2009

Spin diffusion and dynamics studies of the channel forming membrane proteins by solid-state nuclear magnetic resonance

Wenbin Luo
Iowa State University

Spin diffusion and dynamics studies of the channel forming membrane proteins by solid-state nuclear magnetic resonance

by

Wenbin Luo

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemistry

Program of Study Committee:
Mei Hong, Major Professor
Emily Smith
Klaus Schmidt-Rohr
Amy Andreotti
Yan Zhao

Iowa State University

Ames, Iowa

2010

Table of Contents

Acknowledgements v

Abstract vi

Chapter 1. Introduction to Membrane Protein Structures 1

Membrane Protein and Ion Channel 1

Factors that Affect Membrane Protein Structure 5

Thesis Organization 7

Copyright Permissions 7

References 8

Chapter 2. Solid State NMR Techniques 12

Nuclear Spin Interactions 12

Methods to Determine Intermolecular Packing 15

Spin Diffusion Theory 20

References 24

Chapter 3. Large Structure Rearrangement of Colicin Ia Channel Domain after Membrane Binding from 2D ^{13}C Spin Diffusion NMR 26

Abstract 26

Introduction 27

Materials and Methods 28

Results and Discussion 30

Acknowledgements 42

References 42

Chapter 4. A 1D Sensitivity-Enhanced ^{1}H Spin Diffusion Experiment for Determining Membrane Protein Topology 45

Abstract 45

Introduction 45

Materials and Methods 47

Results and Discussion 48

Acknowledgements 60

References 60

Chapter 5. Determination of the Oligomeric Number and Intermolecular Distances of Membrane Protein Assemblies by Anisotropic ^{1}H-Driven Spin Diffusion NMR Spectroscopy 62

Abstract 62

Introduction 63

Materials and Methods 65

Results and Discussion 72

Conclusion 88

Acknowledgements 89

References 89

Chapter 6. Sidechain Conformation of the M2 Transmembrane Peptide Proton Channel of Influenza A Virus from ^{19}F Solid-State NMR 92

Abstract 92

Introduction 93

Experimental Methods 95

Results 98

Discussion 107

Conclusion 113

Acknowledgments 114

References 114

Chapter 7. Immobilization of Membrane Proteins by Eukaryote-Mimetic Lipid Membranes for Solid-State NMR Structure Determination 117

Abstract 117

Introduction 118

Results 120

Discussion 126

Materials and Methods 129

Acknowledgements 130

Supporting Information 131

References 134

Chapter 8. Water-Protein Interaction of the Influenza M2 Proton Channel in lipid Bilayers from Solid-State NMR 139

Abstract 139

Introduction 140

Result and Discussion 141

Materials and Methods 149

Acknowledgement 153

References 153

Supporting Information 158

Appendix A. Viral Membrane Sample Preparation Protocols 166

Viral Membrane Vesicle Preparation 166

M2 Peptide Reconstitution 166

Amantadine Binding 167

Appendix B. CODEX Simulation Code 168

Appendix C. 3D ^{1}H Spin Diffusion Simulation Code 175

Acknowledgements

I wish to thank all those people who guided me and helped me in the last five years. Foremost, my PhD advisor, Dr. Mei Hong, has been an excellent mentor to teach me the knowledge about solid state NMR, who pulled me through various difficulties. It was because of her unyielding support and persistent help that I could finish my PhD study and develop into a NMR researcher. More than an advisor, she is also the academic mother and a model in life to me.

I would also like to thank Dr. Klaus Schmidt-Rohr for many helpful theoretical lessons and discussions in our group meetings. Thank you to my committee members, Dr. Emily Smith, Dr. Amy Andreotti, and Dr. Yan Zhao for the guidance on my study. I also want to thank many members of the Iowa State University Department of Chemistry staff, Dr. Dave Scott and Dr. Shu Xu in the NMR facility, Bev, Carlene, Renee and Lynette in the Chemistry Office, for the great assistance throughout my years.

I am grateful for my fellow coworkers in the Hong group and also Schmidt-Rohr group. I would especially like to thank Dr. Ming Tang for constant help with NMR and for being a great friend in my ISU days. Dr. Rajeswari Mani and Dr. Xiaolan Yao are appreciated because of their training me with various aspects of NMR. Thanks to Dr. Sarah Cady, Dr. Tim Doherty, Dr. Xiaowen Fang, Dr. Qiang Chen, Dr. Shenhui Li, Yanyan Hu Yongchao Su, Yuan zhang, Marilu Perez and Fanghao Hu for their tremendous help and provide such a good working environment in graduate school.

Finally, I thank Gang Wang for various help. I would especially like to thank my parents for unconditional love and support throughout my life, especially their help with taking care of my daughter Mengyao Luo, to whom I lovingly dedicate this thesis.

Abstract

Solid-state nuclear magnetic resonance (SSNMR) is an important tool for the structure, function and dynamics study of many chemical and biological systems, especially powerful in studying membrane proteins, whose structures have been difficult to analyze by traditional x-ray crystallography or solution NMR techniques. In this thesis, various NMR techniques are used to study the structure and dynamics of membrane proteins within lipid bilayers.

The main technique applied in this thesis is spin diffusion experiments. We study the structural rearrangement upon membrane binding of colicin Ia by the proton-driven ^{13}C spin diffusion (PDSD) ^{13}C-^{13}C 2D correlation experiment. Membrane bound colicin Ia turns out to have a more extended structure compared to the soluble state. Then a 1D ^{1}H detected ^{1}H spin diffusion experiment is developed to provide the same membrane protein topology information as the 2D ^{13}C detected version, but with significant sensitivity enhancement. We demonstrated this new technique on the colicin Ia channel-forming domain and achieved about 200 fold time saving. Further, the data analysis method is developed to extract the intermolecular distance as long as 12 Å from ^{19}F spin diffusion experiment CODEX, where the oligomeric state is obtained at the same time. Demonstrated on the M2 proton channel system, this method is applied to extract the intermolecular distances between a key residue Trp41 in different states of the M2 proton channel. Finally, the water accessibility of the M2 proton channel in different states is studied by the ^{1}H spin diffusion experiment and 3D low resolution models are proposed for this proton channel system by simulating the ^{1}H spin diffusion process between the water and protein.

The second focus of this thesis is the dynamics of the M2 peptide in a complex membrane system. Compared to the single component model lipid bilayers, this composite membrane is shown to reduce the rotational rate of the membrane protein by 2 orders of magnitude, which is explained by a rotational diffusion model. The advantage of this immobilization is the ability to acquire high resolution SSNMR spectra at physiological temperatures.

Chapter 1

Introduction to Membrane Protein Structures

1.1 Membrane Proteins and Ion Channels

Membrane proteins are of significant biological importance(*1*), such as cell stability(*2, 3*), signal transduction(*4, 5*), ion transport(*6-8*) and ligand binding(*9*). Membrane proteins amount to about 25% of all human proteins, however their structures have been underdetermined, evidenced by the fact that less than 1% of the solved protein structures deposited in the PDB bank are membrane proteins(*10*). Membrane protein structures are difficult to determine because generally membrane proteins are difficult to crystallize and are insoluble in water. As an alternative where the other common techniques have limitations, solid state NMR is a powerful tool to study membrane protein structure and dynamics within the biological membranes, therefore provides directly biological relevant information(*11, 12*).

Membrane proteins are associated with biological membranes. Cell membranes consist of a phospholipid bilayer, membrane proteins and small molecules, with composition varing from cell to cell and also within the cell. Generally, the cell membrane has a hydrophobic membrane interior and a polar region of the membrane surface surrounded by the aqueous environment. The major components of lipid membranes are phosphatidylcholine (PC), phosphatidylethanolamine (PE), sphingomyelin (SM), phosphatidylserine (PS), phosphatidylinositol (PI), phosphatidylglycerol (PG) and cholesterol (Chol) with their structures shown in Figure 1. PC and PE are zwitterionic with a negative charge on the phosphate group and a positive charge on the amine. As one of the most biologically prevalent, PC is a widely used lipid for biological studies of membrane proteins. PE has a small head group, therefore increases the membrane flexibility by modifying the surface curvature. PI, PS and PG have negatively charged head groups to interact with the charged residues in membrane proteins, and affect membrane protein activities through electrostatic or hydrogen-bonding interactions. As a significant component

in mammalian cell membranes, cholesterol has a planar steroid ring that reduces the flexibility of the lipid chains and increases the viscosity of the lipid bilayers.

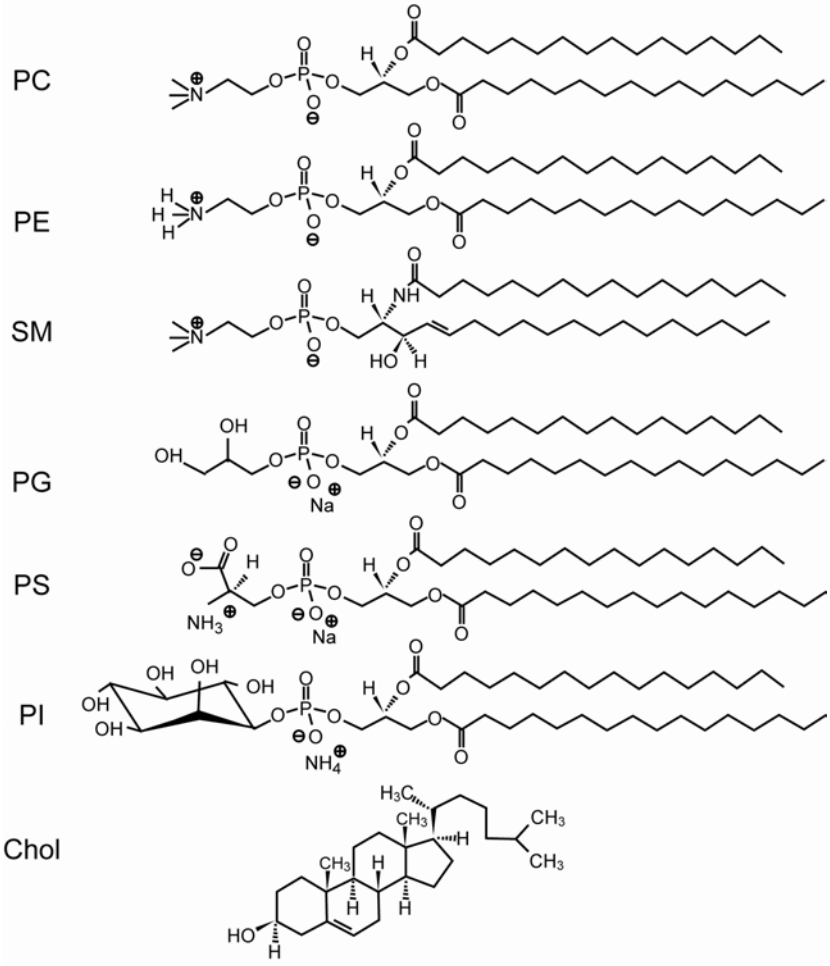


Figure 1. Structures of lipids and cholesterol

Lipid membranes have abundant dynamics and hence affect the membrane proteins. Lipids can diffuse laterally and rotationally and interact with other lipids and with the proteins. The rate of rotational diffusion is affected by the lipid composition and temperature, with orders of magnitude in the rate difference(*13*). The diffusion rate is significantly affected by the phase transition temperature of the specific lipid, above which the membrane is in the liquid crystalline phase and below in the gel phase. In the liquid crystalline phase, lipids and membrane proteins are highly dynamic. In the gel phase, the lipid motion is slow. The phase transition temperature is affected by the lipid composition, lipid chain length and the proportion of saturated fatty acids(*13*). The higher the saturated fatty acids proportion, the

longer the lipid acyl chain, the higher the phase transition temperature. The lipid composition of biological membranes can be regulated by microorganisms and cells to reach a similar fluidity under various growth conditions. Membrane dynamics affect membrane protein dynamics and structures significantly(*14*).

Two major biological functions of membrane proteins are transporting molecules through the lipid membrane(*6-8*) and signal transduction(*4, 5*). As complexes of integral membrane proteins transporting ions through the membranes, ion channels form a hydrophilic pathway for ions to cross the hydrophobic barrier of the lipid bilayers. Ion channels can be turned on or off by chemical or physical modulators, and the conformational change is called gating(*15*). At the open state, the specific ion will flow down the electrochemical gradient, selective against the size and charge of the permeate. Ion channels are important in many biological processes, such as cardiac, skeletal and smooth muscle contraction, eptithelial transport of nutrients and ions(*16*) (*7, 17*). For example, channels are the most prominent components of the nervous system. Many toxins work by modulating ion channel conductance and/or kinetics of the prey and thus shut down their nervous systems. Further, ion channels are often the target for new drugs(*18-20*).

Two main types of ion channels are voltage-gated ion channels(*17*) and ligand-gated ion channels. One of the voltage-gated ion channel examples is formed by the channel-forming domain of colicin Ia(*21, 22*). There are three domains in the colicin Ia protein: the receptor domain binds the outer membrane receptor on the target cell, the N terminus translocation domain delivers the channel-forming domain at the C terminus from the bacterial outer membrane to the inner membrane(*23*). After spontaneous insertion into the inner membrane of sensitive bacteria cells from the aqueous environment, the channel-forming domain opens as a voltage-gated non-selective channel that depletes the cell membrane potential and leads to bacteria death(*24*) (see Figure 2). We are interested in the mechanism of action of this protein.

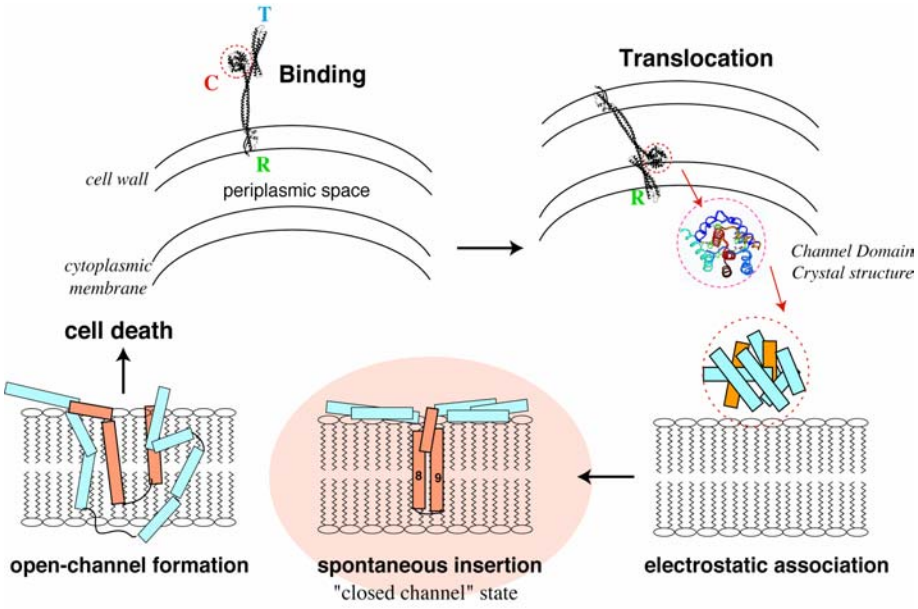


Figure 2. Voltage-gated ion channel forming mechanism of colicin Ia.

The M2 proton channel of influenza virus is a pH-gated ion channel. This channel is the target of anti-flu drug amantadine to interrupt the influenza virus replication(*25, 26*). The proton channel is regulated by the proton concentration outside of the virus: it is closed at neutral pH but opens in the acidic environment (low pH_{out}) of the endosome after viral entry into the host cell(*27*). The acidification of the viral interior initiates the release of the viral RNA into the host cell, causing infection. Binding of amantadine blocks the proton channel and also the life circle of the influenza virus(*28*) (see Figure 3). Therefore elucidating the structure of the M2 proton channel and its function mechanism is important for new anti-flu drug screen.

