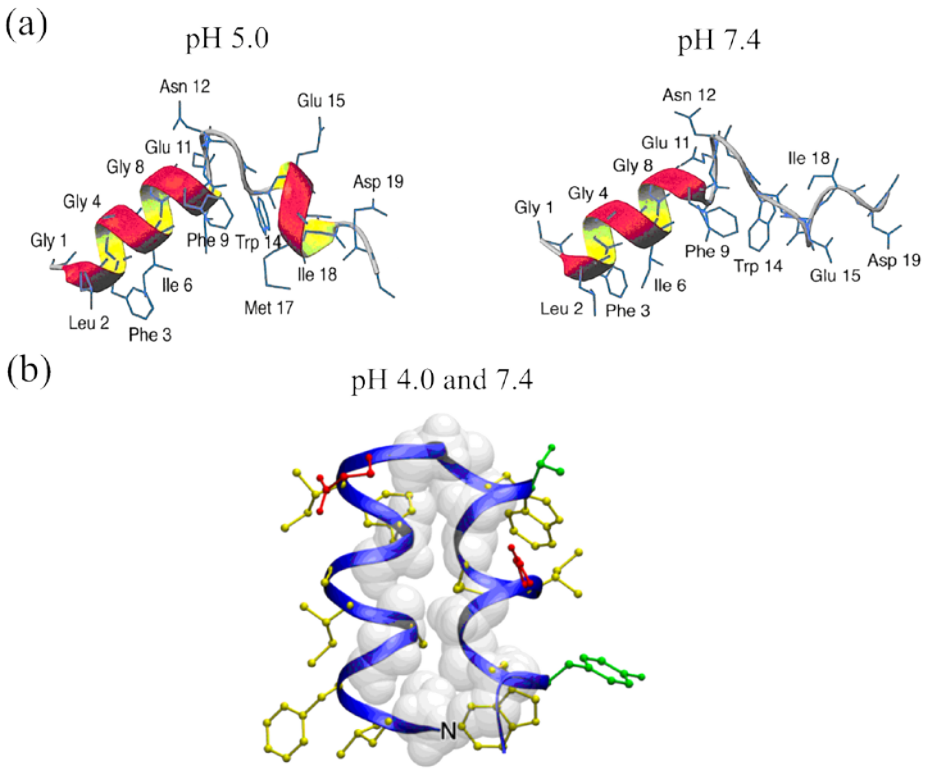

Figure 2.1 Structures of the influenza hemagglutinin fusion peptides H3_20 (a) and H1_23 (b) in DPC detergents[2, 3].

Additional liquid state NMR studies have been conducted on the wild type H1_23 and its mutant H1_23_G8A where the residue Gly8 was substituted by Ala[5]. Relaxation data indicate that the wild type H1_23 adopts a closed structure (helical hairpin) in DPC micelles at pH 7. However, at pH 4, H1_23 adopts a major (~80% population) closed structure and a minor (~20% population) open structure where the two structures are rapidly exchanging with each other at a rate of ~40 kHz. For the mutant H1_23_G8A in detergents at pH 7, it adopts a small fraction (~15%) of closed structure and a large fraction of at least two open structures which are classified as extended and L-shaped, respectively (Figure 2.2). The topology of these N-helix/turn/C-helix structures can be described by the interhelical angle which is defined in such a

way that each helix axis is represented as a vector from the N- to the C-terminus and the angle between the two vectors refers to the interhelical angle. The closed, extended, and L-shaped structures of the H1_23_G8A have interhelical angles of $159^{o} \pm 1^{o}$, $73^{o} \pm 11^{o}$, and $110^{o} \pm 6^{o}$, respectively[5]. It has previously been reported that cells expressing the G8A mutant of HA2 had little membrane fusion activity[6].

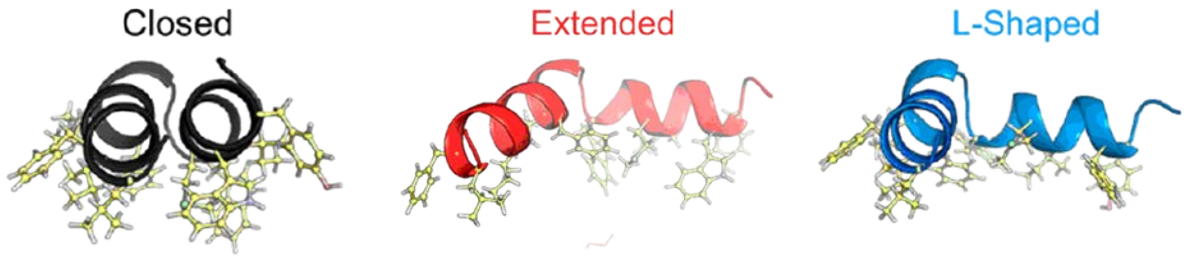

Figure 2.2 N-helix/turn/C-helix structures of the H1_23_G8A mutant in DPC micelles at pH 7[5].

For the H3_20 peptide in lipid membranes, previous SSNMR studies have shown that it adopts predominantly helical structure in membranes without cholesterol and predominantly β sheet structure in membranes containing ~30% (mole fraction) cholesterol (e.g. LM3 membrane)[7]. In membranes without cholesterol, H3_20 adopts an N-helix/turn/C-helix motif at both pH 5.0 and 7.4. Two different conformations of the turn are observed at pH 5.0 since there are two distinct ^{13}C chemical shift sets for Glu11 in the turn region (Figure 2.3)[8]. The fusion activity of IFP has a strong pH dependence. Previous fluorescence and hemolytic studies indicate that IFP induces more rapid and extensive lipid mixing (membrane fusion) at pH 5.0 than at physiological pH 7.4 for cholesterol-free liposomes[9-11]. However, for cholesterol-containing

liposomes, IFP induces lipid mixing at low pHs such as 5.0 but not at high pHs such as 7.4[9, 11].

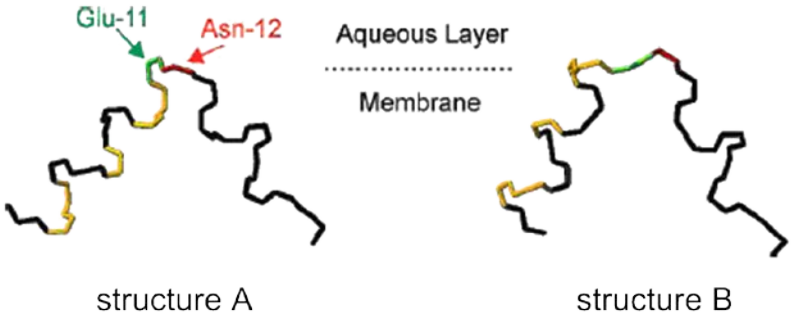

Figure 2.3 Backbone structures of the membrane-associated H3_20 peptide at pH 5.0. Residue E11 is in green, N12 is in red, the hydrophobic residues L2, F3, I6, F9, and I10 are in gold[8].

In the present study, ^{13}C-^{15}N REDOR SSNMR is employed to quantitatively probe the population and interhelical proximity of the H3_20 structures in lipid membranes.

2.2 Sample preparation

Peptide sequence and synthesis

The H3_20 peptide, which has the sequence GLFGAIAGFIENGWEGMIDGGGKKKKG with the underlined part representing the wild type 20-residue IFP of serotype H3, was chemically synthesized by Fmoc solid phase peptide synthesis (SPPS) with a specific ^{13}CO label at residue G16 and a specific ^{15}N label at residue F9[12]. The Fmoc SPPS protocol is described in detail in Appendix II. After the synthesis, H3_20 was purified by reversed-phase high pressure liquid chromatography (HPLC) using a preparative C4 column. The peptide purity was checked by mass spectrometry such as matrix-assisted laser desorption ionization

(MALDI) and electrospray ionization (ESI). Comparison of peak intensities in the mass spectrum indicated the peptide was >95% pure.

REDOR sample preparation

The REDOR samples were prepared in a time sequence of (1) dissolving 40 μmol of 1,2-di-O-tetradecyl-sn-glycero-3-phosphocholine (DTPC) and 10 μmol of 1,2-di-O-tetradecyl-sn-glycero-3-[phospho-(1'-rac-glycerol)] (DTPG) lipids in 3~4 mL chloroform; (2) removing the chloroform using N_2 gas in the fume hood; (3) drying the lipid film in a vacuum desicator overnight; (4) suspending the lipid film in ~2 mL pH 5.1 buffer which contains 5 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and 10 mM 2-(N-morpholino)ethanesulfonic acid (MES); (5) performing ten cycles of "freeze/thaw", where the "freeze" in liquid N_2 breaks apart the lipid vesicles and the "thaw" in water bath at room temperature drives the lipid molecules to reform the vesicles; (6) performing ~20 extrusions using a filter paper with pores of 100 nm in diameter (each time only 1 mL lipid solution was extruded due to the volume limit of the syringe); (7) dissolving 2 μmol of the H3_20 peptide in ~30 mL pH 5.1 buffer; (8) adding the peptide solution to the extruded lipid vesicle solution dropwise with gentle shaking (avoid violent shaking in case of peptide oxidation at the methionine site); (9) mixing the peptide/lipid solution overnight on a rotator; (10) ultracentrifugation at ~100,000 g for 4 hours; (11) lyophilizing the pellet overnight; (12) adding 10 μL pH 5.1 buffer to the bottom of a 4mm rotor, packing the lyophilized pellet in the rotor, adding another 10 μL buffer on top of the sample, and finally allowing overnight sample hydration before running the REDOR experiment. The DTPC:DTPG (4:1) lipid composition reflects a small fraction of

negatively charged lipids and a significant fraction of phosphatidylcholine lipids in the membranes of host epithelial cells of the influenza virus[13]. Ether- rather than ester-linked lipids were used to reduce the natural abundance (n.a.) ^{13}CO signal in the REDOR spectra.

2.3 ^{13}C-^{15}N REDOR experiments

REDOR pulse sequence and parameters

The ^{13}C-^{15}N REDOR pulse sequence consists of (1) ^{1}H $\pi/2$ pulse which rotates ^{1}H magnetization from z axis to the xy plane, (2) simultaneous ^{1}H and ^{13}C CP pulses which transfer magnetization from ^{1}H to ^{13}C to enhance the ^{13}C signal intensity, (3) ^{13}C π pulses at the end of each rotor period except the last one without (S_0 experiment) and with (S_1 experiment) ^{15}N π pulses in the middle of each rotor period, in the meantime, two pulse phase modulation (TPPM) ^{1}H decoupling π pulses are turned on till the end of ^{13}C signal acquisition (Figure 2.4). The period right after CP but prior to ^{13}C acquisition is called dephasing time τ. The REDOR experimental conditions included 10 kHz MAS frequency, sample cooling gas temperature of -50 ^{o}C with actual sample temperature of approximately -30 ^{o}C, 50 kHz ^{1}H $\pi/2$ pulse and CP pulse, 80 kHz ^{1}H decoupling pulses during both the dephasing and acquisition periods, 61-65 kHz ramped ^{13}C CP pulse, 8.0 μs ^{13}C π pulses, and 10.0 μs ^{15}N π pulses. The pulse delay

was 1 s for τ = 2 ms, 8 ms, and 16 ms, 1.5 s for τ = 24 ms and 32 ms, and 2 s for τ = 40 ms and 48 ms, respectively. The data collection time is typically 2~3 hours for τ = 2 and 8 ms, 6~12 hours for τ = 16 ms, ~1 day for τ = 24 ms, ~1.5 days for τ = 32 ms, ~2 days for τ = 40 ms, and 2~3 days for τ = 48 ms, respectively. The pulses were optimized using the lyophilized setup peptide I4 which had the sequence Ac-AEAAAKEAAAKEAAAKA-NH_2 with N-terminal acetylation and C-terminal amidation. The I4 peptide was synthesized with a ^{13}CO label at residue A9 and an ^{15}N label at residue A13. Previous SSNMR studies have shown that I4 is predominantly (83±6%) α helical at A9[14]. The internuclear distance between A9 ^{13}CO and A13 ^{15}N of I4 is ~4.1 Å in the α helical structure which corresponds to a ^{13}C-^{15}N dipolar coupling of ~44 Hz (d(Hz) = 3066/r^3(Å))[15].

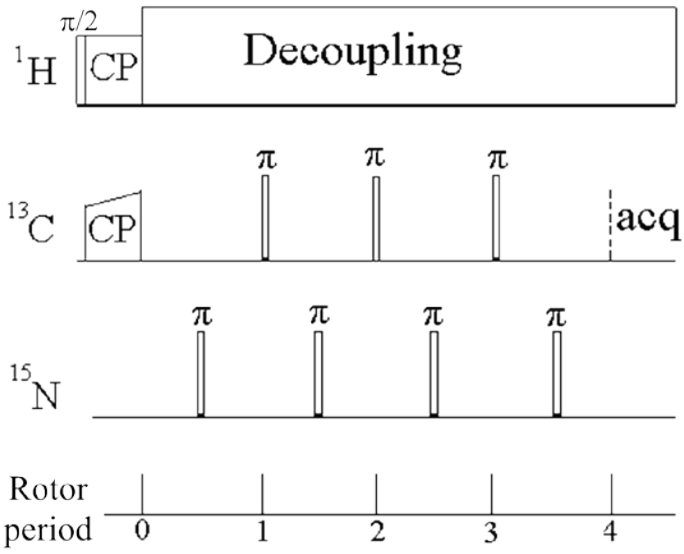

Figure 2.4 ^{13}C-^{15}N REDOR pulse sequence.

As mentioned before, the H3_20 peptide was chemically synthesized with a specific ^{13}CO label at G16 (G16C) and a specific ^{15}N label at F9 (F9N). This labeling scheme is based on earlier solution NMR studies. The internuclear distance between F9N and G16C is ~11.5 Å in the open structure of H3_20 at pH 5.0 in detergents and ~3.9 Å in the closed structure of H1_23 at pH 4.0 in detergents[2, 3]. If H3_20 adopts an open structure in the membrane where the F9N and G16C are ~11.5 Å apart with a corresponding ^{13}C-^{15}N dipolar coupling $d \approx 2$ Hz, there will be no ^{13}C-^{15}N REDOR buildup of $\Delta S/S_0$ vs τ. However, if H3_20 adopts a closed structure in the membrane where the F9N and G16C are ~3.9 Å apart with a corresponding ^{13}C-^{15}N dipolar coupling $d \approx 52$ Hz, significant buildup of $\Delta S/S_0$ vs τ from 2 ms to 48 ms will be observed, as displayed in Figure 2.5.

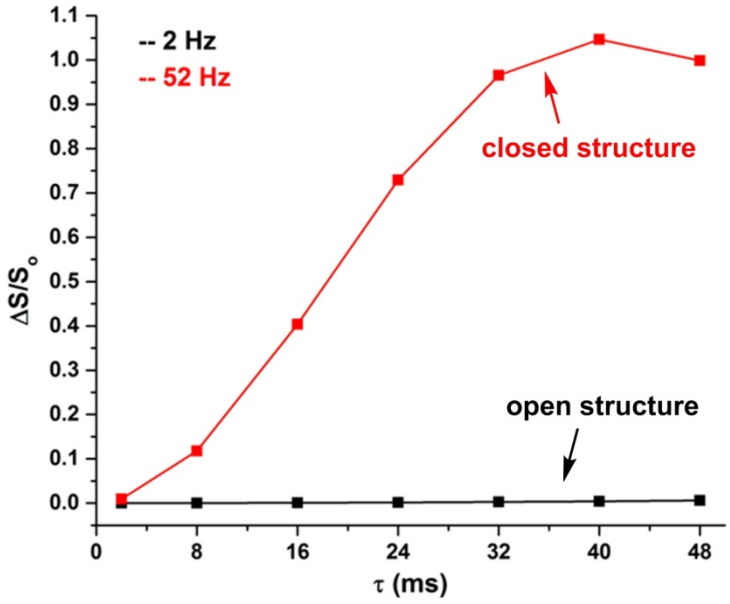

Figure 2.5 ^{13}C-^{15}N REDOR dephasing curves of $\Delta S/S_0$ vs τ simulated by SIMPSON for ^{13}C-^{15}N dipolar couplings of 2 Hz (black) and 52 Hz (red), where 2 Hz corresponds to an open structure and 52 Hz corresponds to a closed structure of H3_20, respectively.

REDOR spectra and dephasing curve of $\Delta S/S_0$ vs τ

To detect whether H3_20 adopts predominantly open or closed structures in the membrane, ^{13}C-^{15}N REDOR was conducted under experimental conditions described earlier. The experimental dephasing curve of $\Delta S/S_0$ vs τ was displayed in Figure 2.6a. The S_0, S_1, and ΔS spectra for τ = 32 ms were displayed in Figure 2.6b, where ΔS spectrum refers to the difference spectrum of S_0 - S_1. All the S_0 and S_1 spectra at each τ were processed with 20 Hz exponential line broadening and baseline correction. The $(\Delta S/S_0)^{exp}$ were calculated from the S_0 and S_1 intensities integrated in a 3 ppm window centered at the ^{13}CO peak shift. The G16 ^{13}CO peak shift of 177.1 ppm is consistent with helical conformation of H3_20[16]. The experimental errors of S_0 and S_1 were calculated as the standard deviation of the integrals of twelve noise regions with each in a 3 ppm window in the S_0 and S_1 spectra, respectively.

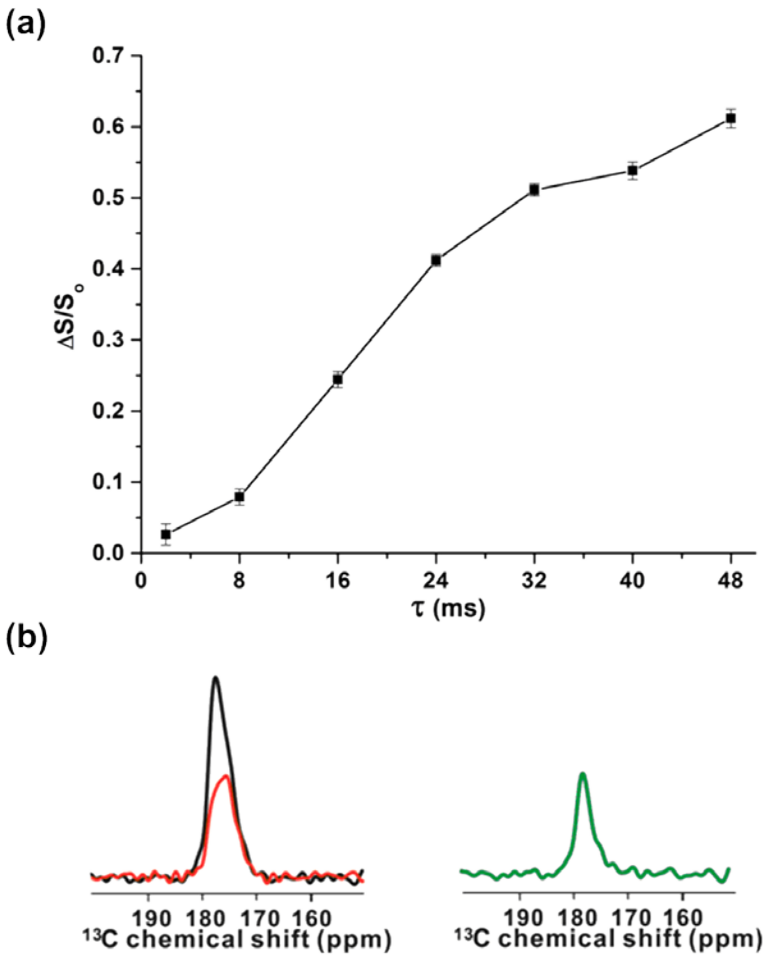

Figure 2.6 (a) Experimental curve of $\Delta S/S_0$ vs τ of the H3_20 peptide; (b) ^{13}CO peaks in the S_0 (black), S_1 (red), and ΔS (green) spectra for τ = 32 ms.

2.4 Result discussion and conclusions

Natural abundance calibration of $(\Delta S/S_0)^{exp}$

The significant buildup of $(\Delta S/S_0)^{exp}$ vs τ of H3_20 supports that H3_20 adopts predominantly closed rather than open structure. SIMPSON simulations were performed to quantitatively analyze the structure models of H3_20 in the membrane. To make the data analysis more accurate, natural abundance (n.a.) contributions to

$(\Delta S/S_o)^{exp}$ from unlabeled ^{13}CO signals were removed to yield $(\Delta S/S_o)^{lab}$ which refers to $\Delta S/S_o$ only for the labeled G16 ^{13}CO signal. The n.a. correction was done as follows:

$$S_o^{\exp} = S_o^{lab} + S_o^{na} = 0.99 + 0.011x26 = 1.276 \quad (2.1)$$

where S_o^{lab} refers to the labeled G16 ^{13}CO S_o signal which equals 0.99 and S_o^{na} refers to the n.a. ^{13}CO S_o signal for the 26 unlabeled residues in H3_20.

$$S_1^{\exp} = S_1^{lab} + S_1^{na} = S_1^{lab} + \sum_{k=1}^{26} S_{1k}^{na} \quad (2.2)$$

where *k* represents the number of the 26 unlabeled residues in H3_20.

For each unlabeled residue,

$$(\frac{\Delta S}{S_o})^{na} = (\frac{S_o^{na} - S_1^{na}}{S_o^{na}}) = \frac{0.011 - S_1^{na}}{0.011} \quad (2.3)$$

$$S_1^{na} = 0.011 - 0.011x(\frac{\Delta S}{S_o})^{na} \quad (2.4)$$

For all the 26 unlabeled residues,

$$\sum_{k=1}^{26} S_{1k}^{na} = \sum_{k=1}^{26} [0.011 - 0.011x(\frac{\Delta S}{S_o})^{na}] = 0.286 - 0.011x \sum_{k=1}^{26} (\frac{\Delta S}{S_o})_k^{na} \quad (2.5)$$

Substitution of equation 2.5 into 2.2 yields

$$S_1^{\exp} = S_1^{lab} + 0.286 - 0.011x \sum_{k=1}^{26} (\frac{\Delta S}{S_o})_k^{na} \quad (2.6)$$

Therefore,

$$(\frac{\Delta S}{S_o})^{\exp} = \frac{S_o^{\exp} - S_1^{\exp}}{S_o^{\exp}} = \frac{1.276 - S_1^{lab} - 0.286 + 0.011x\sum_{k=1}^{26}(\frac{\Delta S}{S_o})_k^{na}}{1.276}$$

$$= \frac{0.99 - S_1^{lab} + 0.011x\sum_{k=1}^{26}(\frac{\Delta S}{S_o})_k^{na}}{1.276} \quad (2.7)$$

Rearrangement of equation 2.7 yields

$$S_1^{lab} = 0.99 - 1.276x(\frac{\Delta S}{S_0})^{\exp} + 0.011x\sum_{k=1}^{26}(\frac{\Delta S}{S_0})_k^{na} \quad (2.8)$$

Finally,

$$(\frac{\Delta S}{S_o})^{lab} = \frac{S_o^{lab} - S_1^{lab}}{S_o^{lab}} = \frac{0.99 - S_1^{lab}}{0.99} = \frac{1.276}{0.99}x(\frac{\Delta S}{S_o})^{\exp} - \frac{0.011}{0.99}x\sum_{k=1}^{26}(\frac{\Delta S}{S_o})_k^{na} \quad (2.9)$$

To simulate $(\frac{\Delta S}{S_o})_k^{na}$ by SIMPSON, the ^{13}CO-^{15}N dipolar coupling needs to be provided and thus the ^{13}CO-^{15}N distance needs to be estimated. Since the $(\Delta S/S_0)^{\exp}$ vs τ of H3_20 has a significant buildup (Figure 2.6a), the peptide may adopt a closed structure in which the labeled G16 ^{13}CO and F9 ^{15}N nuclei are close to each other. Therefore, a reasonable way to estimate the distance between the n.a. ^{13}CO and labeled ^{15}N in H3_20 is to measure the distance in the closed structure of H1_23[3]. The ^{13}CO-^{15}N distance can be either mathematically calculated using the atomic coordinates of H1_23 or directly measured using a structure visualizing software such as PyMol. Table 2.1 displays the distances calculated using the atomic coordinates of

H1_23. For other ^{13}CO-^{15}N distances not listed in Table 2.1, they are beyond 8 Å and the corresponding $(\frac{\Delta S}{S_o})^{na} \approx 0$.

Table 2.1 ^{13}CO-^{15}N distances in the closed structure of H1_23[3].

^{13}CO-^{15}N spin pair	*r* (Å)
F3CO-F9N	7.38
G4CO-F9N	5.40
A5CO-F9N	3.81
I6CO-F9N	3.74
A7CO-F9N	3.14
G8CO-F9N	1.33
F9CO-F9N	2.45
I10CO-F9N	4.69
E11CO-F9N	5.64
G12CO-F9N[a]	5.27
G13CO-F9N	5.23
W14CO-F9N	6.58
T15CO-F9N[b]	6.52
M17CO-F9N	5.36
I18CO-F9N	7.73
D19CO-F9N	7.47
G20CO-F9N	7.28

[a] The 12th residue is Gly in H1_23 and Asn in H3_20.

[b] The 15th residue is Thr in H1_23 and Glu in H3_20.

The ^{13}CO-^{15}N dipolar coupling *d* in Hz is calculated via the equation $d = 3066/r^3$, where *r* is the ^{13}CO-^{15}N distance in Å. After obtaining the dipolar coupling *d*, SIMPSON simulations were performed to produce the $(\frac{\Delta S}{S_o})^{na}$ from τ = 2 ms to τ = 48 ms for each

^{13}CO-^{15}N spin pair listed in Table 2.1. The sum of $(\frac{\Delta S}{S_o})^{na}$ for each τ for all the ^{13}CO-^{15}N spin pairs in Table 2.1 approximately equaled the $\sum_{k=1}^{26}(\frac{\Delta S}{S_0})_k^{na}$ in equation 2.9. At last, the experimental value of $(\frac{\Delta S}{S_o})^{\exp}$ and simulated value of $\sum_{k=1}^{26}(\frac{\Delta S}{S_0})_k^{na}$ were plugged in equation 2.9 to yield the $(\frac{\Delta S}{S_o})^{lab}$ for each τ. The buildup curve of $(\frac{\Delta S}{S_o})^{lab}$ vs τ after the n.a. correction is displayed in Figure 2.7. The $(\frac{\Delta S}{S_o})^{lab}$ was typically higher than $(\frac{\Delta S}{S_o})^{\exp}$ but the difference was less than 10% for each τ from 2 ms to 48 ms.

SIMPSON simulations of $(\Delta S/S_o)^{lab}$ vs τ

Quantitative analysis of $(\frac{\Delta S}{S_o})^{lab}$ vs τ was first done for a single closed structure model. SIMPSON simulations were performed for a range of ^{13}CO-^{15}N dipolar coupling *d*. For each *d*, the $\chi^2(d)$ was calculated via

$$\chi^2(d) = \sum_{\tau} \frac{[(\frac{\Delta S}{S_o})^{lab}(\tau) - (\frac{\Delta S}{S_o})^{sim}(d,\tau)]^2}{\sigma(\tau)^2} \qquad (2.10)$$

where $\chi^2(d)$ was summed over the seven values of τ from 2 ms to 48 ms and $\sigma(\tau)$ was the uncertainty of $(\frac{\Delta S}{S_o})^{lab}$. The best-fit *d* was 32 Hz with a minimum $\chi^2(d)$ of 436. The $\chi^2(d)$ of 436 was much larger than the number of degrees of fitting ν of 6 where ν

= the number of fitted data points - the number of fitting parameters. The poor fitting of this structure model is displayed in Figure 2.7.

Secondly, a "closed/open" structure model was considered. In this model, H3_20 adopts an open structure with a fraction of f_O and a ^{13}CO-^{15}N dipolar coupling of d_O and a closed structure with a fraction of f_C and a ^{13}CO-^{15}N dipolar coupling of d_C, where $f_O + f_C = 1$ and $d_O \approx 0$ with consequent $(\frac{\Delta S}{S_o})^o \approx 0$. SIMPSON simulations were performed for a range of ^{13}CO-^{15}N dipolar coupling d_C and a range of the fraction f_C from 0.01 to 1.00 in an increment of 0.01. For each d_C and f_C, the $\chi^2(d_c,f_c)$ was calculated via

$$\chi^2(d_c,f_c) = \sum_{\tau} \frac{[(\frac{\Delta S}{S_o})^{lab}(\tau) - f_c x (\frac{\Delta S}{S_o})^c(d_c,\tau)]^2}{\sigma(\tau)^2} \qquad (2.11)$$

where $\chi^2(d_c,f_c)$ was summed over the seven values of τ. The best-fit d_c = 51 Hz, f_c = 0.64, $\chi^2(d_c,f_c)$ = 33. However, the χ^2 = 33 was still significantly larger than the number of degrees of fitting ν = 5. Therefore, the "closed/open" model was also statistically unreasonable. The poor fitting of this model is displayed in Figure 2.7.

At last, a "closed/semi-closed" structure model was considered. In this model, H3_20 adopts a closed structure with a fraction of f_C and a ^{13}CO-^{15}N dipolar coupling of d_C and a semi-closed structure with a fraction of f_S and a ^{13}CO-^{15}N dipolar coupling of

d_S and $f_C + f_S = 1$. SIMPSON simulations were then performed for a range of d_C, d_S , and f_C. The $\chi^2(d_c,d_s,f_c)$ was calculated via

$$\chi^2(d_c,d_s,f_c) = \sum_{\tau} \frac{[(\frac{\Delta S}{S_o})^{lab}(\tau) - f_c x(\frac{\Delta S}{S_o})^{c}(d_c,\tau) - (1-f_c)x(\frac{\Delta S}{S_o})^{s}(d_s,\tau)]^2}{\sigma(\tau)^2} \qquad (2.11)$$

The best-fit simulation yielded d_C = 65.4 Hz, d_S = 21.2 Hz, f_C = 0.41, and $\chi^2(d_c,d_s,f_c)$ = 5.7. The χ^2 = 5.7 was close to the number of degrees of fitting ν = 4. The good fitting of this model can be also visualized in Figure 2.7. One measure of the uncertainty of the best-fit value is the deviation which causes an increase in χ^2 by 1. For instance, the uncertainty of d_C was the deviation from 65.4 Hz that led to χ^2 = 6.7 while keeping d_S = 21.2 Hz and f_C = 0.41. Application of this method yielded f_C = 0.41 ± 0.01, d_C = 65.4 ± 1.7 Hz with corresponding ^{13}CO-^{15}N distance r_C = 3.61 ± 0.03 Å, and d_S = 21.2 ± 0.5 Hz with corresponding r_S = 5.25 ± 0.04 Å. These results indicate that the membrane-associated H3_20 peptide adopts ~40% closed structure and ~60% semi-closed structure at pH 5.1[17]. The two structures are displayed in Figure 2.8. To check the reproducibility of the results, a second sample was prepared in the same way and the new REDOR data were acquired under the same experimental conditions. SIMPSON simulations yielded the best-fit χ^2 = 0.83, d_C = 63.2 Hz with r_C = 3.65 Å, d_S = 19.0 Hz with r_S = 5.44 Å, and f_C = 0.38. These values were comparable to those for the first sample.

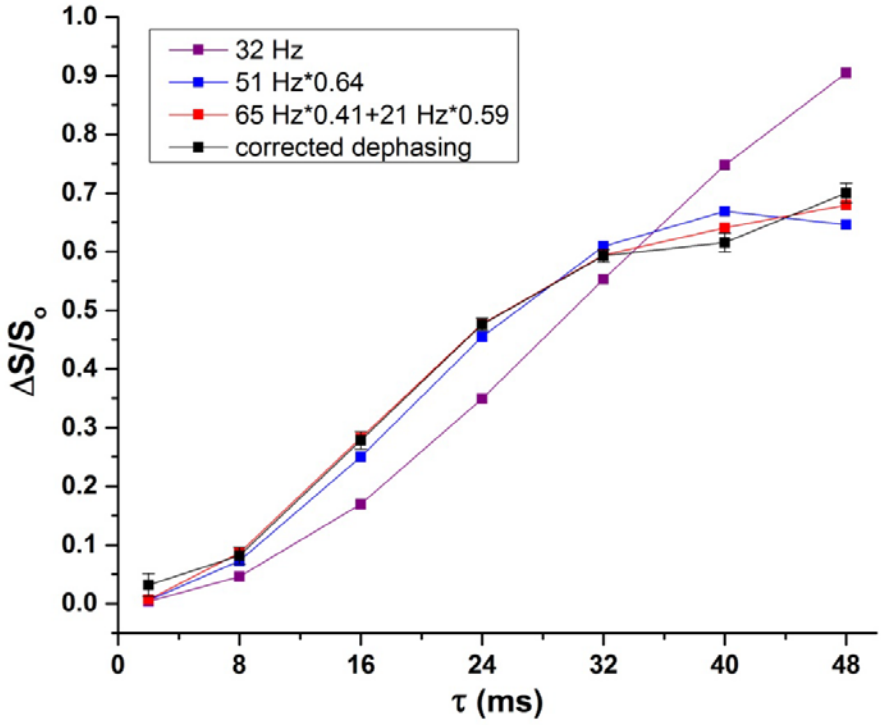

Figure 2.7 Plots of $(\frac{\Delta S}{S_o})^{lab}$ vs τ (black) and best-fit $(\frac{\Delta S}{S_o})^{sim}$ vs τ for 32 Hz (purple) which corresponds to the "single closed structure" model, 51 Hz x 0.64 (blue) which corresponds to the "closed/open" model, and 65 Hz x 0.41 + 21 Hz x 0.59 (red) which corresponds to the "closed/semi-closed" model, respectively.

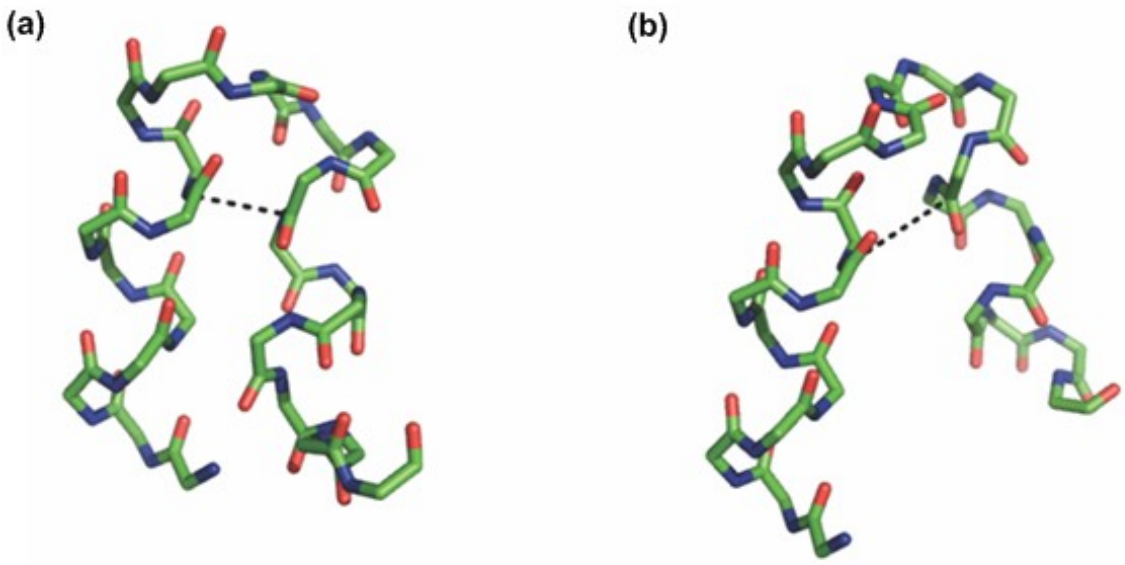

Figure 2.8 Backbone structural models of H3_20 in lipid membranes at pH 5.1. The carbon, nitrogen, and oxygen atoms are in green, blue, and red, respectively. The G16 ^{13}CO-F9 ^{15}N internuclear distance is 3.6 Å in the closed structure (a) and 5.2 Å in the semi-closed structure (b).

Intramolecular vs intermolecular dipolar interactions

In the above data analysis, the ^{13}CO-^{15}N dipolar coupling was assigned to be an intramolecular interaction. This assignment is supported by several pieces of evidence. First, the best-fit results of the REDOR data of H3_20 indicate that ~40% of the peptide molecules have a G16 ^{13}CO-F9 ^{15}N distance of 3.6 Å which is very close to the intramolecular distance (3.9 Å) observed in the closed structure of H1_23 in detergents[3]. Second, an earlier solid-state NMR study has supported an N-helix/turn/C-helix structure of H3_20 in membranes[8]. If the experimentally observed large dephasing of H3_20 were due to intermolecular dipolar interactions, large aggregates of H3_20 would probably be required. However, another solid-state NMR study of H3_20 indicates that the peptide molecules have significant motion at the

ambient temperature which is not consistent with a large aggregate of the peptide[18]. As we freeze the sample in liquid N_2 and then keep cooling it with -50 ^{o}C N_2 gas throughout the REDOR experiments, we do not expect the formation of large aggregates of H3_20 during the experiments. Third, although the formation of antiparallel homodimers of H3_20 could explain the significant dephasing of H3_20 in membranes, this model is probably incorrect since there is strong evidence for a turn in the middle of the H3_20 sequence, as discussed earlier. Furthermore, previous solution NMR studies of H1_23 in detergents have ruled out the antiparallel homodimeric arrangements of the peptide molecules since the experimentally observed interhelical nuclear Overhauser effects (NOEs) between the N-terminal and C-terminal residues have been proven to be intramolecular rather than intermolecular[3]. Fourth, earlier EPR studies have been done on a construct consisting of residues 1-127 of the HA2 subunit of the influenza hemagglutinin protein[19]. This construct is trimeric and induces vesicle fusion under acidic conditions similar to the full length HA. EPR spectra of the spin-labeled construct showed that the fusion peptide region is likely to exist as a monomer in the membrane at both neutral and acidic pH conditions.

To further confirm whether the ^{13}CO-^{15}N dipolar coupling is intra- or intermolecular, a control REDOR experiment was run with a sample containing 50% labeled and 50% unlabeled H3_20 in the membrane at pH 5.0. If the ^{13}CO-^{15}N dipolar interaction is predominantly intramolecular, the dephasing buildup of $\frac{\Delta S}{S_o}$ vs τ for the 100% labeled H3_20 sample will be comparable to that for the 50% labeled H3_20

sample. By contrast, if the ^{13}CO-^{15}N dipolar interaction is predominantly intermolecular, the dephasing buildup rate and extent for the 100% labeled sample will be different than those for the 50% labeled sample. For instance, as we consider a dimer model in which two peptide molecules aggregate as a dimer and the G16 ^{13}CO in molecule is dipolar coupled to the F9 ^{15}N in the other molecule, there will be three different types of dimer consisting of (i) two labeled molecules; (ii) one labeled molecule and one unlabeled molecule; and (iii) two unlabeled molecules, respectively (Figure 2.9a). The probability of (i), (ii), and (iii) is 25%, 50%, and 25%, respectively. Suppose there are 25 dimers of (i), 50 dimers of (ii), and 25 dimers of (iii), the fractional contribution to the overall G16 ^{13}CO signal intensity will be 0.50 from (i) and 0.50 from (ii). If the ^{13}CO-^{15}N dipolar interaction is predominantly intermolecular, then ~1/2 of the G16 ^{13}CO signal from (i) will be dephased in the REDOR experiment, i.e. ~1/4 of the overall G16 ^{13}CO signal will be dephased, as opposed to the 100% labeled sample for which ~1/2 of the overall G16 ^{13}CO signal is dephased. In this case, the $(\frac{\Delta S}{S_o})^{lab}$ for the 50% labeled sample will be approximately half of that for the 100% labeled sample for any given dephasing time τ, where $(\frac{\Delta S}{S_o})^{lab}$ is the dephasing for the labeled G16 ^{13}CO signal only (Figure 2.9b). However, the REDOR data show that the 50% and 100% labeled samples had very

similar buildups of $(\frac{\Delta S}{S_o})^{lab}$ vs τ (Figure 2.10), supporting that the ^{13}CO-^{15}N dipolar coupling is predominantly intra- rather than intermolecular.

Conclusions

In summary, our REDOR studies show that H3_20 adopts majorly closed rather than open structures in the membrane. It is important to distinguish between the open and closed structures since they lead to different peptide-lipid interaction models. For the open "boomerang" structure, the hydrophobic side chains are majorly in the interior of the structure, creating a hydrophobic pocket that allows a deep insertion of the fusion peptide into the membrane hydrocarbon core[2]. However, for the closed structure, the hydrophobic side chains are majorly on one side which interacts with the lipid membrane while the Gly residues are on the other side which is exposed to the solvent[3].

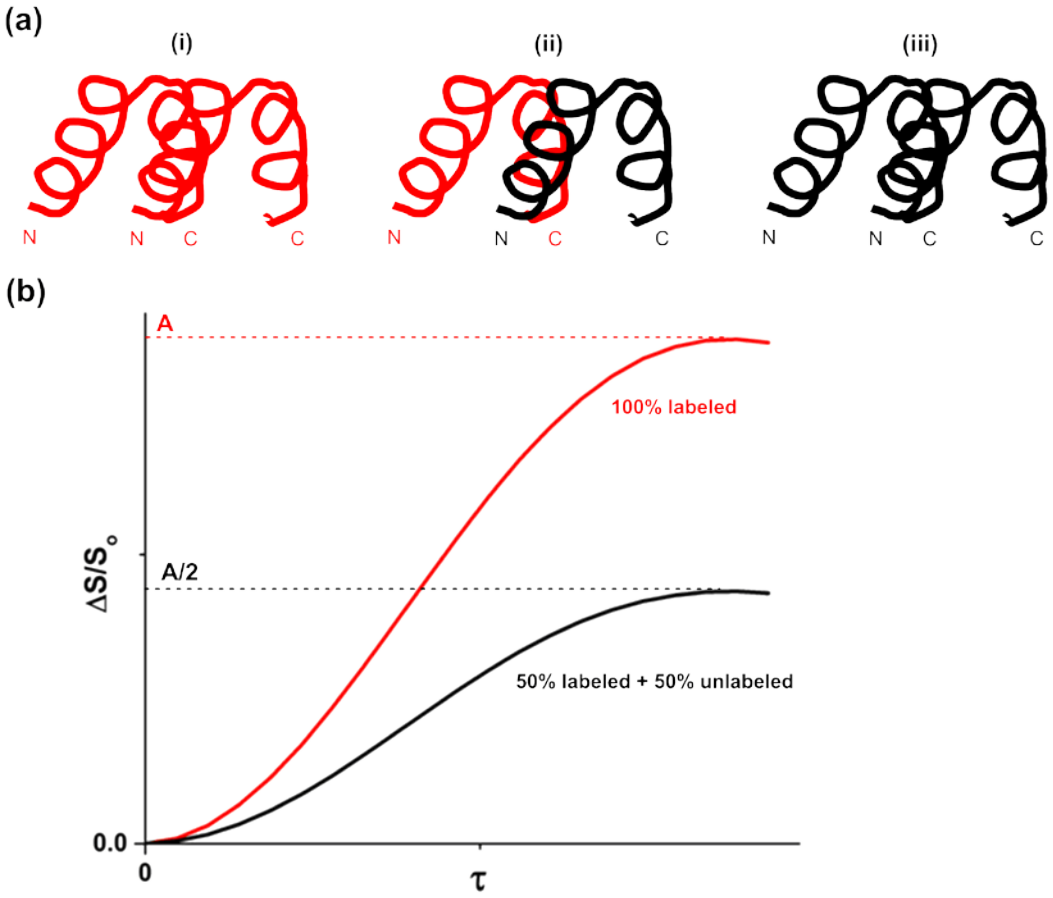

Figure 2.9 (a) Possible aggregates of H3_20 for the dimer model for the "50% labeled + 50% unlabeled" sample. The G16 ^{13}CO and F9 ^{15}N labeled H3_20 is in red while the unlabeled H3_20 is in black. (b) Theoretical ^{13}CO-^{15}N REDOR dephasing plots of $(\frac{\Delta S}{S_o})^{lab}$ vs τ for the 100% labeled H3_20 sample (red) and the "50% labeled + 50% unlabeled" H3_20 sample (black). The $(\frac{\Delta S}{S_o})^{lab}$ for the 100% labeled sample is twice that for the 50% labeled sample for any given dephasing time τ.

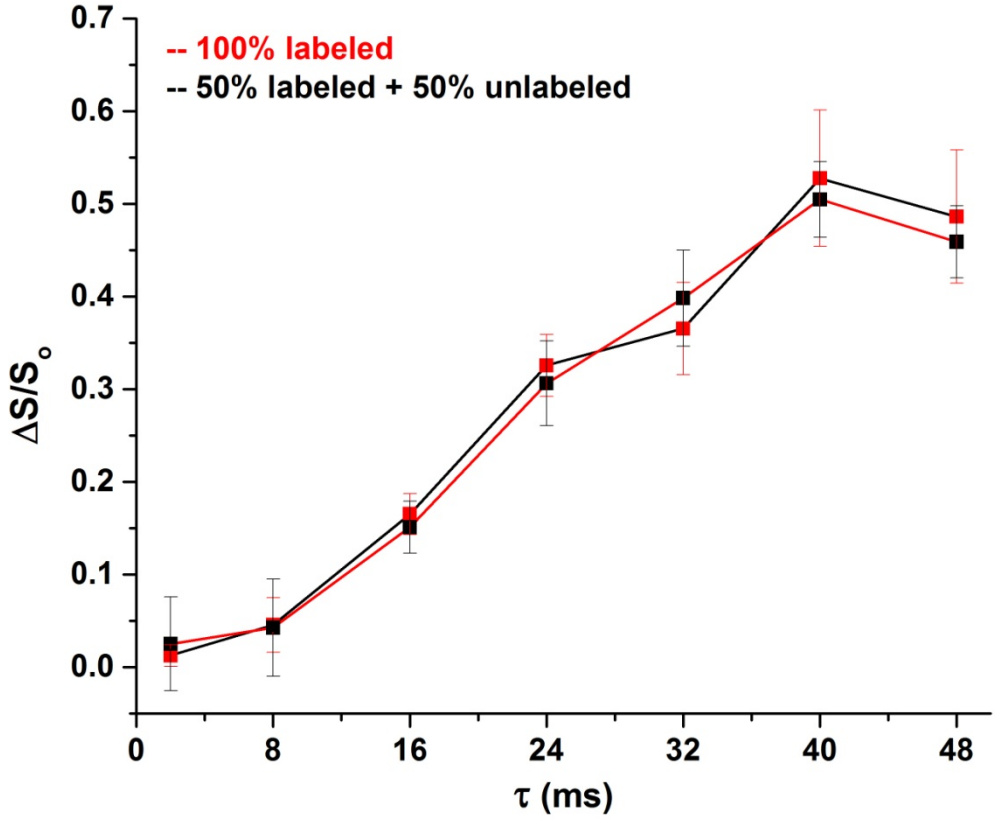

Figure 2.10 ^{13}CO-^{15}N REDOR plots of $(\frac{\Delta S}{S_o})^{lab}$ vs τ after n.a. correction for the membrane sample containing 100% labeled H3_20 (red) and the one containing 50% labeled and 50% unlabeled H3_20 (black) at pH 5.0 (courtesy from Ujjayini Ghosh, the REDOR data for the two samples were acquired under the same optimized experimental conditions which were different than those for Figure 2.6).

REFERENCES

REFERENCES

1. Gething, M.J., et al., *STUDIES ON THE MECHANISM OF MEMBRANE-FUSION - SITE-SPECIFIC MUTAGENESIS OF THE HEMAGGLUTININ OF INFLUENZA-VIRUS.* Journal of Cell Biology, 1986. **102**(1): p. 11-23.

2. Han, X., et al., *Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin.* Nature Structural Biology, 2001. **8**(8): p. 715-720.

3. Lorieau, J.L., J.M. Louis, and A. Bax, *The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(25): p. 11341-11346.

4. Lorieau, J.L., J.M. Louis, and A. Bax, *Impact of Influenza Hemagglutinin Fusion Peptide Length and Viral Subtype on Its Structure and Dynamics.* Biopolymers, 2013. **99**(3): p. 189-195.

5. Lorieau, J.L., et al., *pH-triggered, activated-state conformations of the influenza hemagglutinin fusion peptide revealed by NMR.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(49): p. 19994-19999.

6. Steinhauer, D.A., et al., *STUDIES OF THE MEMBRANE-FUSION ACTIVITIES OF FUSION PEPTIDE MUTANTS OF INFLUENZA-VIRUS HEMAGGLUTININ.* Journal of Virology, 1995. **69**(11): p. 6643-6651.

7. Wasniewski, C.M., et al., *Solid-state nuclear magnetic resonance studies of HIV and influenza fusion peptide orientations in membrane bilayers using stacked glass plate samples.* Chemistry and Physics of Lipids, 2004. **132**(1): p. 89-100.

8. Sun, Y. and D.P. Weliky, *C-13-C-13 Correlation Spectroscopy of Membrane-Associated Influenza Virus Fusion Peptide Strongly Supports a Helix-Turn-Helix Motif and Two Turn Conformations.* Journal of the American Chemical Society, 2009. **131**(37): p. 13228-13229.

9. Korte, T., et al., *Role of the Glu residues of the influenza hemagglutinin fusion peptide in the pH dependence of fusion activity.* Virology, 2001. **289**(2): p. 353-361.

10. Han, X. and L.K. Tamm, *A host-guest system to study structure-function relationships of membrane fusion peptides.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(24): p. 13097-13102.

11. Wharton, S.A., et al., *MEMBRANE-FUSION BY PEPTIDE ANALOGS OF INFLUENZA-VIRUS HEMAGGLUTININ.* Journal of General Virology, 1988. **69**: p. 1847-1857.

12. Chan, W.C. and P.D. White, *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*. 2000: Oxford University Press.

13. Worman, H.J., et al., *RELATIONSHIP BETWEEN LIPID FLUIDITY AND WATER PERMEABILITY OF BOVINE TRACHEAL EPITHELIAL-CELL APICAL MEMBRANES.* Biochemistry, 1986. **25**(7): p. 1549-1555.

14. Long, H.W. and R. Tycko, *Biopolymer conformational distributions from solid-state NMR: alpha-helix and 3(10)-helix contents of a helical peptide.* Journal of the American Chemical Society, 1998. **120**(28): p. 7039-7048.

15. Zheng, Z., et al., *Conformational flexibility and strand arrangements of the membrane-associated HIV fusion peptide trimer probed by solid-state NMR spectroscopy.* Biochemistry, 2006. **45**(43): p. 12960-12975.

16. Zhang, H.Y., S. Neal, and D.S. Wishart, *RefDB: A database of uniformly referenced protein chemical shifts.* Journal of Biomolecular Nmr, 2003. **25**(3): p. 173-195.

17. Ghosh, U., L. Xie, and D.P. Weliky, *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.* Journal of Biomolecular NMR, 2013. **55**(2): p. 139-146.

18. Bodner, M.L., et al., *Temperature dependence and resonance assignment of C-13 NMR spectra of selectively and uniformly labeled fusion peptides associated with membranes.* Magnetic Resonance in Chemistry, 2004. **42**(2): p. 187-194.

19. Macosko, J.C., C.H. Kim, and Y.K. Shin, *The membrane topology of the fusion peptide region of influenza hemagglutinin determined by spin-labeling EPR.* Journal of Molecular Biology, 1997. **267**(5): p. 1139-1148.

Chapter 3 - Development of ^{13}C-^{2}H REDOR

3.1 Probe design

There are different types of REDOR experiments such as ^{13}C-^{15}N, ^{13}C-^{2}H, and ^{31}P-^{19}F REDOR. As noted earlier, REDOR is a very useful SSNMR technique for measuring distances between two heteronuclei and was first developed by Gullion and Schaefer[1]. In the past, our group has been using ^{13}C-^{15}N, ^{13}C-^{31}P, and ^{13}C-^{19}F REDOR to study the structure, membrane topology, and insertion depth of viral fusion peptides in the membrane[2-4]. To the date I joined the group, there was no probe for ^{13}C-^{2}H REDOR. The goal of developing ^{13}C-^{2}H REDOR in our laboratory is to study the membrane locations of peptides using ^{13}CO-labeled peptide and ^{2}H-labeled lipid or cholesterol. One significant advantage of the ^{2}H labeling scheme is that there will be no perturbation of the membrane bilayer no matter what fraction of lipids and/or cholesterol in the membrane are deuterated since ^{1}H and ^{2}H are chemically equivalent.

Low-power and high-power tuning

As we build a new multi-channel probe, the first step is to get all the desired channels tunable, e.g., for a 400 MHz NMR spectrometer, the ^{13}C channel should be tunable at ~100 MHz. There are two types of tuning with one called low-power tuning and the other called high-power tuning. Low-power tuning is used to check the tunable frequency range of a specific channel without applying r.f. pulses to the probe whereas high-power tuning refers to the tuning when applying r.f. pulses to the probe. The cable

connections for low-power tuning are displayed in Figure 3.1. Besides, the "SEC/DIV" knob on the oscilloscope needs to be adjusted to change the time scale to "CH1 X" such that the resonance peaks at different frequencies can be displayed on the oscilloscope.

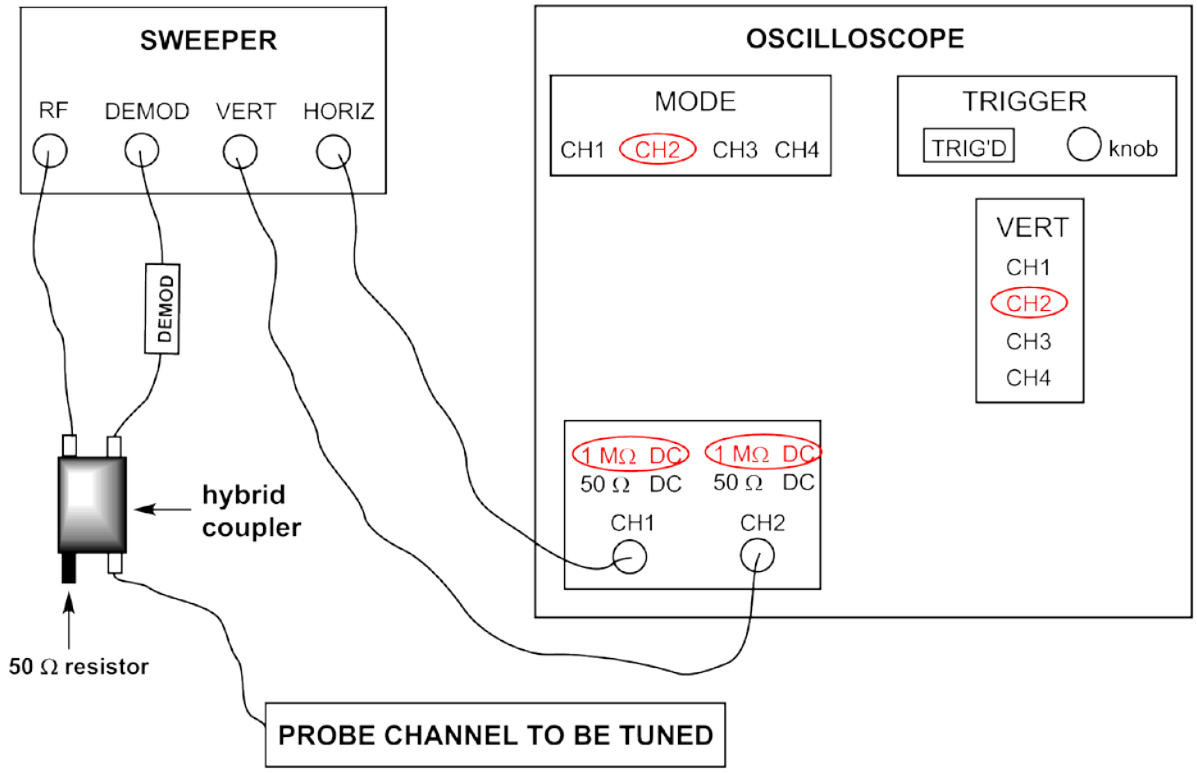

Figure 3.1 Cable connections for low-power tuning. Only part of the sweeper and oscilloscope are displayed. The probe is connected to the sweeper via a hybrid coupler (also known as reflection bridge or magic T). For the low-power tuning, the oscilloscope needs to be adjusted such that all the red circled lights are on.

Tuning configuration

For each probe, there are specific tuning configurations for different NMR experiments, e.g. for the probe (probe # 7058 from Chemagnetics) we used to develop ^{13}C-^{2}H REDOR, the triple resonance tuning configuration for ^{1}H, ^{13}C, and ^{2}H provided by the manufacturer is as follows:

33 pf series plug-in

15 pf, 6 t trap plug-in

SC Low channel receiver platform (Y channel)

36 pf Mid channel receiver platform (X channel)

2.2" Low tune tube (Y channel)

5.4" Mid tune tube (X channel)

where pf represents picofarad and is a unit of capacitance, 6 t refers to 6 turns in the trap plug-in, and SC represents short circuit. For a tune tube, it consists of two parts with part (i) made of copper and part (ii) made of both copper and dielectric plastic. In the above tuning configuration, 2.2" and 5.4" refer to the length of the copper region in part (ii) of the tune tube. Figure 3.2 displays the pictures of probe, tune tube, tuning rod, series plug-in, and trap plug-in.

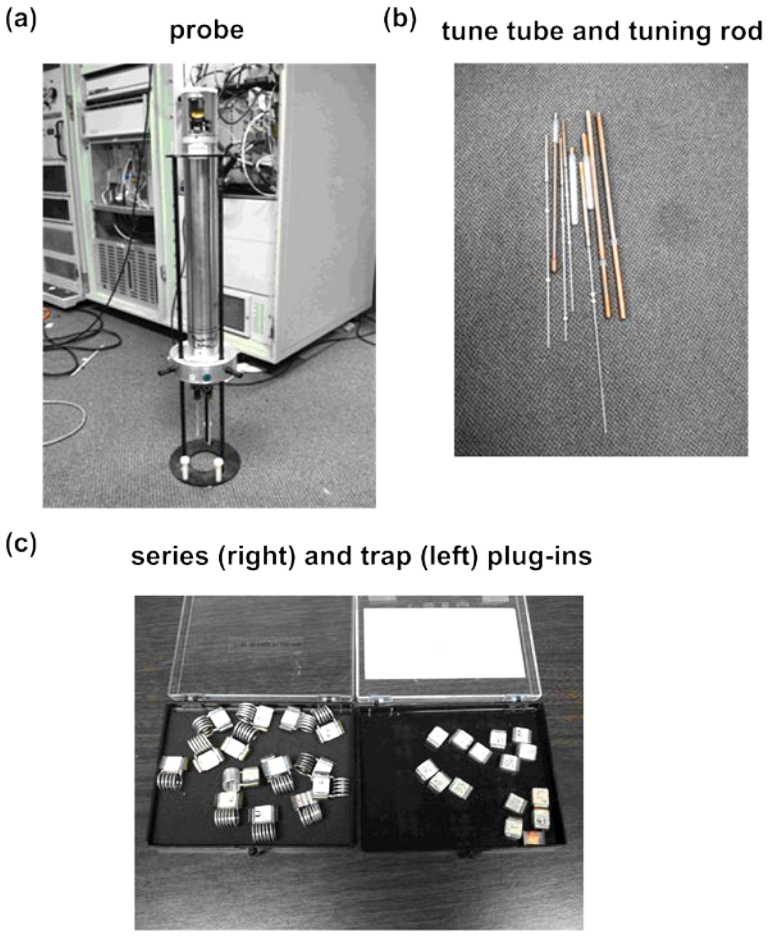

Figure 3.2 (a) Triple-resonance solid state NMR probe. (b) Tune tube and tuning rod. (c) Series plug-in and trap plug-in.

Since there are usually several capacitors and inductors in a probe, each channel in the probe can be considered as an LC circuit. As a r.f. pulse passes through an LC circuit, the capacitive reactance X_C, inductive reactance X_L, and total impedance Z the pulse experiences are defined by

$$X_C = -(\frac{1}{2\pi\nu C})i \tag{3.1}$$

$$X_L = (2\pi\nu L)i \tag{3.2}$$

$$Z = X_C + X_L = (2\pi\nu L - \frac{1}{2\pi\nu C})i \qquad (3.3)$$

where ν is the frequency of the pulse, *C* is the capacitance of the capacitor, *L* is the inductance of the inductor, and *i* refers to the square root of -1. According to equations 3.1-3.3, we can mathematically show that the total impedance *Z* = 0 when $\nu = \frac{1}{2\pi\sqrt{LC}}$, where $\frac{1}{2\pi\sqrt{LC}}$ is the so-called resonance frequency of an LC circuit. To minimize the total impedance *Z* that the pulse experiences, specific components such as plug-ins and tune tube need to be selected such that the corresponding capacitance *C* and inductance *L* will fulfill the condition of $\nu = \frac{1}{2\pi\sqrt{LC}}$, where ν is the transmitter frequency of the pulse. Although the tuning configuration for each specific NMR experiment is provided by the probe manufacturer, the final configuration that works may be different than the one provided. For instance, the tuning configuration for ^{1}H, ^{13}C, and ^{2}H that really worked as we developed the ^{13}C-^{2}H REDOR probe was as follows:

82 pf series plug-in

15 pf, 6 t trap plug-in

33 pf Low channel receiver platform (Y channel)

36 pf Mid channel receiver platform (X channel)

3.9" Low tune tube (Y channel)

5.4" Mid tune tube (X channel)

which was substantially different than the one provided by the probe manufacturer. Note that the tuning configuration may change again as you modify the probe, e.g. replacement of the copper ribbon connecting the stator and ^{13}C channel (X channel) may cause detuning of the channel and thus a new series plug-in may be needed to get the channel tunable again.

During the low-power tuning, we adjust the tuning rod and match to make the inductance *L* and capacitance *C* fulfill the resonance condition of $\nu = \frac{1}{2\pi\sqrt{LC}}$ to minimize the reflected voltage V_r from the probe at a specific frequency ν and detect V_r on the oscilloscope. If a channel can be tuned at a specific frequency (e.g. the transmitter frequency to be set on the spectrometer) to produce a deep and sharp peak near the zero voltage line on the oscilloscope at the low-power tuning, it means the channel can be also well tuned for that specific frequency at the high-power tuning, e.g. the forward-to-reverse voltage ratio of the pulse after the high-power tuning can be 10 or higher. Note that the forward voltage of the pulse can be accurately measured only after the channel has been well tuned.

3.2 Setup peptide I4

Criteria for REDOR setup compound

To optimize the REDOR pulse program, a proper setup compound is required. There are several criteria for an ideal REDOR setup compound described as follows:

(*a*) The setup compound should have the same type of observed nucleus as the samples to be studied. Since the peptide samples to be studied by ^{13}C-^{2}H REDOR are

^{13}CO labeled, the setup compound should be also ^{13}CO labeled. If the setup compound is $^{13}CH_3$ labeled, the optimized pulse parameters for the setup compound are not necessarily optimized for the ^{13}CO labeled samples since $^{13}CH_3$ and ^{13}CO have different ^{13}C chemical shift anisotropies (CSA). The amide ^{13}CO has a CSA of ~150 ppm whereas $^{13}CH_3$ has a CSA of ~30 ppm. As a result, an optimized ^{13}C r.f. pulse power for $^{13}CH_3$ is not necessarily sufficient for ^{13}CO.

(*b*) The setup compound should have a single spin pair with a known internuclear distance, in which case the data analysis as we optimize the pulses will be simplified. We can first calculate the dipolar coupling using the internuclear distance and then run SIMPSON simulations to get the theoretical dephasing buildup plot for that dipolar coupling. At last, we compare the theoretical dephasing buildup with the experimental one to tell whether the pulses (particularly the π pulses during the dephasing time) have been optimized or not.

(*c*) The dipolar interaction for the setup compound should be predominantly intra- rather than intermolecular. In the latter case, it could be challenging to calculate the intermolecular distance and there could be a distribution of intermolecular distances, which complicate the data analysis as we optimize the pulse program.

(*d*) It is advisable to choose a setup compound which has an intermediate dipolar coupling, e.g., 30 or 40 Hz. If we choose one with a very small dipolar coupling such as 6 Hz, the SIMPSON-simulated ^{13}C-^{2}H REDOR dephasing is only 10% for dephasing time τ = 40 ms. By contrast, if we choose one with a very large dipolar coupling such as

200 Hz, the SIMPSON-simulated ^{13}C-^{2}H REDOR dephasing is 65% for dephasing time τ = 16 ms, which is already reaching the maximum dephasing of 2/3. In this case, the dephasing buildup may not be sensitive to the pulse optimization, i.e., even if the pulses (particularly the π pulses during the dephasing time) are not optimized, there might be a significant dephasing buildup.

I4 peptide sequence and isotopic labels

Based on the above criteria, we synthesized an I4 peptide as the setup compound for ^{13}C-^{2}H REDOR. The I4 peptide was synthesized by Fmoc SPPS using rink amide resin and thus the C-terminus of the peptide was amidated. I4 has a sequence of AEAAAKEAAAKEAAAKAW with a ^{13}CO label at residue A9 and a C_{α}-^{2}H label at residue A8. Earlier SSNMR studies of the lyophilized I4 peptide have shown that A9 has an α-helicity of (83 ± 6)%[5]. In the α-helical structure, the internuclear distance between A8 C_{α}-^{2}H and A9 ^{13}CO is ~5.0 Å which corresponds to a ^{13}C-^{2}H dipolar coupling of ~37 Hz. The ^{13}C-^{2}H REDOR pulse program optimization using the setup peptide I4 is described in detail in Section 3.3. For clarification purposes, the setup peptide I4 is denoted as $I4\_A8_{D}A9_{C}$ in the following sections and chapters.

3.3 Pulse sequence optimization

In general, there are three commonly used REDOR pulse sequences (Figure 3.3)[6]. For each REDOR pulse sequence in Figure 3.3, it starts with CP from ^{1}H to nucleus S to enhance the S signal, then a dephasing period during which the dipolar

coupling is decoupled in the S_0 experiment and recoupled in the S_1 experiment, and at last an acquisition period. Windowless ^{1}H decoupling pulses (i.e. there is no interpulse time delay) are applied throughout the dephasing and acquisition periods.

For the pulse sequence in Figure 3.3a, a single S π pulse is applied in the middle of the dephasing period to refocus the isotropic chemical shift and produce an echo at

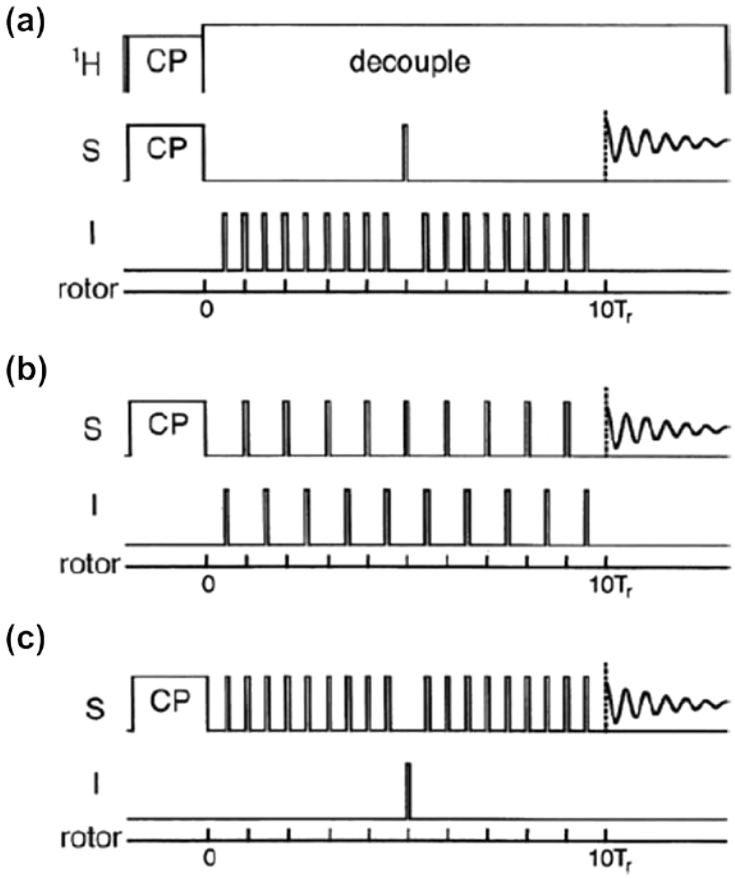

Figure 3.3 Three commonly used REDOR pulse sequences. Each sequence starts with CP from ^{1}H to the observed nucleus S to enhance the S signal, followed by a dephasing period (only ten rotor cycles are shown here) and at last an acquisition period. ^{1}H decoupling pulses are applied throughout the dephasing and acquisition periods[6].

the beginning of the acquisition period. For the S_o experiment, there is no π pulse applied on the I channel and thus the dipolar coupling is decoupled by MAS. For the S_1 experiment, in addition to the single π pulse applied on the S channel in the middle of the dephasing period, there are also π pulses applied on the I channel in the middle and at the end of each rotor cycle during the dephasing period except for the end of the rotor cycle when the S-channel π pulse is applied. Application of the π pulse train on the I channel recouples the dipolar coupling and thus causes dipolar dephasing of the signal during the S_1 experiment.