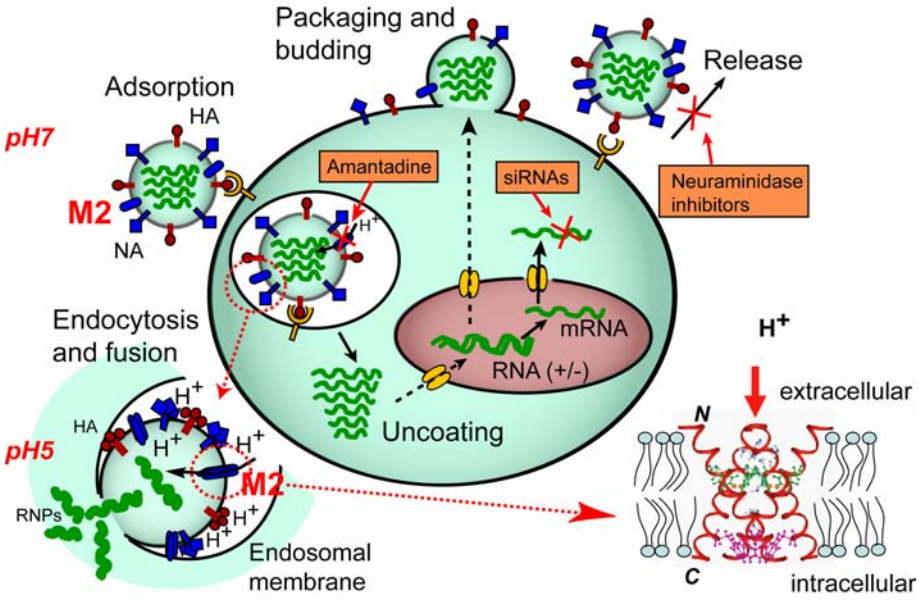


Figure 3. Life circle of influenza virus.

1.2 Factors that Affect Membrane Protein Structures

1.2.1 Solvent effect

In order to adapt to the aqueous solution or the amphiphilic environment of lipid bilayers that are distinctly different, membrane proteins may have conformational change inside and outside the membrane. An example is colicin Ia channel forming domain(*29*). In the water soluble state, the hydrophobic domains of colicin Ia are embedded within the hydrophilic surface and the structure is close to globular(*21*). In the membrane bound state, the hydrophilic parts spread at the membrane surface(*24*), possibly with the charged residues like Arg or Lys contacting the water or the headgroup charges at the surface of the lipid bilayers. The hydrophobic domains insert into or span the membrane interior, with the structure rearranged into an umbrella shape. This way the total energy can be lower and the structure is stabilized.

1.2.2 Bilayer thickness

Different membrane compositions have different membrane thickness and viscosities, therefore different membranes affect membrane protein structure and dynamics significantly (*30*). The length of the lipid acyl chains is a frequent variable in biophysical studies, because membrane proteins may have structural plasticity by becoming more extended or compact to match the membrane hydrophobic thickness. An example is that M2 transmembrane domain adjusts the tilt angle when forming the proton channel in response to membranes with different thickness(*31*). Compared to the thinner DLPC bilayers with 3 nm thickness, where M2 is tilted by 35 degrees(*14*), M2 may tilt to a smaller angle in thicker membrane such as POPC bilayers or viral membranes(*14, 31, 32*). In addition, some residues show backbone conformational changes between membranes with different thickness, which may be coupled to the adaptation to the hydrophobic mismatch.

1.2.3 pH

pH change is an important factor that affects membrane protein structures by changing the protein charged state, conformation of key residues and hydrogen bonds. An example is when M2 proton channel is switched on and off by different pH values(*26*), and also correspondingly the channel overall structure changes(*25*). At neutral pH, the four His37 residues form a dimer of dimer in this proton channel(*33*). When the pH is below pH 6.3 which is the pKa for adding the additional proton to the dimer of dimer state of the four His37 residues(*33*), the M2 proton channel will be open by forming a bigger pore and all the helices will be straight(*32*). When the pH is above pH 8.2 which is the pKa for depleting a proton to break the dimer of dimer state of the four His37 residues (*33*), the M2 channel will be closed as a tighter pore where some of the helices may be kinked from the middle(*34*). The key residues of His37 and Trp41 will have backbone conformational changes as well as sidechain rotamer changes(*33*), so that the channel can be open to conduct proton(*25*).

1.2.4 Ligand binding

Ligand-protein interaction is often described as a lock-key interaction(*35*). Ligand may specifically interact with membrane proteins and affect the protein structure through

various interactions such as electrostatic interactions. Therefore ligand-protein interactions are important to study, and they are especially interesting for the drug inhibition mechanism and for drug screening(*36, 37*). For example, when the anti-flu drug amantadine binds to the M2 channel, M2 helices tend to form a kink near the middle of the channel(*34*). At the apo state under neutral pH, M2 helices have conformation heterogeneity with both kinked and straight helices present(*32*). Upon amantadine binding, the kinked conformation of the helix backbone is favored and dominates. So ligand binding may improve the conformation homogeneity(*32*),

1.3 Thesis Organization

SSNMR studies conducted on two membrane proteins, colicin Ia and influenza A virus M2 peptide, are presented in this thesis. Chapter 2 describes the solid state NMR theory and techniques related to this work. Spin diffusion theory is briefly reviewed. Then the structural rearrangement upon membrane binding of colicin Ia is studied by the proton-driven ^{13}C spin diffusion (PDSD) ^{13}C-^{13}C 2D correlation experiment (chapter 3). In chapter 4, a 1D ^{1}H detected ^{1}H spin diffusion experiment is described to provide the same topology information as the 2D ^{13}C detected version, but with significant sensitivity enhancement. Demonstrated on the M2 proton channel system, the data analysis of ^{19}F spin diffusion process is developed to probe intermolecular distance as long as 12 Å, while the oligomeric state is obtained from the spin diffusion curve (chapter 5). Subsequently this method is applied to extract the intermolecular distance between a key residue Trp41 in different states of the M2 proton channel (chapter 6). In chapter 7, the dynamics of the M2 peptide in a complex membrane system is shown to be reduced significantly compared to the single component model membrane system and is explained by a rotational diffusion model. Finally, the water accessibility of the M2 proton channel in different states is studied by the ^{1}H spin diffusion experiment and a 3D low resolution model is proposed for this proton channel system.

1.4 Copyright permission

Chapters 3, 4, 5, 6 and 7 are reprints of published papers. Permissions have been obtained from the following publishing groups.

Chapters 3, 5, 6 and 7 American Chemical Society

Chapter 4 Elsevier

References

1. Sachs, J. N., and Engelman, D. M. (2006) Introduction to the membrane protein reviews: the interplay of structure, dynamics, and environment in membrane protein function, Annu. Rev. Biochem. 75, 707-712.

2. Phillips, R., Ursell, T., Wiggins, P., and Sens, P. (2009) Emerging roles for lipids in shaping membrane-protein function, Nature 459, 379-385.

3. Langosch, D., and Arkin, I. T. (2009) Interaction and conformational dynamics of membrane-spanning protein helices, Protein. Sci. 18, 1343-1358.

4. Bao, L., Redondo, C., Findlay, J. B., Walker, J. H., and Ponnambalam, S. (2009) Deciphering soluble and membrane protein function using yeast systems (Review), Mol. Membr. Biol. 26, 127-135.

5. Bhattacharyya, R. P., Remenyi, A., Yeh, B. J., and Lim, W. A. (2006) Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits, Annu. Rev. Biochem. 75, 655-680.

6. Minor, D. L., Jr. (2001) Potassium channels: life in the post-structural world, Curr. Opin. Struct. Biol. 11, 408-414.

7. Dubyak, G. R. (2004) Ion homeostasis, channels, and transporters: an update on cellular mechanisms, Adv. Physiol. Educ. 28, 143-154.

8. DeFelice, L. J., and Goswami, T. (2007) Transporters as channels, Annu. Rev. Physiol. 69, 87-112.

9. Engel, A., and Gaub, H. E. (2008) Structure and mechanics of membrane proteins, Annu. Rev. Biochem. 77, 127-148.

10. Raman, P., Cherezov, V., and Caffrey, M. (2006) The Membrane Protein Data Bank, Cell. Mol. Life. Sci. 63, 36-51.

11. Hong, M. (2006) Oligomeric structure, dynamics, and orientation of membrane proteins from solid-state NMR, Structure 14, 1731-1740.

12. Hong, M. (2007) Structure, topology, and dynamics of membrane peptides and proteins from solid-state NMR Spectroscopy, J. Phys. Chem. B 111, 10340-10351.

13. Saffman, P. G., and Delbruck, M. (1975) Brownian-Motion in Biological-Membranes, Proc. Natl. Acad. Sci. USA 72, 3111-3113.

14. Cady, S. D., Goodman, C., Tatko, C. D., DeGrado, W. F., and Hong, M. (2007) Determining the orientation of uniaxially rotating membrane proteins using unoriented samples: A H-2, C-13, and N-15 solid-state NMR investigation of the dynamics and orientation of a transmembrane helical bundle, J. Am. Chem. Soc. 129, 5719-5729.

15. Chanda, B., Asamoah, O. K., Blunck, R., Roux, B., and Bezanilla, F. (2005) Gating charge displacement in voltage-gated ion channels involves limited transmembrane movement, Nature 436, 852-856.

16. Lehnart, S. E., Ackerman, M. J., Benson, D. W., Jr., Brugada, R., Clancy, C. E., Donahue, J. K., George, A. L., Jr., Grant, A. O., Groft, S. C., January, C. T., Lathrop, D. A., Lederer, W. J., Makielski, J. C., Mohler, P. J., Moss, A., Nerbonne, J. M., Olson, T. M., Przywara, D. A., Towbin, J. A., Wang, L. H., and Marks, A. R. (2007) Inherited arrhythmias: a National Heart, Lung, and Blood Institute and Office of Rare Diseases workshop consensus report about the diagnosis, phenotyping, molecular mechanisms, and therapeutic approaches for primary cardiomyopathies of gene mutations affecting ion channel function, Circulation 116, 2325-2345.

17. Bezanilla, F. (2005) Voltage-gated ion channels, IEEE Trans Nanobioscience 4, 34-48.

18. Errington, A. C., Stohr, T., and Lees, G. (2005) Voltage gated ion channels: targets for anticonvulsant drugs, Curr. Top. Med. Chem. 5, 15-30.

19. Li, S., Gosling, M., Poll, C. T., Westwick, J., and Cox, B. (2004) Therapeutic scope of modulation of non-voltage-gated cation channels, Drug Discov. Today 9, 1045-1054.

20. Akar, F. G., and Tomaselli, G. F. (2005) Ion channels as novel therapeutic targets in heart failure, Ann. Med. 37, 44-54.

21. Wiener, M., Freymann, D., Ghosh, P., and Stroud, R. M. (1997) Crystal structure of colicin Ia, Nature 385, 461-464.

22. Stroud, R. (1995) Ion channel forming colicins, Curr. Opin. Struct. Biol. 5, 514-520.

23. Zakharov, S. D., and Cramer, W. A. (2004) On the mechanism and pathway of colicin import across the E. Coli outer membrane, Front Biosci 9, 1311-1317.

24. Zakharov, S. D., and Cramer, W. A. (2002) Colicin crystal structures: pathways and mechanisms for colicin insertion into membranes, Biochim. Biophys. Acta 1565, 333-346.

25. Cady, S. D., Luo, W. B., Hu, F. H., and Hong, M. (2009) Structure and Function of the Influenza A M2 Proton Channel, Biochemistry 48, 7356-7364.

26. Holsinger, L. J., Nichani, D., Pinto, L. H., and Lamb, R. A. (1994) Influenza A virus M2 ion channel protein: a structure-function analysis, J. Virol. 68, 1551-1563.

27. Pinto, L. H., Holsinger, L. J., and Lamb, R. A. (1992) Influenza-Virus M2 Protein Has Ion Channel Activity, Cell 69, 517-528.

28. Miller, C. (2008) Ion channels: Coughing up flu's proton channels, Nature 451, 532-533.

29. Luo, W. B., Yao, X. L., and Hong, M. (2005) Large structure rearrangement of colicin la channel domain after membrane binding from 2D C-13 spin diffusion NMR, J. Am. Chem. Soc. 127, 6402-6408.

30. Niemela, P. S., Hyvonen, M. T., and Vattulainen, I. (2008) Atom-scale molecular interactions in lipid raft mixtures, Biochim. Biophys. Acta.

31. Duong-Ly, K. C., Nanda, V., Degrad, W. F., and Howard, K. P. (2005) The conformation of the pore region of the M2 proton channel depends on lipid bilayer environment, Protein Science 14, 856-861.

32. Luo, W. B., Cady, S. D., and Hong, M. (2009) Immobilization of the Influenza A M2 Transmembrane Peptide in Virus Envelope-Mimetic Lipid Membranes: A Solid-State NMR Investigation, Biochemistry 48, 6361-6368.

33. Hu, J., Fu, R., Nishimura, K., Zhang, L., Zhou, H.-X., Busath, D. D., Vijayvergiya, V., and Cross, T. A. (2006) Histidines, heart of the hydrogen ion channel from influenza A virus: Toward an understanding of conductance and proton selectivity, Proc. Natl. Acad. Sci. USA 103, 6865-6870.

34. Hu, J., Asbury, T., Achuthan, S., Li, C. G., Bertram, R., Quine, J. R., Fu, R. Q., and Cross, T. A. (2007) Backbone structure of the amantadine-blocked trans-membrane domain M2 proton channel from influenza A virus, Biophys. J. 92, 4335-4343.

35. Brewerton, S. C. (2008) The use of protein-ligand interaction fingerprints in docking, Curr. Opin. Drug. Discov. Devel. 11, 356-364.

36. Glen, R. C., and Allen, S. C. (2003) Ligand-protein docking: cancer research at the interface between biology and chemistry, Curr. Med. Chem. 10, 763-767.

37. Good, A. (2001) Structure-based virtual screening protocols, Curr. Opin. Drug Discov. Devel. 4, 301-307.

Chapter 2

Solid State NMR Techniques

2.1 Nuclear Spin Interactions

NMR is a powerful tool for structure determinations of organic and inorganic molecules and biological assemblies. Combining magic angle spinning (MAS)(*1, 2*) with isotopic labeling(*3, 4*) to provide high resolution spectra, solid state NMR probes membrane protein structures in native lipid bilayers at the atomic level. The key interactions in solid state NMR of interest in this work are described below(*5, 6*).