For the pulse sequence in Figure 3.3b, π pulses are applied on the S channel at the end of each rotor cycle during the dephasing period. The I-channel π pulses in the middle of each rotor cycle are applied during the S_1 experiment but not during the S_o experiment.

For the pulse sequence in Figure 3.3c, π pulses are applied on the S channel in the middle and at the end of each rotor cycle except for the midpoint of the dephasing period. The I-channel π pulse in the midpoint of the dephasing period is applied during the S_1 experiment but not during the S_o experiment. This pulse sequence can be used when the I spin has a very large anisotropic interaction (e.g. quadrupolar interaction). The I-channel π pulse can be a composite π pulse which may better compensate for the limited r.f. pulse power and large nuclear resonance offsets than a simple π pulse do, where a composite π pulse consists of multiple simple pulses with different phases and the net flip angle of the composite pulse is 180^o. This pulse sequence has been used

for ^{13}C-^{2}H REDOR and the advantage of applying a single rather than multiple ^{2}H π pulses is that it can minimize the effect of ^{2}H pulse imperfections[7].

All the three REDOR pulse sequences in Figure 3.3 have the same optimization procedure described as follows (take ^{13}C-^{2}H REDOR for example):

(a) MAS setup using KBr

pulse program = "1 pulse" (one pulse with a phase cycling of x, -x, y, -y), Br transmitter frequency = 99.8942000 MHz, spinning rate = 4 kHz, temperature = -50 ^{o}C.

We need to first tune the Br channel and then adjust the magic angle rod. As we adjust the magic angle rod, we can either click "Repeat Scan" in the Spinsight window and then keep adjusting the magic angle rod until we observe as many and strong rotary echoes as possible in the Br FID or click "RS and Process" and then keep adjusting the magic angle rod until we observe as many and strong peaks as possible in the Br spectrum (Figure 3.4).

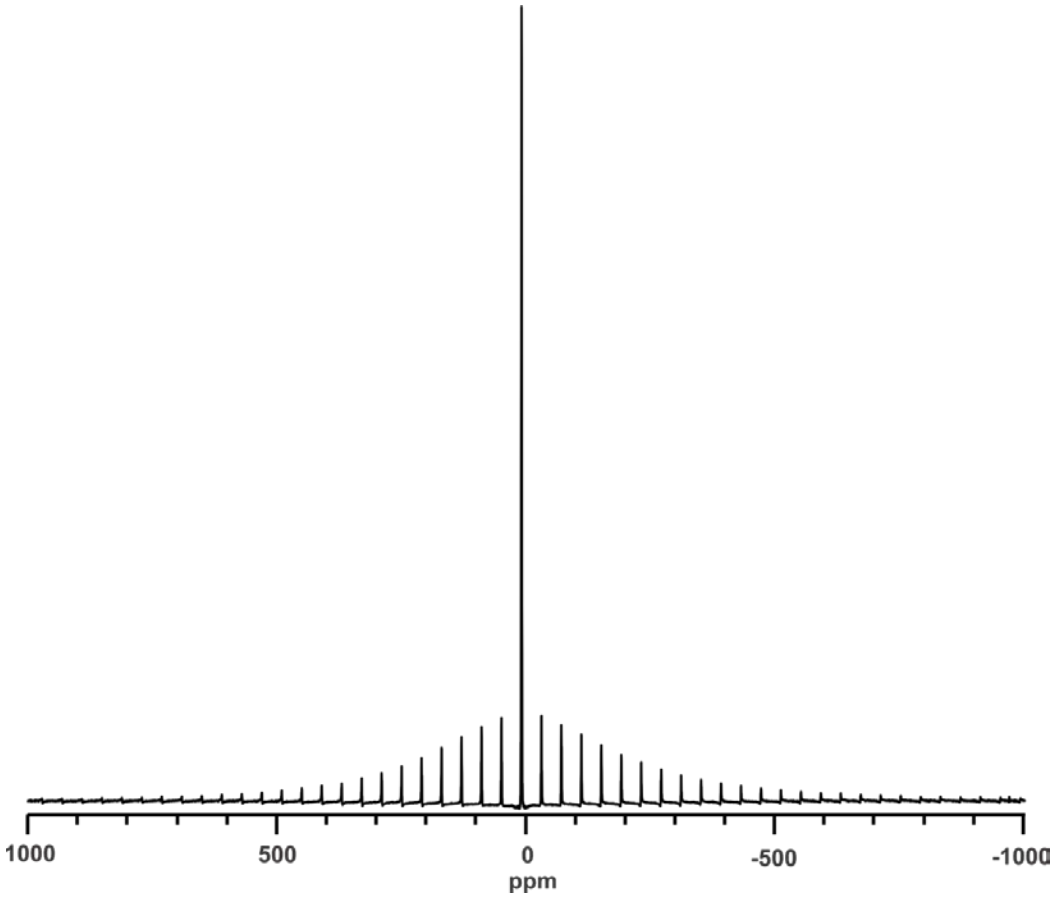

Figure 3.4 Br spectrum (64 scans) after MAS setup, spinning rate = 4 kHz, -50 ^{o}C. The spectrum was processed with 20 Hz Gaussian line broadening.

(b) ^{13}C chemical shift referencing using adamantane

pulse program = "1 pda" (one pulse with a phase cycling of x, -x, y, -y and ^{1}H decoupling), ^{13}C transmitter frequency = 100.2677200 MHz, ^{1}H transmitter frequency = 398.6976190 MHz, spinning rate = 4 kHz (other spinning rates such as 8 and 10 kHz also work and they produce very similar ^{13}C chemical shifts within 0.1 ppm deviation), temperature = -50 ^{o}C. The ^{13}C transmitter frequency is chosen in such a way that the ^{13}C peaks of interest (e.g. ^{13}CO) of the samples to be studied (not adamantane) are

close to 0 ppm (e.g. 15 ppm) before doing the chemical shift referencing, in which case the ^{13}C resonance offsets are small and thus the ^{13}C pulses can work efficiently. The ^{13}C spectrum of adamantane under 4 kHz MAS is shown in Figure 3.5. The peak at -119.0 ppm refers to the methylene ^{13}C peak and the peak at -127.9 ppm refers to the methine ^{13}C peak. For all the membrane samples of peptides to be studied, the ^{13}C shifts will be externally referenced to the methylene peak of adamantane at 40.5 ppm such that they can be directly compared with the ^{13}C shifts of proteins in aqueous solution[8, 9]. Since the methylene ^{13}C shift of adamantane in Figue 3.5 is -119.0 ppm, the corresponding ^{13}C chemical shift referencing value = 40.5 ppm - (-119.0 ppm) = 159.5 ppm. In this case, if the peptide ^{13}CO peak shift is 20.0 ppm before referencing, the peak shift after referencing will be 179.5 ppm.

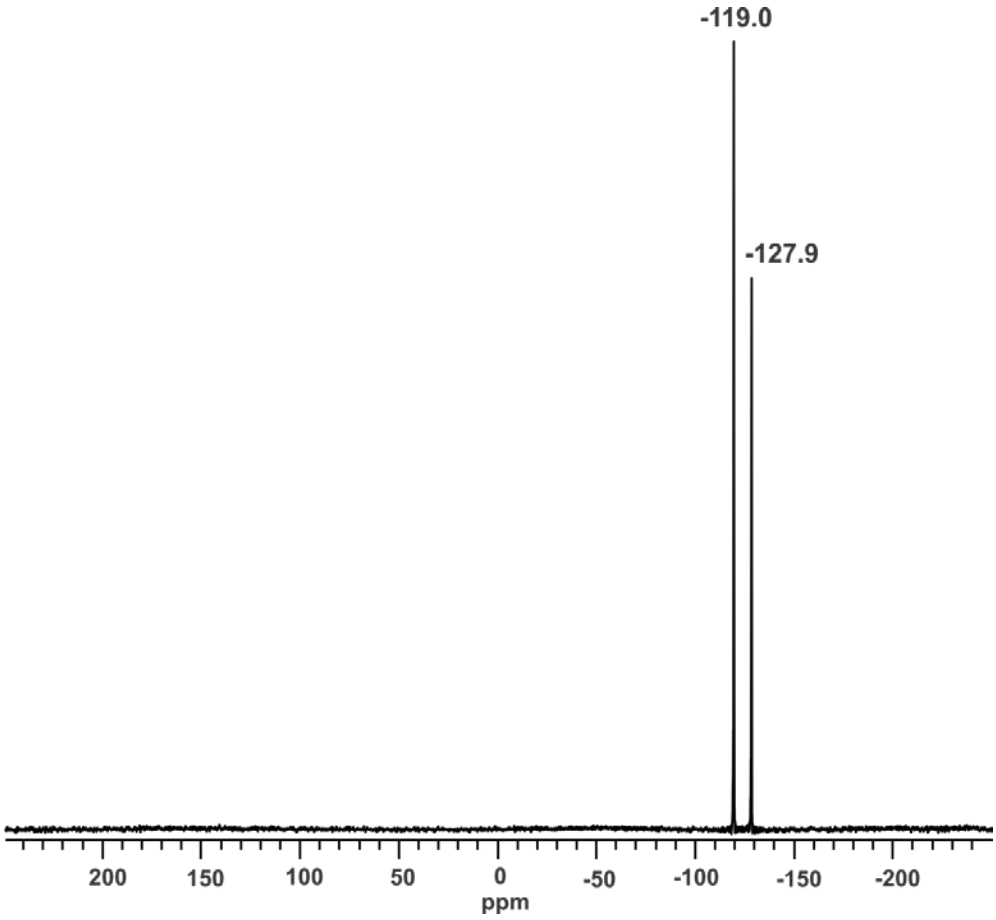

Figure 3.5 ^{13}C spectrum of adamantane under 4 kHz MAS, -50 ^{o}C. The left peak at -119.0 ppm corresponds to the methylene ^{13}C whereas the right peak at -127.9 ppm corresponds to the methine ^{13}C.

To correctly view the ^{13}C shifts before referencing, we need to make sure 0 ppm is right in the middle of the chemical shift axis in the Spinsight window, if not, click "Analysis"→"Reference", then click the spectrum to pop up the "Set Reference" window in which you set the "Reference point" as 0.5 and "Reference value" as 0, and at last click "Set"→ "OK" to finish. To correctly view the ^{13}C shifts after referencing, click "Analysis"→"Reference" in the Spinsight window, then click the spectrum to pop up the "Set Reference" window where you set the "Reference point" as 0.5 and "Reference

value" as the referencing value of adamantane (e.g. 159.5), and finally click "Set"→ "OK" to finish.

(c) CP optimization

sample = I4_A8$_D$A9$_C$, pulse program = "cp_ramp", ^{13}C transmitter frequency = 100.2677200 MHz, ^{1}H transmitter frequency = 398.6976190 MHz, spinning rate = 10 kHz, temperature = -50 ^{o}C.

step 1: ^{1}H π/2 pulse optimization

^{1}H π/2 pulse is the first pulse in the REDOR sequence and it rotates the ^{1}H magnetization from z axis to the xy plane. We first set the ^{1}H π/2 pulse width (pw90H) as 10 μs and then array the pulse power aH from 0.50 to 0.70 in an increment of 0.02. Comparison of the ^{13}C spectra (Figure 3.6) indicates that pw90H of 10 μs and aH of 0.58 correspond to a ^{1}H π pulse. The ^{1}H π pulse produces a zero transverse magnetization of ^{1}H and thus there is no magnetization transfer from ^{1}H to ^{13}C during CP, which results in a zero ^{13}CO signal in the spectrum. To set the ^{1}H π/2 pulse, we fix aH at 0.58 and change pw90H from 10 μs to 5 μs. The Rabi frequency of the ^{1}H π/2 pulse = $\frac{\gamma B_1}{2\pi} = \frac{1}{4x(pw90H)}$ = 50 kHz (the pulse flip angle = γB_1 x (pw90H) = π/2, refer to Chapter 1 for more details). An alternative way to set the ^{1}H π/2 pulse is to fix the pulse power aH and array the pulse width pw90H.

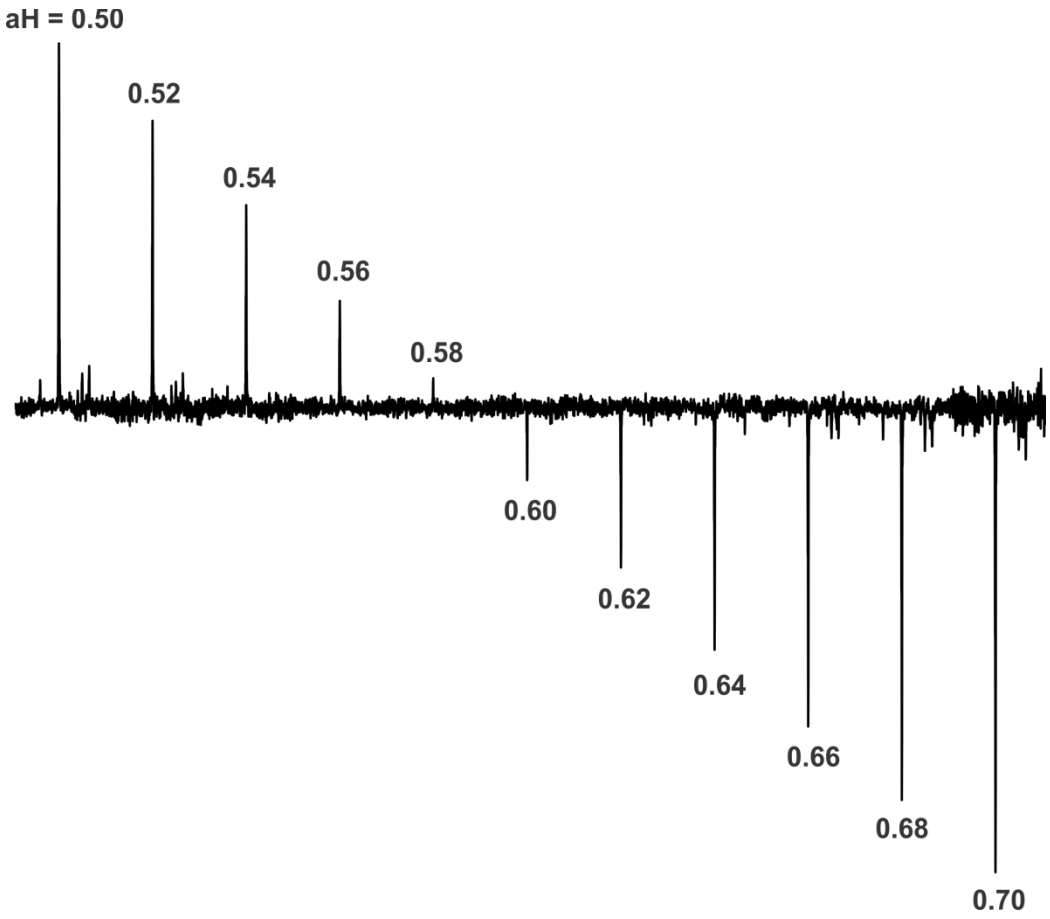

Figure 3.6 ^{13}C spectra of I4_A8$_D$A9$_C$ for "aH" array from 0.50 to 0.70 in an increment of 0.02, 10 kHz MAS, -50 ^{o}C.

step 2: ^{1}H and ^{13}C CP pulse optimization

After optimizing the ^{1}H $\pi/2$ pulse, we then set the ^{1}H CP pulse power (aHcp) the same as aH. Although aHcp can be different than aH, we usually set them the same such that we can easily measure the forward voltage of the ^{1}H $\pi/2$ pulse. If we set them differently, it will be difficult to distinguish the ^{1}H $\pi/2$ pulse from the CP pulse due to the short pulse width (5 μs) compared to the much longer CP contact time (1~2 ms) and thus difficult to measure the forward voltage of the $\pi/2$ pulse. The forward voltage of the

^{1}H $\pi/2$ pulse is important since we need it to set the ^{1}H decoupling power, which will be discussed later in more details.

As we set aHcp = aH = 0.58, the next step is to array the ^{13}C CP pulse power (aXcp) to find out what power produces the maximum ^{13}CO signal intensity. After that, aXcpmod is arrayed to find out the best ^{13}C CP ramp which can maximize the magnetization transfer from ^{1}H to ^{13}C (refer to Chapter 1 for more details about CP ramp). The optimized aXcp and aXcpmod are 0.40 and 0.03, respectively.

step 3: CP contact time optimization

The last step of CP optimization is to array the contact time to maximize the ^{13}CO signal intensity. If the CP contact time is too short, the magnetization transfer from ^{1}H to ^{13}C will be incomplete which results in weaker ^{13}CO signal intensities. If the CP contact time is too long, the ^{13}CO signal intensities will also decrease due to the ^{1}H $T_{1\rho}$ relaxation which refers to the ^{1}H magnetization decay from M_o to $(B_1/B_o)M_o$, where M_o is the ^{1}H magnetization magnitude established under the external magnetic field $\mathbf{B_o}$ and $(B_1/B_o)M_o$ is the ^{1}H magnetization magnitude at the thermal equilibrium under the ^{1}H CP pulse field $\mathbf{B_1}$. Comparison of ^{13}CO signal intensities indicates that the optimized CP contact time is 1.5 ms.

In the ^{13}C-^{2}H REDOR pulse sequence, ^{1}H decoupling pulses are applied right after CP to decouple all ^{1}H-related couplings such as ^{13}C-^{1}H J-coupling, ^{13}C-^{1}H dipolar coupling, and ^{1}H-^{1}H dipolar coupling, etc. We can set the ^{1}H decoupling power (aHdec) based on the forward voltage of the ^{1}H $\pi/2$ pulse. For example, if the ^{1}H decoupling power is set as 75 kHz, we need to adjust the parameter aHdec until the forward voltage of the ^{1}H decoupling pulses is 1.5 times that of the ^{1}H $\pi/2$ pulse which has a Rabi frequency of 50 kHz. For the forward voltage of a pulse, recall that it can be correctly measured only after the channel has been well tuned. Figure 3.7 shows the ^{13}C spectrum of I4_A8$_D$A9$_C$ after optimizing CP and setting ^{1}H decoupling power as 75 kHz. The A9 ^{13}CO shift of 178.8 ppm is consistent with the α-helical structure of I4[9].

(d) ^{13}C π pulse optimization

sample = I4_A8$_D$A9$_C$, pulse program = "cp_zfilter", ^{13}C transmitter frequency = 100.2677200 MHz, ^{1}H transmitter frequency = 398.6976190 MHz, spinning rate = 10 kHz, temperature = -50 ^{o}C. The pulse sequence of cp_zfilter is shown in Figure 3.8. When both the ^{13}C $\pi/2$ and π pulses are set correctly, a zero ^{13}C signal will be observed. However, when the ^{13}C π pulse is set as a $\pi/2$ pulse, a maximum ^{13}C signal

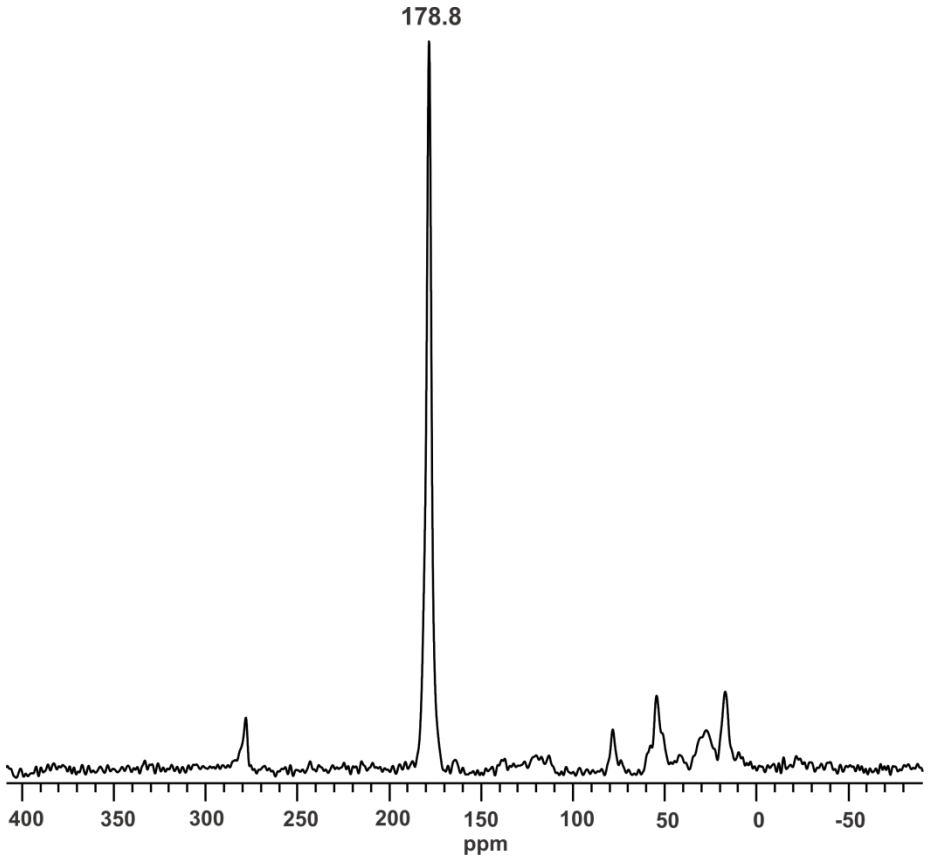

Figure 3.7 ^{13}C spectrum of I4_A8$_D$A9$_C$ after CP optimization under 10 kHz MAS, -50 ^{o}C. The A9 ^{13}CO chemical shift after referencing is 178.8 ppm.

will be observed. To optimize the ^{13}C pulses, we first set the ^{13}C $\pi/2$ pulse width (pw90X) and π pulse width (pw180X) both as 4.0 μs and then array the power (aX) from 0.30 to 0.45 in an increment of 0.01. The maximum ^{13}CO signal intensity is observed for aX = 0.39, i.e. the pulse width 4.0 μs and power 0.39 correspond to a ^{13}C $\pi/2$ pulse. The next step is to fix pw90X = 4.0 μs and aX = 0.39 and then arry pw180X. A zero ^{13}CO signal intensity is observed for pw180X = 8.3 μs. Therefore, the pulse width 8.3 μs and power 0.39 correspond to a ^{13}C π pulse.

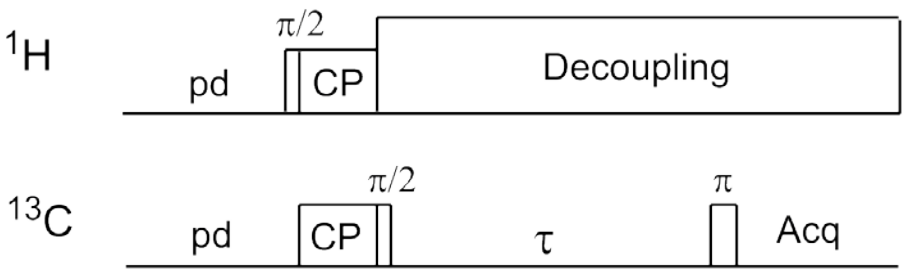

Figure 3.8 Pulse sequence of cp_zfilter. The phase cycle is "x, -x" for ^{1}H $\pi/2$ pulse, "-y, -y, -x, -x" for ^{13}C CP pulse, "x, x, -y, -y" for ^{13}C $\pi/2$ pulse, and "-x, -x, y, y" for ^{13}C π pulse, respectively. The ^{1}H CP and decoupling pulses have a fixed phase of "y".

(e) ^{2}H π pulse optimization

sample = I4_A8$_D$A9$_C$, pulse program = "redorxy8xypi_pm" (refer to Figure 3.3b), ^{1}H transmitter frequency = 398.6976190 MHz, ^{13}C transmitter frequency = 100.2677200 MHz, ^{2}H transmitter frequency = 61.2030000 MHz, dephasing time τ = 32 ms (other values such as 24 ms also work), spinning rate = 10 kHz, temperature = -50 ^{o}C. The ^{2}H transmitter frequency is set as 61.2030000 MHz such that the lipid ^{2}H quadrupolar spectra are centered at ~0 ppm. To optimize the ^{2}H π pulses, we first set the pulse power (aY180) as 0.90 and then array the pulse width (pw180Y) to find out which value produces the maximum dephasing. For τ = 32 ms, the maximum dephasing, 0.464 (6), was observed when pw180Y = 4.7 μs. Therefore, pw180Y and aY180 are set as 4.7 μs and 0.90, respectively. To check whether pw180Y = 4.7 μs and aY180 = 0.90 correspond to a ^{2}H π pulse or not, ^{2}H "1pulse" experiment of D_2O at room temperature (18.9 ^{o}C) without spinning is carried out to calibrate the ^{2}H pulses. For direct

comparison, the ^{2}H amplifier for "1pulse" experiment of D_2O is the same as that for ^{13}C-^{2}H REDOR. As we set pulse power as 0.90, the maximum ^{2}H signal of D_2O is observed for pulse width of 2.5 μs, i.e. the pulse power of 0.90 and pulse width of 2.5 μs correspond to a ^{2}H $\pi/2$ pulse. Therefore, pw180Y = 4.7 μs and aY180 = 0.90 do not exactly correspond to a ^{2}H 180^{o} pulse but a 169^{o} pulse. This deviation could be due to the temperature change since the ^{2}H pulse is calibrated using D_2O at 18.9 ^{o}C whereas the ^{13}C-^{2}H REDOR pulses are optimized at -50 ^{o}C.

At last, we set pw180Y = 4.7 μs and aY180 = 0.90 and then run ^{13}C-^{2}H REDOR experiments of I4_A8$_D$A9$_C$ for dephasing time τ = 2, 8, 16, 24, 32, 40, 48, 56, 64, 72, and 80 ms, respectively. The experimental dephasing buildup of $\Delta S/S_o$ vs τ is shown in Figure 3.9.

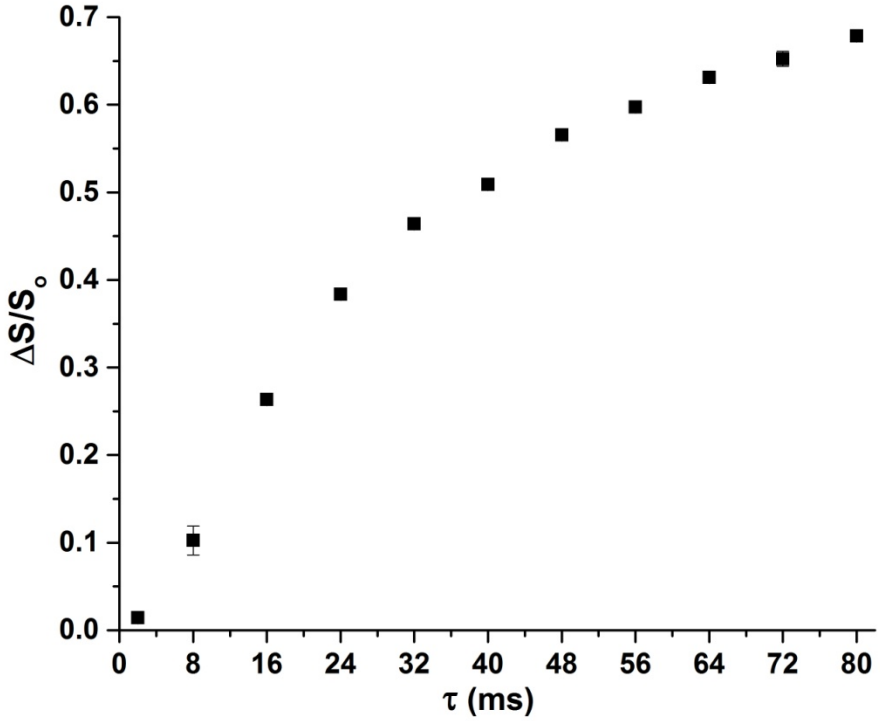

Figure 3.9 ^{13}C-^{2}H REDOR experimental dephasing plot of $\Delta S/S_0$ vs τ of I4_A8$_D$A9$_C$ under 10 kHz MAS and -50 ^{o}C.

Comparison of REDOR pulse sequences

As discussed earlier, there are three commonly used REDOR pulse sequences (Figure 3.3). For the ^{13}C-^{2}H REDOR experiments of I4_A8$_D$A9$_C$ described above, the pulse sequence is "redorxy8xypi_pm" (Figure 3.3b), where "xy8" refers to the X/Y (i.e. $^{13}C/^{2}H$) π pulse phase cycle of "x, y, x, y, y, x, y, x" and "pm" refers to TPPM ^{1}H decoupling. The π pulse phase cycling is applied to compensate for pulse imperfections and produce REDOR dephasings which are not sensitive to the resonance offsets[10]. In addition to "redorxy8xypi_pm", we also tested the "redorxy8ypi_pm" (Figure 3.3a) and

"redorxy4xpi_pm" (Figure 3.3c) pulse sequences and compared them to "redorxy8xypi_pm" in terms of signal sensitivity and REDOR dephasing buildup. Figure 3.10 shows the REDOR dephasing buildups of I4_A8$_D$A9$_C$ for "redorxy8xypi_pm" (black) and "redorxy8ypi_pm" (red). Compared to "redorxy8xypi_pm", "redorxy8ypi_pm" produces a slower dephasing buildup. The reason for that could be due to the ^{2}H pulse imperfections including limited pulse power relative to the large resonance offsets of ^{2}H. For a given dephasing time τ, the number of ^{2}H π pulses for "redorxy8ypi_pm" is twice that for "redorxy8xypi_pm". Therefore, the accumulated ^{2}H π pulse imperfections are more significant for "redorxy8ypi_pm" which cause a less efficient inversion of ^{2}H spins and consequently a slower dephasing buildup. Figure 3.11 shows the REDOR S_0 spectra (^{13}CO region) of I4_A8$_D$A9$_C$ for "redorxy8xypi_pm" and "redorxy4xpi_pm". The ^{13}C signal sensitivity for "redorxy4xpi_pm" is much lower than that for "redorxy8xypi_pm". One possible reason for it could be that the anisotropic chemical shift interactions are not averaged out by MAS due to the application of ^{13}C π pulses in the middle and at the end of each rotor period T_r. As shown in Figure 3.12, assuming the ^{13}C magnetization is initially along the x axis for dephasing time $\tau = 0$ (i), the two individual ^{13}C magnetic dipole moments *a* and *b* will precess under anisotropic chemical shift fields from $\tau = 0$ to $\tau = T_r/2$ (ii). When a ^{13}C π_x pulse is applied at $\tau = T_r/2$,

the magnetic moments *a* and *b* will be flipped to the opposite side about the x axis (iii). In the meantime, the anisotropic chemical shift field directions will be inverted at $\tau = T_r/2$

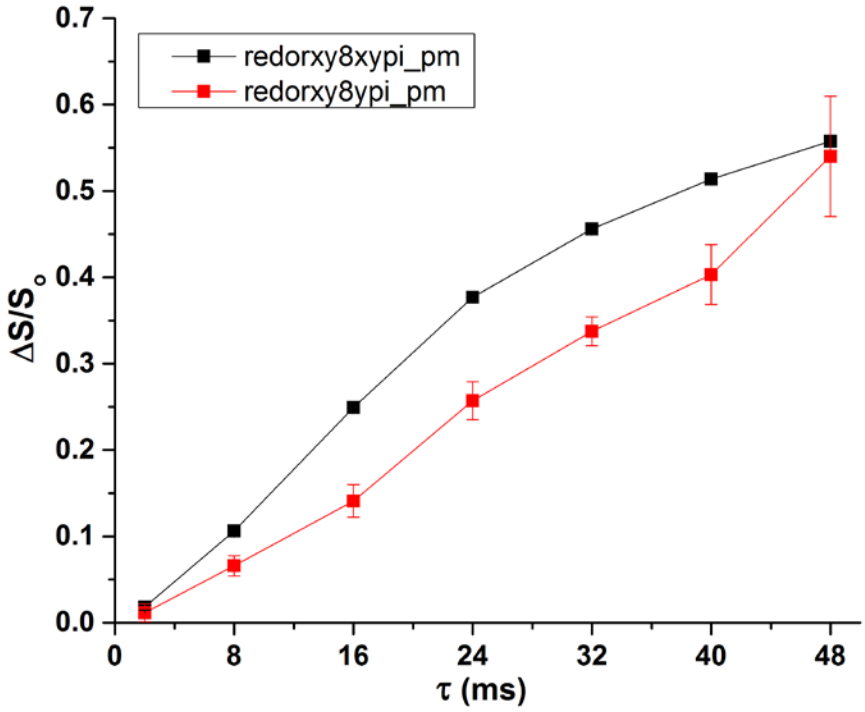

Figure 3.10 ^{13}C-^{2}H REDOR experimental dephasing plots of $\Delta S/S_0$ vs τ of $I4\_A8_DA9_C$ for the "redorxy8xypi_pm" (black) and "redorxy8ypi_pm" (red) pulse sequences under 10 kHz MAS and -50 ^{O}C.