Zeeman Interaction

The vector sum of the individual magnetic moments associated with all the nuclei in a sample gives the net magnetization $\vec{M}$:

$$\vec{M} = \sum_i \vec{\mu}_i \qquad (2.1)$$

Each nuclear magnetic moment is related to the nuclear spin $\vec{I}_i$ and the gyromagnetic ratio γ characteristic for a given type of nucleus. γ is positive for the ^{1}H, ^{2}H, ^{13}C, ^{19}F and ^{31}P and negative for ^{29}Si and ^{15}N.

$$\vec{\mu}_i = \gamma \hbar \vec{I}_i \qquad (2.2)$$

The net magnetization $\vec{M}$ is related to the net nuclear spin angular momentum $\vec{J}$ of the sample.

$$\vec{M} = \gamma \vec{J} \qquad (2.3)$$

If the nuclei are placed in a uniform magnetic field $\vec{B}$ as in an NMR experiment, this torque is given by

$$\vec{N} = \vec{M} \times \vec{B} \qquad (2.4)$$

The torque $\vec{N}$ changes the net nuclear spin angular momentum $\vec{J}$ as follows:

$$\vec{N} = \frac{d}{dt}\vec{J} \tag{2.5}$$

Combining equations 2.4 and 2.5, we can write the equation describing the motion of the magnetization vector $\vec{M}$ in the $\vec{B}$ field.

$$\frac{d\vec{M}}{dt} = \gamma\vec{M} \times \vec{B} \tag{2.6}$$

The equation predicts that $\vec{M}$ precesses about a fixed $\vec{B}$ at a constant rate $\omega = -\gamma B$. In NMR, the applied magnetic field is generally labeled $\vec{B}_0$ and is taken to be along z of the laboratory frame of reference, i.e. $\vec{B}$ = (0, 0, B_0) in the above equations. Then the frequency with which the magnetization precesses about this field $\vec{B}_0$ is defined as ω_0, termed as Larmor frequency.

$$\omega_0 = -\gamma B_0 \tag{2.7}$$

The Hamiltonian is the energy operator for the system, and its form depends on the nuclear spin. For an isolated spin free of other interactions in a static uniform magnetic field, the Hamiltonian is referred as the Zeeman Hamiltonian. As the interaction between the magnetic dipole moment $\vec{\mu} = \gamma\hbar\vec{I}$ of the spin $\vec{I}$ and the external magnetic field $\vec{B}_0$, the Zeeman interaction is the strongest interaction experienced by the nuclear spin in the magnetic field.

$$\hat{H} = -\vec{\mu} \cdot \vec{B}_0 \tag{2.8}$$

where $\hat{\mu}$ is the nuclear magnetic moment operator which is related to the nuclear spin operator $\hat{I}$ according to (2.2)

$$\hat{H}_Z = -\gamma\vec{B}_0 \cdot \hbar\hat{I} = -\gamma\hbar B_0 \hat{I}_Z \tag{2.9}$$

Dipolar Coupling

The dipolar coupling Hamiltonian acting on a spin *I* in Cartesian tensorial form is

$$\hat{H}_{dd} = -2\hat{I} \cdot D \cdot \hat{S} \tag{2.10}$$

The spin *S* is the source of the local field at spin *I*. The term *D* is the dipole-coupling tensor, with the principal values of *-d/2*, *-d/2*, *+d*, where *d* is the dipolar-coupling constant (in unit of rad·s^{-1}).

$$d = \hbar(\frac{\mu_0}{4\pi})\frac{1}{r^3}\gamma_I\gamma_S \tag{2.11}$$

The heteronuclear dipolar-coupling spin pair Hamiltonian is:

$$\hat{H}_{dd}^{hetero} = -d(3\cos^2\theta - 1)\hat{I}_Z\hat{S}_Z \tag{2.12}$$

The homonuclear dipolar-coupling spin pair Hamiltonian is:

$$\hat{H}_{dd}^{homo} = -d\cdot\frac{1}{2}(3\cos^2\theta - 1)[3\hat{I}_Z\hat{S}_Z - \hat{I}\cdot\hat{S}\ ] \tag{2.13}$$

Chemical Shielding and Chemical Shift Anisotropy (CSA)

The chemical shielding Hamiltonian acting on a spin *I* is (in frequency units)

$$\hat{H}_{CS} = \gamma\hat{I}\cdot\sigma\cdot B_0 \tag{2.14}$$

σ is the chemical shielding tensor and can be decomposed into a symmetric σ^s and an antisymmetric σ^{as} component.

$$\sigma\ = \sigma^s + \sigma^{as} \tag{2.15}$$

Only the symmetric part σ^s has significant effects on the NMR spectra, thus only this part will be considered. The chemical shift tensor σ can be defined in a certain axis frame so that the shielding tensor is diagonal. This frame is called the principal axis frame, designated as 'PAF', or x^{PAF}, y^{PAF}, z^{PAF}. The diagonal terms of σ^{PAF} are the principal values of the shielding tensor. The three principal values of the shielding tensor σ^{PAF} are often expressed as the isotropic value σ_{iso}, the anisotropy δ_{cs} and the asymmetry parameter η_{cs}. These quantities are defined from the principal values as follows:

$$\sigma_{iso} = \frac{1}{3}(\sigma_{xx}^{PAF} + \sigma_{yy}^{PAF} + \sigma_{zz}^{PAF}) \tag{2.16}$$

$$\delta_{cs} = \sigma_{zz}^{PAF} - \sigma_{iso} \tag{2.17}$$

$$\eta_{cs} = \frac{(\sigma_{yy}^{PAF} - \sigma_{xx}^{PAF})}{\delta_{cs}} \tag{2.18}$$

The frequency of the observed NMR signal from the chemical shielding ω_{cs} is related to polar angle (θ, ϕ) between the B_0 and the shielding tensor principal axis frame $b_0^{PAF} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. For the non-axially symmetric case, the chemical shift frequency can be expressed in terms of the isotropic component $\omega_{iso} = \sigma_{iso} \cdot \omega_0$, shielding anisotropy and asymmetry, relative to the Lamor frequency ω_0.

$$\begin{aligned} \omega_{cs}(\theta,\phi) &= -\omega_0(\sigma_{xx}^{PAF}\sin^2\theta\cos^2\phi + \sigma_{yy}^{PAF}\sin^2\theta\sin^2\phi + \sigma_{zz}^{PAF}\cos^2\theta) \\ &= \omega_{iso} + \frac{1}{2}\delta_{cs}(3\cos^2\theta - 1 - \eta\sin^2\theta\cos 2\phi) \end{aligned} \tag{2.19}$$

For a shielding tensor with axial symmetry, $\sigma_{xx}^{PAF} = \sigma_{yy}^{PAF}$, η=0, and this equation can be simplified.

$$\omega_{cs}(\theta) = \omega_0\sigma_{zz}^{PAF}\frac{1}{2}(3\cos^2\theta - 1) = \omega_{iso} + \frac{1}{2}\delta_{cs}(3\cos^2\theta - 1) \tag{2.20}$$

2.2 Methods to Determine Intermolecular Packing

2.2.1 REDOR

Internuclear distances can be obtained by MAS NMR by measuring the dipolar coupling between nuclear spins, since the dipolar interaction is inversely proportional to the cube of the inter-nuclear distance (see equation 2.11). Rotational-echo double-resonance (REDOR) (*7, 8*) is a versatile and straightforward technique to recouple the heternuclear dipolar coupling under MAS using rotor-synchronized π pulses (Figure 1).

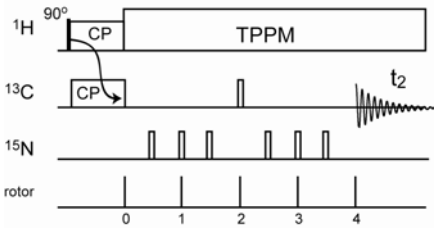


Figure 1. The pulse sequence of the REDOR experiment.

Under magic angle spinning at a frequency ω_r with α, β defined as the azimuthal angle and polar angles between the inter-nuclear vector and the spinning axis, the dipolar frequency is(*7*)

$$\omega_d(\alpha,\beta,t) = \pm\frac{1}{2}d\left\{\sin^2\beta\cos 2(\alpha+\omega_r t) - \sqrt{2}\sin 2\beta\cos(\alpha+\omega_r t)\right\} \qquad (2.21)$$

This equation shows that the average value of the dipolar interaction over each rotor cycle is zero. If there are no ^{15}N π pulses, the ^{13}C-^{15}N dipolar coupling has no effect on the detected ^{13}C signal amplitude. At the end of the period, the ^{13}C isotropic chemical shift is refocused by the ^{13}C π pulse. The local field has a positive average value for the first half of each rotor period and a negative average value for the second half of each rotor period caused by the spatial modulation of the dipolar interaction under MAS. Consequently, there is dipolar dephasing of ^{13}C spins by ^{15}N during the first half of the rotor period, but this dephasing is reversed during the second half of the rotor period, resulting in an averaged value of zero for the overall dipolar interaction of each rotor period. This ^{15}N pulse free spectrum S_0 has full signal and only accounts for the ^{13}C T_2 decay.

In the second experiment for the S spectrum acquisition, π pulses are applied to both the ^{13}C and ^{15}N spins, with ^{15}N π pulses applied every half rotor period. The dipolar coupling will be reintroduced by changing the relative sign of the local dipolar field. This is the spin modulation by the additional ^{15}N π pulses, since the dipolar coupling is a product of the spatial and spin functions. Then the average local field experienced by a ^{13}C spin is now positive during the second half of the rotor period and thus is summed to be positive for the first and second rotor periods, not zero as before, resulting in net dipolar dephasing. Since the averaged dipolar field for a ^{13}C spin is positive, the ^{13}C spin magnetization does not refocus due to the reintroduced ^{13}C-^{15}N dipolar coupling and this S spectrum has reduced signal in the REDOR experiment.

The net dephasing angle for one rotor period can be calculated from the $\omega_d = \pm d(t)$.

$$\phi = \int_0^{t_r/2}\omega_d(t)dt - \int_{t_r/2}^{t_r}\omega_d(t)dt \qquad (2.22)$$

The negative sign for the second term is due to the ^{15}N π pulse. The dipolar evolution for one rotor period is as follows and the phase accumulation is the same for every rotor period.

$$\phi = \frac{t_r d}{\pi}\sqrt{2}\sin 2\beta \sin\alpha \tag{2.23}$$

Then for N_c rotor periods, the total phase accumulation is calculated by

$$\phi = \frac{N_c t_r d}{\pi}\sqrt{2}\sin 2\beta \sin\alpha \tag{2.24}$$

The reduced signal in S spectrum is

$$S = S_i e^{-N_c t_r / T_2}\cos\phi \tag{2.25}$$

The ratio of the reduced (S) and full (S_0) signals for a single ^{13}C-^{15}N spin pair is

$$S / S_0 = \cos\phi \tag{2.26}$$

For a powder sample, all values of α and β must be summed over all inter-nuclear orientations, yielding the powder averaged S/S_0 value in REDOR experiment with respect to the geometrical weighting factor $\sin\beta$.

$$S/S_0 = \frac{1}{\pi}\int_0^{\pi}\int_0^{\pi}\cos\phi \sin\beta d\alpha d\beta \tag{2.27}$$

2.2.2 Spin Diffusion

2.2.2.1 ^{1}H spin diffusion

Two types of ^{1}H spin diffusion experiments were applied in this work: the ^{1}H-^{13}C HETCOR(*9, 10*) and the CHHC experiment(*11*). The first method probes the topology of membrane protein based on the proximity of the protein to the membrane or water. The CHHC experiment provides long-range distance restraints between segments of proteins which are not available from the direct ^{13}C-^{13}C transfer pathway due to weak ^{13}C-^{13}C homonuclear dipolar couplings.

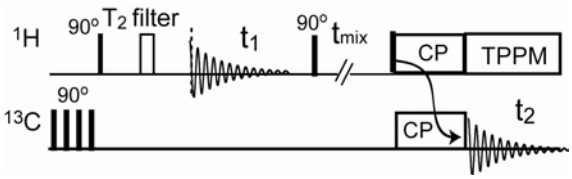


Figure 2. The pulse sequence of the 2D ^{1}H-^{13}C HETCOR experiment.

The ^{1}H-^{13}C HETCOR experiment (Figure 2) starts with a ^{1}H T_2 filter ranging from 0.8 ms to 2 ms which selects the mobile components and destroys the rigid protein ^{1}H magnetization. The ^{1}H signals of water and lipid in the membrane are encoded in the ω_1 dimension and are allowed to transfer to the protein during the mixing time $\mathbf{t_{mix}}$. Then protein ^{1}H signal will be transferred to protein ^{13}C spins by cross polarization (CP) and the ^{13}C signal will be detected. A series of mixing times was applied from 1 ms to 225 ms in this work to yield build up curves for quantitative distance analysis.

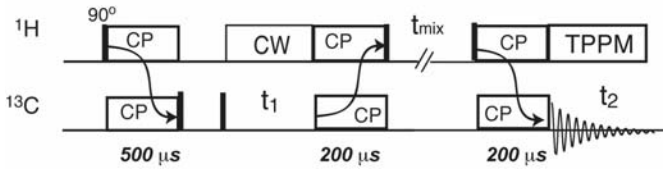


Figure 3. The pulse sequence of the 2D CHHC experiment. The z filter time is normally 3 ms and the CP contact times are indicated in the diagram.

In the CHHC experiment, the ^{13}C spins are polarized by CP and the z filter formed by two ^{13}C 90 degree pulses removes all transverse ^{1}H magnetization. After the ^{13}C t_1 evolution, the protein ^{13}C signal is transferred to protein ^{1}H by a short CP of 200 μs, followed by a ^{1}H-^{1}H spin diffusion period where medium-to-long range contacts are established. Then the protein ^{1}H magnetization is transferred back to the protein ^{13}C spins to be detected in t_2. The advantage of this experiment is to use the strong ^{1}H-^{1}H homonuclear dipolar coupling to provide longer distances within short mixing times. The disadvantage of this experiment is that the sensitivity is reduced by the three CP steps and it requires high ^{13}C labeling levels.

2.2.2.2 ^{13}C spin diffusion

A robust ^{13}C spin diffusion technique applied in this work is the ^{1}H driven ^{13}C spin diffusion (PDSD) experiment(*12*) which provides the correlations mainly between directly bonded ^{13}C-^{13}C spins (Figure 4).

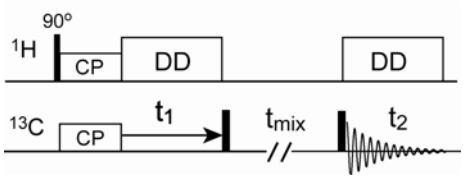


Figure 4. The pulse sequence of the 2D ^{13}C-^{13}C PDSD experiment.

In the PDSD experiment, the ^{13}C spins are polarized by CP and the ^{13}C chemical shifts are encoded in the ω_1 dimension. Then two ^{13}C 90° pulses keep the ^{13}C magnetization along z for spin diffusion to take place with the help of ^{1}H-^{1}H homonuclear dipolar coupling and ^{1}H-^{13}C heternuclear dipolar coupling, followed by the ^{13}C detection. Since there is no proton decoupling during t_{mix}, there is no strong restriction about the length of the mixing time except for the T_1 relaxation consideration and the typical t_{mix} used in this work is 30 to 40 ms. This experiment is straightforward to set up and provides both the bonded and non-bonded information useful for amino acid type assignment.

2.2.2.3 ^{19}F spin diffusion

The pulse sequence of the ^{19}F centerband-only detection of exchange (CODEX) experiment (*13*) is shown in Figure 5. Two rotor-synchronized π pulse trains recouple the CSA interaction under MAS. During the mixing time (τ_m), ^{19}F spin diffusion changes the chemical shift frequency and prevents complete refocusing of the stimulated echo. To correct for the T_1 relaxation during t_m, a z-filter (τ_z) is added after the second π pulse train. Two experiments are carried out: the one with the desired τ_m and a short τ_z (10 μs) is called the exchange experiment (S) and the other one is the reference experiment (S_0) with the interchanged τ_m and τ_z. The normalized intensity of S/S_0 is measured as a function of mixing time until it reaches a plateau.

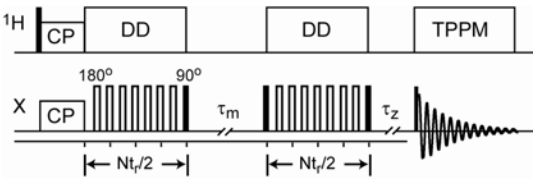


Figure 5. The pulse sequence of the CODEX experiment.

The CODEX experiment is a useful tool for elucidating the oligomeric structure of membrane proteins because it provides both spin counting and intermolecular distances from the same spin diffusion curve(*14*). Between the two π pulse trains, ^{1}H driven ^{19}F spin diffusion occurs between orientationally inequivalent spins, changing the CSA frequency, therefore the echo cannot be completed at the end of the second π pulse train. At long exchange mixing times, the initial magnetization is equally distributed among *n* orientations in the cluster, reducing the CODEX echo intensity to *1/n*. Thus the equilibrium value of the exchange intensity gives the oligomeric number of proteins. At the same time, intermolecular distances can be extracted from the time-dependent CODEX intensity decay based on the fact that the homonuclear ^{19}F-^{19}F dipolar coupling that drives the magnetization exchange scales with distance as $1/r^3$.

2.3 Spin Diffusion Theory

As shown above, spin diffusion experiments are robust and powerful for membrane protein structure determination and the theory will be reviewed below.

2.3.1 Principles

As typical exchange experiments, spin diffusion experiments consist of an selection period or evolution, a mixing time and a detection period(*5*). During the selection period, a spatially inhomogeneous distribution of z magnetization is created to prepare for spin diffusion. Generally, the magnetization of the desired component or components will be selected and the rest will be depleted. This magnetization difference drives the exchange process, and the source magnetization will be transferred along the gradient to the sink. During this process, the integral magnetization redistribution can be monitored by NMR

spectra through the sink component detection. The efficiency of this exchange process provides information about the domain size or the proximity between the source and the sink, and the intermolecular distances may be extracted from the spin diffusion rate.

2.3.2 Selection by Differences

In order to create a magnetization difference during the selection period, various properties can be explored. The magnetization selection can be based on differences in T_2 or T1ρ relaxation times (as in the 2D ^{1}H-^{13}C HETCOR experiment), or based on the CSA (as in the CODEX experiment) caused by different molecular orientations, or based on isotropic chemical shifts (as in PDSD experiment) caused by different chemical environments. (*5*).

2.3.3 Analysis of ^{1}H spin diffusion in three dimensions

The treatment summarized here mathematically describes the spin diffusion process in a model system of membrane-protein and water system as an example, but the principle is general. The diffusion equation for z magnetization $M(\vec{r},t_m)$ is described in (2.28)(*5*), assuming a semi-infinite two-phase system (A and B phases) with a uniform ^{1}H spin density.

$$\frac{\partial M(\vec{r}t_m)}{\partial t_m} = \vec{\nabla}\cdot\left\{D(\vec{r})\vec{\nabla}M(\vec{r},t_m)\right\} = \frac{\partial}{\partial x}\left\{D(\vec{r})\frac{\partial}{\partial x}M(\vec{r},t_m)\right\} + \frac{\partial}{\partial y}\left\{D(\vec{r})\frac{\partial}{\partial y}M(\vec{r},t_m)\right\} + \frac{\partial}{\partial z}\left\{D(\vec{r})\frac{\partial}{\partial z}M(\vec{r},t_m)\right\}$$

(2.28)

For this two-phase system, the magnetization is initially homogenously distributed in phase A, and zero in phase B. The complete exchange intensity $I_B(t_m\rightarrow\infty)$ is at the 100% level. For short mixing times, everywhere along the A-B interface the dependence of the magnetization density $M(x_\perp,t_m)$ on the spatial coordinate $x_\perp$ perpendicular to the interface is as follows:

$$I_B(t_m) = \int_B M(\vec{r},t_m)d^3r \cong S_{A-B}^{tot}\int_0^\infty M(x_\perp,t_m)dx_\perp \qquad (2.29)$$

S_{A-B}^{tot} is the total area of all interfaces between the phase A and B. For spatially constant diffusivity D, the magnetization profile perpendicular to the interface is

$$M(x_{\perp}, t_m) \cong M_0 / 2erfc(x_{\perp} / \sqrt{4Dt_m}) \tag{2.30}$$

If t_m is sufficiently short and the domains are semi-infinite, equation (2.30) is a solution of equation (2.28). With $\int_0^{\infty} erfc(\tilde{x})d\tilde{x} = 1/\sqrt{\pi}$, we obtain

$$I_B(t_m) = S_{A-B}^{tot} M_0 \sqrt{Dt_m / \pi} \tag{2.31}$$

The complete-exchange intensity is given by equation (2.32).

$$I_B(t_m \to \infty) = f_B I^{tot} = f_B M_0 f_A V^{tot} \tag{2.32}$$

Combining (2.31) and (2.32),

$$\frac{I_B(t_m)}{I_B(t_m \to \infty)} = \sqrt{Dt_m / \pi} \frac{1}{f_A f_B} \frac{S_{A-B}^{tot}}{V^{tot}} - O(\sqrt{t_m^2}) \tag{2.33}$$

Assuming the quadratic term $O(\sqrt{t_m^2})$ can be ignored, $f_A \approx 1$ and f_B is small, equation (2.33) is reduced to

$$\frac{I_B(t_m)}{I_B(t_m \to \infty)} \approx \sqrt{\frac{Dt_m}{\pi}} \frac{S_{A-B}^{tot}}{V_B} \tag{2.34}$$

If the system is the water-protein magnetization transfer system, the diffusion of magnetization is described in (2.35), replacing D with the effective diffusion coefficient D_{eff}.

$$\frac{M_P(t_m)}{M_P(t_m \to \infty)} \approx \sqrt{\frac{D_{eff}}{\pi}} \frac{S_{WP}}{V_P} \sqrt{t_m} \tag{2.35}$$

Therefore at the limit of $M_P(t_m^s) = M_P(t_m \to \infty)$

$$\sqrt{t_m^s} = \sqrt{\frac{\pi}{D_{eff}}} \frac{V_P}{S_{WP}} \tag{2.36}$$

Here $\sqrt{t_m^s}$ is proportional to the volume to surface ratio and is related to the molecular dimension of the protein. t_m^s can be experimentally extracted from the initial transfer efficiency for the water-protein magnetization exchange system. From this link, a 3D protein model with proper molecular dimension can be built to describe the water-protein magnetization transfer process, where the water magnetization is allowed to transfer to the protein within the three-dimensional spin network(*15*) as follows:

$$M_{x,y,z}(t_m + \Delta t_m) = M_{x,y,z}(t_m) + \sum_i \frac{D_{ij}\Delta t_m}{d^2}(M_i(t_m) - M_{x,y,z}(t_m)) \qquad (2.37)$$

Here $M(t_m)$ monitors the magnetization of the six neighboring cubes (x+1,y,z; x-1,y,z; x, y+1,z; x,y-1,z; x,y,z+1; x,y,z-1) and the D_{ij} stand for the corresponding diffusion coefficients. The indirect magnetization transfer pathway from water to lipids to the protein is neglected to first approximation since the magnetization transfer via the lipid is a slow process due to the small water-lipid, lipid-lipid and lipid-protein diffusion coefficients. We treat the lipid as an inert system and the lipid magnetization remains zero at all times. The magnetization of water is kept to be 1 in the whole simulation, assuming a large water pool and fast water bulk diffusion. With the magnetization of all the other components set to 0 initially, the water magnetization is allowed to transfer within the spin network. The relative protein magnetization was integrated and read out every 250 microseconds from the simulation. The relevant diffusion coefficients are the water-protein diffusion coefficient D_{WP} of 0.008 nm^2/ms(*15*), and the protein-protein diffusion coefficient D_{PP} of 0.3 nm^2/ms(*16*). When the proper protein model is built with a good estimation of the channel size, the simulation curve will fit the experiment data.

Another independent alternative treatment is also carried out, to verify the limit of validity of the infinite water reservoir simulation. In this second treatment, the water magnetization starts as 1 and is allowed to change along as spin diffusion to the protein and lipids without restrictions. The water-protein system is set to be periodic in the z dimension. The water layers are set to be 12 nm thick while each bilayers is 4.4 nm thick. The simulation

result is identical to the earlier calculation described in the paragraph above. The water-water diffusion coefficient D_{WW} of 3 nm^2/ms(*16*) gives reasonable fit to the experiment data.

References

(1) Andrew, E. R., Bradbury, A., and Eades, R. G. (1959) Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation. Nature 183, 1802-1803.

(2) Lowe, I. J. (1959) Free Induction Decays of Rotating Solids. Phys. Rev. Lett. 2, 285-287.

(3) Hong, M., and Jakes, K. (1999) Selective and extensive C-13 labeling of a membrane protein for solid-state NMR investigations. J. Biomol. NMR 14, 71-74.

(4) Hong, M. (1999) Determination of multiple phi-torsion angles in proteins by selective and extensive C-13 labeling and two-dimensional solid-state NMR. J. Magn. Reson. 139, 389-401.

(5) Schmidt-Rohr, K., and Spiess, H. W. (1994) Multidimensional Solid-State NMR and Polymers, 1 ed., Academic Press, London.

(6) Duer, M. J. (2004) Introduction to Solid-State NMR Spectroscopy, 1 ed., Blackwell Science, Oxford.

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

(8) Gullion, T., and Schaefer, J. (1989) Rotational echo double resonance NMR. J. Magn. Reson. 81, 196-200.

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

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

(11) Lange, A., Luca, S., and Baldus, M. (2002) Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. J. Am. Chem. Soc. 124, 9704-9705.

(12) Luo, W. B., Yao, X. L., and Hong, M. (2005) Large structure rearrangement of colicin la channel domain after membrane binding from 2D C-13 spin diffusion NMR. J. Am. Chem. Soc. 127, 6402-6408.

(13) Schmidt-Rohr, K., deAzevedo, E. R., and Bonagamba, T. J. (2002) Centerband-Only Detection of Exchange (CODEX): Efficient NMR Analysis of Slow Motions in Solids, in Encyclopedia of NMR (Grant, D. M., and Harris, R. K., Eds.), John Wiley & Sons, Chichester.

(14) Luo, W., and Hong, M. (2006) Determination of the oligomeric number and intermolecular distances of membrane protein assemblies by anisotropic H-1-driven spin diffusion NMR spectroscopy. J. Am. Chem. Soc. 128, 7242-7251.

(15) Ader, C., Schneider, R., Seidel, K., Etzkorn, M., Becker, S., and Baldus, M. (2009) Structural Rearrangements of Membrane Proteins Probed by Water-Edited Solid-State NMR Spectroscopy. J. Am. Chem. Soc. 131, 170-176.

(16) Huster, D., Yao, X. L., and Hong, M. (2002) Membrane Protein Topology Probed by 1H Spin Diffusion from Lipids Using Solid-State NMR Spectroscopy. J. Am. Chem. Soc. 124, 874-883.

Chapter 3

Large Structure Rearrangement of Colicin Ia Channel Domain after Membrane Binding from 2D ^{13}C Spin Diffusion NMR

A paper published in Journal of the American Chemical Society
2005 vol. 127 (17) pp. 6402-6408

Wenbin Luo, Xiaolan Yao and Mei Hong

Abstract

One of the main mechanisms of membrane protein folding is by spontaneous insertion into the lipid bilayer from the aqueous environment. The bacterial toxin, colicin Ia, is one such protein. To shed light on the conformational changes involved in this dramatic transfer from the polar to the hydrophobic milieu, we carried out 2D magic-angle spinning ^{13}C NMR experiments on the water-soluble and membranebound states of the channel-forming domain of colicin Ia. Proton-driven ^{13}C spin diffusion spectra of selectively ^{13}C-labeled protein show unequivocal attenuation of cross-peaks after membrane binding. This attenuation can be assigned to distance increases but not reduction of the diffusion coefficient. Analysis of the statistics of the interhelical and intrahelical ^{13}C-^{13}C distances in the soluble protein structure indicates that the observed cross-peak reduction is well correlated with a high percentage of short interhelical contacts in the soluble protein. This suggests that colicin Ia channel domain becomes open and extended upon membrane binding, thus lengthening interhelical distances. In comparison, cross-peaks with similar intensities between the two states are dominated by intrahelical contacts in the soluble state. This suggests that the membrane-bound structure of colicin Ia channel domain may be described as a "molten globule", in which the helical secondary structure is retained while the tertiary structure is unfolded. This study demonstrates that ^{13}C spin diffusion NMR is a valuable tool for obtaining qualitative long-range distance constraints on membrane protein folding.

Introduction

The determination of the three-dimensional structure of solid proteins by NMR has made dramatic progress in the last few years due to advances in the preparation of ordered and microcrystalline proteins [1], multi-dimensional resonance assignment techniques [2-5], and distance [6] and torsion angle [7,8] determination methods for uniformly ^{13}C, ^{15}N-labeled proteins [9,10]. Short-range distances (< 5 Å) that restrain both backbone and local sidechain conformations can now be measured accurately. However, long-range distances (>> 5 Å) are still difficult to determine, calling for development of methods to measure dipolar couplings between spins with high gyromagnetic ratios [11,12].

In fortunate cases, information on very long distances may be deduced unambiguously using semi-quantitative approaches such as spin diffusion. We show here that a classical 2D ^{13}C correlation technique, ^{1}H-driven ^{13}C spin diffusion (PDSD), yields fresh new insight into the three-dimensional fold of a large membrane protein, the channel-forming domain of colicin Ia (25,082 Da). A bacterial toxin, colicin Ia exerts its toxic effect by spontaneously inserting into the inner membrane of sensitive bacteria cells from the aqueous environment and opening a voltage-gated channel that depletes the membrane potential of the cell [13]. How the protein changes its structure to adapt to both the polar and hydrophobic milieu is a fascinating and fundamental biophysical problem. The water-soluble structure of colicin Ia has been determined by X-ray crystallography [14], and shows the channel domain to be a compact ten α-helix globule. However, the high-resolution structure of the membrane-bound (MB) state is still unknown. Our earlier ^{13}C chemical shift analysis indicated that the protein backbone largely preserves its α-helical conformation upon membrane binding [15]. Various biophysical measurements indicated that homologous colicins adopt an open conformation upon membrane binding, with a hydrophobic helical hairpin well embedded in the lipid bilayer [16-23]. A schematic of this large conformational change is shown in Fig. 1. However, to date, no direct distance measurements for the unfolding of colicin Ia channel domain has been reported.

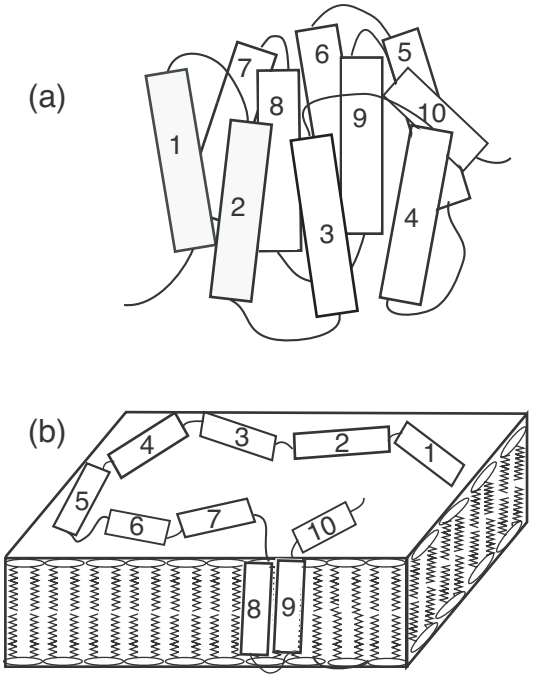


Figure 1. Cartoons of protein refolding from a globular structure with hydrophobic residues inside (a), to an open topology with the hydrophobic residues exposed to the lipid molecules (b).

Here we report 2D ^{13}C PDSD data that support an extended topology for the membrane-bound state of colicin Ia channel domain. We found that inter-residue cross peaks are significantly attenuated in the membrane-bound protein compared to the soluble protein. Moreover, the intensity reduction is more pronounced for inter-residue cross peaks with a high fraction of inter-helical contacts in the soluble state. Given the similar secondary structure of the two states of the protein, this suggests that membrane binding lengthens the inter-helical distances while retaining the intra-helical distances.