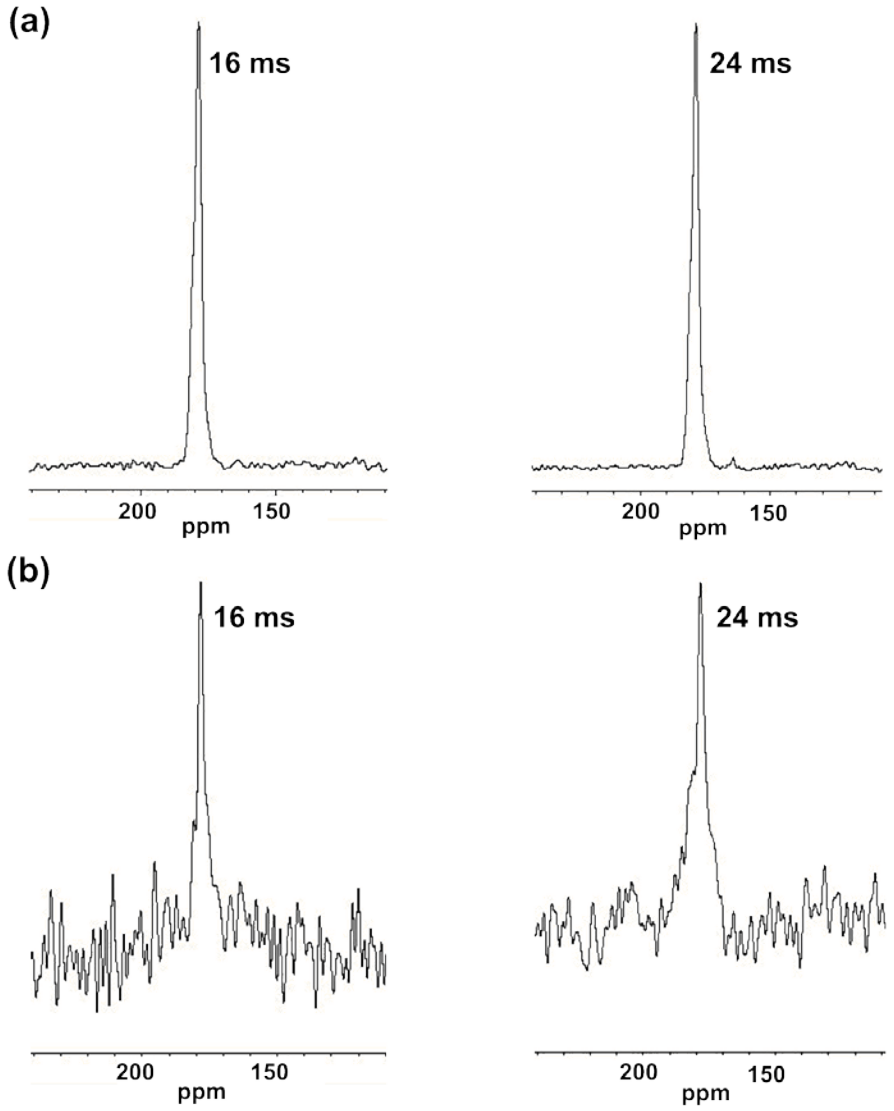

Figure 3.11 ^{13}C-^{2}H REDOR S_0 spectra of I4_A8$_D$A9$_C$ for the "redorxy8xypi_pm" (a) and "redorxy4xpi_pm" (b) pulse sequences under 10 kHz MAS and -50 ^{o}C. The number of acquisitions is 400 and 1000 for the 16 ms and 24 ms S_0 spectra in panel a and 6875 and 13331 for the 16 ms and 24 ms S_0 spectra in panel b. All spectra are processed with 20 Hz Gaussian line broadening and baseline correction.

due to MAS. Therefore, the magnetic moments *a* and *b* will precess in the opposite direction from $\tau = T_r/2$ to $\tau = T_r$ as they do from $\tau = 0$ to $\tau = T_r/2$ (iii). As a result, the magnetic moments *a* and *b* will not be refocused along the x axis at $\tau = T_r$ (iv). As

dephasing time τ goes on, the individual ^{13}C magnetic dipole moments will become randomly oriented and thus their vector sum (i.e. ^{13}C magnetization) will decrease in magnitude as a function of τ. This is probably why the ^{13}C signal sensitivity for "redorxy4xpi_pm" is so poor compared to that for "redorxy8xypi_pm".

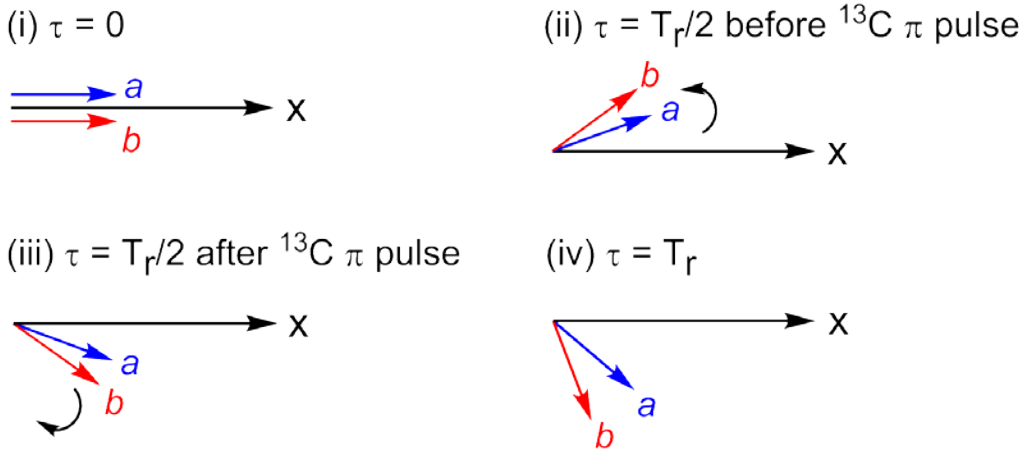

Figure 3.12 Evolution of two individual ^{13}C magnetic dipole moments *a* and *b* due to anisotropic chemical shift interactions under MAS when applying ^{13}C π pulses in the middle and at the end of each rotor period T_r.

In summary, the best pulse sequence for ^{13}C-^{2}H REDOR in terms of signal sensitivity and dephasing buildup is "redorxy8xypi_pm" and this pulse sequence is used for the peptide ^{13}CO to lipid ^{2}H REDOR experiments to be discussed in chapter 4. For the ^{2}H π pulses, although composite pulses may better compensate for the large ^{2}H quadrupolar anisotropy than simple pulses do, the ~100 kHz fields of ^{2}H simple π pulses used in our ^{13}C-^{2}H REDOR experiments are strong enough to produce efficient ^{2}H spin inversion even in the presence of non-motionally averaged ^{2}H quadrupolar

anisotropy (see Figure 3.13 and more discussion in chapter 4). For the "redorxy4xpi_pm" pulse sequence, it could be advantageous in that only a single ^{2}H π pulse is applied during the dephasing time and thus there is no accumulated pulse imperfection. However, this pulse sequence may be beneficial for alkyl ^{13}C (e.g. $^{13}CH_3$) to ^{2}H REDOR but not for ^{13}CO to ^{2}H REDOR since alkyl ^{13}C has a small CSA (e.g. ~30 ppm for $^{13}CH_3$) whereas ^{13}CO has a large CSA (~150 ppm). The anisotropic chemical shift interactions that are averaged out under MAS will be reintroduced by the ^{13}C π pulses applied in the middle and at the end of each rotor period during the dephasing time. As a result, the ^{13}CO signal sensitivity is very low due to the large CSA (Figure 3.11).

3.4 I4 data fitting

Natural abundance calibration of $(\Delta S/S_0)^{exp}$

In section 3.3, we have discussed the ^{13}C-^{2}H REDOR pulse sequence optimization using the setup peptide I4_A8$_D$A9$_C$. The experimental dephasing curve of $\Delta S/S_0$ vs τ is shown in Figure 3.9. To obtain $\Delta S/S_0$ vs τ for the labeled A9 ^{13}CO only, the natural abundance contribution to the experimental $\Delta S/S_0$ is removed in a similar way as described in chapter 2. The labeled $\Delta S/S_0$ vs τ for A9 ^{13}CO is shown in Figure 3.13 (black squares). For all τ, the difference between the experimental and labeled $\Delta S/S_0$ is ≤ 0.05.

SIMPSON simulations of $(\Delta S/S_o)^{lab}$ *vs* τ

In the α-helical structure of I4_A8$_D$A9$_C$, the internuclear distance *r* between A8 C_α-^{2}H and A9 ^{13}CO is ~5.0 Å which corresponds to a ^{13}C-^{2}H dipolar coupling *d* ≈ 37 Hz via the relationship *d*(Hz)=4642/*r*^{3}(Å)[10]. The theoretical $\Delta S/S_o$ vs τ for *d* = 37 Hz is simulated by SIMPSON for realistic MAS frequency, ^{2}H r.f. field, and ^{2}H quadrupolar coupling, etc. As shown in Figure 3.13, the simulated $\Delta S/S_o$ buildup for *d* = 37 Hz has a sigmoidal shape whereas the labeled $\Delta S/S_o$ buildup for A9 ^{13}CO has an exponential shape. In addition, the simulated $\Delta S/S_o$ has a rapid buildup with a maximum value of 2/3 whereas the labeled $\Delta S/S_o$ has a slower buildup with values greater than 2/3 at longer τ. The maximum value of 2/3 for simulated $\Delta S/S_o$ reflects 1/3 fractional population of ^{2}H spins in the *m*=0 state which cannot be inverted by ^{2}H π pulses and thus have no contribution to the dipolar dephasing of ^{13}CO.

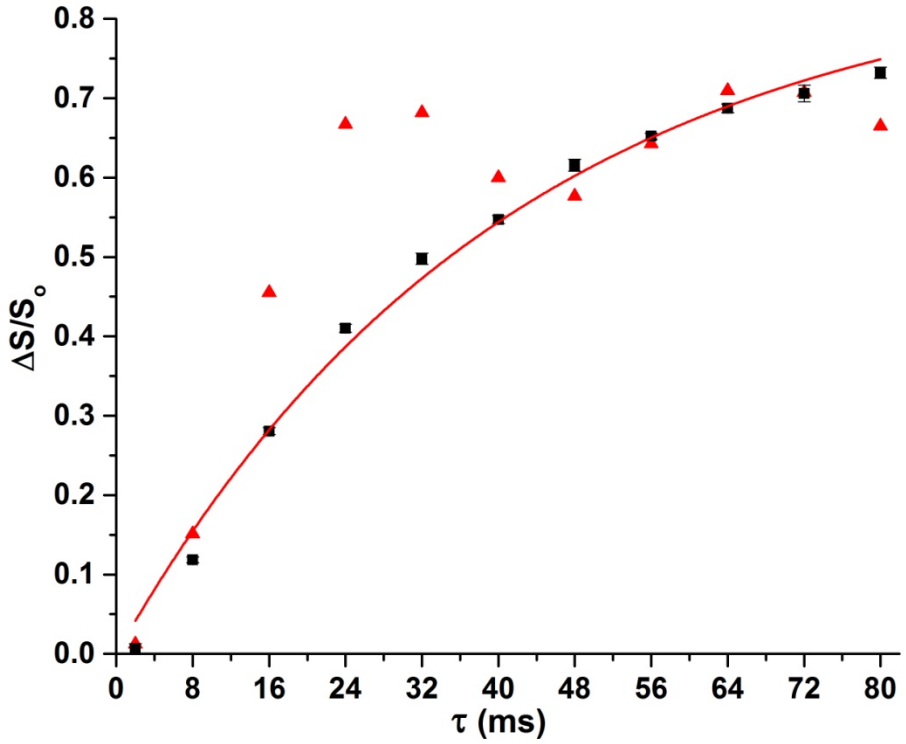

Figure 3.13 ^{13}C-^{2}H REDOR dephasing curves of $\Delta S/S_0$ vs τ of I4_A8$_D$A9$_C$ for the pulse sequence "redorxy8xypi_pm" under 10 kHz MAS and -50 ^{o}C. Black squares are labeled $\Delta S/S_0$ for A9 ^{13}CO after removal of the natural abundance contribution. Red triangles are SIMPSON simulated $\Delta S/S_0$ for d = 37 Hz. Red line is the best-fit curve using the fitting function Ax(1-e$^{-\beta\tau}$).

The differences between simulated and labeled $\Delta S/S_0$ buildups could be due to the T_1 relaxation between ^{2}H spin states in the experiment which is not considered in the SIMPSON simulation. In the absence of relaxation, a ^{13}CO coupled to a m=±1 ^{2}H experiences dipolar evolution for the full dephasing period τ whereas a ^{13}CO coupled to a m=0 ^{2}H experiences no evolution. With relaxation, a ^{13}CO coupled to a m=±1 ^{2}H

experiences dipolar evolution for part of τ but no evolution for the remainder of τ due to the relaxation that changes the ^{2}H in the m=±1 state to the m=0 state. The relaxation will result in a slower $\Delta S/S_0$ buildup with long-τ values close to 1, which is reflected by the $\Delta S/S_0$ buildup of I4_A8$_D$A9$_C$ (Figure 3.13).

Exponential fitting of $(\Delta S/S_0)^{lab}$ vs τ

Given the exponential buildup shape, the labeled $\Delta S/S_0$ vs τ of I4_A8$_D$A9$_C$ is fitted using the function $Ax(1-e^{-\beta\tau})$. A is the fraction of ^{13}CO nuclei within the r < 9 Å detection limit and $\beta = \alpha d$, where α is a coefficient and d is the ^{13}CO-^{2}H dipolar coupling. The best-fit values are A = 0.87(5) and β = 24(2) Hz. Since d = 37 Hz for I4_A8$_D$A9$_C$, the coefficient $\alpha = \beta/d$ = 0.65(5). Since $\Delta S/S_0$ of a spin 1/2-spin 1/2 pair has a single universal dependence on λ where $\lambda = \tau d$, we expect that $\Delta S/S_0$ of a spin 1/2-spin 1 pair also has a universal dependence on λ and thus the coefficient α = 0.65 for I4_A8$_D$A9$_C$ may be also used for other different samples[10].

REFERENCES

REFERENCES

1. Gullion, T. and J. Schaefer, *Rotational-Echo Double-Resonance NMR.* Journal of Magnetic Resonance, 1989. **81**(1): p. 196-200.

2. Ghosh, U., L. Xie, and D.P. Weliky, *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.* Journal of Biomolecular NMR, 2013. **55**(2): p. 139-146.

3. Schmick, S.D. and D.P. Weliky, *Major Antiparallel and Minor Parallel beta Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide.* Biochemistry, 2010. **49**(50): p. 10623-10635.

4. Qiang, W., Y. Sun, and D.P. Weliky, *A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(36): p. 15314-15319.

5. Long, H.W. and R. Tycko, *Biopolymer conformational distributions from solid-state NMR: alpha-helix and 3(10)-helix contents of a helical peptide.* Journal of the American Chemical Society, 1998. **120**(28): p. 7039-7048.

6. Gullion, T., *Rotational-Echo, Double-Resonance NMR.* Modern Magnetic Resonance, 2008: p. 713-718.

7. Cady, S.D., et al., *Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers.* Nature, 2010. **463**(7281): p. 689-U127.

8. Morcombe, C.R. and K.W. Zilm, *Chemical shift referencing in MAS solid state NMR.* Journal of Magnetic Resonance, 2003. **162**(2): p. 479-486.

9. Zhang, H.Y., S. Neal, and D.S. Wishart, *RefDB: A database of uniformly referenced protein chemical shifts.* Journal of Biomolecular Nmr, 2003. **25**(3): p. 173-195.

10. Gullion, T., *Introduction to rotational-echo, double-resonance NMR.* Concepts in Magnetic Resonance, 1998. **10**(5): p. 277-289.

Chapter 4 - Membrane location studies of KALP and HIV fusion peptide (HFP)

4.1 Background

The insertion depth (i.e. membrane location) of protein and peptide residues in lipid membranes is an important feature of membrane-bound proteins and peptides. For instance, the membrane insertion depth of HFP is crucial for its membrane fusion activity[1]. Earlier lipid mixing assays and REDOR SSNMR studies together have shown a strong positive correlation between the membrane insertion depth and fusion activity of HFP[1-3]. The residue-specific locations of proteins and peptides in membranes can be detected by a variety of techniques including fluorescence spectroscopy, EPR, and SSNMR (e.g. PRE, spin diffusion, and REDOR), which have been discussed in detail in Chapter 1. In this chapter, we will focus on the membrane location studies of peptides using REDOR SSNMR.

As previously described in Chapter 3, our group have been using ^{13}C-^{31}P and ^{13}C-^{19}F REDOR to probe the membrane locations of peptides. For peptide ^{13}C to lipid ^{31}P REDOR, it detects membrane locations of peptides with respect to the membrane surface. For peptide ^{13}C to lipid ^{19}F REDOR, it detects membrane locations of peptides with respect to the membrane center when using the C16-^{19}F labeled dipalmitoylphosphatidylcholine (DPPC) lipid that contains a ^{1}H→^{19}F substitution at C16 in one of the two palmitoyl chains (Figure 4.1b). However, a disadvantage of using the C16-^{19}F labeled DPPC lipid is that the membrane bilayer integrity may be disrupted.

Earlier studies have shown that 100% monofluorinated DPPC lipids with a $^{1}H \rightarrow ^{19}F$ substitution at C16 in one of the two acyl chains form an interdigitated bilayer rather than a regular bilayer (Figure 4.1)[4]. The C-F bond at C16 in DPPC has a very strong electric dipole that is energetically more stable in a polar environment and drives the formation of the interdigitated lipid bilayer where the C-F bond may contact the polar headgroup region of the DPPC lipids.

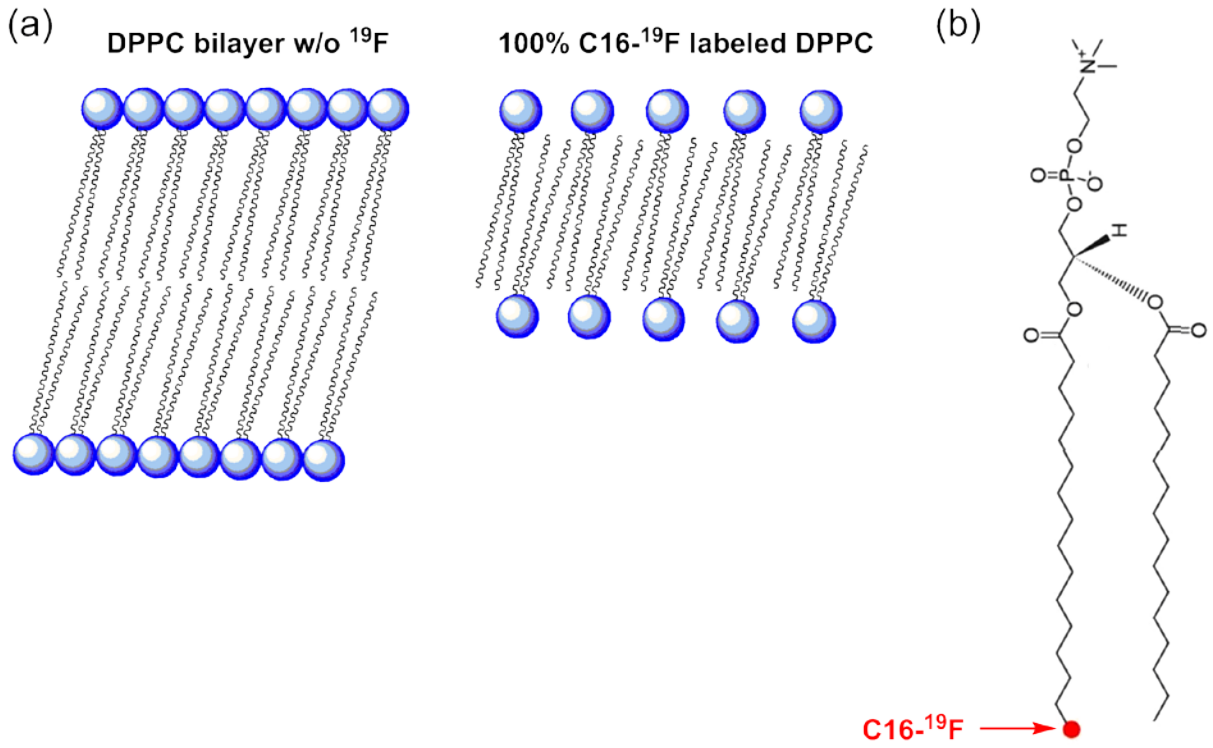

Figure 4.1 (a) Regular lipid bilayer (left) consisting of unlabeled DPPC and interdigitated bilayer (right) consisting of labeled DPPC that contains a $^{1}H \rightarrow ^{19}F$ substitution at C16 in one of the two palmitoyl chains. (b) Structure of C16-^{19}F labeled DPPC lipid.

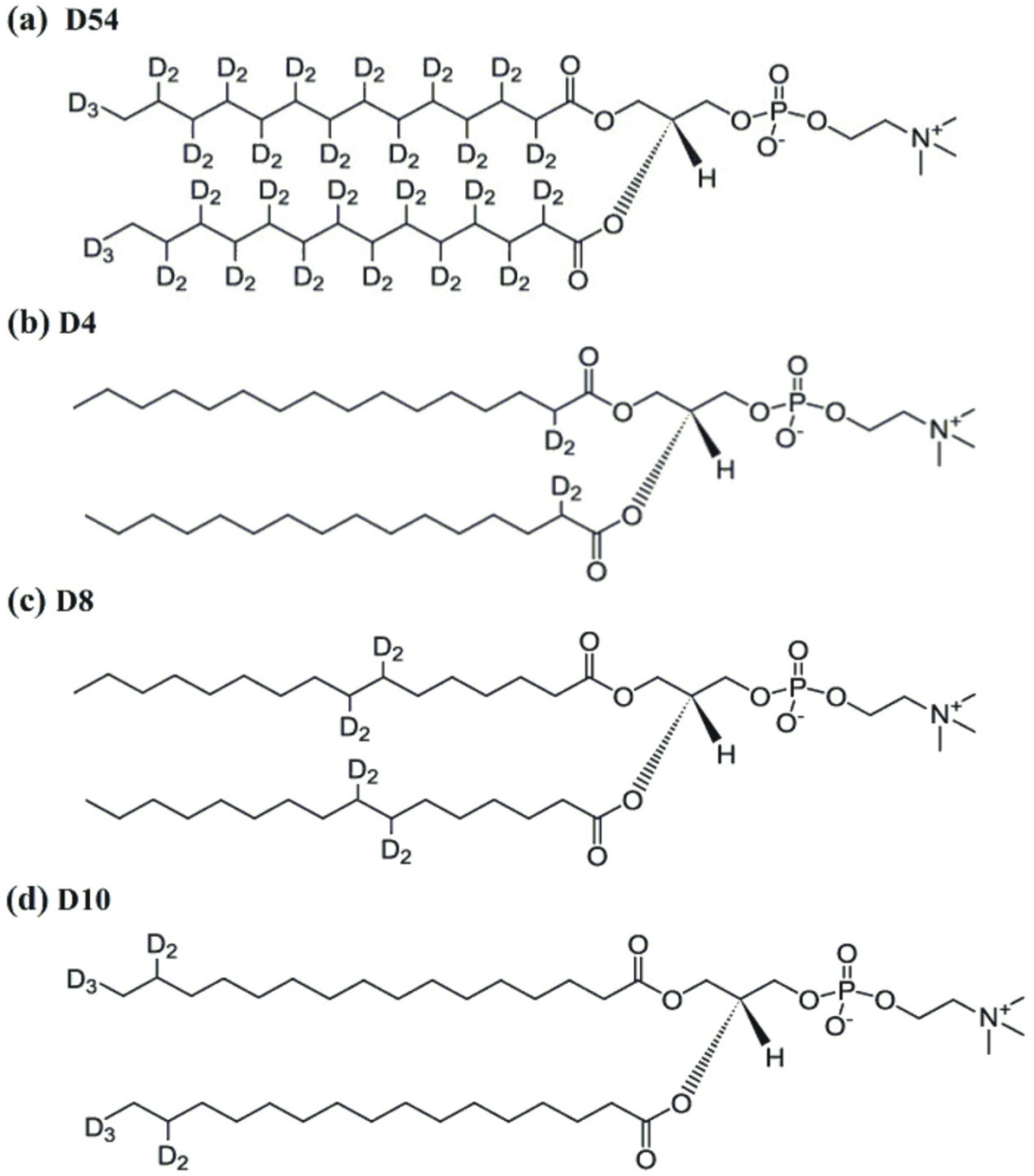

Figure 4.2 (a) D54: dimyristoylphosphatidylcholine perdeuterated in the myristoyl chains. (b–d) D4, D8, and D10: dipalmitoylphosphatidylcholine deuterated at palmitoyl carbons 2; 7 and 8; and 15 and 16, respectively[5].

By contrast, for peptide ^{13}C to lipid ^{2}H REDOR, there is no membrane perturbation no matter what fraction of lipids are deuterated since ^{1}H and ^{2}H are chemically equivalent. For the lipids in Figure 4.2, the perdeuterated DMPC lipid (D54)

is used to qualitatively determine whether the ^{13}CO labeled residue is inserted in the membrane hydrocarbon core whereas the selectively deuterated DPPC lipids (D4, D8, and D10) are used to quantitatively detect the residue-specific membrane locations of peptides.

4.2 Peptide synthesis and sample preparation

Peptide sequence and synthesis

The membrane-associated peptides KALP_A11$_C$, HFP_G5$_C$, and HFP_F8$_C$ were manually synthesized by Fmoc SPPS. KALP is a designed membrane-spanning α-helical peptide and KALP_A11$_C$ has the sequence GKKLALALALALALALALALKKA with a ^{13}CO label at A11, acetylation at the N-terminus, and amidation at the C-terminus[6]. HFP_G5$_C$ and HFP_F8$_C$ have the sequence AVGIGALFLGFLGAAGSTMGARSWKKKKKKA with a ^{13}CO label at either G5 or F8. The underlined residues are the 23 N-terminal residues of HIV gp41 fusion protein. A tryptophan residue is added as a 280 nm chromophore for peptide quantification and a non-native six-lysine tag is added to increase the aqueous solubility of HFP. The synthesized peptides were purified by reversed-phase HPLC using a preparative C4 column. The peptide purities were checked by mass spectrometry (MALDI and ESI). Comparison of peak intensities in the mass spectra indicated the peptides were >95% pure.

REDOR sample preparation

For the REDOR samples, each KALP sample was prepared using 1 μmol KALP and 50 μmol lipids whereas each HFP sample was prepared using 2 μmol HFP and 50 μmol lipids. The KALP_A11$_C$ and HFP_F8$_C$ samples were prepared in a time sequence of (1) dissolving the peptide and lipid (either D4, D8, D10, or D54) together in a mixture of 2,2,2-trifluoroethanol, chloroform, and 1,1,1,3,3,3-hexafluoroisopropanol in a 2:3:2 volume ratio. The initial co-solubilization in organic solvents minimized the fraction of kinetically trapped peptide on the membrane surface; (2) removing the organic solvents using N_2 gas in the fume hood; (3) drying the peptide-lipid film in a vacuum desicator overnight; (4) suspending the peptide-lipid film in 2~3 mL pH 7.4 buffer which contains 5 mM HEPES and 10 mM MES; (5) performing ten cycles of "freeze/thaw", where the "freeze" in liquid N_2 breaks apart the lipid vesicles and the "thaw" in water bath at room temperature drives the lipid molecules to reform vesicles in the buffer; (6) adding more buffer until the centrifuge tube is full (note that a balance centrifuge tube is needed and the mass difference between the balance and sample tubes should be within 100 mg); (7) ultracentrifugation at 270,000*g* for 4 hours (UV measurement at 280 nm showed little peptide in the supernatant liquid after ultracentrifugation); (8) lyophilizing the pellet in a freeze dryer overnight; (9) adding 10 μL pH 7.4 buffer to the bottom of a 4mm rotor, packing the lyophilized pellet in the rotor, adding another 10 μL buffer on top of the sample, and finally allowing overnight sample hydration before running ^{13}C-^{2}H REDOR experiments.

The HFP_$G5_C$ samples were prepared in the same procedure as described above except that the lipid membrane consisted of either 100% D4 or 80% D8 (or D10) + 20% DTPG, where the PC:PG (4:1) lipid composition reflected a small fraction (~10%) of negatively charged lipids and a significant fraction (~85%) of zwitterionic lipids in HIV host cell membranes[7].

4.3 ^{13}C-^{2}H REDOR experiments and ^{2}H T_1 measurements

Part I: ^{13}C-^{2}H REDOR experiments

The ^{13}C-^{2}H REDOR pulse sequence (Figure 4.3) consists of (1) ^{1}H $\pi/2$ pulse which rotates ^{1}H magnetization from z axis to the xy plane, (2) simultaneous ^{1}H and ^{13}C CP pulses which transfer magnetization from ^{1}H to ^{13}C to enhance the ^{13}C signal intensity, (3) ^{13}C π pulses at the end of each rotor period except the last one without (S_o experiment) and with (S_1 experiment) ^{2}H π pulses in the middle of each rotor period during the dephasing time τ, and (4) ^{13}C signal acquisition. TPPM ^{1}H decoupling π pulses are applied during τ and the acquisition period. The REDOR experimental conditions included 10 kHz MAS frequency, sample cooling gas temperature of -50 ^{o}C with actual sample temperature of approximately -30 ^{o}C, 5.0 μs ^{1}H $\pi/2$ pulse, 1.5 ms CP with 50 kHz ^{1}H and 62-66 kHz ramped ^{13}C fields, 8.3 μs ^{13}C π pulses, 5.1 μs ^{2}H π pulses, and 75 kHz TPPM ^{1}H decoupling fields. The recycle delay was 1 s for τ = 2 ms, 8 ms, and 16 ms, 1.5 s for τ = 24 ms and 32 ms, and 2 s for τ = 40 ms and 48 ms,

respectively. The pulses were optimized using the setup peptide I4_A8$_D$A9$_C$ as described in chapter 3.

KALP_A11$_C$ samples

Figure 4.4 displays the ^{13}CO regions of the S_0 and S_1 spectra of KALP_A11$_C$ for τ = 40 ms in D4, D8, D10, and D10 lipid membranes, respectively. The strong signal at 178.7 ppm is dominated by the labeled A11 ^{13}CO (~0.80 molar fraction) and this peak

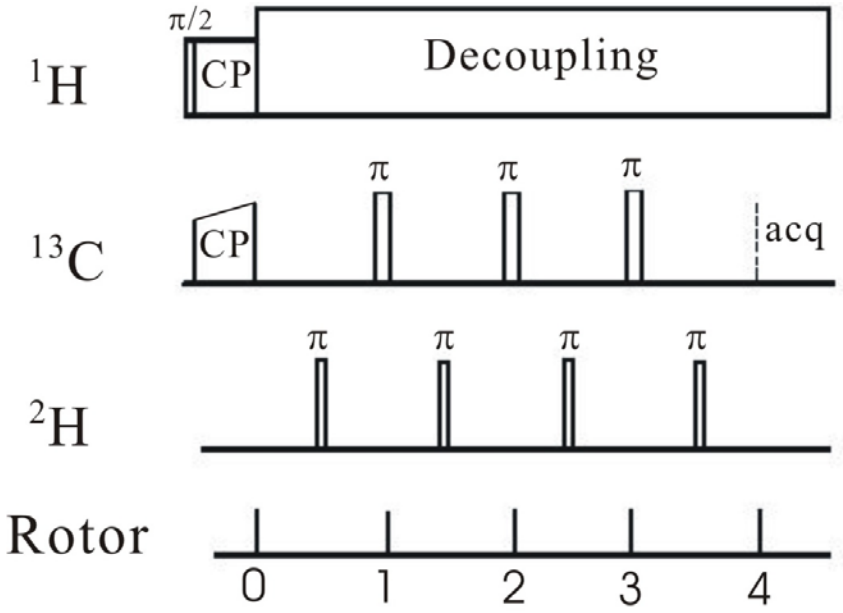

Figure 4.3 ^{13}C-^{2}H REDOR pulse sequence.

shift is consistent with helical Ala in a peptide or protein[8]. The fraction 0.80 is determined by spin counting. For KALP_A11$_C$, the labeled ^{13}CO signal is 0.99 and the unlabeled (i.e. natural abundance) ^{13}CO signal is 0.24 (0.011 x 22), thus the fractional contribution from the labeled A11 ^{13}CO is ~0.80. The shoulder at 175.7 ppm is

dominated by natural abundance lipid ^{13}CO nuclei. Comparison of the S_0 and S_1 spectra for τ = 40 ms with those for τ = 2 ms for KALP_A11$_C$ in D4 membrane (Figure 4.5) shows that the lipid signal (175~176 ppm) is less intense relative to the peptide signal (~179 ppm) at longer τ, which indicates a shorter lipid T_2[5]. For the D4 membrane, there is a significant lipid signal dephasing even for the very short τ = 2 ms since the carbonyl carbon (C1) is very close to the deuterons at C2 (~2 Å apart) in each acyl chain.

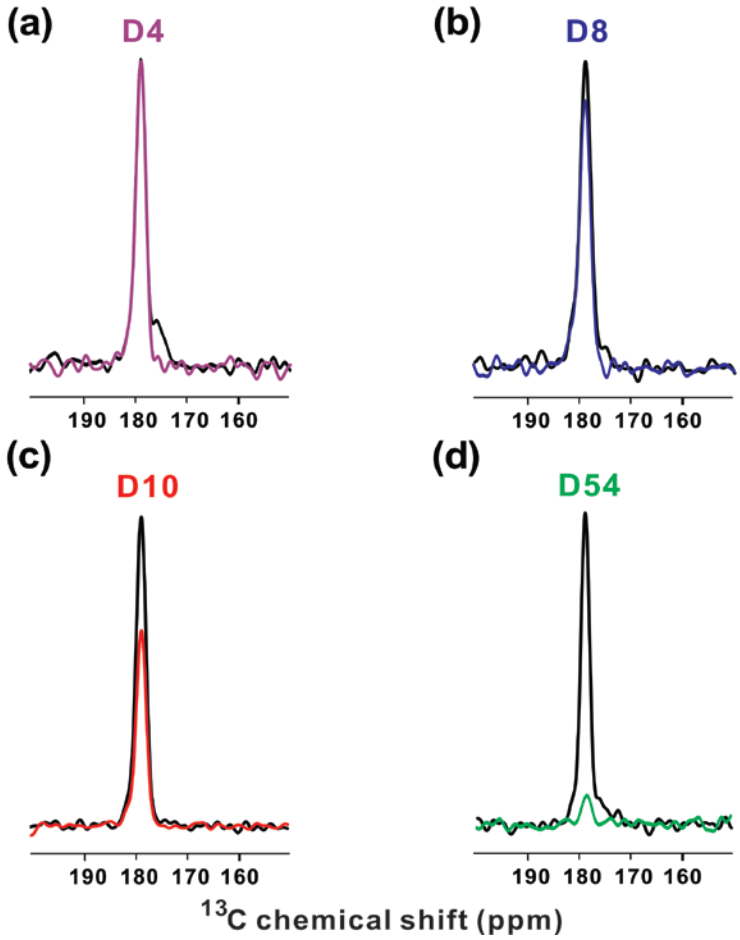

Figure 4.4 ^{13}C-^{2}H REDOR S_0 (black) and S_1 (colored) spectra for τ = 40 ms for KALP_A11$_C$ in D4, D8, D10, and D54 membranes, respectively. All spectra were processed using 100 Hz Gaussian line broadening and baseline correction.