Material and Methods

Preparation of soluble and membrane-bound colicin Ia channel domain

All ^{13}C and ^{15}N labeled compounds, including $^{15}NH_4Cl$, ^{15}N-Glu, ^{15}N-Gln, [1,6-^{13}C] glucose, and [U-^{13}C] glucose, were purchased from Cambridge Isotope Laboratory (Andover, MA). His_6 tagged colicin Ia channel domain was expressed from pKSJ120-containing *E. coli*

BL21 (DE3) cells in a modified M9 medium containing appropriate isotopic labels, and was purified by His-bind metal chelation resin (Novagen) as described before [15,24]. TEASE protocol was used for isotopic labeling [24], where the labeled ^{13}C and ^{15}N precursors are supplemented with ten unlabeled amino acids from the citric acid cycle (Glu, Gln, Arg, Pro, Asn, Asp, Lys, Ile, Met, Thr). In this way, only the amino acids from the glycolysis pathway and the pentose phosphate pathway (Gly, Ser, Cys, His, Ala, Val, Leu, Trp, Phe, Tyr) are labeled. This simplifies the ^{13}C spectra and allows straightforward peak assignment to the amino acid type. The ^{13}C precursor was either [1, 6-^{13}C] glucose or [U-^{13}C] glucose. For the former, the labeled ^{13}C sites are Ala Cβ, Leu Cα, Cδ1, Cδ2, Ser Cβ, Cys Cβ, Val Cγ1, Cγ2, His Cδ2, C', Phe Cβ, Cγ, Cδ1, Cδ2, Tyr Cβ, Cδ1, Cδ2, Cζ, and Trp Cβ, Cδ2, Cε1, Cε2 [25]. This sample is called 1,6-colicin. The main ^{15}N precursor was ^{15}N-ammonium chloride, supplemented with ^{15}N-Glu and ^{15}N-Gln to reduce dilution of the ^{15}N labeling level by the transamination reaction. The yield of the protein was 20 - 30 mg/L.

The soluble colicin sample was prepared by packing dry, lipid-free, colicin directly into a 4-mm magic-angle spinning (MAS) rotor, then hydrating it to 35% water by mass. The membrane-bound protein sample was prepared by mixing the protein solution with vesicle solutions of POPC and POPG lipid mixtures (Avanti Polar Lipids, Alabaster, AL) to achieve a protein-lipid molar ratio of 1:100. The molar ratio of neutral POPC to anionic POPG lipids was 3:7. The lipid solution, in a citrate buffer of pH 4.8 (0.3 M KCl, 10 mM citrate), was extruded across polycarbonate filter membranes of 100-nm diameter [26] to produce large unilamellar vesicles before protein binding. The mixed proteoliposome solution was ultracentrifuged at 150,000 g for 2 hours using a Beckman swinging-bucket rotor (SW60 Ti) to obtain a membrane pellet. The supernatant contained less than 5% unbound protein, as measured by a photometric assay [27]. The membrane protein pellet was lyophilized, packed into a 4 mm MAS rotor, and hydrated to 35% water by mass.

Solid-state NMR

2D PDSD experiments were carried out on a Bruker DSX-400 spectrometer (Karlsruhe, Germany) operating at a resonance frequency of 100.71 MHz for ^{13}C. A double-

resonance MAS probe with a 4-mm spinning module was used. The ^{1}H radiofrequency (rf) field strengths for heteronuclear TPPM decoupling [28] were 70 kHz. ^{1}H and ^{13}C 90º pulse lengths were typically 3.5 μs and 5.0 μs, respectively. Cross polarization (CP) contact time was 0.2-0.5 ms. Spinning speeds were 5 kHz for 1,6-colicin and 6 kHz for the partial uniformly labeled colicin (pU-colicin) to avoid rotational resonance effects [29]. The spectra were collected with 160 - 320 scans per t_1 slice, a spectral window of 20 kHz for the indirect dimension, and a maximum t_1 evolution time of 6.8 ms. The ^{13}C mixing times ranged from 50 ms to 500 ms. For the ^{1}H spin diffusion experiment, the second and third ^{1}H-^{13}C CP contact time was 200 μs. The ^{1}H spin-diffusion mixing times ranged from 150 μs to 350 μs.

Results and Discussion

Equivalence of aligned and powder samples for uniaxially mobile molecules

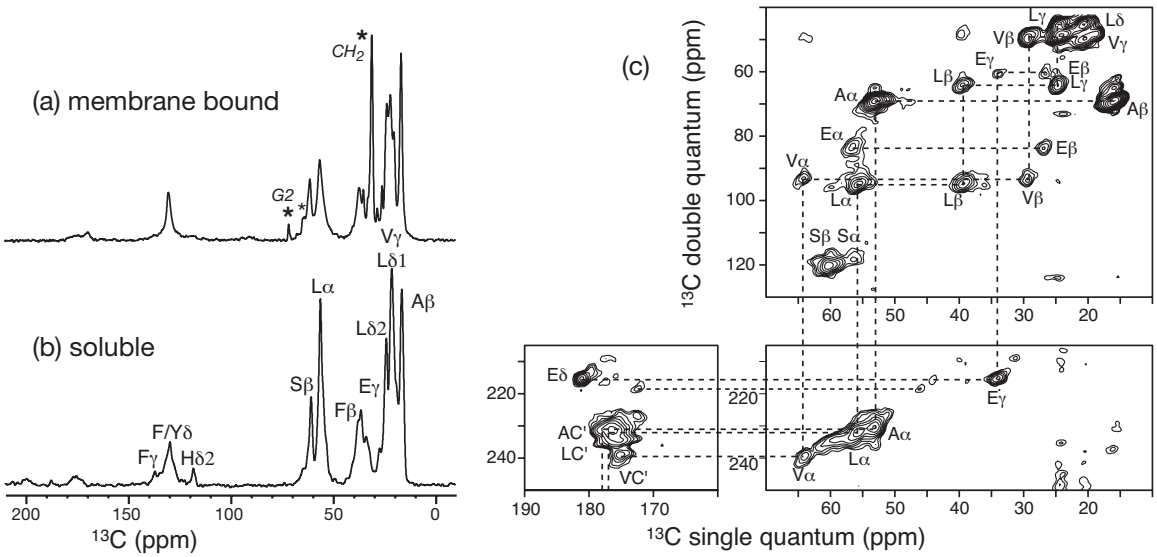


Figure 2. ^{13}C assignment of colicin Ia channel domain. (a-b) 1D ^{13}C CP-MAS spectra of 1,6-colicin in the (a) membrane-bound and (b) soluble state. Lipid peaks are indicated by asterisks. (c) 2D dipolar INADEQUATE spectrum of soluble pU-colicin.

Fig. 2(a-b) shows the 1D ^{13}C CP-MAS spectra of the membrane-bound and soluble 1,6-colicin. The membrane-bound protein (a) does not exhibit excessively high lipid peaks due to the use of a short CP contact time, 200 μs. The most visible lipid background signal is the 32 ppm $(CH_2)_n$ peak, which is resolved from the protein signals. Fig. 2c shows the 2D INADEQUATE spectrum of the soluble pU-colicin. The ^{13}C peaks were readily assigned to amino acid types based on the characteristic chemical shifts and the connectivity pattern in the 2D spectrum. The clean suppression of the citric-acid-cycle amino acids was demonstrated by the weakness of the carbonyl and several sidechain signals. Quantitative analysis of 1D direct-polarization spectrum (Supporting Information Fig. S1) indicates that the $^{13}C'$ intensity is consistent with the predicted labeling of only His residues in the protein. If the citric-acid cycle is not suppressed, a much higher $^{13}C'$ peak of twice the Leu Cα signal would be expected. This was not observed. For methyl-containing Thr, if scrambling occurred, its Cβ site (~67 ppm) would be labeled at a higher level than Cγ (~20 ppm) and thus two correlation peaks would be expected in the 2D INADEQUATE spectrum. No such peaks were observed. Similarly, for the only other methyl-containing amino acid of the citric acid cycle, Ile, the lack of Cγ1-Cγ2 correlation peaks in the 2D spectrum and the lack of a resolved Cδ signal (~10 ppm) in the 1D spectrum rule out scrambling. Thus, the methyl carbon signals of interest for the spin diffusion analysis below result purely from three glycolysis amino acids: Leu, Val and Ala.

Fig. 3(a-b) compares the room-temperature 2D PDSD spectra of 1,6-colicin between the soluble and the membrane-bound states after a mixing time of 400 ms. It can be seen that many cross peaks present in the soluble protein spectrum are missing in the membrane protein spectrum. For example, the F/Yδ-Lδ1/Vγ1 peak (129.5 ppm, 21.4 ppm) and the F/Yβ-Lδ1/Vγ1 peak (36.2 ppm, 21.4 ppm) disappeared in the membrane protein spectrum. The 1D cross sections, taken from the average of the column and row at the corresponding frequencies, are shown for F/Yδ 129.5 □□□, F/Yβ 36.2 □p□, and Aβ 16.0 □□□ (Fig. 3d-f). They indicate that the changes in the cross peak intensities are often an unambiguous yes (for soluble colicin) and no (for membrane-bound colicin) situation.

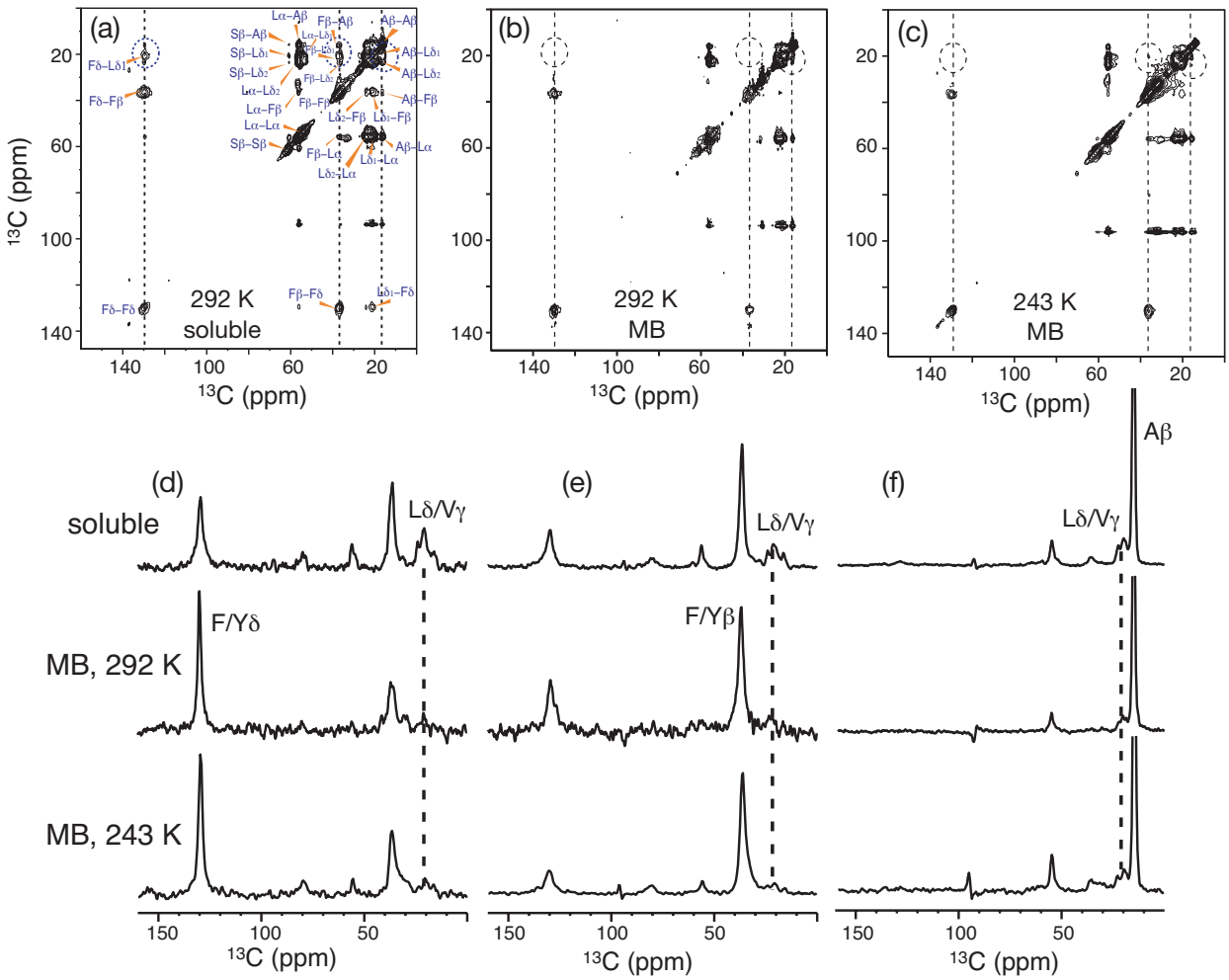


Figure 3. 2D PDSD spectra of 1,6-colicin with a mixing time of 400 ms. (a) Soluble protein, with assignments indicated. (b-c) Membrane-bound (MB) protein at 292 K and 243 K. Dashed circles highlight cross-peak intensity differences. (d-f) Selected cross sections from the 2D spectra of the soluble protein (top row), the room-temperature membrane protein (middle row), and the low-temperature membrane protein (bottom row). The cross sections are the average of the corresponding column and row. (d) F/Yδ slice at 129.5 ppm. (e) F/Yβ slice at 36.2 ppm. (f) Aβ slice at 16.0 ppm. Note the significant drop of the cross peak intensities of the membrane bound protein at both temperatures.

The reduction of cross peak intensities can result from reduction of the spin diffusion coefficients and/or from distance increases in the membrane-bound state. The main factors

influencing the diffusion coefficient are ^{1}H decoupling and motion, since both reduce the ^{1}H-^{1}H and ^{1}H-^{13}C dipolar couplings, which in turn affect the overlap between the two ^{13}C single-quantum signals. This overlap is described by the integral $f_{ij}(0) = \int d\omega f_i(\omega) f_j(\omega)$ [30], where $f_{i,j}(\omega)$ is the single-quantum spectrum of each peak. A well-known example of motion-facilitated ^{13}C spin diffusion through an increase of the overlap integral is adamantane [31], a plastic crystal with large-amplitude motion that permitted the observation of spin diffusion even among ^{13}C sites at natural abundance. Our previous C-H dipolar coupling measurements showed that membrane-bound colicin exhibits larger-amplitude segmental motions than the soluble state due to thermal motions of the lipid bilayer. The C-H order parameters were 0.88 - 0.93 for the backbone and 0.60 - 0.75 for the sidechain, which are smaller than the order parameters of 0.97 – 1.0 for the backbone and 0.87-0.88 for the sidechain in the soluble state [32]. For ^{13}C chemical shift differences of 2 – 10 kHz (20 – 100 ppm), this enhanced although still moderate segmental motion may actually increase the overlap integral and thus facilitate spin diffusion. Indeed, for intra-residue cross peaks with fixed ^{13}C-^{13}C distances, the cross-peak intensities are higher in the membrane-bound state than in the soluble state (see below), indicating that the segmental motions of the membrane-bound protein increases rather than decreases the ^{13}C diffusion coefficient.

To eliminate the possibilities of other motional factors such as aromatic ring flips complicating the distance analysis, we carried out the 2D ^{13}C spin diffusion experiment for the membrane-bound colicin at 243 K, well below the gel-to-liquid crystalline phase transition temperature of the lipids (271 K). 2D ^{13}C-^{1}H LG-CP spectrum of the membrane-bound colicin Ia channel domain at 243 K shows either the same or even slightly larger dipolar couplings than the soluble protein at room temperature (Supporting Information Fig. S2). Thus, direct distance comparison can be made between the membrane protein at 243 K and the soluble protein at room temperature. The low-temperature PDSD spectrum of the membrane-bound colicin Ia channel domain is shown in Fig. 3c. The attenuation of inter-residue cross peaks persisted at low temperature. Thus, the cross-peak reduction can indeed be attributed to distance increases in the membrane-bound protein.