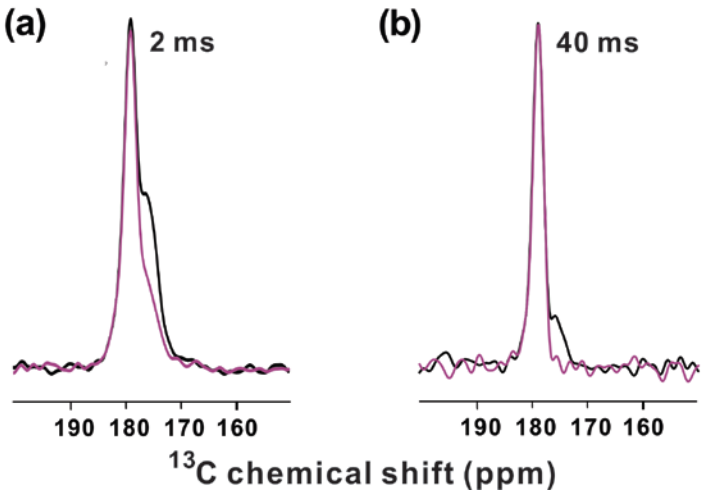

Figure 4.5 ^{13}C-^{2}H REDOR S_0 (black) and S_1 (purple) spectra for τ = 2 ms (a) and τ = 40 ms (b) for KALP_A11$_C$ in D4 lipid membrane.

The ^{13}C-^{2}H REDOR experimental dephasing plots of $\Delta S/S_0$ vs τ for KALP_A11$_C$ in D4, D8, D10, and D54 membranes are displayed in Figure 4.6 and will be discussed in section 4.4.

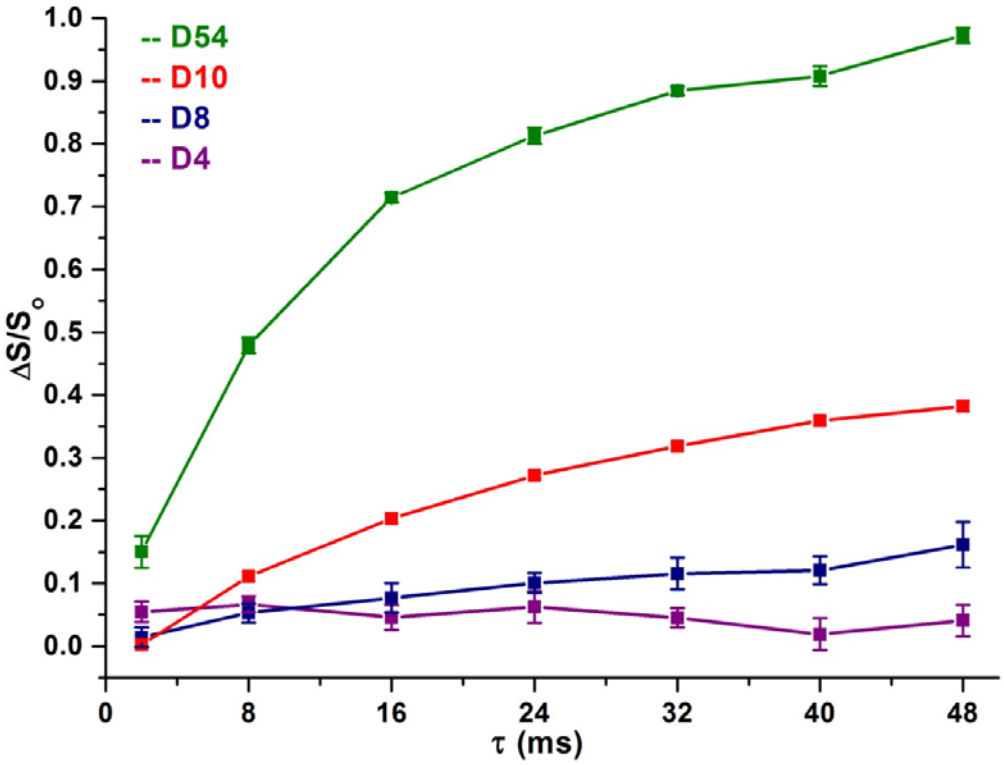

Figure 4.6 ^{13}C-^{2}H REDOR dephasing plots of $\Delta S/S_0$ vs τ for KALP_A11$_C$ in D4 (purple), D8 (blue), D10 (red), and D54 (green) membranes, respectively. The $\Delta S/S_0$ at each τ was calculated using S_0 and S_1 ^{13}CO intensities determined over a 3.0 ppm integration width.

HFP_F8$_C$ samples

Figure 4.7 displays the ^{13}CO regions of the S_0 and S_1 spectra of HFP_F8$_C$ for τ = 40 ms in D4, D8, D10, and D10 lipid membranes, respectively. The F8 ^{13}CO peak at 174 ppm is consistent with β-sheet structure and is partially overlapped with the natural abundance lipid ^{13}CO signal at 175~176 ppm[8]. The experimental dephasing plots of $\Delta S/S_0$ vs τ for HFP_F8$_C$ in membranes are displayed in Figure 4.8.

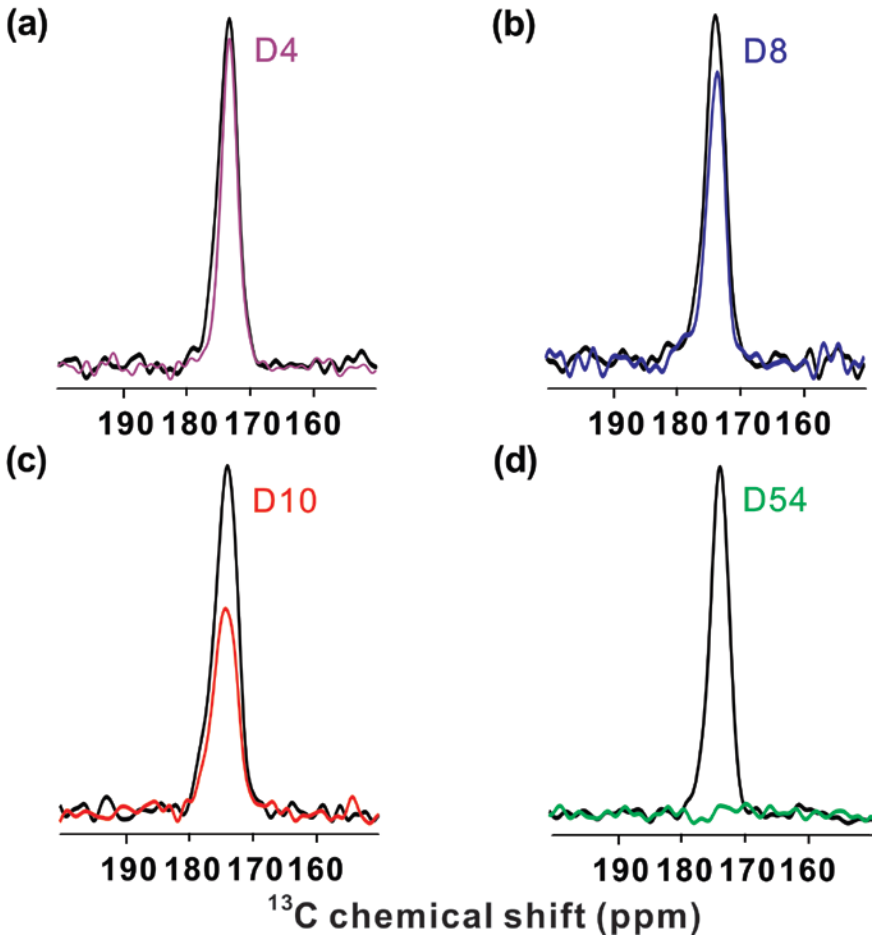

Figure 4.7 ^{13}C-^{2}H REDOR S_0 (black) and S_1 (colored) spectra for τ = 40 ms for HFP_F8$_C$ in D4, D8, D10, and D54 membranes, respectively. All spectra were processed using 100 Hz Gaussian line broadening and baseline correction.

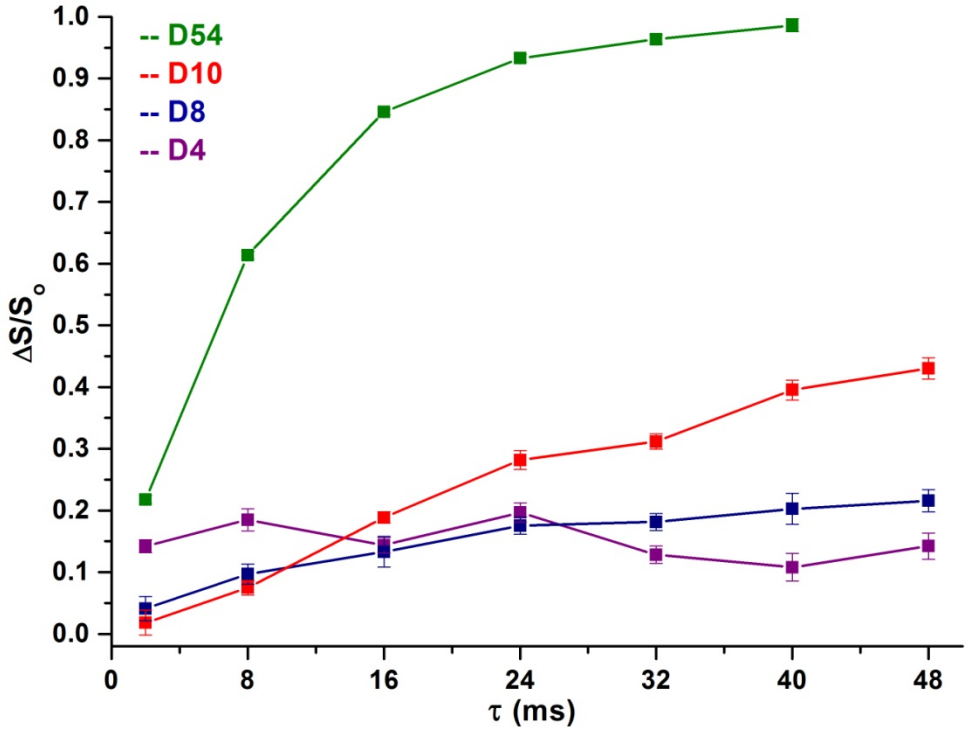

Figure 4.8 ^{13}C-^{2}H REDOR dephasing plots of $\Delta S/S_0$ vs τ for HFP_F8$_C$ in D4 (purple), D8 (blue), D10 (red), and D54 (green) membranes, respectively. The $\Delta S/S_0$ at each τ was calculated using S_0 and S_1 ^{13}CO intensities determined over a 3.0 ppm integration width.

HFP_G5$_C$ samples

Figure 4.9 displays the ^{13}CO regions of the S_0 and S_1 spectra for τ = 40 ms for HFP_G5$_C$ in D4 (100%), D8+DTPG (4:1), and D10+DTPG (4:1) membranes, respectively. The peak at 171 ppm is dominated by the labeled G5 ^{13}CO and is consistent with β-sheet G5 in the peptide. The shoulder at 175 ppm in the S_0 spectrum of Figure 4.9a may have contributions from labeled ^{13}CO of helical G5, n.a. ^{13}CO of

unlabeled residues, and n.a. ^{13}CO of D4 lipid. The experimental dephasing plots of $\Delta S/S_0$ vs τ for HFP_G5$_C$ in membranes are displayed in Figure 4.10.

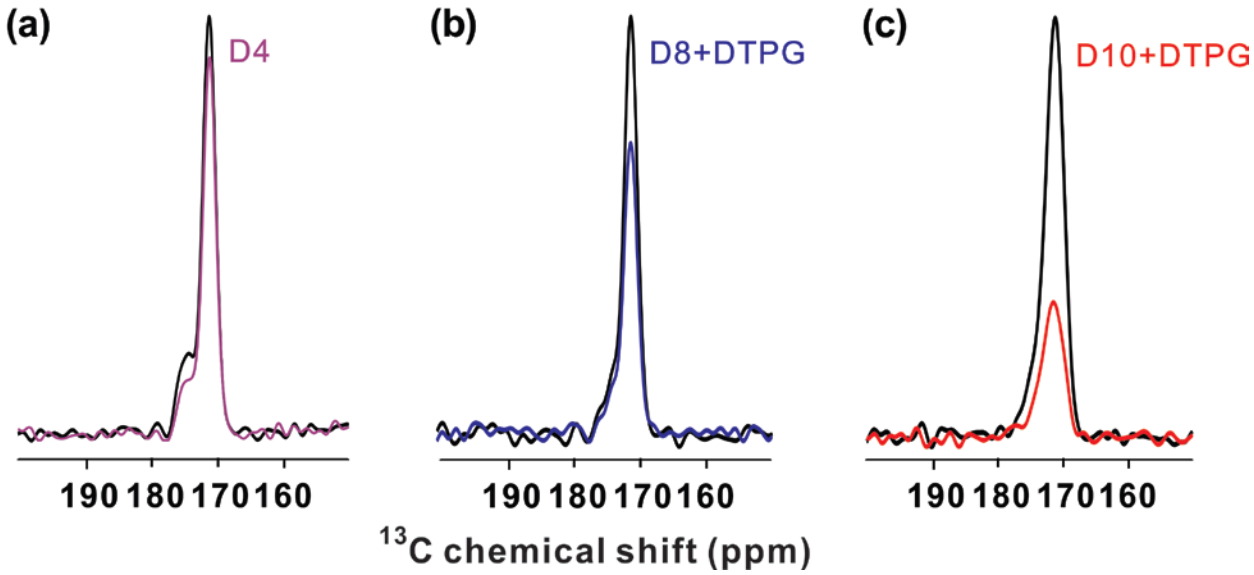

Figure 4.9 ^{13}C-^{2}H REDOR S_0 (black) and S_1 (colored) spectra for τ = 40 ms for HFP_G5$_C$ in "100% D4" (a), "80% D8 + 20% DTPG" (b), and "80% D10 + 20% DTPG" (c) membranes, respectively. All spectra were processed using 100 Hz Gaussian line broadening and baseline correction.

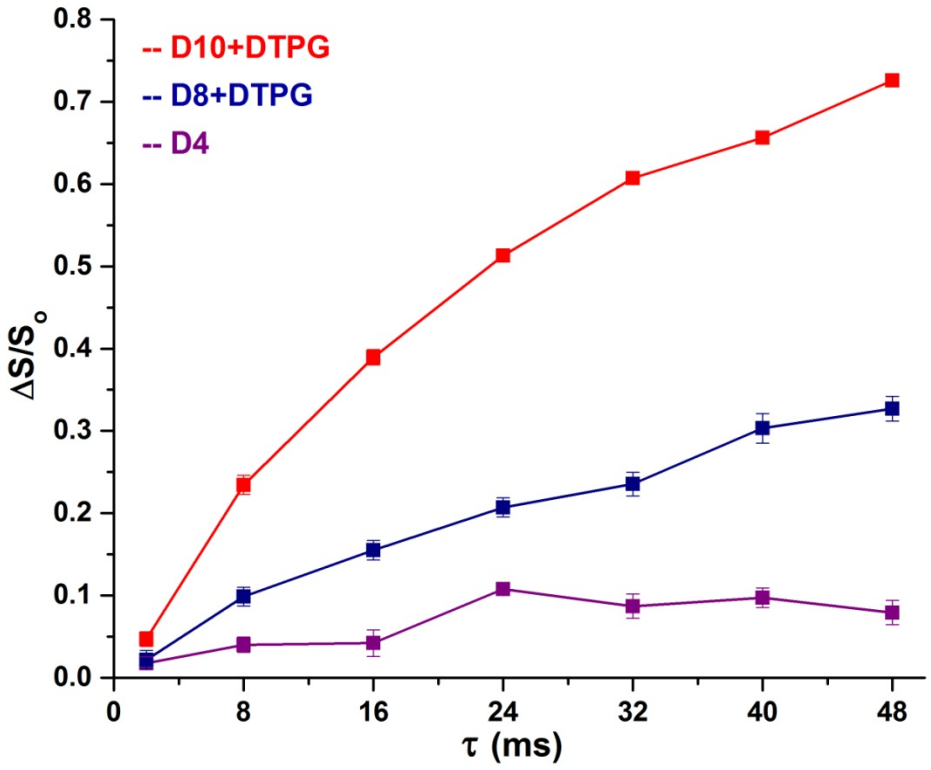

Figure 4.10 ^{13}C-^{2}H REDOR dephasing plots of $\Delta S/S_0$ vs τ for HFP_G5$_C$ in "100% D4" (purple), "80% D8 + 20% DTPG" (blue), and "80% D10 + 20% DTPG" (red) membranes, respectively. The $\Delta S/S_0$ at each τ was calculated using S_0 and S_1 ^{13}CO intensities determined over a 3.0 ppm integration width.

Part II: ^{2}H T_1 measurements

To investigate the relaxation effects on ^{13}C-^{2}H REDOR dephasing buildups of $\Delta S/S_0$ vs τ, the lipid ^{2}H T_1 relaxation times were measured by t1D_ir under both static and 10 kHz MAS conditions for HFP_F8$_C$ samples in D8, D10, and D54 membranes, respectively. The t1D_ir pulse sequence is inversion-recovery followed by a quadrupolar echo, $\pi - \tau_1 - (\pi/2)_x - \tau_2 - (\pi/2)_y - \tau_3$ – detect (Figure 4.11). ^{2}H spectra were acquired

for different τ_1 with fixed τ_2 and τ_3 (τ_2 was set as 100 μs such that the two $\pi/2$ pulses were rotor-synchronized under 10 kHz MAS).

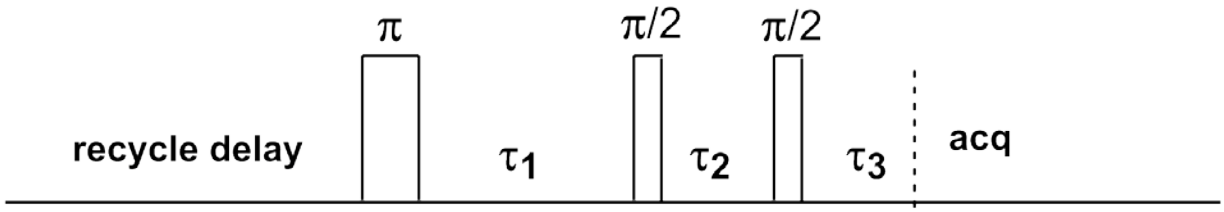

Figure 4.11 "t1D_ir" pulse sequence used for ^{2}H T_1 measurements. The phase/phase cycle is "x" for the π pulse, "x, -x, y, -y" for the first $\pi/2$ pulse, and "y, y, x, x" for the second $\pi/2$ pulse.

^{2}H spectra under static conditions (no spinning)

The static ^{2}H spectra were acquired at -50 ^{o}C for τ_1 = 0.1 ms through 300 ms in an increment of 20 ms for HFP_F8$_C$ in D8, D10, and D54 membrane, respectively. The number of acquisitions was set the same for each τ_1. Ideally, all the ^{2}H spectra for different τ_1 should be acquired in the same array. ^{2}H spectra were also acquired for HFP_F8$_C$ in D4 but the signal-to-noise was very poor, which was probably due to (a) small population of deuterons compared to the D8, D10, and D54 samples and (b) large quadrupolar anisotropy causing spectral broadening and thus lower signal-to-noise ratio.

As we process the ^{2}H FID data, we need to do "data shift" to move the ^{2}H quadrupolar echo which is the most intense signal in the FID to time zero before performing Fourier transformation (FT). For instance, if the dwell time which is the interval between two nearby sampling points is 2 μs and the ^{2}H quadrupolar echo in the

FID appears at 22 μs, we need to set "data shift" = -11 (i.e. shift the FID to the left by 11 data points) in the process panel before performing FT. For the same sample (e.g. HFP_F8$_C$ in D10), all ^{2}H spectra were processed using the same parameters (e.g. zero- and 1st-order phasing, data shift, and line broadening). The static ^{2}H spectra for HFP_F8$_C$ in D8, D10, and D54 membranes are displayed in Figure 4.12, 4.13, and 4.14, respectively.

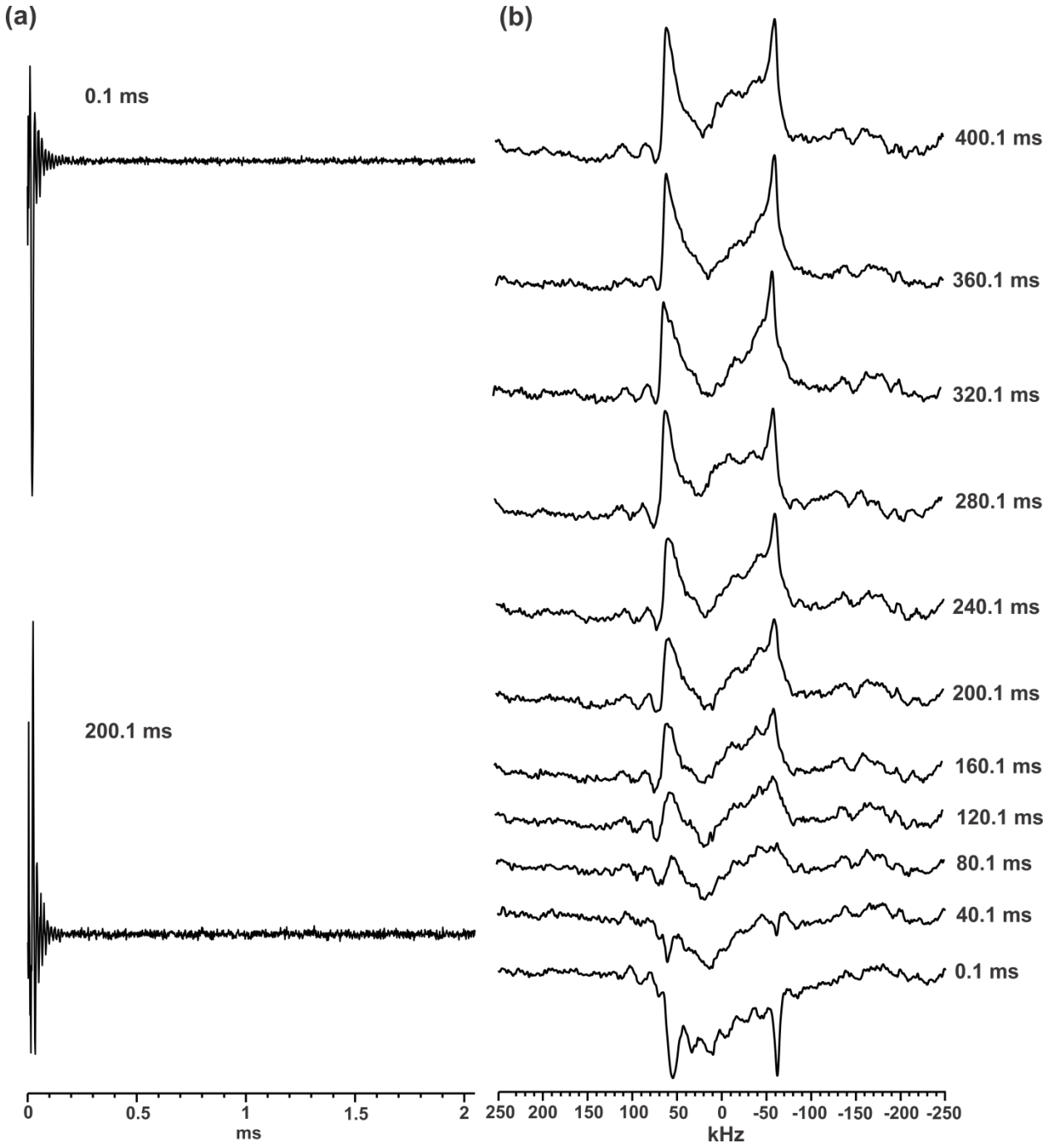

Figure 4.12 "t1D_ir" experiments of HFP_F8$_C$ in D8 membrane at -50 ^{o}C under static condition. For each τ_1, the number of acquisition = 16000. (a) ^{2}H FID for τ_1 = 0.1 ms and 200.1 ms; (b) ^{2}H spectra for τ_1 = 0.1 ms through 400.1 ms in an increment of 40 ms. All spectra were processed using 2000 Hz Gaussian line broadening and data shift of -12. For simplicity, spectra for τ_1 = 20.1 ms through 380.1 ms in an increment of 40 ms are not shown here.

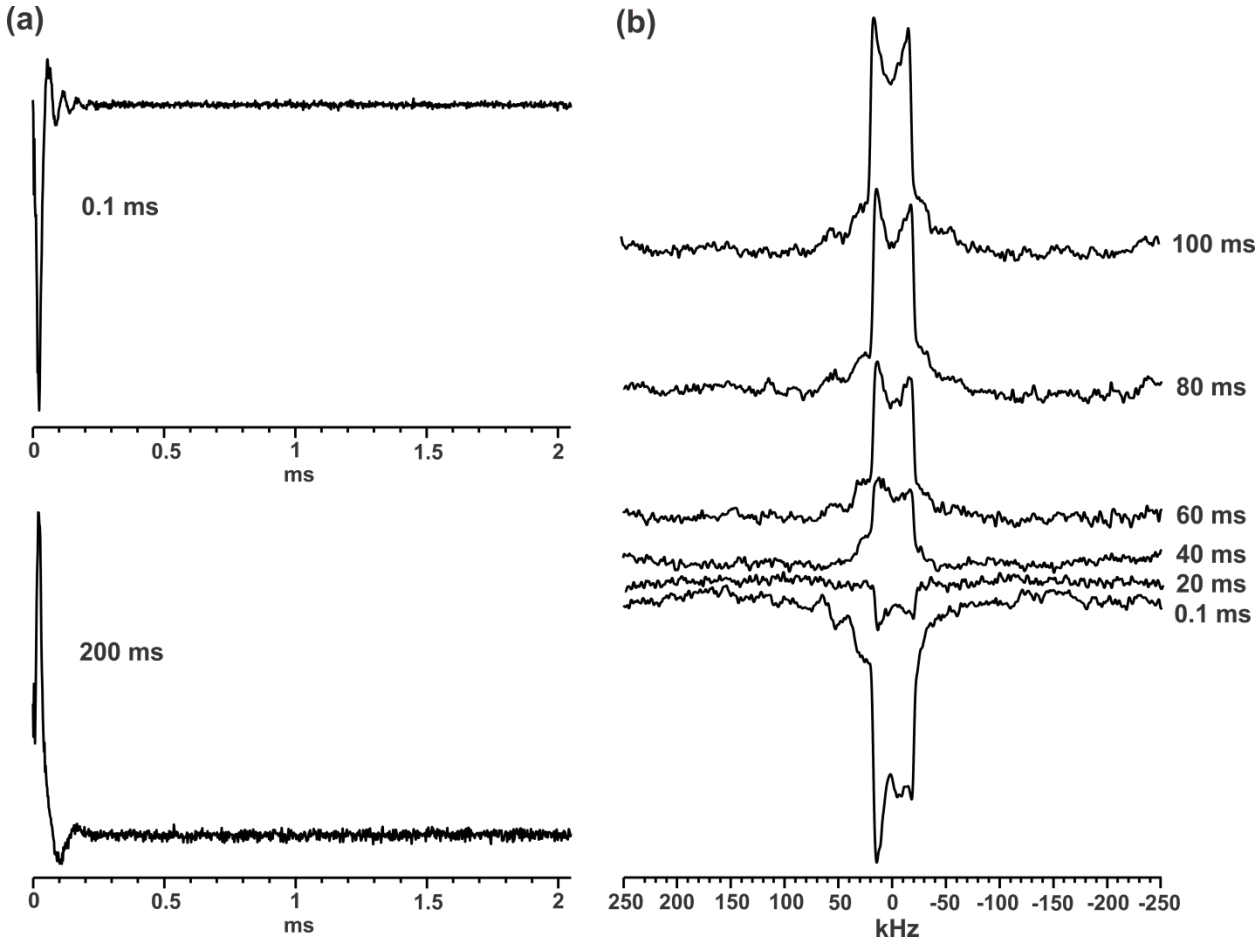

Figure 4.13 "t1D_ir" experiments of HFP_F8$_C$ in D10 membrane at -50 ^{o}C under static condition. For each τ_1, the number of acquisition = 3000. (a) ^{2}H FID for τ_1 = 0.1 ms and 200 ms; (b) ^{2}H spectra for τ_1 = 0.1 ms through 100 ms. All spectra were processed using 2000 Hz Gaussian line broadening, data shift of -11, and baseline correction of order 3. For simplicity, spectra for τ_1 = 120 ms through 300 ms are not shown.

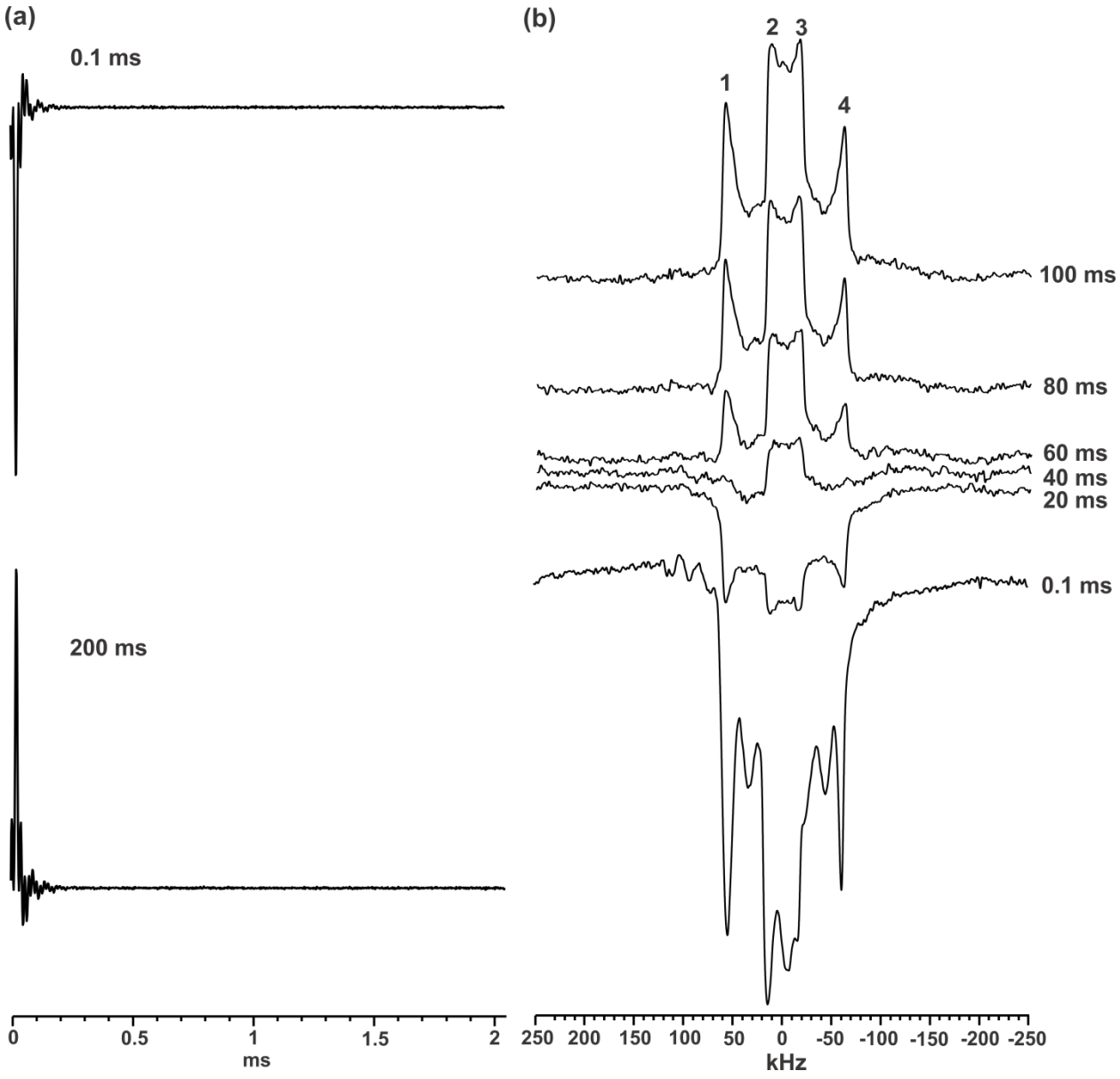

Figure 4.14 "t1D_ir" experiments of HFP_F8$_C$ in D54 membrane at -50 ^{o}C under static condition. For each τ_1, the number of acquisition = 2000. (a) ^{2}H FID for τ_1 = 0.1 ms and 200 ms; (b) ^{2}H spectra for τ_1 = 0.1 ms through 100 ms. All spectra were processed using 2000 Hz Gaussian line broadening and data shift of -12. The outer horns (1 and 4) and inner horns (2 and 3) were diagnostic of -CD_3 and -CD_2- in D54 membrane, respectively. For simplicity, spectra for τ_1 = 120 through 300 ms are not shown.

^{2}H spectra under 10 kHz MAS

The ^{2}H spectra under MAS for HFP_F8$_C$ in D8, D10, and D54 membranes were acquired under the same experimental conditions as those for static ^{2}H spectra described above except for 10 kHz MAS. The ^{2}H spectra under MAS for HFP_F8$_C$ in D8, D10, and D54 membranes are displayed in Figure 4.15, 4.16 and 4.17, respectively.