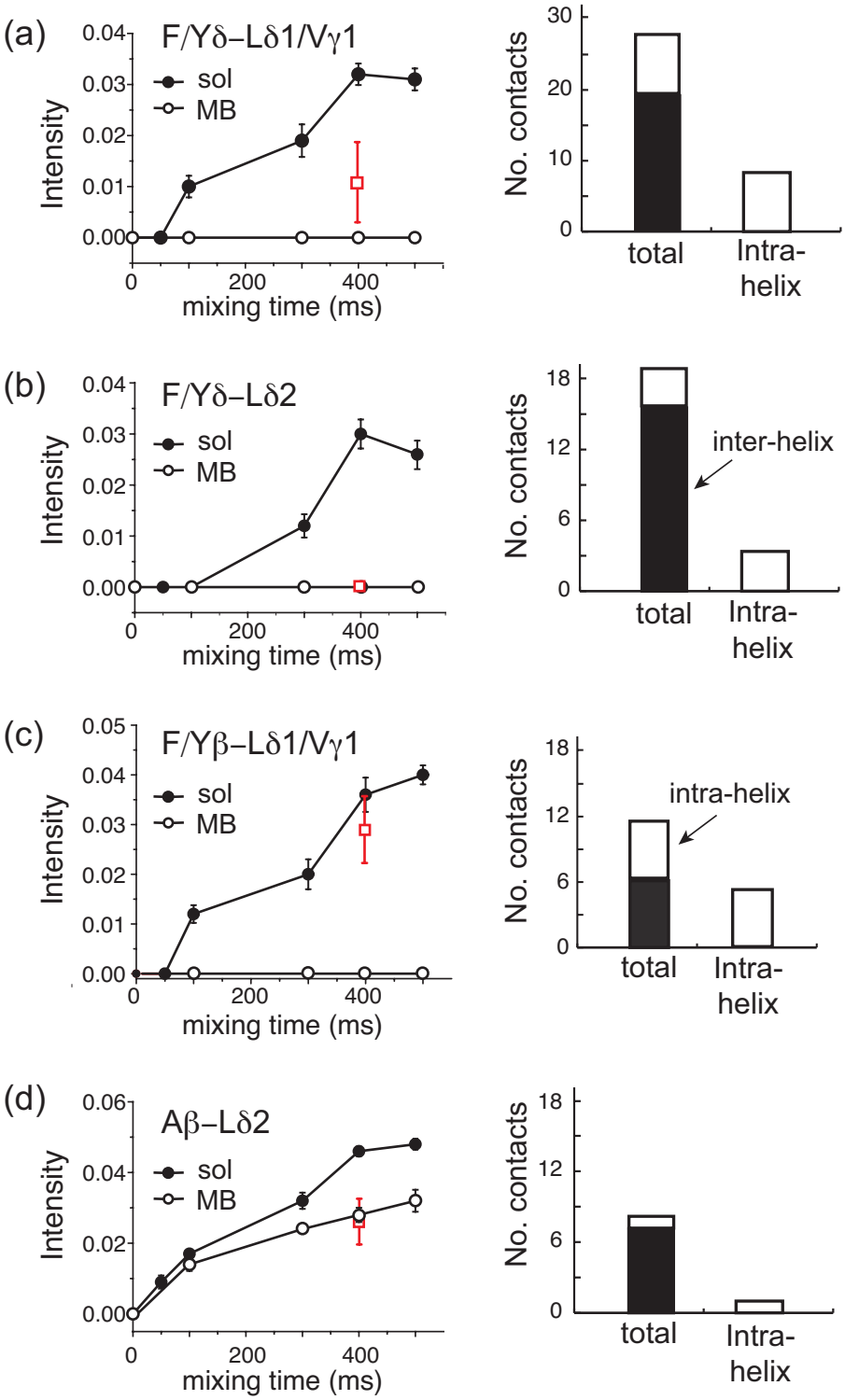


Figure 4. ^{13}C spin diffusion buildup curves for inter-residue cross peaks (left column) and the corresponding numbers of intra-helical (open) and inter-helical (filled) contacts within

5.5 Å in the soluble colicin structure (right column). The cross peaks are (a) F/Yδ-Lδ1/Vγ1, (b) F/Yδ-Lδ2, (c) F/Yβ- Lδ1/Vγ1, and (d) Aβ-Lδ2. Filled circles represent the soluble colicin data at 292 K. Open circles and red open squares represent the membrane-bound colicin data at 292 K and 243 K, respectively.

To quantitatively compare the cross-peak intensity buildup between the two states of colicin Ia channel domain, we measured a series of 2D PDSD spectra with varying mixing times. The cross peak intensities, which are the sum of the two symmetric peaks ($I_{AB} + I_{BA}$) in the 2D spectra, were normalized by the total intensity of the four rows and columns that contain the cross peaks. Two types of buildup behaviors are observed. In the first, the membrane-bound protein exhibits no detectable or weak signals for all mixing times up to 500 ms, while the soluble protein shows much higher cross peaks that increase with the mixing time. This is the case, for example, for F/Yδ-Lδ2 and Aβ-Lδ2 (Fig. 4b, d). The second type of cross peaks exhibits similar or higher intensities for the membrane-bound colicin. This is the case for Aβ-Lα and F/Yδ-F/Yβpeaks (Fig. 5). The intra-residue F/Yδ-F/Yβ cross peak provides an important control for assessing the diffusion coefficients of the two states of the protein. This intra-residue F/Yδ-F/Yβ cross peak (Fig. 5b) is 5-10 times higher than the inter-residue cross peaks in Fig. 4, consistent with the fact that the two-bond distance is fixed at 2.53 Å, much shorter than any potential inter-residue contacts. Because the distance is fixed, any difference in the cross-peak buildup curves between the soluble and membrane-bound protein must result from difference in the diffusion coefficients. The membrane-bound colicin shows higher F/Yδ-F/Yβ intensities for most mixing times, confirming that the enhanced segmental motions of the membrane protein increases the overlap between the ^{13}C single-quantum lineshapes. Thus, spin diffusion is actually more efficient in the membrane-bound state than in the soluble state, and the reduced cross peaks of the membrane-bound colicin must result from distance increases.

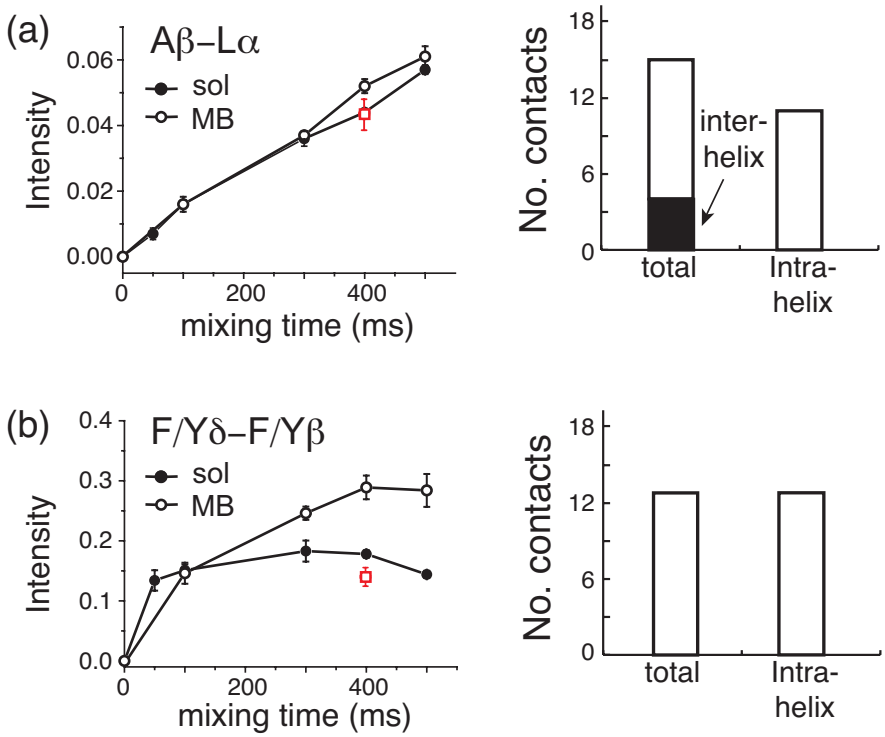


Figure 5. ^{13}C spin diffusion buildup curves (left column) and the numbers of inter-helical and intra-helical contacts within 5.5 Å in the soluble colicin structure. (a) Aβ-Lα. (b) F/Yδ-F/Yβ. Symbols are the same as Figure 4. At room temperature, the membrane-bound protein exhibits similar or faster spin diffusion than the soluble protein due to the predominance of intra-helical contacts and the large diffusion coefficient.

To understand the exact nature of this distance increase, we consider the number of inter-helical and intra-helical contacts within a cut-off distance of 5.5 Å in soluble colicin. For example, 18 Fδ-Lδ2 distances within 5.5 Å were found in the soluble protein structure, 15 of which are inter-helical while 3 are intra-helical. The statistics are shown in Fig. 4 and Fig. 5 as bar graphs for each cross peak. For unresolved peaks such as Phe and Tyr Cδ, and Lδ1 and Vγ1, all possible contacts are considered. A distance of 5.5 Å was estimated to be the upper limit detectable by ^{13}C spin diffusion within 500 ms [9,33]. Increasing the cut-off distance to 7.0 Å did not change the relative weight of the inter-helical and intra-helical contacts. Sidechain-sidechain cross peaks such as F/Yδ-Lδn and Aβ-Lδ2 appear to have

more inter-helical contacts than intra-helical ones, while backbone-sidechain cross peaks such as Aβ-Lα are dominated by intra-helical contacts.

Fig. 4 and Fig. 5 show a clear correlation between the amount of cross-peak intensity reduction and the percentage of inter-helical distances: the larger the fraction of inter-helical contacts, the larger the drop of the cross peak intensity in the membrane-bound state. Since we have previously shown that the helicity of colicin Ia channel domain is not significantly changed by membrane binding [15], this suggests that the intensity decrease results from an increase of the inter-helical distances after membrane binding. Indeed, for the Aβ-Lαpeak (Fig. 5a), which results mostly (11 out of 15) from intra-helical contacts in the soluble protein structure, the buildup curves are similar between the two states, confirming that intra-helical distances are largely unaffected by membrane binding. Combined, these suggest that membrane binding does not change the secondary structure or intra-helical distances of colicin Ia channel domain, but significantly loosens the tertiary structure, lengthening the inter-helical distances.

The Aβ-Lδ2 cross peak (Fig. 4d) shows a smaller gap between the soluble and the MB state, despite the fact that only one out of eight contacts found within 5.5 Å is intra-helical and should have weakened the signal of the MB protein more substantially. This may result from relayed transfer of the Aβ magnetization first to Lα then to Lδ2. Eleven out of fifteen Aβ-Lα contacts are intra-helical (Fig. 5a), thus Aβ-Lδ2 diffusion is likely facilitated by this relayed mechanism.

Figures 4 and 5 also show the 400 ms PDSD cross-peak intensities of the membrane-bound colicin at 243 K. The most dramatic change is seen for the intra-residue F/Yδ-F/Yβ cross peak (Fig. 5b), which now have lower intensity than the soluble protein, in contrast to the room-temperature result. This further confirms the hypothesis that reduced motion decreases the overlap integral and slows down spin diffusion in the membrane-bound colicin. Among the five inter-residue cross peaks examined, four (F/Yδ-Lδ1/Vγ1, F/Yδ-Lδ2, Aβ-Lδ2, Aβ-Lα) are either unaffected or only slightly modified by the temperature change, and the reduction of cross-peak intensity compared to the soluble protein is retained. The only significant deviation from the room-temperature data occurred for the F/Yβ-Lδ1/Vγ1 cross

peak (Fig. 4c), which increased in intensity but is still lower than the soluble protein data. This is consistent with the fact that F/Yβ-Lδ1/Vγ1 has the largest fraction of intra-helical contacts among the aromatic-methyl cross peaks (Fig. 4a-c), and thus the membrane-bound protein retains the largest number of distances within spin-diffusion reach.

Overall, the PDSD data indicates that membrane binding increases the inter-helical distances while retaining the intra-helical ones. This implies an open and extended structure for the membrane-bound colicin Ia channel domain, which is consistent with previous solid-state NMR studies of the secondary structure, dynamics, and topology of colicin Ia channel domain. ^{13}C isotropic and anisotropic chemical shifts of the soluble and membrane-bound colicin Ia channel domain do not differ significantly, suggesting that the helicity of the protein is similar between the two states [15]. However, the membrane-bound protein exhibits much larger-amplitude segmental motions than the soluble protein [32], which indicates that the protein adopts a looser tertiary structure upon membrane binding and is able to interact extensively with the lipid molecules. ^{1}H spin diffusion from lipids and water to protein indicated that a substantial portion of the protein is located at the surface of the bilayer while a small component is deeply embedded in the membrane [22,34]. These data support an "umbrella" model for the membrane-bound colicin Ia channel domain [20,35]. The present ^{13}C spin diffusion data show that the surface of this "umbrella" is extended and possibly away from the hydrophobic "stem", thus reducing the cross peak intensities in the spectra.

The different spin diffusion behavior of the soluble and membrane-bound colicin persists when partial uniform ^{13}C labeling was used (Fig. 6). However, the significant peak overlap makes assignment of inter-residue cross peaks difficult, thus the simplification offered by 1,6-^{13}C labeling is crucial for concluding distance elongation in the membrane-bound protein.

It is interesting to compare ^{13}C spin diffusion with ^{1}H spin diffusion, which is far more efficient due to the 16-fold stronger dipolar coupling. ^{1}H spin diffusion can be detected through ^{13}C, using the CHHC technique developed recently for uniformly ^{13}C labeled proteins [36,37]. The ^{1}H spin diffusion buildup curves for a number of inter-residue cross peaks in the soluble 1,6-colicin are shown in Supporting Information Fig. S3. The cross peak

intensities are either comparable to or higher than the corresponding PDSD cross peaks, thus in principle ^{1}H spin diffusion can detect longer distances. However, the sensitivity of CHHC is much lower, as shown by the larger error bars. This is because only 30% of all carbons in colicin Ia channel domain are labeled with the TEASE [1, 6-^{13}C] glucose scheme. Thus at long ^{1}H mixing times, ~70% of the ^{1}H magnetization is transferred to ^{1}H sites bonded to ^{12}C. Due to this inherent sensitivity limitation, the CHHC technique is not well suited for selectively labeled proteins.