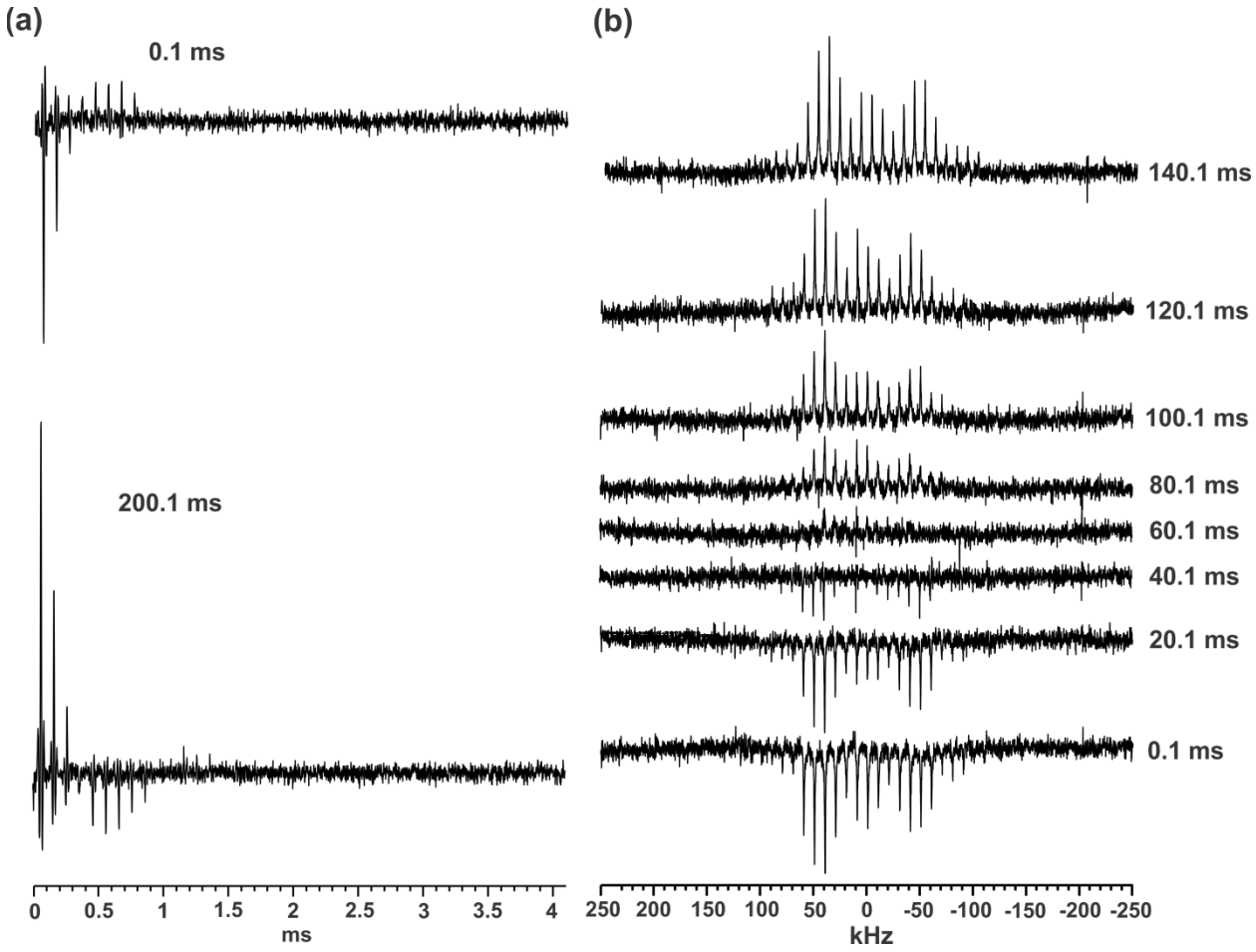

Figure 4.15 "t1D_ir" experiments of HFP_F8$_C$ in D8 membrane at -50 ^{o}C under 10 kHz MAS. For each τ_1, the number of acquisition = 15000. (a) ^{2}H FID for τ_1 = 0.1 ms and 200.1 ms; (b) ^{2}H spectra for τ_1 = 0.1 ms through 140.1 ms. All spectra were processed using 200 Hz Gaussian line broadening, data shift of -32, and baseline correction of order 3. For simplicity, spectra for τ_1 = 160.1 through 400.1 ms are not displayed.

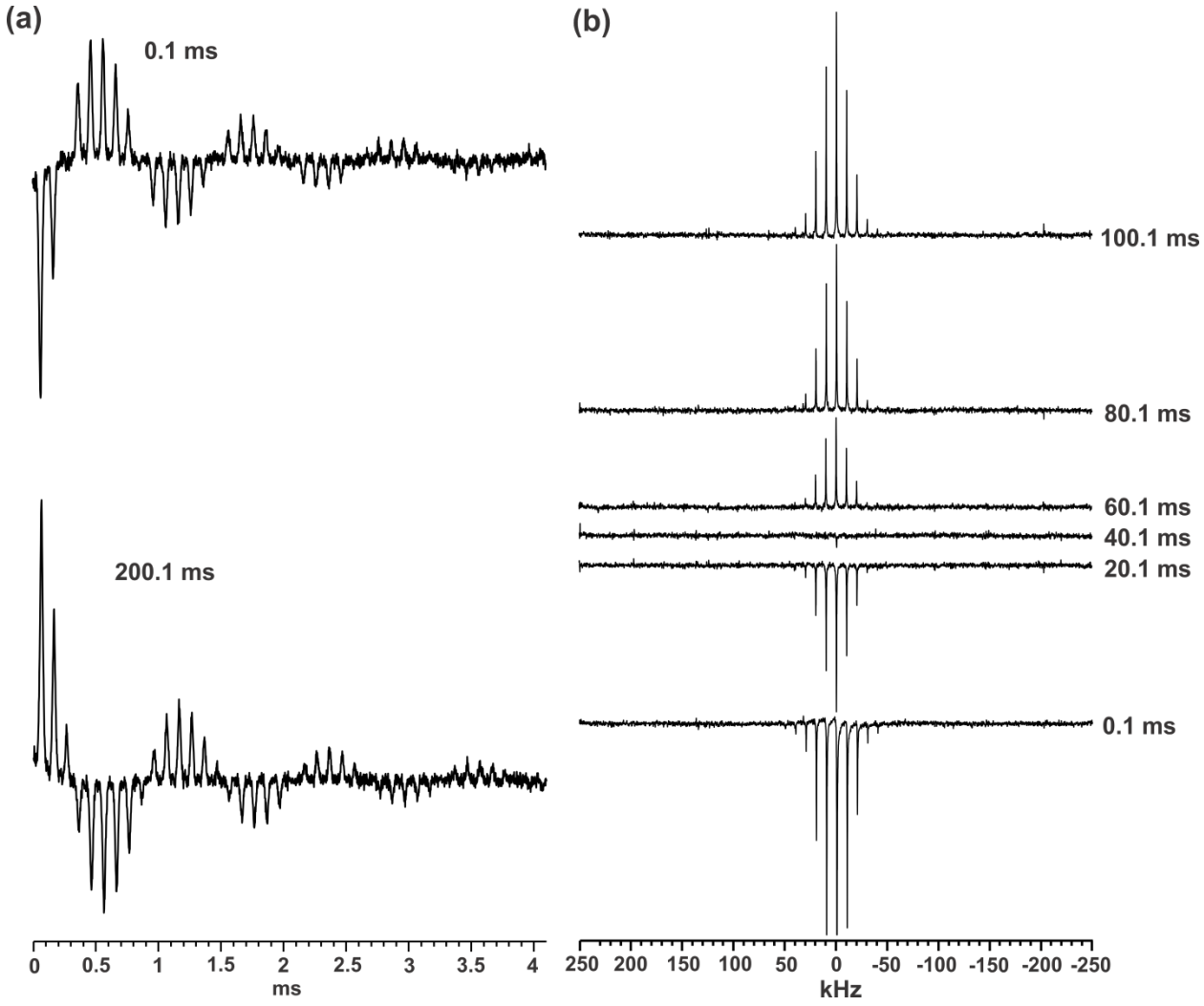

Figure 4.16 "t1D_ir" experiments of HFP_F8$_C$ in D10 membrane at -50 ^{o}C under 10 kHz MAS. For each τ_1, the number of acquisition = 5000. (a) ^{2}H FID for τ_1 = 0.1 ms and 200.1 ms; (b) ^{2}H spectra for τ_1 = 0.1 ms through 100.1 ms. All spectra were processed using 200 Hz Gaussian line broadening, data shift of -31, and baseline correction of order 10. For simplicity, spectra for τ_1 = 120.1 through 300.1 ms are not displayed.

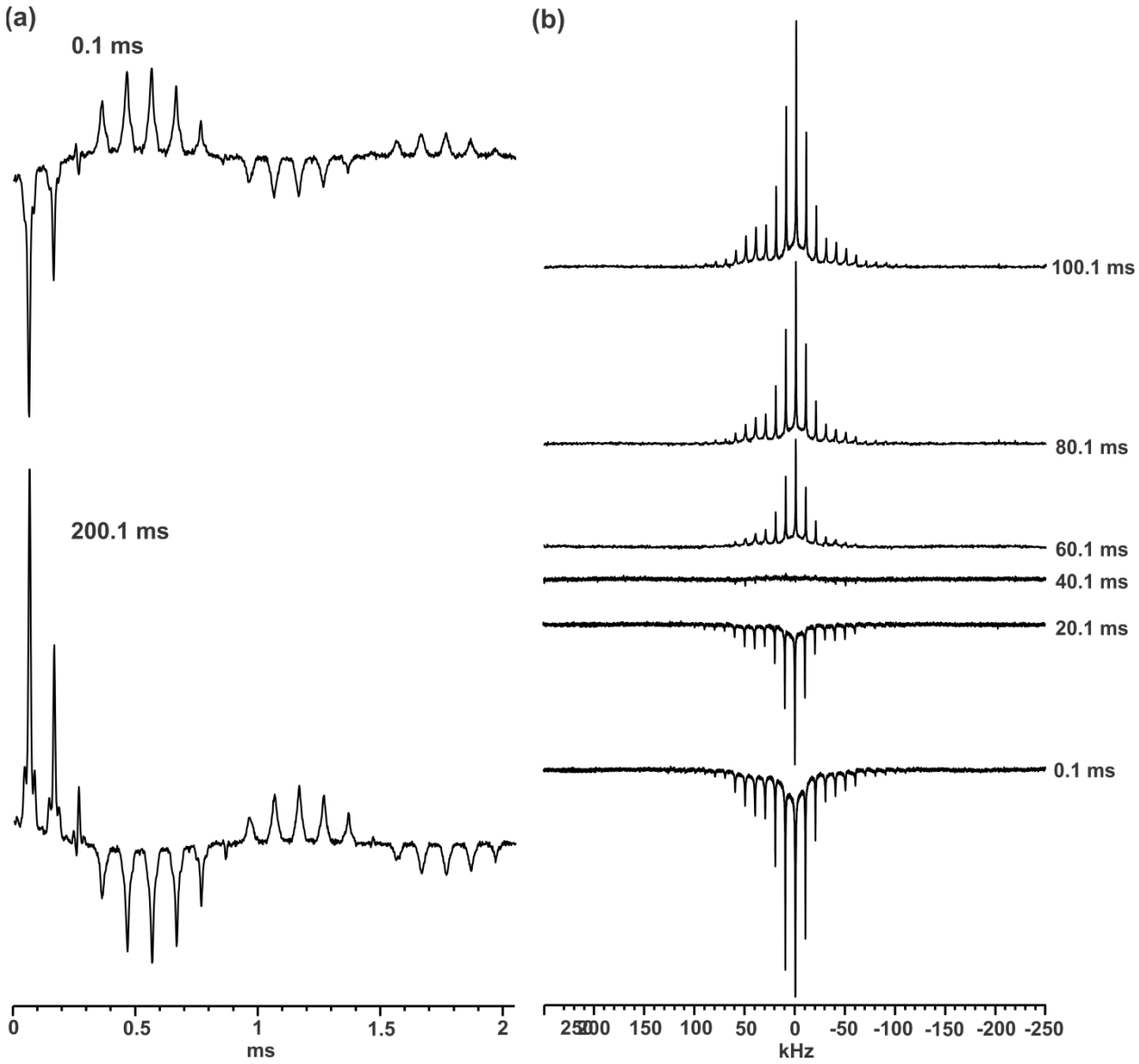

Figure 4.17 "t1D_ir" experiments of HFP_F8$_C$ in D54 membrane at -50 ^{o}C under 10 kHz MAS. For each τ_1, the number of acquisition = 3000. (a) ^{2}H FID for τ_1 = 0.1 ms and 200.1 ms; (b) ^{2}H spectra for τ_1 = 0.1 ms through 100.1 ms. All spectra were processed using 200 Hz Gaussian line broadening and data shift of -12. For simplicity, spectra for τ_1 = 120.1 ms through 300.1 ms are not displayed.

Data fitting

(1) *under static condition*

(1a) For HFP_F8$_C$ in D8 membrane, there were two horns (Pake doublet, see Chapter 1 for more details) in each ^{2}H spectrum, one at 62 kHz and the other at -62 kHz. The ^{2}H intensity for each τ_1 was calculated as the sum of the integrations over a 100 ppm (~6 kHz) width for both horns. The ^{2}H signals correspond to the -CD_2- deuterons in D8 membrane. The spectral error σ was calculated as the standard deviation of twelve noise integrals over a 100 ppm width for each and the uncertainty of the ^{2}H signal (i.e. sum of the two horn integrals) for each τ_1 was calculated as $\sqrt{2}\,\sigma$.

(1b) For HFP_F8$_C$ in D10 membrane, there were also two horns in each ^{2}H spectrum (Figure 4.12), one at ~10 kHz and the other at ~-20 kHz, which correspond to the -CD_3 deuterons in D10 membrane. The ^{2}H intensity and uncertainty for each τ_1 were calculated in the same way as described in (1a). For the -CD_2- deuterons, the ^{2}H intensity for each τ_1 was not calculated due to the poor resolution and low signal-to-noise ratio.

(1c) For HFP_F8$_C$ in D54 membrane, there were two outer horns corresponding to the -CD_2- deuterons and two inner horns corresponding to the -CD_3 deuterons in each ^{2}H spectrum (Figure 4.13). The ^{2}H intensity for -CD_2- for each τ_1 was calculated as the sum of the integrations over a 100 ppm (~6 kHz) width for the two outer horns and the intensity for -CD_3 was calculated as the sum of the integrations over a 100 ppm (~6 kHz)

width for the two inner horns. The uncertainty of each ^{2}H intensity was calculated in the same way as described in (1a).

For each sample, the integrated ^{2}H intensity vs τ_1 was fitted by

$$I(\tau_1) = I_0 + \{\Delta I \times [1 - \exp(-\tau_1 / T_1)]\} \quad (4.1)$$

where I_0, ΔI, and T_1 are fitting parameters and respectively correspond to $I(\tau_1 = 0)$, $[I(\tau_1 = \infty) - I(\tau_1 = 0)]$, and 1/(longitudinal ^{2}H relaxation rate). The best-fit plots are displayed in Figure 4.18 and best-fit T_1 values are listed in Table 4.1.

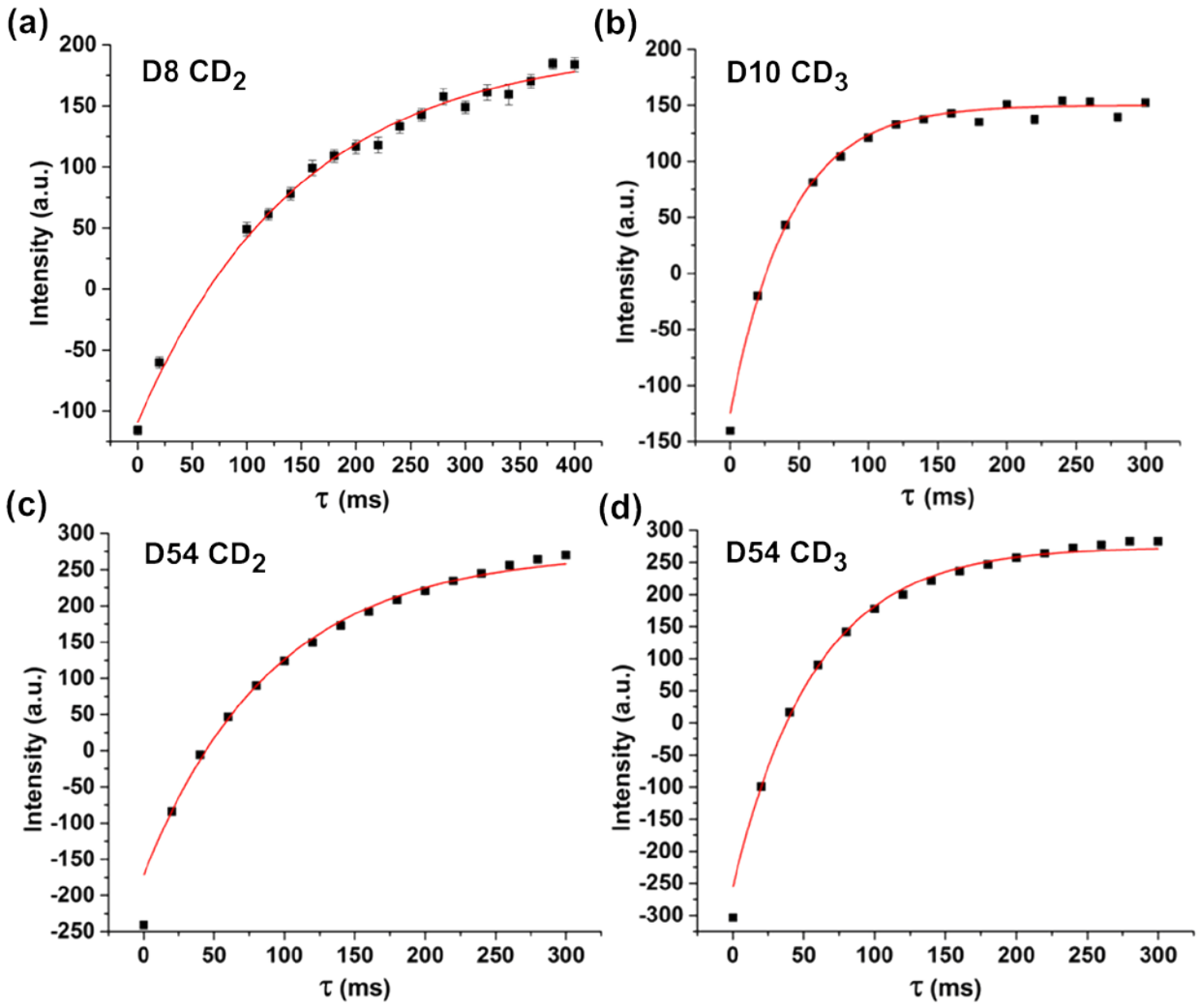

Figure 4.18 "t1D_ir" experimental (black squares with uncertainties) and best-fit (red line) plots of ^{2}H intensity vs τ_1 under static conditions for CD_2 in D8 sample (a), CD_3 in D10 sample (b), and CD_2 and CD_3 in D54 sample (c, d), respectively. a.u. ≡ arbitrary unit.

(2) *under 10 kHz MAS*

(2a) For HFP_F8$_C$ in D8 membrane, each ^{2}H spectrum was split into a series of spinning sidebands with equal spacing of 10 kHz. The ^{2}H intensity for each τ_1 was calculated as the sum of integrations of the centerband and 12 dominant sidebands over a 12 ppm integration width for each (the line width at half maximum (LWHM) for

each sideband is ~12 ppm). The spectral error σ was calculated as the standard deviation of twelve noise integrals over a 12 ppm integration width for each and the uncertainty of the ^{2}H intensity (i.e. sum of the centerband and 12 sideband integrals) for each τ_1 was calculated as $\sqrt{13}\,\sigma$.

(2b) For HFP_F8$_C$ in D10 membrane, each ^{2}H spectrum was also split into a series of spinning sidebands with equal spacing of 10 kHz. The ^{2}H intensity for each τ_1 was calculated as the sum of integrations of the centerband and 6 dominant sidebands over a 6 ppm integration width for each (the LWHM for each sideband is ~6 ppm). The spectral error σ was calculated as the standard deviation of twelve noise integrals over a 6 ppm integration width for each and the uncertainty of the ^{2}H intensity (i.e. sum of the centerband and 6 sideband integrals) for each τ_1 was calculated as $\sqrt{7}\,\sigma$.

(2c) For HFP_F8$_C$ in D54 membrane, the ^{2}H intensity for each τ_1 was calculated as the sum of integrations of the centerband and 10 dominant sidebands over a 6 ppm integration width for each (the LWHM for each sideband is ~6 ppm). The spectral error σ was calculated as the standard deviation of twelve noise integrals over a 6 ppm integration width for each and the uncertainty of the ^{2}H intensity (i.e. sum of the centerband and 10 sideband integrals) for each τ_1 was calculated as $\sqrt{11}\sigma$.

For the D8 and D54 samples, the integrated ^{2}H intensity vs τ_1 was fitted by equation 4.1. For the D10 sample, the integrated ^{2}H intensity vs τ_1 was fitted by

$$I(\tau_1) = I_0 + \{0.4 \times \Delta I \times [1 - \exp(-\tau_1/T_{1(CD2)})]\} + \{0.6 \times \Delta I \times [1 - \exp(-\tau_1/T_{1(CD3)})]\} \quad (4.2)$$

where separate $-CD_2$ and $-CD_3$ contributions to $I(\tau_1)$ were considered with distinct $T_{1(CD2)}$ and $T_{1(CD3)}$. The best-fit plots are displayed in Figure 4.19 and best-fit T_1 values are listed in Table 4.1.

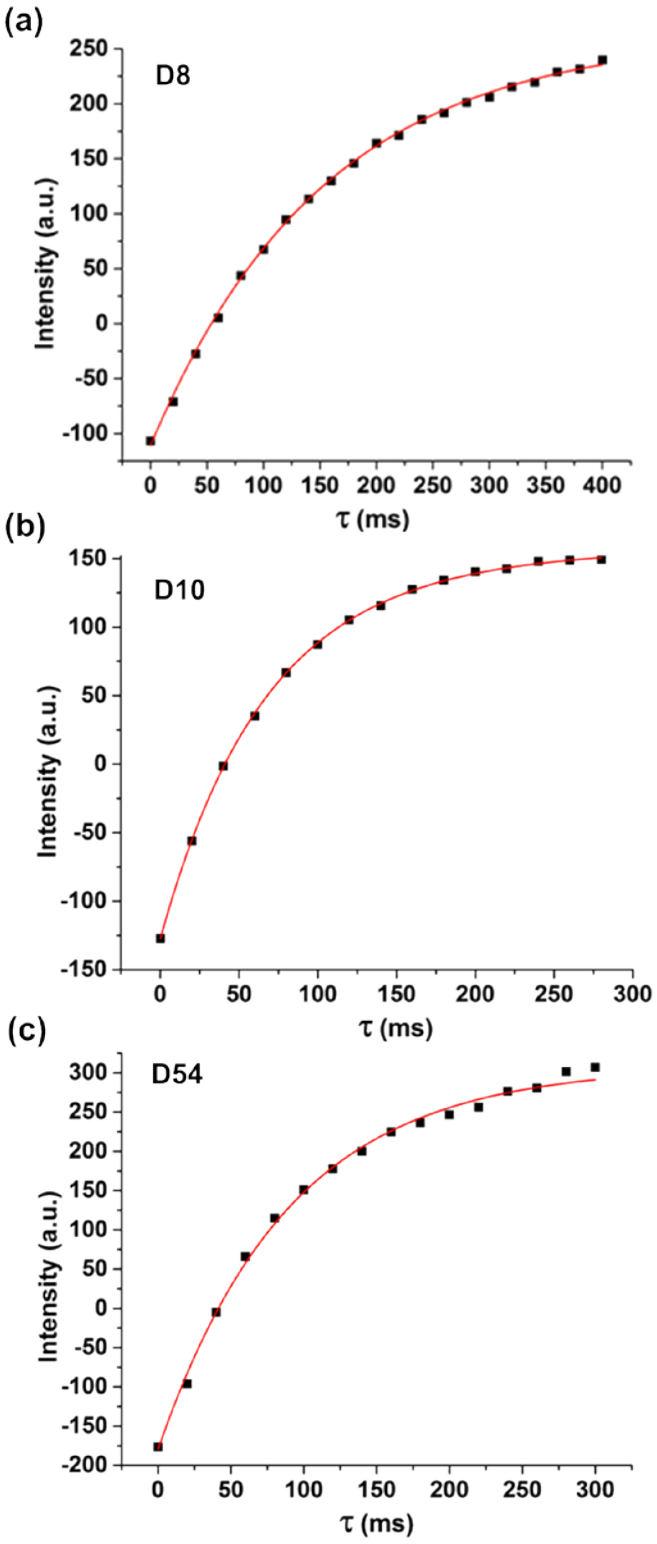

Figure 4.19 "t1D_ir" experimental (black squares with uncertainties) and best-fit (red line) plots of ^{2}H intensity vs τ_1 under 10 kHz MAS for ^{2}H in D8 (a), D10 (b), and D54 samples (c), respectively.

Table 4.1 Best-fit lipid ^{2}H T_1 in ms with uncertainty in parentheses

Sample	Static	10 kHz MAS
D54 + HFP	CD_3: 57(2) CD_2: 91(6)	90(4)
D10 + HFP	CD_3: 43(1) CD_2: not determined	CD_3: 58(3) CD_2: 92(8)
D8 + HFP	CD_2: 148(9)	CD_2: 153(3)

4.4 Result discussion and conclusions

Initial fitting of the REDOR data to determine ^{13}C-^{2}H dipolar coupling *d* was done using SIMPSON simulations. However, similar to the $(\Delta S/S_0)^{exp}$ buildup of $I4\_A8_DA9_C$ which contained a single ^{13}CO-^{2}H spin pair with *r* = 5.0 Å and *d* = 37 Hz (Figure 3.9), the $(\Delta S/S_0)^{exp}$ buildups of all KALP and HFP samples had exponential shape which contrasted with the sigmoidal shape of $(\Delta S/S_0)^{sim}$ buildups. As previously discussed in chapter 3, the differences between $(\Delta S/S_0)^{exp}$ and $(\Delta S/S_0)^{sim}$ buildups could be due to the ^{2}H T_1 relaxation in the experiments which was not considered in the SIMPSON simulations. This possibility was supported by the $\Delta S/S_0$ buildup of $I4\_A8_DA9_C$ (see chapter 3 for more details) and the lipid ^{2}H T_1 values in the 40-150 ms range that were

comparable to the larger values of τ in experiments (Table 4.1). All $(\Delta S/S_0)^{exp}$ buildups were fitted well by the function Ax(1-e$^{-\beta\tau}$). A is the fraction of ^{13}CO nuclei within the r < 9 Å detection limit and β = 0.65xd, where 0.65 is a coefficient determined from the known d = 37 Hz for I4_A8$_D$A9$_C$ (see chapter 3). The best-fit parameters are summarized in Table 4.2 and best-fit plots of $\Delta S/S_0$ vs τ for KALP_A11$_C$, HFP_F8$_C$, and HFP_G5$_C$ are displayed in Figure 4.20a, Figure 4.22a, and Figure 4.22b, respectively. For the distance $r = (4642/d)^{1/3}$, its uncertainty listed in Table 4.2 is calculated by

$$\sigma(r) = 4642^{1/3} x \sigma(d^{-1/3})$$

$$= 4642^{1/3} x(-1/3) x d^{-4/3} x \sigma(d)$$

$$= (-1/3) x (4642/d)^{1/3} x d^{-1} x \sigma(d)$$

$$= (-r/3) x \frac{\sigma(d)}{d} \qquad (4.3)$$

where $\sigma(d)$ is the uncertainty of dipolar coupling d and the - sign can be removed since the uncertainty is usually reported as a positive number (magnitude). The derivation of $\sigma(r)$ is based on the general equation to calculate the uncertainty of A^n:

$$\sigma(A^n) = n \times A^{n-1} \times \sigma(A) \qquad (4.4)$$

where σ(A) is the uncertainty of A. According to equation 4.3, the relative uncertainty of r is only 1/3 of that of d, i.e. $\sigma(r)/r = (1/3) \times \sigma(d)/d$.

Membrane locations of KALP

The A(D54) = 0.96 and r(D54) = 3.3 Å for KALP_A11$_C$ in perdeuterated D54 membrane support a transmembrane KALP topology with van der Waals contact between KALP and the membrane hydrocarbon core. This topology is further supported by $(\Delta S/S_o) \approx 0$ for the D4 sample where the ^{2}Hs are close to the membrane headgroups. For a single membrane location of KALP, we would expect A(D8) ≈ A(D10) and possibly different values of d(D8) and d(D10) and therefore different r(D8) and r(D10). However, the experimentally derived A(D8) and A(D10) were very different while r(D8) ≈ r(D10) ≈ 4 Å. We were unable to interpret these results quantitatively and self-consistently using a single membrane location of α helical KALP with sidechains. We were successful using two distinct membrane locations with major and minor populations. The major population has A11 contact with D10 ^{2}H nuclei whereas the minor population has A11 contact with D8 ^{2}H nuclei (Figure 4.20b). The distance r(D8) and r(D10) were converted to the membrane locations of A11 ^{13}CO relative to the bilayer center using the membrane location models of KALP in Figure 4.21. In an α helical peptide structure, the backbone radius of the helix is 2.3 Å. For KALP with sidechains, the overall α helix radius is approximately 2.3 Å + 3.8 Å = 6.1 Å, where 3.8 Å is the estimated length of leucine sidechain. The distance between D8 ^{2}H and D10 ^{2}H in the DPPC membrane is ~ 9 Å. For the major KALP location (Figure 4.21a), a is the distance between A11 ^{13}CO nucleus and DPPC membrane bilayer center, r is the closest distance between A11 ^{13}CO and lipid ^{2}H nuclei, and $r = r(D10) = (3.8^2 + a^2)^{1/2} =$

Table 4.2 ^{13}C-^{2}H REDOR fitted parameters

Sample	*A*	β(Hz)	*d*(Hz)	*r*(Å)
I4_A8$_D$A9$_C$	0.87(5)	24(2)	37	5.0
KALP_A11$_C$ in D8	0.15(2)	47(10)	72(15)	4.0(3)
KALP_A11$_C$ in D10	0.48(4)	34(5)	52(7)	4.5(2)
KALP_A11$_C$ in D54	0.96(1)	85(4)	131(12)	3.3(1)
HFP_F8$_C$ in D8	0.21(1)	71(10)	109(17)	3.5(2)
HFP_F8$_C$ in D10	0.82(20)	16(5)	25(8)	5.7(6)
HFP_F8$_C$ in D54	0.99(1)	122(1)	188(14)	2.9(1)
HFP_G5$_C$ in D8+DTPG	0.45(5)	27(5)	42(8)	4.8(3)
HFP_G5$_C$ in D10+DTPG	0.85(3)	37(3)	57(6)	4.3(2)

4.5 Å which corresponds to a = 2.4 Å. For the minor KALP location (Figure 4.21b), r = r(D8) = $(3.8^2 + b^2)^{1/2}$ = 4.0 Å which corresponds to b = 1.2 Å and 9 Å – b = 7.8 Å, i.e. the minor A11 ^{13}CO location is 7.8 Å from the DPPC membrane center.

For the membrane location models in Figure 4.21, KALP was considered as a monomer in the membrane. Although earlier molecular dynamics simulations have suggested that KALP and WALP, which is similar to KALP in both sequence and

topology except that the flanking residues are tryptophan rather than lysine, could aggregate and form oligomers, this possibility was ruled out in our experiments because a rapid and complete buildup of $\Delta S/S_O$ vs τ was observed for KALP in D54 membrane (Figure 4.6)[9, 10]. For an oligomer population, we expect that there would be some A11 residues in the oligomer interior. This population would not have van der Waals contact with lipid ^2H and would dephase slowly.

One possible reason for multiple membrane locations of KALP may be the hydrophobic mismatch. The hydrophobic length of KALP is ~26 Å (L4 to L20) whereas the hydrophobic thickness of the DPPC membrane is ~31 Å. To compensate for the hydrophobic mismatch, the lysine sidechains of KALP could snorkel into the DPPC membrane headgroup region to increase the hydrophobic length of KALP, which is consistent with the snorkeling model proposed for charged residues such as lysine and arginine in transmembrane peptides[11, 12]. The different KALP locations may be correlated with different snorkeling geometries in the membrane (Figure 4.16b). There could be additional membrane locations for KALP because the sum A(D8) + A(D10) = 0.65 is significantly smaller than A(D54) = 0.96. Although a fraction of KALP molecules trapped on the membrane surface may explain why A(D8) + A(D10) < 1, this possibility was ruled out since a complete (i.e. 100%) dephasing was observed for KALP in the D54 membrane (Figure 4.6).

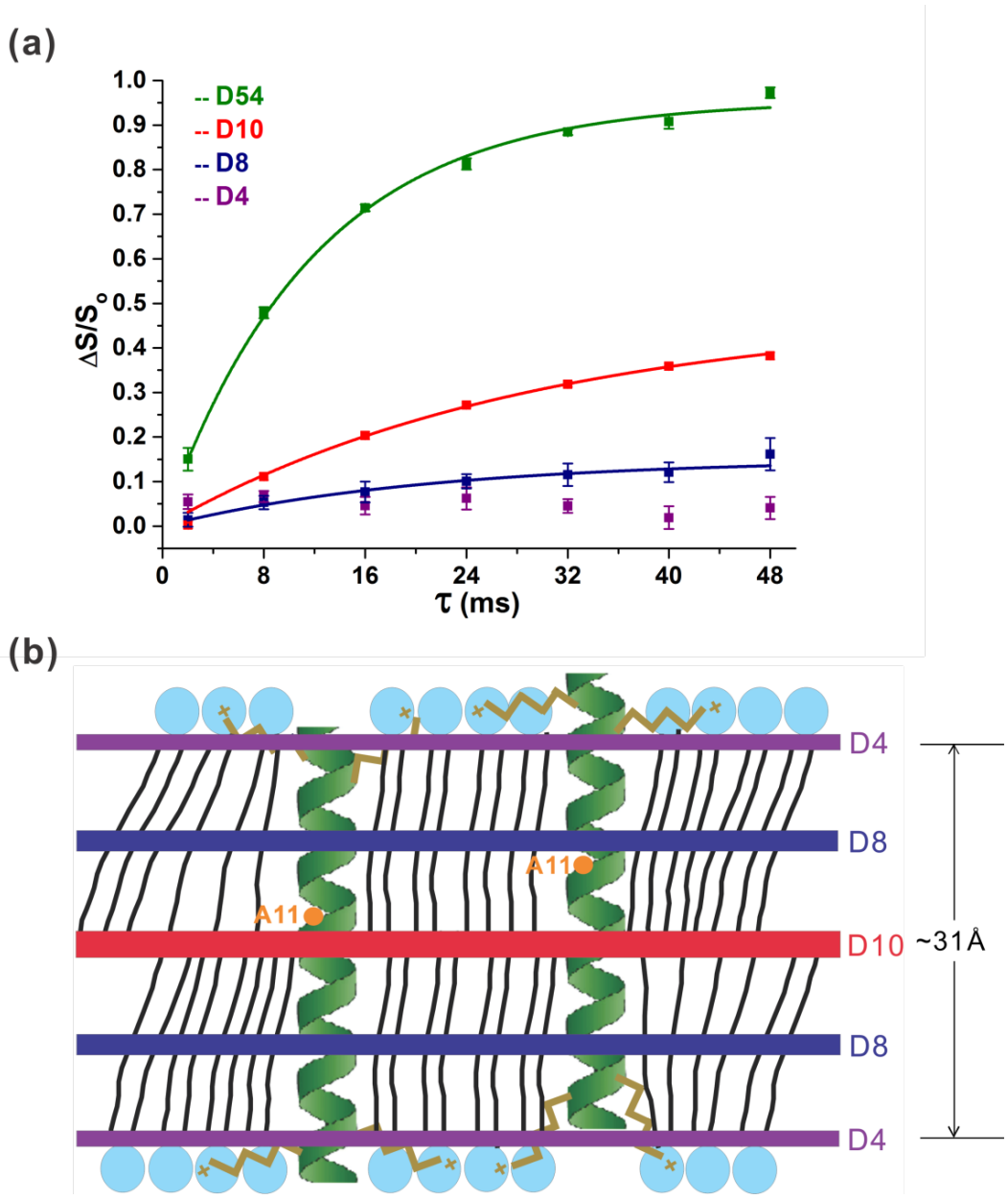

Figure 4.20 (a) Experimental (points with uncertainties) and best-fit (solid line) plots of $\Delta S/S_0$ vs τ for KALP_A11$_C$ in membranes. The D4 data were not fitted. (b) Membrane locations of KALP with major (left) and minor (right) populations. The colored bands, orange dots, and brown lines represent ^{2}H positions, A11 ^{13}CO nuclei, and lysine sidechains, respectively. The thickness of the palmitoyl region of the DPPC membrane is ~31 Å[5].