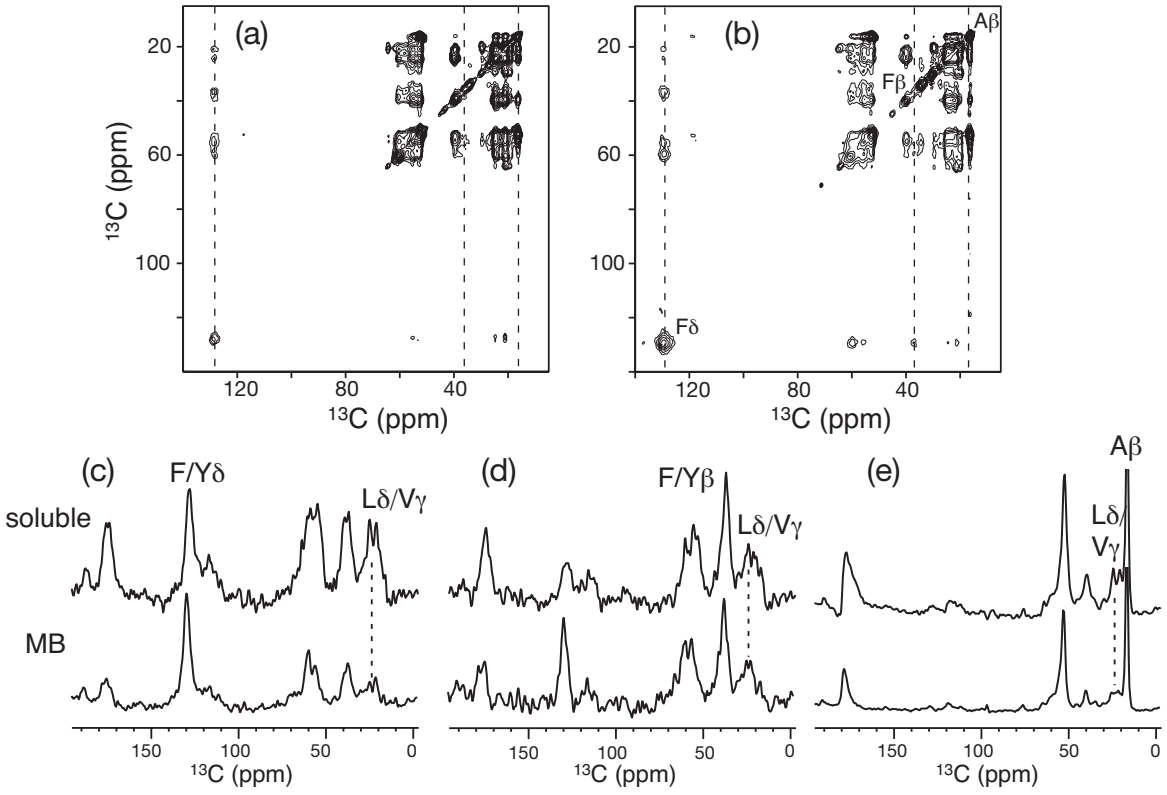


Figure 6. 2D PDSD spectra of pU-colicin in the (a) soluble and (b) membrane-bound state at 292 K. Mixing time: 300 ms. (c-e) ω_1 cross sections from the 2D spectra of the soluble (top row) and the membrane-bound protein (bottom row). (c) F/Yδ slice at 129.5 ppm. (d) F/Yβ slice at 36.2 ppm. (e) Aβ slice at 16.0 ppm.

Fig. 7a shows the crystal structure of the soluble colicin Ia channel domain (PDB accession code: 1CII) and all Fδ-Lδ1 distances within 5.5 Å. A total of 15 distances are

found, a subset of those shown in Fig. 4a. Fig. 7b shows a model of the membrane-bound colicin Ia channel domain to illustrate the possible extension of the protein after membrane binding. The model was obtained by modifying the torsion angles of residues in the loops and turns connecting consecutive helices. This produced an extended helical array [16] for helices 1-7, while putting the hydrophobic helices 8-9 roughly perpendicular to the rest of the protein. This perpendicular orientation is not determined by the current PDSD experiment, but is derived from solid-state NMR ^{1}H spin diffusion [22] and orientation measurements [23], and from trapping of biotinylated residues by trans-side streptavidin in a planar bilayer [38]. The model does not represent the actual three-dimensional structure of the membrane-bound protein, which requires many more distance constraints than currently available; however, it illustrates the dramatic distance increase possible without changing a large number of torsion angles from the soluble protein structure. In this extended helical array, ten of the Fδ-Lδ1 distances increased dramatically, to 15 – 50 Å, while five intra-helical distances remained similar to before. Those residue pairs close in the protein backbone but belong to a helix and its nearest loops are considered intra-helical to avoid skewing the statistics in favor of the inter-helical category. By keeping the intra-helical torsion angles the same as in the soluble protein, we found that the ratio of the lengthened inter-helical distances versus unchanged intra-helical distances is relatively independent of the actual torsion angle changes used for the turn and loop residues.

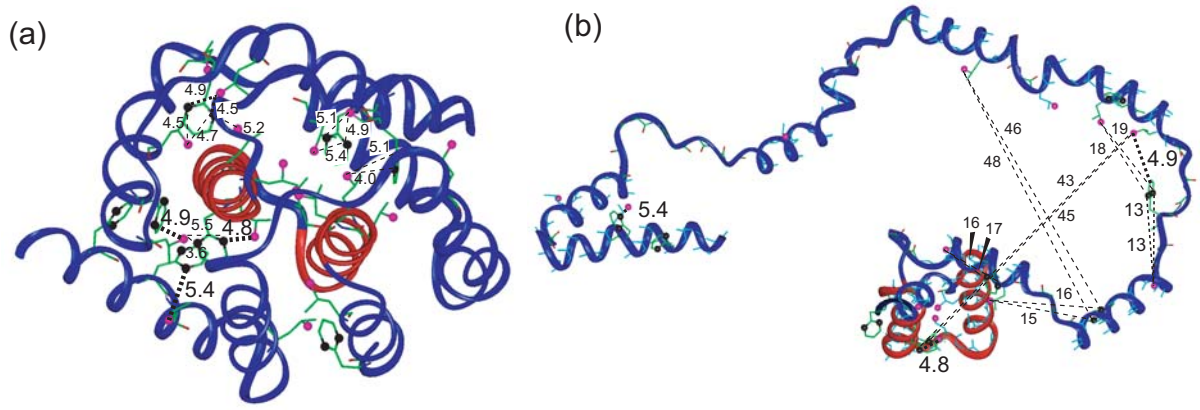


Figure 7. Schematic model of distance elongation and structural rearrangement of colicin Ia channel domain upon membrane binding. (a) Fifteen Fδ – Lδ1 contacts within 5.5 Å are

found in the soluble structure (PDB code: 1CII). (b) Model of the membrane-bound protein, looking down the plane of the bilayer. Only the torsion angles of loop residues are modified. This increased most inter-helical distances (in Å) to above 15 Å.

Site-specific distance constraints for membrane-bound colicins have been previously obtained using fluorescence resonance energy transfer (FRET) from Trp donors to extrinsic fluorescent probes attached to Cys at specific residues. These experiments found clear evidence of distance elongation upon membrane binding of colicin A and colicin E1 [16-18]. Compared to the present 2D solid-state NMR approach, which gives cross peaks for only short distance (< 5.5 Å) contacts, the FRET method probes much longer distances of 15 – 35 Å, and can be site-specific. However, it has the uncertainties of potential perturbation of the extrinsic probe to the protein structure, significant mobility of the probe, and changes of the energy transfer efficiency by the membrane. EPR spectroscopy has also been used to probe conformational changes of colicin E1 channel domain. By monitoring the mobility changes in the EPR spectra of site-directed nitroxide spin labels as a function of time, the insertion depths of the labeled residues and the time course of protein binding to the membrane were inferred [19]. Again, however, a bulky probe is necessary for this technique. Compared to these two methods, spin diffusion NMR requires no extrinsic probe and thus is less perturbing to the protein. It also allows direct internal calibration of cross-peak intensities for membrane-induced mobility changes. Moreover, by better sample preparation protocols and tailored labeling schemes, it is possible to enhance the spectral resolution, thus yielding multiple site-specific distance constraints from a single experiment [9].

In conclusion, the use of sparse sidechain ^{13}C labeling and ^{1}H-driven ^{13}C spin diffusion unambiguously indicate the lengthening of inter-helical distances when colicin Ia channel domain binds to the membrane. The strong attenuation of many inter-residue cross peaks for the membrane-bound sample is well correlated with high percentages of short inter-helical contacts in the soluble state, while cross peaks with similar intensities between the two states have mostly intra-helical contacts. Thus, the membrane-bound structure of colicin Ia channel domain may be described as a “molten globule”, in which individual helices retain

their structure while the packing of the helices is significantly loosened. Spin diffusion NMR is an excellent way of identifying global structural changes of proteins, and is capable of providing long-distance constraints in a qualitative but unambiguous fashion.

Acknowledgement: This work is supported by an NSF CAREER grant (MCB-0093398) and a Sloan Fellowship to M.H.

References

(1) Martin, R. W.; Zilm, K. W. *J. Magn. Reson.* **2003**, *165*, 162-174.

(2) Pauli, J.; Baldus, M.; vanRossum, B.; Groot, H. d.; Oschkinat, H. *ChemBioChem* **2001**, *2*, 272-281.

(3) McDermott, A. E.; Polenova, T.; Bockmann, A.; Zilm, K. W.; Paulsen, E. K.; Martin, R. W.; Montelione, G. T. *J. Biomol. NMR* **2000**, *16*, 209-219.

(4) Hong, M. *J. Biomol. NMR* **1999**, *15*, 1-14.

(5) Detken, A.; Hardy, E. H.; Ernst, M.; Kainosho, M.; Kawakami, T.; Aimoto, S.; Meier, B. H. *J. Biomol. NMR* **2001**, *20*, 203-221.

(6) Jaroniec, C. P.; Filip, C.; Griffin, R. G. *J. Am. Chem. Soc.* **2002**, *124*, 10728-10742.

(7) Rienstra, C. M.; Hohwy, M.; Mueller, L. J.; Jaroniec, C. P.; B, B. R.; Griffin, R. G. *J. Am. Chem. Soc.* **2002**, *124*, 11908-11922.

(8) Antzutkin, O. N. In *Solid-state NMR spectroscopy principles and applications*; M. J. Duer, Ed.; Blackwell Sciences, Inc.: Oxford, 2002; pp 280-390.

(9) Castellani, F.; vanRossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. *Nature* **2002**, *420*, 98-102.

(10) Jaroniec, C. P.; MacPhee, C. E.; Bajaj, V. S.; McMahon, M. T.; Dobson, C. M.; Griffin, R. G. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 711-716.

(11) Schmidt-Rohr, K.; Hong, M. *J. Am. Chem. Soc.* **2003**, *125*, 5648-5649.

(12) Wi, S.; Sinha, N.; Hong, M. *J. Am. Chem. Soc.* **2004**, *126*, 12754-12755.

(13) Cramer, W. A.; Heymann, J. B.; Schendel, S. L.; Deriy, B. N.; Cohen, F. S.; Elkins, P. A.; Stauffacher, C. V. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 611-641.

(14) Wiener, M.; Freymann, D.; Ghosh, P.; Stroud, R. M. *Nature* **1997**, *385*, 461-464.
(15) Huster, D.; Yao, X.; Jakes, K.; Hong, M. *Biochim Biophys Acta* **2002**, *1561*, 159-170.
(16) Zakharov, S. D.; Lindeberg, M.; Griko, Y.; Salamon, Z.; Tollin, G.; Prendergast, F. G.; Cramer, W. A. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4282-4287.
(17) Lindeberg, M.; Zakharov, S. D.; Cramer, W. A. *J. Mol. Biol.* **2000**, *295*, 679-692.
(18) Lakey, J. H.; Baty, D.; Pattus, F. *J. Mol. Biol.* **1991**, *218*, 639-653.
(19) Shin, Y.; Levinthal, C.; Levinthal, F.; Hubbell, W. L. *Science* **1993**, *259*, 960-963.
(20) Massotte, D.; Yamamoto, M.; Scianimanico, S.; Sorokine, O.; Dorsselaer, A. v.; Nakatani, Y.; Ourisson, G.; Pattus, F. *Biochemistry* **1993**, *32*, 13787-13794.
(21) van der Goot, F. G.; Gonzalex-Manas, J. M.; Lakey, J. H.; Pattus, F. *Nature* **1991**, *354*, 408-410.
(22) Huster, D.; Yao, X.; Hong, M. *J. Am. Chem. Soc.* **2002**, *124*, 874-883.
(23) Kim, Y.; Valentine, K.; Opella, S. J.; Schendel, S. L.; Cramer, W. A. *Protein Sci.* **1998**, *7*, 342-348.
(24) Hong, M.; Jakes, K. *J. Biomol. NMR* **1999**, *14*, 71-74.
(25) Hong, M. *J. Magn. Reson.* **1999**, *139*, 389-401.
(26) Hope, M. J.; Bally, M. B.; Webb, G.; Cullis, P. R. *Biochim. Biophys. Acta* **1985**, *812*, 55-65.
(27) Smith, P. K.; Krohn, R. I.; Hermanson, G. T.; Mallia, A. K.; Gartner, F. H.; Provenzano, M. D.; Fujimoto, E. K.; Goeke, N. M.; Olson, B. J.; Klenk, D. C. *Anal. Biochem.* **1985**, *150*, 76-85.
(28) Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G. *J. Chem. Phys.* **1995**, *103*, 6951-6958.
(29) Raleigh, D. P.; Levitt, M. H.; Griffin, R. G. *Chem. Phys. Lett.* **1988**, *146*, 71-76.
(30) Meier, B. H. *Adv. Magn. Opt. Reson.* **1994**, *18*, 1-115.
(31) Bronniman, C. E.; Szeverenyi, N. M.; Maciel, G. E. *J. Chem. Phys.* **1983**, *79*, 3694-3700.
(32) Huster, D.; Xiao, L.; Hong, M. *Biochemistry* **2001**, *40*, 7662-7674.
(33) Castellani, F.; Rossum, B. J. v.; Diehl, A.; Rehbein, K.; Oschkinat, H. *Biochemistry* **2003**, *42*, 11476-11483.

(34) Kumashiro, K. K.; Schmidt-Rohr, K.; Murphy, O. J.; Ouellette, K. L.; Cramer, W. A.; Thompson, L. K. *J. Am. Chem. Soc.* **1998**, *120*, 5043-5051.

(35) Parker, M. W.; Postma, J. P. M.; Pattus, F.; Tucker, A. D.; Tsernoglou, D. *J. Mol. Biol.* **1992**, *224*, 639-657.

(36) Lange, A.; Luca, S.; Baldus, M. *J. Am. Chem. Soc.* **2002**, *124*, 9704-9705.

(37) Lange, A.; Seidel, K.; Verdier, L.; Luca, S.; Baldus, M. *J. Am. Chem. Soc.* **2003**, *125*, 12640-12648.

(38) Kienker, P. K.; Qiu, X.; Slatin, S. L.; Finkelstein, A.; Jakes, K. S. *J. Membrane. Biol.* **1997**, *157*, 27-37.

Chapter 4

A 1D Sensitivity-Enhanced ^{1}H Spin Diffusion Experiment for Determining Membrane Protein Topology

A paper published in Solid State NMR

2006 vol. 29 pp. 163-169

Wenbin Luo and Mei Hong

Abstract

A sensitivity-enhanced 1D ^{1}H spin diffusion experiment, CHH, for determining membrane protein topology is introduced. By transferring the magnetization of the labeled protein ^{13}C to lipid and water protons for detection, the CHH experiment reduces the time of the original 2D ^{13}C-detected experiment by two orders of magnitude. The sensitivity enhancement results from ^{1}H detection and the elimination of the ^{13}C dimension. Consideration of the spin statistics of the membrane sample indicates that the CHH sensitivity depends on the ^{13}C labeling level and the number of protein protons relative to the mobile protons. 5-35% of the theoretical sensitivity was achieved on two extensively ^{13}C labeled proteins. The experimental uncertainties arise from incomplete suppression of the equilibrium ^{1}H magnetization and the magnetization of lipid protons directly bonded to natural-abundance carbons. The technique, demonstrated on colicin Ia channel domain, confirms the presence of a transmembrane domain and the predominance of surface-bound helices.

Introduction

^{1}H spin diffusion between lipids and membrane-bound proteins was recently introduced as an approach for determining the depth of insertion of the protein in lipid bilayers (*1*). Two experimental implementations have been demonstrated. In the original 1D

experiment, the membrane sample is frozen to a temperature at which the lipids cease their fast uniaxial rotation while the inter-bilayer water remains partially mobile. This residual mobile water magnetization is then selected as the source of the ^{1}H magnetization and transferred to the rigid lipid and protein. Detection of the protein ^{13}C or ^{15}N signals as a function of the ^{1}H spin diffusion mixing time yields a buildup curve that can be simulated to yield the distance of the labeled site from the water at the membrane surface. Detection of the natural abundance ^{31}P or ^{13}C signals of the lipids with known distances from the membrane surface allows internal calibration of the ^{1}H spin diffusion coefficients (*1, 2*).