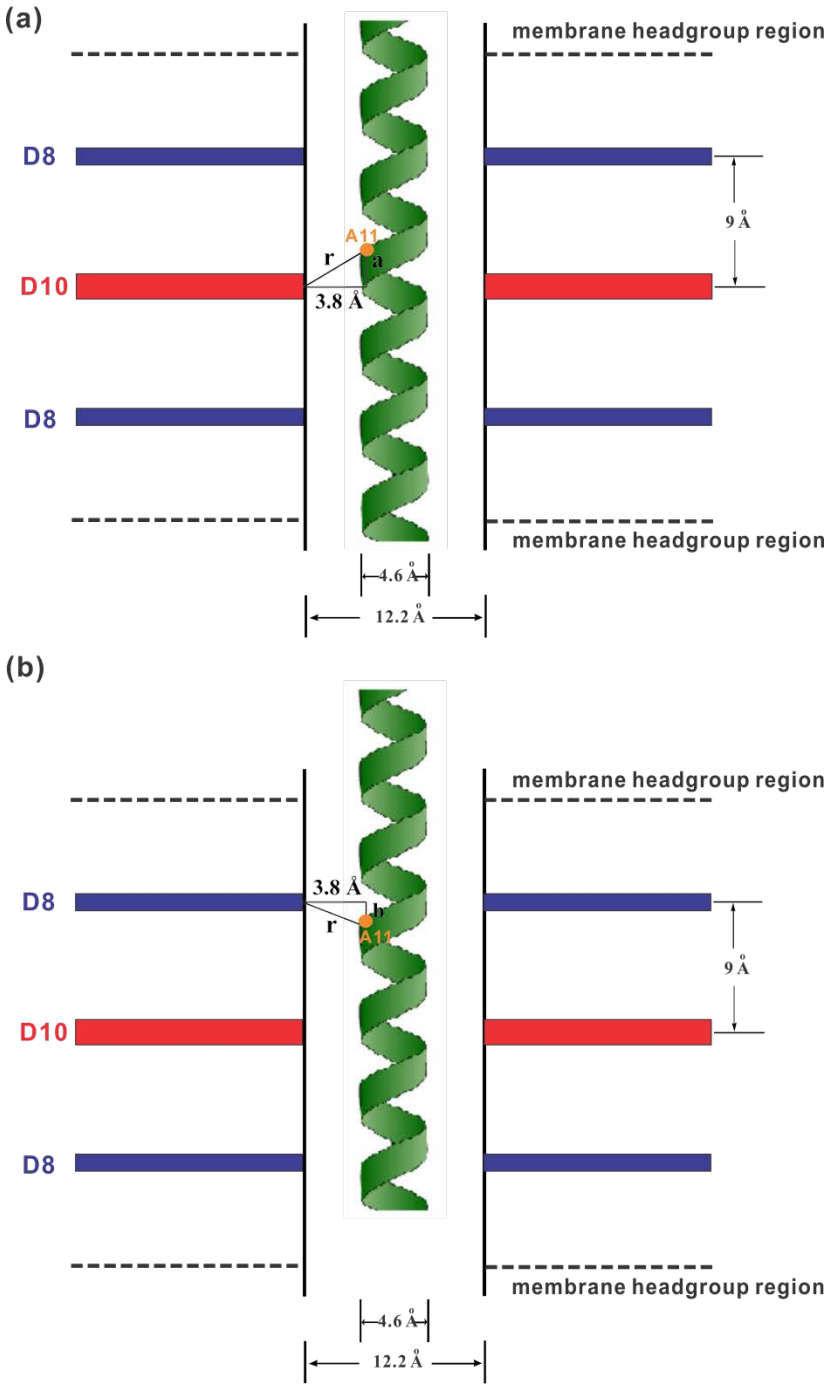

Figure 4.21 Models of α helical KALP in the DPPC membrane. The green ribbon is the KALP backbone and the vertical lines are the full van der Waals extent of the helix including leucine sidechains. The horizontal black dashed lines are the boundaries of the ~31 Å-thick palmitoyl region of the bilayer. The blue and red bands are the ^{2}H locations in the D8 and D10 membranes, respectively. Model (a) is the major KALP location with *a* = 2.4 Å calculated using the best-fit *r* (D10) = 4.5 Å. Model (b) is the minor KALP location with (9 Å – *b*) = 7.8 Å calculated using the best-fit *r* (D8) = 4.0 Å.

Membrane locations of HFP

The A(D54) = 0.99 and r(D54) = 2.9 Å for HFP_F8$_C$ support van der Waals contact between F8 and the DMPC membrane hydrocarbon core for all HFP molecules. Since the structure of DMPC is quite similar to that of DPPC except that DPPC has two more carbons in each acyl chain (Figure 4.2), F8 should also have van der Waals contact with the DPPC membrane hydrocarbon core. Similar contact is expected for G5 because A(D10) = 0.85 and A(D8) + A(D10) > 1. For either HFP_F8$_C$ or HFP_G5$_C$, there are very different values of A(D8) and A(D10), and for HFP_G5$_C$, r(D8) ≈ r(D10) ≈ 4.5 Å. These trends support at least two distinct membrane locations of HFP. The larger A(D10) values are attributed to a major HFP population with deep membrane insertion and HFP contact with D10 ^{2}H nuclei and the smaller A(D8) values are attributed to a minor population with shallow HFP insertion and HFP contact with D8 ^{2}H nuclei (Figure 4.22c, d). The major-to-minor population ratio is ~7:3 as calculated from the τ = 48 ms $(\Delta S/S_0)_{D10}$-to-$(\Delta S/S_0)_{D8}$ ratio for either the HFP_F8$_C$ or HFP_G5$_C$ samples. There is negligible HFP localized to the membrane surface, as evidenced by $(\Delta S/S_0) \approx 0.1$ without buildup for either HFP_F8$_C$ or HFP_G5$_C$ bound to the D4 membrane (Figure 4.8 and 4.10). The multiple membrane locations of HFP are attributed to the distribution of antiparallel β sheet registries[13]. Specifically, the membrane insertion depth of a HFP registry likely depends on the lengths of its contiguous hydrophobic regions and these lengths vary among registries. Deep and shallow insertions may also have a distribution of membrane locations of HFP. The predominant deep insertion of HFP

could significantly perturb the membrane bilayer and lower the activation energy of membrane fusion. This is consistent with the observed strong positive correlation between membrane insertion depth and fusion rate for several HFP constructs[1].

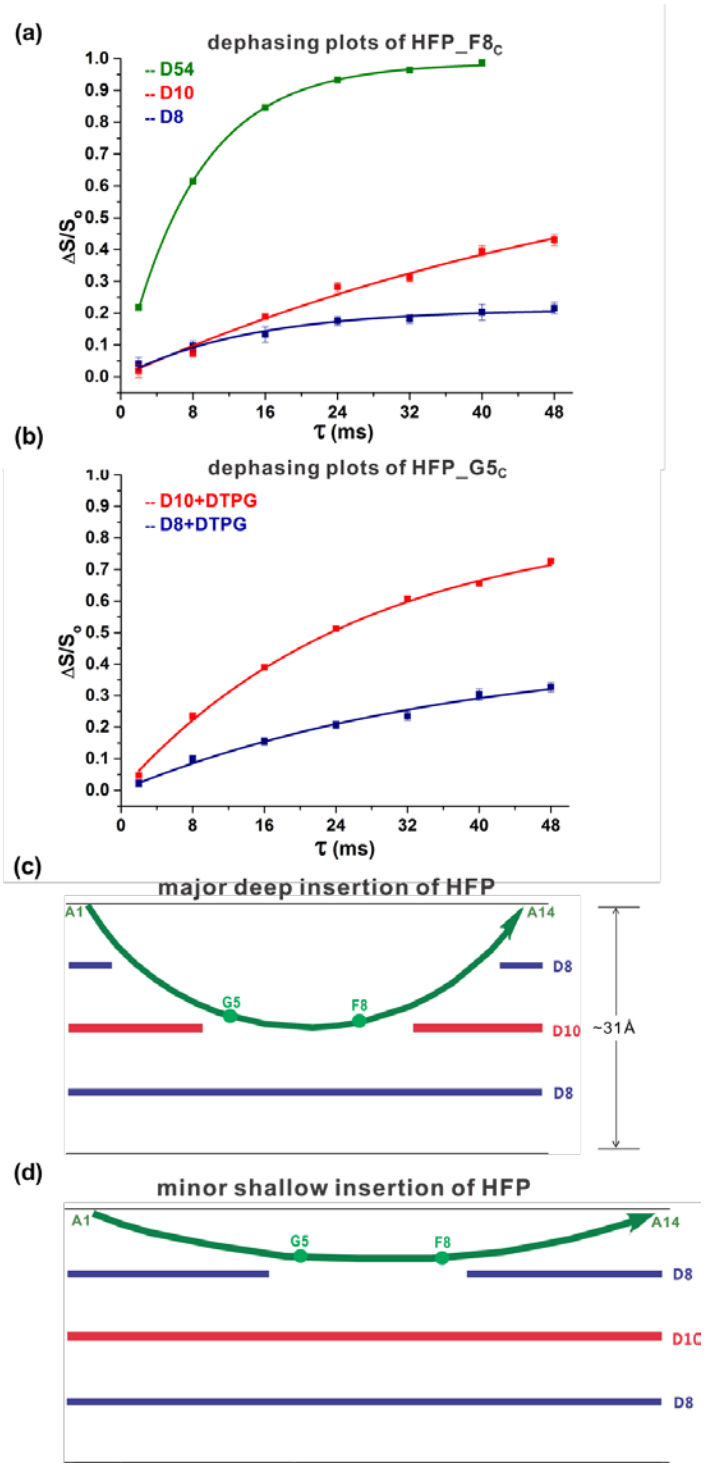

Figure 4.22 ^{13}C-^{2}H REDOR experimental (points with uncertainties) and best-fit (solid line) plots of $\Delta S/S_0$ vs τ for (a) HFP_F8$_C$ and (b) HFP_G5$_C$ in membranes. (c, d) Deep and shallow membrane insertion of HFP. Residues A1 and A14 are close to the membrane surface in concurrence with peptide ^{13}CO-lipid ^{31}P distances of ~5 Å for these residues[1]. For clarity, only one β strand is displayed.

It is noticeable that the best-fit *A* values are in an increasing order of *A*(D4) < *A*(D8) < *A*(D10) < *A*(D54) for both KALP and HFP, where the D4 data were not fitted but *A*(D4) is expected to be the smallest. One possible reason for the increasing order of *A* could be that there is an increasing order of mobility of ^{2}H from D4 to D54 which causes an increasing order of ^{2}H spin inversion efficiency from D4 to D54. In recent years ^{2}H composite π pulses have been used to compensate for pulse imperfections and incomplete spin inversion due to large ^{2}H quadrupolar resonance offsets. Sack and coworkers showed that ^{2}H composite π pulses introduced a two times faster ^{15}N-^{2}H REDOR dephasing buildup rate than simple π pulses did, where the ^{2}H quadrupolar splitting was ~146 kHz and B_1 field was ~40 kHz[14]. On the other hand, Cady and coworkers showed that the ^{2}H spin inversion efficiency in ^{13}C-^{2}H REDOR was only about 70% (i.e. maximum dephasing was ~70% x 2/3) when applying ^{2}H composite π pulse with a B_1 field of ~40 kHz[15]. In contrast to these studies, ^{2}H simple π pulses with a B_1 field of ~100 kHz were used in our ^{13}C-^{2}H REDOR experiments. Although ^{2}H pulse imperfections and incomplete spin inversion can affect the ^{13}C dipolar dephasing, the main reason for the increasing order of *A*(D4) < *A*(D8) < *A*(D10) < *A*(D54) for both KALP and HFP is the multiple membrane locations of the peptide rather than an increasing order of deuteron mobility from D4 to D54, as supported by the following experimental observations:

(I) Buildups of $\Delta S/S_0$ vs τ for other membrane peptide systems do not show an increasing order of *A*(D4) < *A*(D8) < *A*(D10). For example, IFP_L2$_C$ which has the sequence of GLFGAIAGFIENGWEGMIDGGGKKKK with a ^{13}CO label at residue L2 shows an order of *A*(D10) < *A*(D8) (Figure 4.23) whereas HFP_G5$_C$ with a point mutation of L9→R shows an order of *A*(D8) < *A*(D4) < *A*(D10) (Figure 4.24). The order of A for the L9R mutant of HFP_G5$_C$ supports at least two membrane locations of G5, one close to the membrane headgroup region (resulting in the D4 buildup) and the other close to the membrane center (resulting in the D10 buildup).

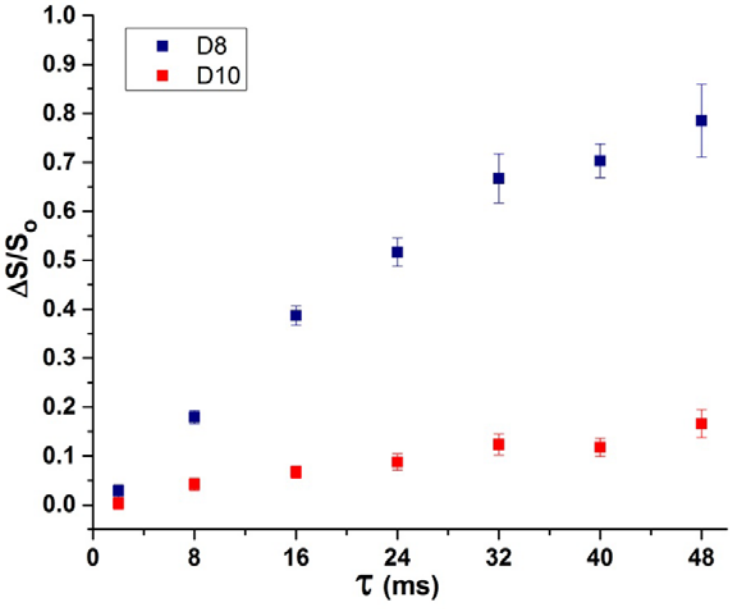

Figure 4.23 ^{13}C-^{2}H REDOR experimental plots $\Delta S/S_0$ vs τ for IFP_L2$_C$ in either D8 (blue) or D10 (red) membrane.

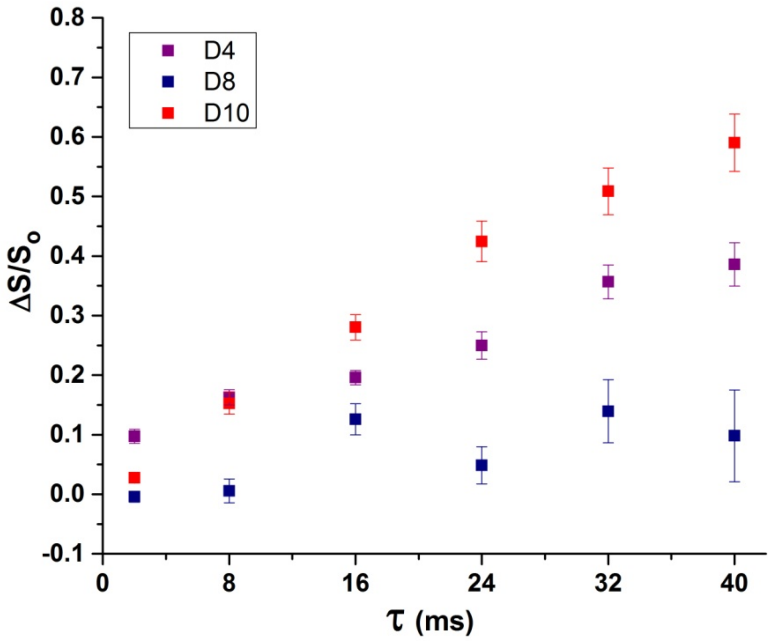

Figure 4.24 ^{13}C-^{2}H REDOR experimental plots $\Delta S/S_0$ vs τ for the L9R mutant of HFP_G5$_C$ in "D4+DPPG" (4:1, purple), "D8+DPPG" (4:1, blue), and "D10+DPPG" (4:1, red) membranes, respectively.

(II) There is a rapid buildup of $\Delta S/S_0$ vs τ for the lyophilized I4 peptide with long-time $\Delta S/S_0$ > 2/3 (Figure 3.13), indicating an efficient ^{2}H spin inversion caused by the ^{2}H simple π pulses with B_1 field of ~100 kHz in the presence of a large non-motionally averaged quadrupolar coupling from the rigid C-^{2}H bond in I4. Therefore, efficient ^{2}H spin inversions are also expected for the hydrated membrane samples.

(III) All samples were cooled by N_2 gas at -50 ^{o}C with actual sample temperature near -30 ^{o}C to minimize motion that would average the ^{13}CO-^{2}H dipolar coupling. At this temperature, the ^{2}H quadrupolar splittings were ~30 kHz for -CD_3 in D10 and D54

samples and ~120 kHz for -CD_2- in D4, D8, D10, and D54 samples (Figure 4.25), which indicated similar -CD_2- mobilities among these lipids. The lipid ^2H quadrupolar splittings were measured under static conditions by "quecho" experiments using the pulse sequence $(\pi/2)_x - \tau_1 - (\pi/2)_y - \tau_2$ – detect.

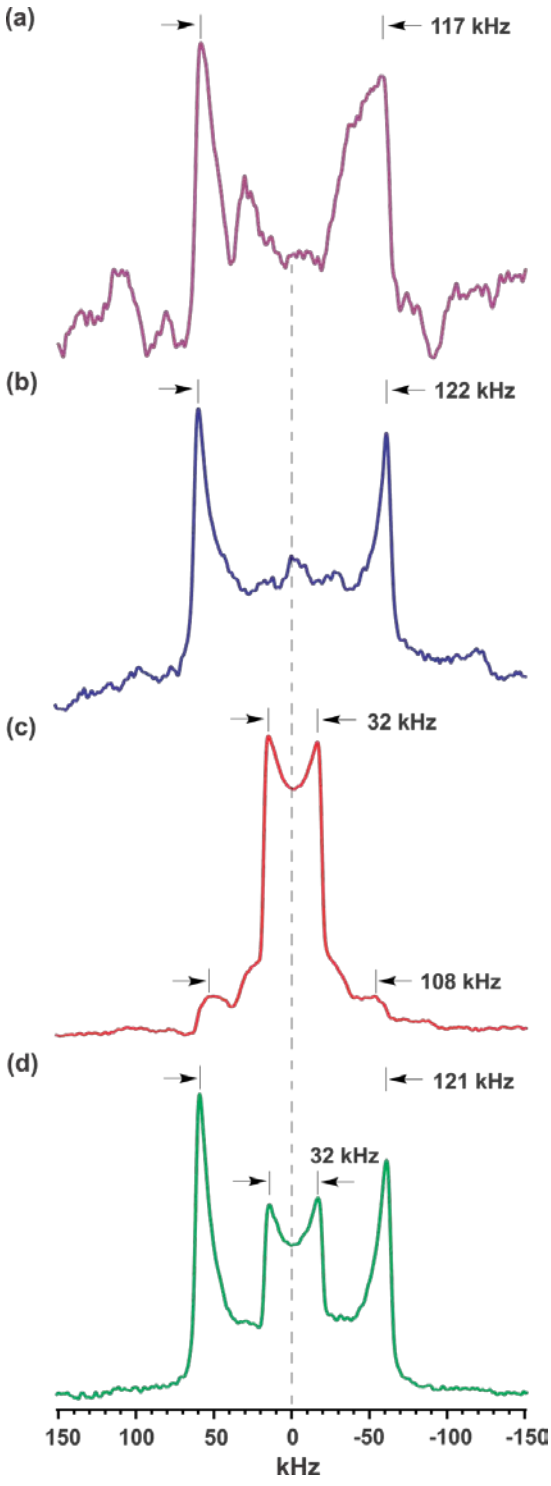

Figure 4.25 Lipid ^{2}H Pake doublets acquired by "quecho" experiments under static conditions for (a) KALP_A11$_C$ in D4 and (b-d) HFP_F8$_C$ in D8 (blue), D10 (red), and D54 (green) membranes, respectively. The number of acquisitions was 14482 for (a), 8000 for (b), 13591 for (c), and 3000 for (d), respectively. All spectra were processed using 2000 Hz Gaussian line broadening and baseline correction.

(IV) Both D10 and D54 samples contain all the -CD_3 groups. However, D10 contains only -CD_2- at C15 whereas D54 contains all the -CD_2- (Figure 4.2). The dephasing buildup is much faster in D54 than that in D10 for both KALP_A11$_C$ and HFP_F8$_C$ (Figures 4.6 and 4.8). These results indicate that the buildup rate for ^{13}CO-CD_2 proximity is comparable to that for ^{13}CO-CD_3 with a similar proximity even if -CD_2- is much more rigid than -CD_3.

Conclusions

In summary, ^{13}CO-^{2}H REDOR SSNMR reveals multiple membrane locations for both the α helical KALP and β sheet HFP peptides. The KALP locations are attributed to the hydrophobic mismatch and consequent snorkeling of lysine sidechains to the membrane headgroup region. The HFP locations are attributed to the distribution of antiparallel β sheet registries. ^{13}C-^{2}H REDOR is expected to be a general method for determining the residue-specific distribution of membrane locations of peptides and proteins and for developing membrane topology models.

REFERENCES

REFERENCES

1. Qiang, W., Y. Sun, and D.P. Weliky, *A strong correlation between fusogenicity and membrane insertion depth of the HIV fusion peptide.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(36): p. 15314-15319.

2. Yang, R., et al., *A trimeric HIV-1 fusion peptide construct which does not self-associate in aqueous solution end which Has 15-fold higher membrane fusion rate.* Journal of the American Chemical Society, 2004. **126**(45): p. 14722-14723.

3. Gabrys, C.M., et al., *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.* Journal of Physical Chemistry A, 2013. **117**(39): p. 9848-9859.

4. Hirsh, D.J., et al., *A new monofluorinated phosphatidylcholine forms interdigitated bilayers.* Biophysical Journal, 1998. **75**(4): p. 1858-1868.

5. Xie, L., et al., *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.* Journal of Biomolecular NMR, 2013. **55**(1): p. 11-17.

6. de Planque, M.R.R., et al., *Sensitivity of single membrane-spanning alpha-helical peptides to hydrophobic mismatch with a lipid bilayer: Effects on backbone structure, orientation, and extent of membrane incorporation.* Biochemistry, 2001. **40**(16): p. 5000-5010.

7. Brugger, B., et al., *The HIV lipidome: A raft with an unusual composition.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(8): p. 2641-2646.

8. Zhang, H.Y., S. Neal, and D.S. Wishart, *RefDB: A database of uniformly referenced protein chemical shifts.* Journal of Biomolecular NMR, 2003. **25**(3): p. 173-195.

9. Kandasamy, S.K. and R.G. Larson, *Molecular dynamics simulations of model trans-membrane peptides in lipid bilayers: A systematic investigation of hydrophobic mismatch.* Biophys. J., 2006. **90**(7): p. 2326-2343.

10. Im, W. and C.L. Brooks, *Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(19): p. 6771-6776.

11. Strandberg, E., et al., *Lipid dependence of membrane anchoring properties and snorkeling behavior of aromatic and charged residues in transmembrane peptides.* Biochemistry, 2002. **41**(23): p. 7190-7198.

12. Segrest, J.P., et al., *Amphipathic Helix Motif - Classes and Properties.* Proteins-Structure Function and Genetics, 1990. **8**(2): p. 103-117.

13. Schmick, S.D. and D.P. Weliky, *Major Antiparallel and Minor Parallel beta Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide.* Biochemistry, 2010. **49**(50): p. 10623-10635.

14. Sack, I., et al., *Deuterium REDOR: Principles and applications for distance measurements.* Journal of Magnetic Resonance, 1999. **138**(1): p. 54-65.

15. Cady, S.D., et al., *Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers.* Nature, 2010. **463**(7281): p. 689-U127.

Chapter 5 - Summary and future work

During the past four years, I have been working on two projects, with one focused on structural characterization of serotype H3 IFP (H3_20, see chapter 2 for the sequence and other details) in membranes and the other focused on membrane location detection of peptides including KALP and HFP.

Summary of experimental results and future work of IFP

For the IFP work, ^{13}C-^{15}N REDOR was used to quantitatively measure the interhelical proximity of H3_20 in membranes at the fusogenic pH 5. SIMPSON simulations of the experimental $\Delta S/S_0$ vs τ after n.a. calibration have show that H3_20 adopts ~40% closed structure and ~60% semi-closed structure in membranes at pH 5.1[1]. The results are contradictory to the previously reported open structure of H3_20 in membranes[2]. For future work, it will be interesting to measure the interhelical proximity of H3_20 in membranes at the physiological, non-fusogenic pH 7.4. We can correlate the structure with fusogenicity by comparing the structures of H3_20 at pH 5.1 and 7.4. Another interesting study is the interhelical proximity measurement of H3_23 in membranes at pH 5.1 and 7.4. Earlier liquid-state NMR studies of H1_20 and H1_23 in detergents have shown that H1_20 adopts predominantly open structure whereas H1_23 adopts predominantly closed structure, indicating that the residues 21 through 23 (i.e. WYG) play a key role in stabilizing the closed structure of the peptide[3]. We can determine whether the residues WYG also cause a greater population of closed structure of IFP in membranes by comparing the structures of H3_20 with those of H3_23 at either pH 5.1 or pH 7.4. Furthermore, lipid mixing assays using fluorescence

spectroscopy may be performed to investigate the fusogenicities of H3_20 and H3_23 at both pH 5.1 and 7.4. The impact of peptide length and structure of IFP on its membrane fusogenicity may be elucidated based on the results of REDOR proximity measurements and lipid mixing assays of H3_20 and H3_23.

In addition to the structural and functional studies of IFP, another interesting work could be the insertion depth and tilt angle measurements of IFP in membranes. The residue-specific insertion depth of IFP can be probed by ^{13}C-^{2}H REDOR using the D10 lipid and IFP which is ^{13}CO labeled at a specific residue (e.g. Leu 2, Ala 5). After the insertion depths of IFP at different residues have been measured, the tilt angle of IFP with respect to the membrane surface can be calculated using the membrane insertion model of IFP in Figure 5.1.

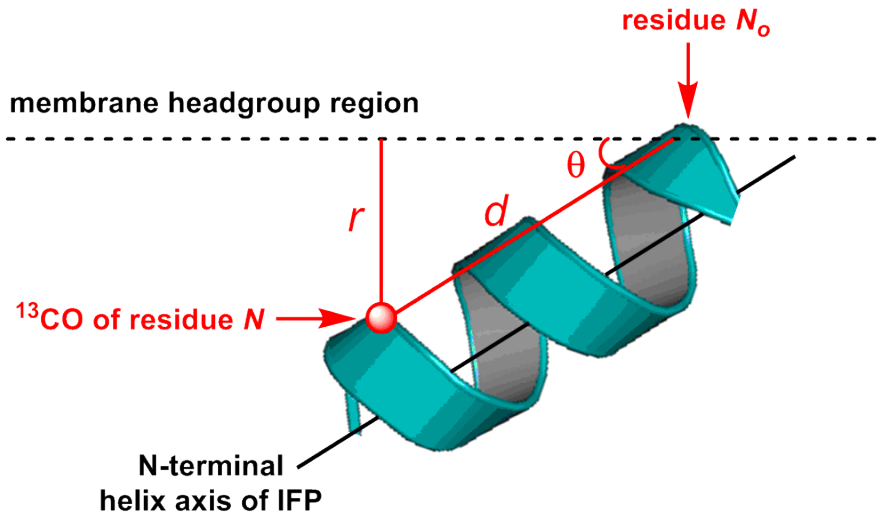

Figure 5.1 Membrane insertion model of the N-terminal helix of IFP. The ^{13}CO labeled residue is numbered *N* and the residue in the membrane headgroup region is numbered N_o. θ is the tilt angle of the N-terminal helix of IFP to the membrane surface, *r* is the distance between the labeled ^{13}CO nucleus and the lipid headgroup, and *d* is the distance from residue *N* to residue N_o along the N-terminal helix axis.

In Figure 5.1, the distance *d* in Å between residue *N* and N_o along the N-terminal helix axis is given by

$$d\ (\text{Å}) = 1.5\text{x}(N_o - N) \qquad (5.1)$$

where 1.5 Å refers to the rise per residue in an α helix. The distance *r* between the labeled ^{13}CO nucleus of residue *N* and the lipid headgroup can be estimated by

$$r\ (\text{Å}) = d\sin\theta = 1.5\text{x}(N_o - N)\sin\theta \qquad (5.2)$$

where θ is the tilt angle of the N-terminal helix of IFP to the membrane surface. However, equation 5.2 needs to be modified since the distance *r* is not linearly proportional to the residue position number *N*. For the N-terminal helix, *r* is oscillating as a function of the residue position number *N* in a periodicity of 3.6 since there are 3.6 residues per turn in an α helix. Therefore, an oscillating term, $A\sin[2\pi(N_o - N)/3.6]$, needs to be introduced to account for the periodicity of *r* and equation 5.2 is modified as

$$r\ (\text{Å}) = d\sin\theta + A\sin[2\pi(N_o - N)/3.6] = 1.5\text{x}(N_o - N)\sin\theta + A\sin[2\pi(N_o - N)/3.6] \qquad (5.3)$$

where *r* can be experimentally derived by ^{13}C-^{2}H REDOR using the D10 lipid. For instance, if the D10 data show that the labeled ^{13}CO nucleus is 5 Å from the membrane bilayer center, the distance *r* will be ~11 Å since the hydrophobic thickness of the DPPC bilayer is ~31 Å[4]. The residue N_o is the one that produces the most rapid ^{13}C-^{2}H REDOR dephasing buildup in the D4 membrane. To obtain the tilt angle θ, the

experimentally derived distance *r* can be fitted as a function of the position number *N* of the ^{13}CO labeled residue using equation 5.3, where θ and A are both fitting parameters.

Summary of experimental results of KALP and HFP and future work of HFP

For the membrane location studies of KALP and HFP, each peptide was ^{13}CO backbone labeled and the lipids were either perdeuterated (D54) or selectively deuterated (D4, D8, and D10) in their acyl chains. ^{13}C-^{2}H REDOR was used to detect the residue-specific membrane locations of KALP and HFP. The data fitting of experimental $\Delta S/S_0$ vs τ showed a membrane location distribution for both KALP and HFP. The multiple locations of KALP are attributed to the hydrophobic mismatch between the membrane and KALP and consequent flexible snorkeling of lysine sidechains to the membrane headgroup region[5]. The multiple locations of HFP are attributed to the distribution of antiparallel β sheet registries in membranes[6]. To our knowledge, this is the first clear experimental support for a distribution of membrane locations of a peptide in the membrane hydrocarbon core. For future work, we may detect the residue-specific membrane locations of HFP in a larger construct (e.g. HFP-Hairpin including the NHR-loop-CHR region) to see whether there is also a membrane location distribution, if so, what are the populations of deep and shallow insertion compared to those for HFP?

In addition to the deuterated lipids, deuterated cholesterol can be also used to study the membrane locations of peptides. Similar to the D10 lipid, cholesterol deuterated at the chain end (cholesterol-d_6, Figure 5.2) can be used to detect the

peptide location with respect to the membrane center. Furthermore, HFP binding preference to the lipid or cholesterol may be determined by comparison of ^{13}C-^{2}H REDOR data for the D10 and cholesterol-d_6 samples. For instance, we can prepare two membrane samples of HFP_$G5_C$, *a* and *b*. For sample *a*, the membrane consists of 30% D10 lipid, 40% undeuterated DPPC lipid, and 30% undeuterated cholesterol. For sample *b*, the membrane consists of 70% undeuterated DPPC lipid and

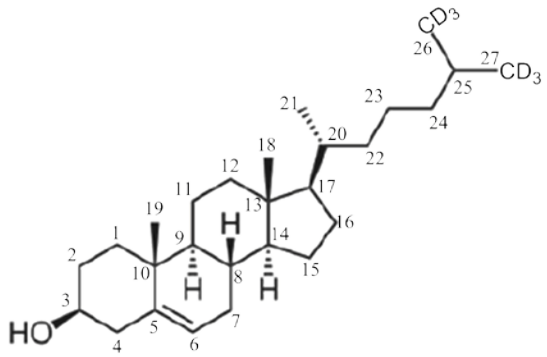

Figure 5.2 Structure of cholesterol-26,26,26,27,27,27-d_6.

30% cholesterol-d_6. After obtaining the experimental ^{13}C-^{2}H REDOR $\Delta S/S_0$ buildups of HFP_$G5_C$ in sample *a* and *b*, we can fit the plots of $\Delta S/S_0$ vs τ by the function Ax$(1-e^{-\beta\tau})$ (refer to section 4.4 in Chapter 4 for more details) and compare the best-fit values of *A* for sample *a* and *b*. If the *A* value for sample *a* is comparable to that for sample *b* (e.g. 65% vs 70%), it supports that HFP has no binding preference to either lipids or cholesterol. If the *A* value for sample *a* is significantly smaller (or larger) than that for sample *b* (e.g. 65% vs 85%), it supports that HFP has binding preference to cholesterol

(or lipids). To further elucidate the HFP binding preference, we may also prepare four other membrane samples of HFP_G5$_C$ by using 20% D10, 20% cholesterol-d$_6$, 10% D10, or 10% cholesterol-d$_6$, respectively, and compare their *A* values.

REFERENCES

REFERENCES

1. Ghosh, U., L. Xie, and D.P. Weliky, *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.* Journal of Biomolecular NMR, 2013. **55**(2): p. 139-146.

2. Han, X., et al., *Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin.* Nature Structural Biology, 2001. **8**(8): p. 715-720.

3. Lorieau, J.L., J.M. Louis, and A. Bax, *Impact of Influenza Hemagglutinin Fusion Peptide Length and Viral Subtype on Its Structure and Dynamics.* Biopolymers, 2013. **99**(3): p. 189-195.

4. Xie, L., et al., *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.* Journal of Biomolecular NMR, 2013. **55**(1): p. 11-17.

5. Kandasamy, S.K. and R.G. Larson, *Molecular dynamics simulations of model trans-membrane peptides in lipid bilayers: A systematic investigation of hydrophobic mismatch.* Biophysical Journal, 2006. **90**(7): p. 2326-2343.

6. Schmick, S.D. and D.P. Weliky, *Major Antiparallel and Minor Parallel beta Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide.* Biochemistry, 2010. **49**(50): p. 10623-10635.

APPENDICES

APPENDIX A

NMR file locations

Figure 2.6

/export/home/hapi0/mb4c/data/Ujjayini/IFP_061312

Figure 2.9

/export/home/khafre0/mb4b/data/Ujjayini/Data_112112/IFP_020413_4

Figure 3.4

/export/home/hapi0/mb4c/data/Li/13C2H/setup_121912/MAS_KBr_122012

Figure 3.5

/export/home/hapi0/mb4c/data/Li/13C2H/setup_121912/adamantane_122012

Figure 3.6

/export/home/hapi0/mb4c/data/Li/13C2H/setup_121912/cpr_aHrf_122012

Figure 3.7

/export/home/hapi0/mb4c/data/Li/13C2H/setup_121912/i4_optimizedCP_122012

Figure 3.9

/export/home/hapi0/mb4c/data/Li/13C2H/i4/i4_071713

Figure 3.10

/export/home/hapi0/xieli4c/data/Li/REDOR/13C2H/i4_031112 (for "redorxy8xypi_pm" data)
/export/home/hapi0/xieli4c/data/Li/REDOR/13C2H/i4_redorxy8ypi_031912 (for "redorxy8ypi_pm" data)

Figure 3.11

/export/home/hapi0/xieli4c/data/Li/REDOR/13C2H/i4_031112 (for "redorxy8xypi_pm" spectra)
/export/home/hapi0/xieli4c/data/Li/REDOR/13C2H/i4_redorxy4xpi_031912 (for "redorxy4xpi_pm" spectra)

Figure 4.4

/export/home/hapi0/mb4c/data/Li/13C2H/KALP/KALP_092412 (KALP_A11$_C$ in D4)

/export/home/hapi0/mb4c/data/Li/13C2H/KALP/KALP_032713 (KALP_A11$_C$ in D8)

/export/home/hapi0/mb4c/data/Li/13C2H/KALP/KALP_122112 (KALP_A11$_C$ in D10)

/export/home/hapi0/mb4c/data/Li/13C2H/KALP/KALP_012213 (KALP_A11$_C$ in D54)

Figure 4.5

/export/home/hapi0/mb4c/data/Li/13C2H/KALP/KALP_092412 (KALP in D4)

Figure 4.6

the same file locations as those for Figure 4.4.

Figure 4.7

/export/home/hapi0/mb4c/data/Li/13C2H/HFP/wt_F8C_120712 (HFP_F8$_C$ in D4)

/export/home/hapi0/mb4c/data/Li/13C2H/HFP/wt_F8C_112912 (HFP_F8$_C$ in D8)

/export/home/hapi0/mb4c/data/Li/13C2H/HFP/wt_F8C_110112 (HFP_F8$_C$ in D10)

/export/home/hapi0/mb4c/data/Li/13C2H/HFP/wt_F8C_112512 (HFP_F8$_C$ in D54)

Figure 4.8

the same file locations as those for Figure 4.7.

Figure 4.9

/export/home/hapi0/mb4c/data/Lihui/13C2H/HFP/HFP_050813 (HFP_G5$_C$ in D4)

/export/home/hapi0/mb4c/data/Lihui/13C2H/HFP/HFP_090313 (HFP_G5$_C$ in 80% D8 and 20% DTPG)

/export/home/hapi0/mb4c/data/Lihui/13C2H/HFP/HFP_053113 (HFP_G5$_C$ in 80% D10 and 20% DTPG)

Figure 4.10

the same file locations as those for Figure 4.9.

Figure 4.12

/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D8_101713/D8_array_110613

Figure 4.13

/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D10_101713/D10_tau0.1_101713
/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D10_101713/D10_tau200_101713
/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D10_101713/D10_array_101713

Figure 4.14

/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D54_102113/D54_tau0.1_102113
/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D54_102113/D54_array_102113

Figure 4.15

/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D8_101713/D8_array_MAS_103013

Figure 4.16

/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D10_101713/D10_array_MAS_102813

Figure 4.17

/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D54_102113/D54_array_MAS_102513

Figure 4.23

/export/home/hapi0/mb4c/data/Shuang/13C2H/IFP/IFP_091113 (IFP in D8)
/export/home/hapi0/mb4c/data/Shuang/13C2H/IFP/IFP_040213 (IFP in D10)

Figure 4.24

/export/home/hapi0/mb4c/data/Lihui/13C2H/HFP/HFP_010214 (L9R mutant of HFP_$G5_C$ in 80% D4 and 20% DPPG)

/export/home/hapi0/mb4c/data/Lihui/13C2H/HFP/HFP_011414 (L9R mutant of HFP_G5$_C$ in 80% D8 and 20% DPPG)

/export/home/hapi0/mb4c/data/Lihui/13C2H/HFP/HFP_012714 (L9R mutant of HFP_G5$_C$ in 80% D10 and 20% DPPG)

Figure 4.25

/export/home/hapi0/mb4c/data/Li/quecho/DPPC/D4_112112 (KALP in D4)

/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D8_101713/D8_quecho_102013 (HFP in D8)

/export/home/hapi0/mb4c/data/Li/quecho/DPPC/D10_112112 (HFP in D10)

/export/home/hapi0/mb4c/data/Li/t1D_ir/Lipids/D54_102113/D54_quecho_102513 (HFP in D54)

APPENDIX B

Fmoc solid phase peptide synthesis (SPPS)

Solid phase peptide synthesis (SPPS) is a peptide synthesis technique first developed by Robert Merrifield[1]. There are two commonly used SPPS protocols, Fmoc and t-Boc SPPS. Compared to t-Boc SPPS which uses TFA as the deprotecting reagent, Fmoc SPPS is milder and safer since it uses piperidine as the deprotecting reagent.

For the Fmoc SPPS protocol, each time we may start the peptide synthesis using either Fmoc (9-fluorenylmethyloxycarbonyl) Wang resin which produces peptides with free C-terminal carboxylic group or rink amide resin which produces peptides with amidated C-terminus. Take the synthesis of 0.2 mmol HFP which has a sequence of AVGIGALFLGFLGAAGSTMGARSWKKKKKKA as an example, if we start with Fmoc-Ala-resin that has a loading capacity of 0.5 mmol/g, we will need to weigh out 0.4 g resin. The Fmoc SPPS procedure for HFP synthesis is described as follows:

(1) Swell the 0.4 g Fmoc-Ala-resin in dichloromethane (DCM) in a reaction vessel for 2 hours on a vortex mixer (Vortex-Genie 2), drain the DCM, swell the resin in DCM for another 2 hours, drain the DCM, and then add N,N-dimethylformamide (DMF) to rinse the resin twice.

(2) Add ~10 mL deprotecting solution to remove the Fmoc protecting group from the residue for 5 minutes, drain the solution, add DMF to rinse the resin once, add another ~10 mL deprotecting solution to remove residual Fmoc for 20 minutes, and then drain the solution. The deprotecting solution is prepared using piperidine and DMF in a volume ratio of 1:4. For instance, a 500 mL deprotecting solution contains 100 mL piperidine and 400 mL DMF.

(3) Rinse the resin using DCM once and then using DMF twice. After that, add ~8 mL freshly-prepared coupling solution which contains 2 mmol Fmoc-Lys(Boc)-OH, where the lysine sidechain is protected by the tert-butyloxycarbonyl (t-Boc) protecting group. To increase the peptide yield, the amount of Fmoc-amino acid added is 10 times the stoichiometry. For isotopically labeled Fmoc-amino acid (e.g. ^{13}CO labeled), the amount added is 5 times the stoichiometry. The coupling solution is prepared using the following materials:

(a) 2 mmol Fmoc-amino acid;

(b) 2 mmol 1-hydroxybenzotriazole (HOBt), the mass is 270 mg;

(c) 2 mmol *2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU)*, the mass is 759 mg;

(d) 700 µL N,N-diisopropylethylamine (DIEA);

(e) add DMF until the total volume is 8~10 mL.

(4) Allow at least 4 hours for the coupling reaction. In general, the minimum time required for coupling reaction is 2 hours for every Fmoc-amino acid except for Cys, Ser, Thr, Met, Trp, and isotopically labeled (e.g. ^{13}CO or ^{15}N labeled) amino acids which require a minimum coupling time of 4 hours. For the first coupling reaction, no matter what Fmoc-amino acid is added, allow at least 4 hours to increase the yield. After the coupling reaction, drain the coupling solution, rinse the resin using DCM once and then using DMF twice, finally add ~10 mL capping solution to cap the N-terminus of any unreacted peptide to which the added Fmoc-amino acid is not coupled. The capping

reaction prevents the unreacted peptide from reacting in the next coupling reaction. For 200 mL capping solution, it consists of

(a) 9.5 mL acetic anhydride;

(b) 4.5 mL DIEA;

(c) 0.4 g HOBt;

(d) add DMF until the total volume is 200 mL.

(5) Allow 5~10 minutes for the capping reaction, drain the capping solution, and then rinse the resin using DCM once and DMF twice.

(6) Add ~10 mL deprotecting solution and vortex for 5 minutes, drain the solution, rinse the resin with DMF once and then add another ~10 mL deprotecting solution and vortex for 20 minutes. In the meantime, prepare the coupling solution.

(7) Drain the deprotecting solution, rinse the resin using DCM once and then using DMF twice. Add ~8 mL freshly-prepared coupling solution which contains 2 mmol of the second residue to be coupled, i.e. Fmo-Lys(Boc)-OH. The coupling solution should be prepared in the same way as described in step (3).

(8) Repeat steps 4-7 until we get the whole sequence. For the last residue, after the coupling reaction, perform capping reaction once and then deprotection twice. After that, perform capping reaction again if the N-terminus of the peptide needs to be acetylated, otherwise, proceed to rinse the resin using DCM 3~4 times, transfer the resin from the reaction vessel to the cleavage tube (more than one tube may be needed, depending on the tube capacity). Use DCM to wash all resin down to the cleavage tube from the reaction vessel. After that, use gentle N_2 gas to remove DCM from the cleavage tube.

Be careful not to blow away the resin. Keep the resin in the cleavage tube and store the tube in a vacuum desiccator overnight to further remove any residual DCM and/or other organic solvents including DMF.

(9) Add 12 mL cleavage solution to the resin in the cleavage tube(s) (if the resin is split into 2 or 3 tubes, then add 6 or 4 mL to each) and gently vortex for ~4 hours. If the cleavage time is too short, incomplete cleavage may result. However, if the cleavage time is too long, oxidation or other side reactions of the deprotected sidechains may result. To cleave 0.2 mmol peptide from the resin, 12 mL cleavage solution needs to be prepared using the following materials:

(a) 600 µL distilled water;

(b) 240 µL triisopropylsilane(TIS);

(c) 240 µL thioanisole;

(d) 240 µL 1,2-ethanedithol;

(e) add trifluoroacetic acid (TFA) until the total volume is 12 mL.

TFA cleaves the peptide from the resin and also removes all sidechain protecting groups from the peptide. For components a-d, they are called “scavengers” and are used to protect the peptide from oxidation and other side reactions after all sidechain protecting groups of the peptide have been removed by TFA. The cleavage solution is sometimes also called "cocktail solution".

(10) After the cleavage reaction, drain the 12 mL cleavage solution to three 50 mL conical vials (4 mL for each), add 1~2 mL TFA to wash the resin and then drain the TFA to the vials.

(11) Add ~40 mL cold diethyl ether (4 ^{o}C) to each vial to extract the peptide from the cleavage solution. A 10:1 volume ratio of cold diethyl ether to cleavage solution is used to maximize the peptide extraction. After adding the ether, we will observe white precipitate immediately, i.e. the peptide precipitates out due to a poor solubility in cold diethyl ether. During the extraction, do not shake the peptide-ether mixture in case of peptide oxidation (e.g. methionine may get oxidized by O_2 or oxidized ether). After the extraction, centrifuge the peptide-ether mixture at a speed of 4000 rpm (revolutions per minute) or above for ~5 minutes. After that, dump the supernatant, add ~10 mL fresh cold diethyl ether to the peptide pellet, perform ~2 min sonication to help dissolve the ether-soluble impurities, centrifuge for another 5 minutes, dump the supernatant, and finally put the vials in the hood to let the peptide pellet air dry.

(12) Add 10~20 mL distilled water containing 0.1% (by volume) TFA to each vial to dissolve the peptide pellet. Again, do not shake the vial in case of peptide oxidation. It may take a few hours to dissolve the peptide pellet. If the pellet does not get dissolved after several hours or even a day, add a few mL acetonitrile containing 0.1% TFA to help dissolve the peptide.

(13) After the peptide has been completely dissolved, perform reversed-phase HPLC using a C4 or C18 column to purify the peptide.

(14) At last, use MALDI or ESI mass spectrometry to identify the peptide purity and molecular weight.

The following is a list of the mass of 2 mmol Fmoc-amino acid (the mass takes into account the sidechain protecting group if applicable). The Fmoc-amino acids are commercially available from either “Peptides International” or “Novabiochem” company.

Lys (K) : 937.2 mg

Trp (W) : 1053.2 mg

Ser (S) : 794.8 mg

Arg (R) : 1297.6 mg

Ala (A) : 658.8 mg (monohydrate, i.e. Fmoc-Ala-OH·H_2O)

Gly (G) : 594.8 mg

Met (M) : 742.8 mg

Thr (T) : 1139.3 mg

Leu (L) : 706.8 mg

Ile (I) : 706.8 mg

Val (V) : 678.8 mg

Phe (F) : 774.9 mg

Glu (E) : 851.0 mg

APPENDIX C

HPLC purification of peptides

High-performance liquid chromatography (HPLC, sometimes referred to as high-pressure liquid chromatography) is a technique used to separate different components in a mixture and quantify each component. There are two types of HPLC. One is referred to as normal phase HPLC for which the stationary phase (e.g. unmodified silica resins) is polar and the mobile phase is non-polar. The other is referred to as reversed-phase HPLC for which the stationary phase is non-polar and the mobile phase is polar. Examples of non-polar stationary phase include C4 and C18 columns where alkyl chains containing four or eighteen carbons are covalently bonded to the resins to create a hydrophobic stationary phase. For reversed-phase HPLC, hydrophobic molecules in the mobile phase tend to bind to the column whereas hydrophilic molecules do not bind to the column and are eluted first.

In our work, the peptides synthesized by Fmoc SPPS were purified by reversed-phase HPLC using a C4 column. The C4 column is filled with resins (5 μm in diameter) to which butyl groups are covalently bonded. The inner diameter of the column is 19 mm and the length is 300 mm. To purify the peptides, two solvents, A and B, are used as the mobile phase. Solvent A is degased distilled water containing 0.1% TFA and solvent B is the mixture of 10% degased distilled water and 90% acetonitrile containing 0.1% TFA. Degased distilled water is used to avoid air bubbles in the pump system. 0.1% TFA is added to each solvent to increase the peptide solubility in the mobile phase. Different peptide samples may require different HPLC gradient programs for purification. An ideal HPLC gradient program should have the following criteria:

(a) The product peak should be well resolved from other peaks. If there is severe peak overlapping, a slower gradient buildup rate can be used to solve this issue.

(b) The product peak should not be too far from the nearby peaks. For instance, if the product peak is 10 mins apart from its nearest peak, it will be a waste of time to wait such a long period. If this issue appears, a faster gradient buildup rate can be used to solve it.

(c) After the product peak ends, the gradient can be increased to the maximum % B (e.g. 75% or 100% B) to elute all residual components from the column. It is unnecessary to wait a long period (e.g. 10 mins) following the product peak to increase the gradient to the maximum % B.

(d) After the gradient goes back to the initial % B from the maximum % B, allow 3~5 mins to equilibrate the column.

Take HFP peptide purification using a C4 column for example, the time sequence of the HPLC gradient program used was as follows:

t = 0.0 min, % B = 37.0%;

t = 2.0 min, % B = 37.0%;

t = 6.0 min, % B = 38.0%;

t = 11.0 min, % B = 38.0%;

t = 20.0 min, % B = 42.5%;

t = 20.5 min, % B = 75.0%;

t = 29.5 min, % B = 75.0%;

t = 30.0 min, % B = 37.0%;

t = 33.0 min, % B = 37.0%.

where *t* was the elution time (sometimes also referred to as retention time) and the flow rate was 8.0 mL/min. This gradient program is displayed in Figure C1 (blue dotted line). The HPLC chromatogram of "Absorbance at 214 nm vs Retention time" in Figure C1 (black solid line) was obtained from the purification of HFP which had the sequence of AVGIGALFLGFLGAAGSTMGARSWKKKKKKA. The UV detection wavelength was chosen as 214 nm since the peptide amide bonds had strong absorbance at this wavelength. The wavelength of 280 nm may be also used since tryptophan has strong absorbance at 280 nm and the corresponding molar absorptivity (also referred to as extinction coefficient) is ~5600 $M^{-1}cm^{-1}$. In the HPLC chromatogram, the first peak appeared at $t \approx 6.5$ min and corresponded to the unretained components which were eluted directly with the mobile phase from the column. The time $t \approx 6.5$ min (referred to as dead time) was the time required for the mobile phase to pass through the column. For the HFP peak, it appeared at $t \approx 19.0$ min. Since it took ~6.5 min to pass through the column to reach the detector, the actual time when HFP was eluted from the column should be ~12.5 min. According to the linear gradient program described above, % B ≈ 38.8 for $t \approx 12.5$ min, thus the minimum % B for HFP elution from the column was ~38.8. For the last peak in the chromatogram, which corresponded to the most hydrophobic components (e.g. peptide with uncleaved Fmoc group at its N-terminus), it appeared at $t \approx 27.0$ min. Given that the dead time was ~6.5 min, the actual elution time of the last peak was ~20.5 min, which was the time when % B was increased to 75.0%.

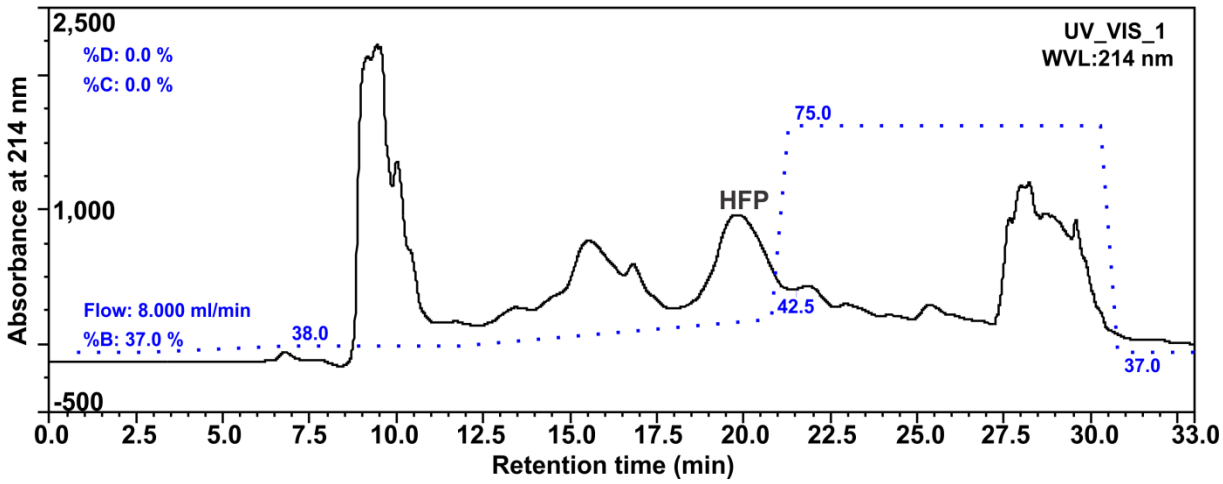

Figure C1 HPLC chromatogram of HFP purification.

The peak centered at $t \approx 20.0$ min in Figure C1 was diagnostic of HFP based on the MALDI mass spectrum. MALDI is a soft ionization source used in mass spectrometry. Examples of commonly used matrices for MALDI include 2,5-dihydroxybenzoic acid (DHB) and α-cyano-4-hydroxycinnamic acid (CHCA). The matrices can heavily absorb UV laser light at specific wavelengths (e.g. 337 and 355 nm) and cause sample desorption, during which a hot plume is produced and the analyte molecules (e.g. peptides) are ionized (more accurately protonated) in the hot plume[2]. For the positive-ion mode of MALDI, it is the protonated species $M+H^+$ that is detected. However, sometimes other species such as $M+Na^+$ and $M+K^+$ may be also detected. The MALDI mass spectrum in Figure C2 shows a predominant peak with m/z = 3148.8, which corresponds to HFP having a sequence of AVGIGALFLGFLGAAGSTMGARSWKKKKKKA and a theoretical molecular weight of 3150.8 Da. The peak with m/z = 3171.1 is diagnostic of HFP binding to a Na^+ ion instead of a proton.

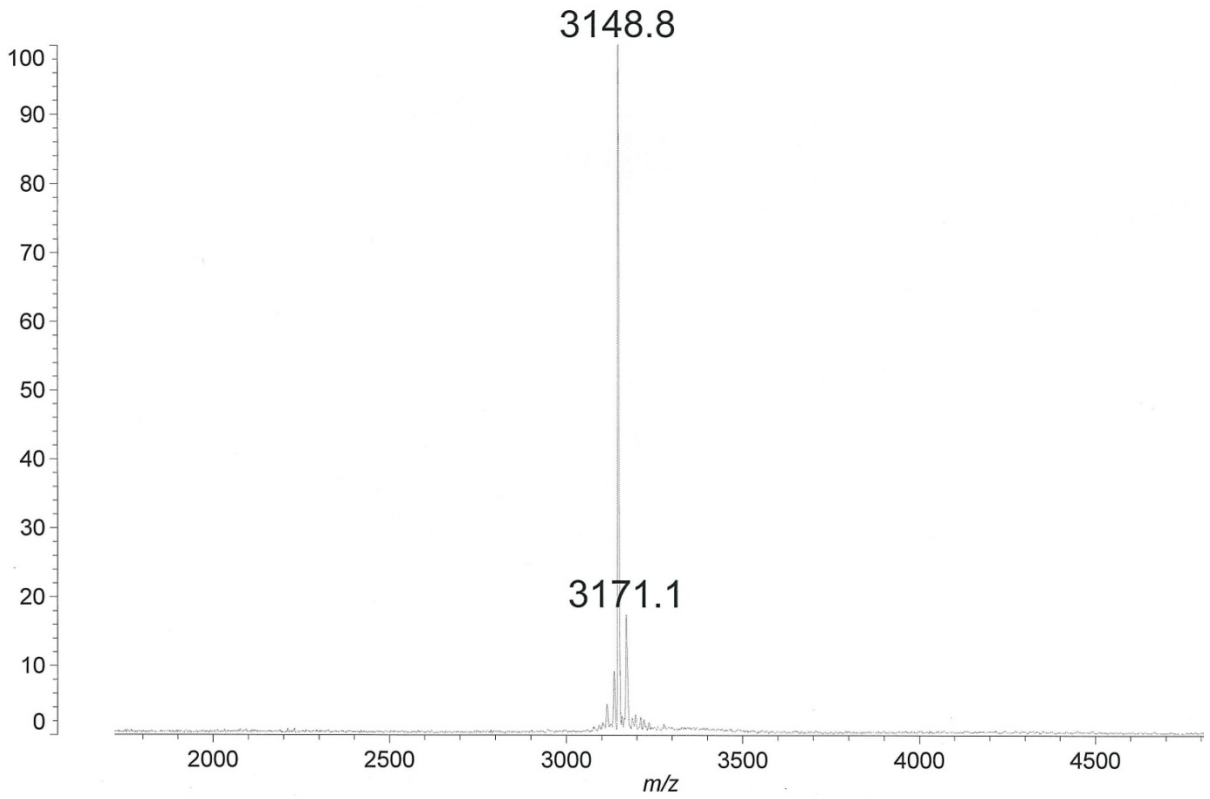

Figure C2 MALDI mass spectrum of HFP after purification. The theoretical molecular weight of HFP is 3150.8 Da.

APPENDIX D

NMR troubleshooting

1. Spinning problems

(a) If the rotor cannot spin at all or even cannot completely insert in the stator (also referred to as spinning module), check the alignment of the rotor, front bearing, and bottom bearing. If the rotor can only insert halfway in the stator, it is either the solenoid coil is twisted or the bottom bearing is not aligned with the coil. To find out which one is the case, we need to take apart the stator. If the coil is severely twisted, a new coil is needed. If the coil is fine but the bottom bearing is not aligned with the coil, we need to first heat up the leg of the coil using a tip heater and then adjust the leg to realign the coil and bottom bearing until the rotor can insert in the stator completely and freely.

(b) If the rotor can spin in the stator but the MAS speed reading is not stable in the "Auto" mode, first switch to the "Manual" mode, if the MAS controller can read the spinning speed stably, it means there is a communication issue rather than a rotor or stator issue. To solve the communication issue, we can restart the MAS controller and reopen the accessory control panel (i.e. ACC panel) on the host computer. If the speed reading is not stable in the "Manual" mode either, there are several possibilities described as follows:

(1) The sample is not tightly packed in the rotor. If this is the case, repack the sample and use bottom and top spacers with proper lengths.

(2) The black marking on the rotor is not clear. If this is the case, remark the rotor.

(3) The drive tip of the rotor is partially damaged. If this is the case, use a new drive tip.

(4) The front or bottom bearing in the stator will get scratched and its inner diameter will become bigger after long-time use, which may cause unbalancing of the rotor while spinning. If this is the case, replace the front or bottom bearing with a new one.

(5) The tiny holes at the bottom of the stator may get clogged, in which case the drive gas will be partially blocked and thus unstable rotor spinning will result. To solve this issue, the holes at the stator bottom need to be cleaned either manually or by ultrasonication.

(6) The tip of the optical fiber is damaged or improperly positioned. If this is the case, cut the tip of the optical fiber or reposition it until a stable speed reading is obtained.

(7) The bearing gas parameters in the ACC panel are not properly set. The typical values for a 4 mm MAS probe are given in Figure D1.

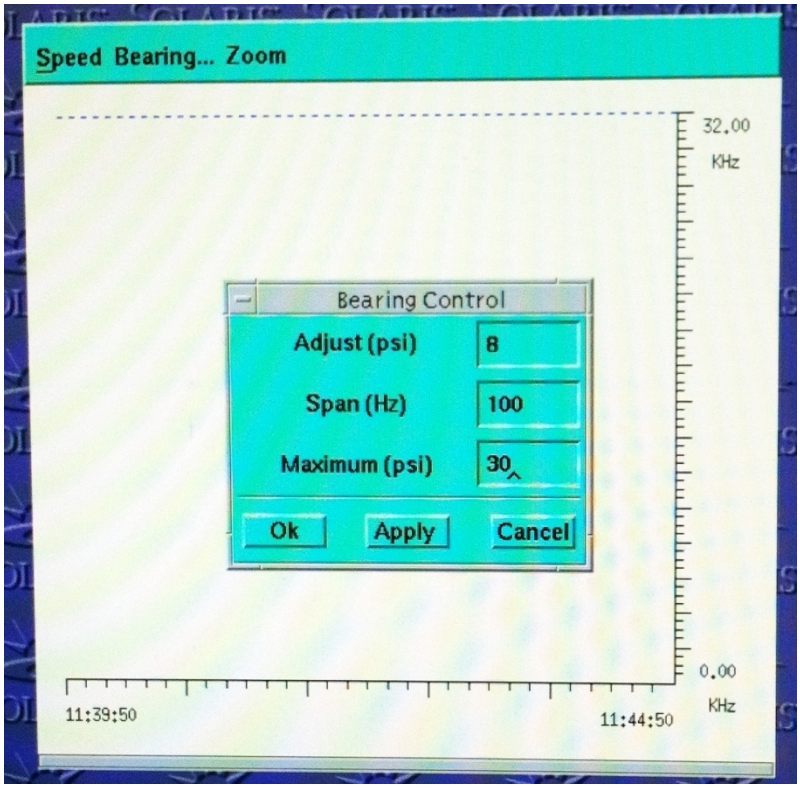

Figure D1 Typical bearing parameters in the ACC panel for a 4 mm MAS probe.

2. Tuning problems

(a) The channel is not tunable for the low-power tuning.

If a channel is untunable for the low-power tuning, i.e. no resonance is observed at the targeted frequency on the oscilloscope, it is definitely untunable for the high-power tuning (i.e. tuning in experiments where high-power r.f. pulses are applied). If a resonance is observed near the targeted frequency but not at that frequency no matter how we adjust the broad tuning rod and match, there are several parts including the series plug-in, trap plug-in, tuning tube, and tuning rod that we may change to move the resonance to the targeted frequency. For instance, if the targeted frequency is 100 MHz while the resonance only appears in the frequency range of 90-95 MHz, a series or trap plug-in with a lower capacitance may be used to move the resonance to 100 MHz. If no resonance is observed either at or near the targeted frequency, it is likely that there is a contact problem in that channel. In this case, we need to check whether some connection is loose, for example, the copper ribbon connecting the solenoid coil in the stator and the ^{1}H (or X) channel may not be soldered properly and thus resoldering may be needed, probably with a new copper ribbon. For the ^{1}H channel, there is a little screw on the match rod (Figure D2), if this screw is loose or lost, the ^{1}H channel will lose its resonance and become untunable.

(b) The channel is tunable for the low-power tuning but not for the high-power tuning.

If a channel is tunable at the transmitter frequency for the low-power tuning but not for the high-power tuning, there are usually two possibilities described as follows:

(1) The tuning rod is way off during the high-power tuning. When the tuning rod is way

Figure D2 Little screw (red circled) on the ^{1}H match that plays a key role for ^{1}H tuning.

off the correct position, in which case the probe resonance frequency is way off the transmitter frequency of the r.f. pulse, the transmitter frequency will be beyond the tunable frequency range achieved by adjusting the match. Therefore, the reflected voltage (also referred to as reverse voltage) of the r.f. pulse will not go down significantly no matter how we adjust the match while the tuning rod is way off. In this case, we can switch back to the low-power tuning and first adjust the tuning rod to move the resonance back to the transmitter frequency and then adjust the match until the resonance peak at the transmitter frequency becomes as sharp as possible, in which case the reflected voltage is close to zero. After that, do the high-power tuning again and the probe should be tunable now.

(2) The probe is arcing. When too much r.f. pulse power is input to the probe, the significant heating may induce carbonization at or near the connection spots and cause probe arcing. When severe arcing occurs, sparkle may be observed at the arcing spot and huge spikes may be observed in the FID as well. Figure D3 displays the REDOR S_0 and S_1 FIDs (panel a) in the presence of severe arcing at the carbonized rim near one leg of the solenoid coil (panel b). In this case, the rim needs to be cleaned thoroughly to remove all black carbon. After that, the probe may become tunable again.

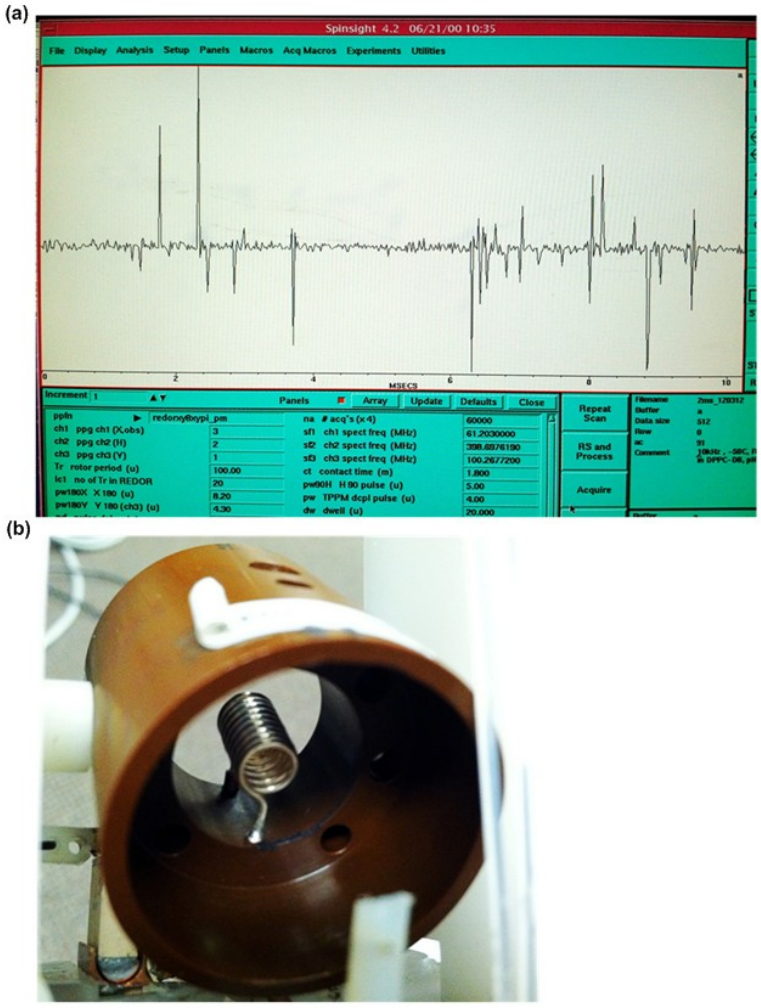

Figure D3 (a) REDOR S_0 and S_1 FIDs in the presence of severe probe arcing; (b) Stator showing the carbonized rim near one leg of the solenoid coil where the arcing occurs.

REFERENCES

REFERENCES

1. Merrifield, R.B., *SOLID PHASE PEPTIDE SYNTHESIS .1. SYNTHESIS OF A TETRAPEPTIDE.* Journal of the American Chemical Society, 1963. **85**(14): p. 2149-&.

2. *http://en.wikipedia.org/wiki/Matrix-assisted_laser_desorption/ionization.*