In the second version of the ^{1}H spin diffusion experiment (*3*), the membrane sample is kept at ambient temperature where the lipids are in the liquid-crystalline state. The multiple mobile ^{1}H magnetization sources are resolved in the indirect dimension of a 2D spectrum by ^{1}H chemical shift evolution. Following ^{1}H evolution, the mobile ^{1}H magnetization is transferred to the rigid protein protons during a mixing time. The amount of the transferred magnetization, which depends on the proximity between the lipid moiety and the protein, is finally detected through the protein ^{13}C signals in the direct dimension. The ^{1}H-^{13}C cross peak intensities in the 2D spectra as a function of the spin diffusion mixing time give rise to the buildup curve, from which the shortest distance separation between the protein and the mobile proton source can be determined semi-quantitatively.

The key difference between the room-temperature 2D ^{1}H spin diffusion experiment and the low-temperature 1D experiment is that in the 2D experiment, the abundant motion of the fluid lipid bilayer makes its ^{1}H spin diffusion coefficient substantially smaller than that of the rigid protein, thus as soon as the lipid ^{1}H magnetization is transferred across the interface to the protein, the magnetization equilibrates rapidly in the rigid protein. As a result, the shape of the ^{1}H buildup curve is primarily determined by the shortest separation between the lipid source proton and the protein, and doesn't depend on the exact location of the specific ^{13}C or ^{15}N label in the protein. While less quantitative and lacking distance site resolution compared to the 1D experiment, the 2D experiment is well suited to extensively labeled proteins where single site resolution is not available to begin with. It can yield the global topology of the protein, such as the presence or absence of transmembrane domains, in large polytopic membrane proteins (*3, 4*).

Despite its utility, the need for collecting a series of 2D spectra as a function of mixing time poses challenges in sensitivity and experimental stability. Since the intensities of several 2D spectra need to be compared, a long block of experimental time is necessary. Spinning hydrated membrane samples at ambient temperature for an extended period of time is prone to cause dehydration and thus changes in the ^{1}H T_1 relaxation times. In addition, fluctuation in the radio frequency powers can affect the 2D intensity from one spectrum to another. For these reasons, it is desirable to increase the sensitivity of the room-temperature 2D spin diffusion experiment.

Indeed, sensitivity enhancement should be possible if one considers the fact that the 2D experiment puts the high-resolution ^{1}H chemical shift spectrum in the indirect dimension while the low-sensitivity ^{13}C chemical shift spectrum in the direct dimension. This means that not only the detection sensitivity is low, but also a large number of t_1 slices is necessary to resolve the different ^{1}H signals of the mobile lipid and water. Thus, a switch of the two dimensions would clearly be advantageous. Moreover, since the room-temperature spin diffusion experiment inherently does not resolve the depths of different ^{13}C sites within the protein, the ^{13}C dimension can in fact be removed altogether. In this paper, we demonstrate this simplified 1D ^{1}H-detected spin diffusion experiment. We designate this experiment CHH, to represent the magnetization pathway of protein ^{13}C –> protein ^{1}H –> lipid and water ^{1}H. In analogy, the 2D spin diffusion experiment will be called HHC, representing the reverse magnetization pathway of lipid and water ^{1}H –> protein ^{1}H –> protein ^{13}C.

Materials and Methods

Sample preparation

Colicin Ia channel domain (MW: 25 kDa) was expressed as described before (*5, 6*) and labeled using the TEASE protocol (*7*), where the labeled carbon precursor is supplemented with unlabeled ten amino acids from the citric acid cycle to suppress the labeling of these mostly polar amino acids. 1,6-colicin was expressed using [1,6-^{13}C] glucose as the main precursor, resulting in the following ^{13}C-labeled sites: Ala Cβ, Leu Cα, Cδ1, Cδ2, Ser Cβ, Val Cγ1, Cγ2, His Cδ2, C', Phe Cβ, Cγ, Cδ1, Cδ2, Tyr Cβ, Cδ1, Cδ2, Cζ, and

Trp Cβ, Cδ2, Cε1, Cε2 (6, 8). pU-colicin was expressed using [U-^{13}C] glucose as the ^{13}C-labeled precursor, resulting in uniform ^{13}C labeling of Ala, Leu, Ser, Val, Gly, His, Phe, Tyr and Trp. Both protein samples were uniformly ^{15}N labeled using ^{15}N-NH_4Cl.

The labeled colicin Ia channel domain was reconstituted into large unilamellar vesicles of POPC/POPG (3:7 molar ratio) membranes at pH 4.8 as described previously (*6*). The protein/lipid molar ratio is 1:100. The membrane pellet obtained after ultracentrifugation was lyophilized, packed into a 4 mm MAS rotor, and hydrated to 35% water by mass.

Solid-state NMR experiments

All NMR experiments were carried out on a Bruker DSX-400 spectrometer (Karlsruhe, Germany) operating at a resonance frequency of 400.49 MHz for ^{1}H and 100.71 MHz for ^{13}C. A double-resonance MAS probe with a 4-mm spinning module was used. The ^{1}H radiofrequency (rf) field strengths for TPPM decoupling (*9*) were ~65 kHz. Typical ^{13}C and ^{1}H 90˚ pulse lengths were 5 μs and 4 μs. The recycle delay ranged from 1.8 s to 3 s. Cross-polarization (CP) contact times were 200 μs for the first CP and 150 μs for the second CP in the 1D CHH experiment. All experiments were conducted at room temperature (293 K) at a spinning speed of 5 kHz. The 1D CHH spectra were collected with mixing times of 1 μs to 625 ms. The pU-colicin required 32 or 64 scans per mixing time, while the 1,6-colicin used 128 or 256 scans per mixing time. Peak intensities were plotted as a function of the mixing time and corrected for ^{1}H T_1 relaxation as measured from a separate T_1 inversion recovery experiment. 2D HHC spectra were acquired with 64 or 128 scans per t_1 slice, 176 t_1 slices, and a ^{1}H spectral width of 5 kHz. ^{1}H mixing times of 25, 100 and 225 ms, were used for the 2D experiments on pU-colicin. States detection was used to obtain pure-phase spectra (*10*).

Results and Discussion

CHH Pulse sequence

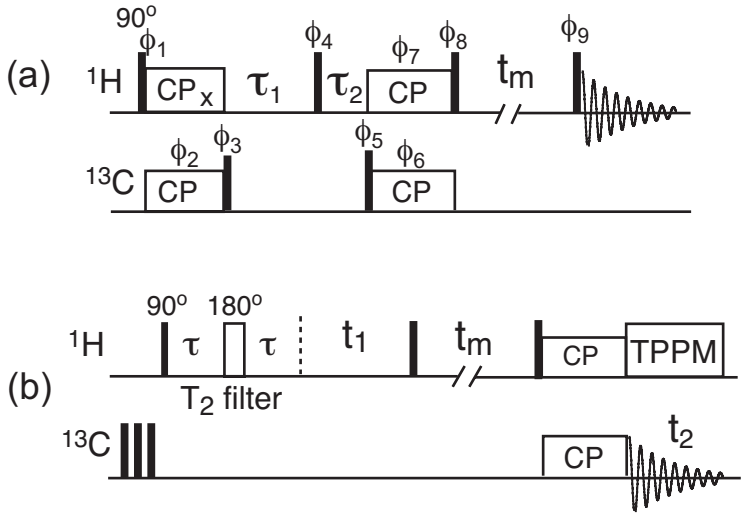


Figure 1. Pulse sequences for (a) the 1D CHH experiment and (b) the 2D HHC experiment. In (a), τ_1 = 20-30 ms and τ_2 = 5 ms. CP times were 200 μs for the first CP and 150 μs for the second CP. In (b), τ = 1 ms. Pulse sequence (a) used 32- or 64-step phase cycles. The 64-step phase cycles are: ϕ_1 = 1 3; ϕ_2 = 0 0 1 1 2 2 3 3; ϕ_3 = 1 1 2 2 3 3 0 0, 3 3 0 0 1 1 2 2; ϕ_4 = 1 3; ϕ_5 = 3 3 0 0 1 1 2 2; ϕ_6 = (0 0 1 1 2 2 3 3) x 2, (2 2 3 3 0 0 1 1) x 2; ϕ_7 = 0 0 1 1 2 2 3 3; ϕ_8 = (1 1 2 2 3 3 0 0) x 4, (3 3 0 0 1 1 2 2) x 4; ϕ_9 = 3 3 0 0 1 1 2 2; ϕ_{10} = $\overline{R}RRR\overline{R}RR\overline{R}R$, where R = 0 2 1 3 2 0 3 1, and $\overline{R}$ = 2 0 3 1 0 2 1 3. (0 = x, 1=y, 2=-x, 3=-y).

The pulse sequence for the sensitivity-enhanced 1D CHH experiment with ^{1}H detection is shown in Figure 1a. After ^{1}H-^{13}C CP, a 90° ^{13}C pulse stores the ^{13}C magnetization along the z-axis while the ^{1}H magnetization of both protein and lipids is destroyed by a T_2 filter. The ^{1}H T_2 filter includes two delay periods, τ_1 and τ_2, separated by a 90° pulse. A τ_1 value of 20 – 30 ms and τ_2 of 5 ms were used. The purpose of the 90° pulse is to rotate the ^{1}H magnetization that has undergone T_1 relaxation back to the transverse plane, to be destroyed by the second delay τ_2. At the end of the ^{1}H T_2 filter, mostly only ^{13}C magnetization of the labeled sites in the protein remains. This is returned to the transverse plane by a ^{13}C 90° pulse, then transferred to the protein protons by a short ^{13}C-^{1}H CP step. The protein ^{1}H magnetization, flipped to the z-axis by a 90° pulse, then spin-diffuses to the mobile lipid and water, whose signals are detected without homonuclear decoupling. The

intensity of the lipid and water ^{1}H signals transferred from the protein protons is monitored as a function of the spin diffusion mixing time. The resulting buildup curve is fit to give a semi-quantitative distance, which represents the shortest distance between the protein and the detected lipid or water protons.

For comparison, the pulse sequence for the 2D ^{13}C-detected ^{1}H spin diffusion experiment (*3*) is shown in Figure 1b. The key difference from the 1D experiment is that the site-resolved lipid and water ^{1}H chemical shift is encoded in the indirect dimension, while the result of spin diffusion is detected via the protein ^{13}C labels in the direct dimension. However, since this spin diffusion experiment, conducted at room temperature where there is significant mobility difference between the soft membrane and the rigid protein, does not give site-specific distances between individual protein ^{13}C sites and the lipid moiety (*3*), the ^{13}C resolution in the direct dimension is not useful. Thus, the 1D CHH experiment eliminates this ^{13}C dimension altogether and detects the ^{1}H spectrum instead. Due to the high mobility of the lipid and water protons in the membrane, it is possible to conduct straightforward ^{1}H detection, without multiple-pulse homonuclear decoupling. Thus the sensitivity enhancement over ^{13}C detection is simply proportional to $(\gamma_H/\gamma_C)^{3/2}=8$. Combined with the reduction of the dimensionality, this results in an experimental time saving of two orders of magnitude, as we show below.

Sensitivity enhancement of the 1D CHH over the 2D HHC experiments

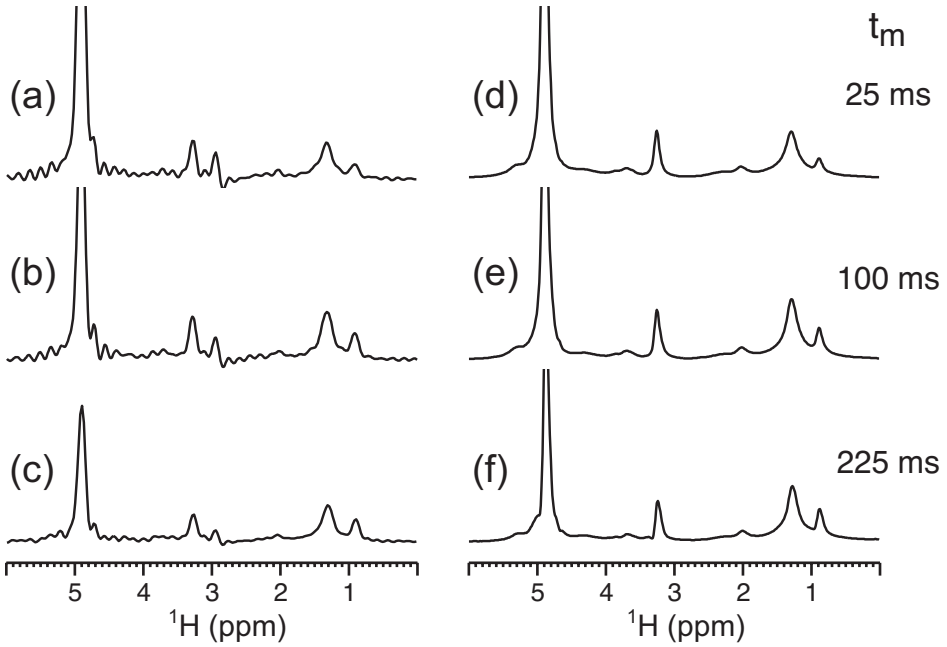


Figure 2. Comparison of the ^{1}H spectra of pU-colicin from the 2D HHC experiment (a-c) and from the sensitivity-enhanced 1D CHH experiment (d-f) with the indicated mixing times. In each column, the spectra are drawn to scale after taking into account the number of scans. (a-c) F1 projection of the aliphatic region of the 2D HHC spectra. The peak at 3 ppm is a zero-frequency artifact. These 2D spectra were acquired with 11264 scans (b) and 22528 scans (a, c). (d-f) 1D CHH spectra, acquired with 64 scans. All spectra were processed with 10 Hz exponential line broadening.

Figure 2 compares the F1 projection of the 2D HHC spectra of pU-colicin with the ^{1}H-detected 1D CHH spectra. The spectra of three mixing times: 25 ms, 100 ms, and 225 ms, are shown. The ^{1}H spectra were processed with an exponential line broadening of 10 Hz. It can be seen that the 2D HHC projections is noisier than the 1D CHH spectra due to the lower sensitivity of ^{13}C detection, and residual truncation wiggles for the water signal remain due to the limited evolution time of the ^{1}H dimension. The 2D spectra were acquired with 22528 scans for the 25 ms and 225 ms mixing time points (~ 13 hours), and half that amount for the 100 ms time point. In comparison, the 1D spectra were acquired with 64 scans (~ 2 minutes), corresponding to a time saving of 180 - 350 fold.

Information on the populations of residues on the membrane surface versus the bilayer center can be obtained from the relative height of the water ^1H signal and the CH_2 signal. The fact that the water signal is much higher indicates that colicin Ia channel domain is predominantly located on the membrane surface. To verify that the 1D CHH experiment reproduces the ^1H spin diffusion behavior seen in the 2D experiment, Figure 3(a-b) superimposes the complete buildup curves (open symbols) of pU-colicin from the 1D CHH spectra with the corresponding 2D data at three mixing times (filled symbols). The buildup intensities were corrected for ^1H T_1 relaxation and normalized by the equilibrium values at long mixing times. It can be seen that the 1D buildup intensities superimpose well with the 2D data points, confirming that the same distance information is obtained. Specifically, the CH_2 and CH_3 buildup curves correspond to a short protein-lipid separation of ~2 Å, indicating that the protein contains a small but non-negligible fraction of residues that are transmembrane (3).

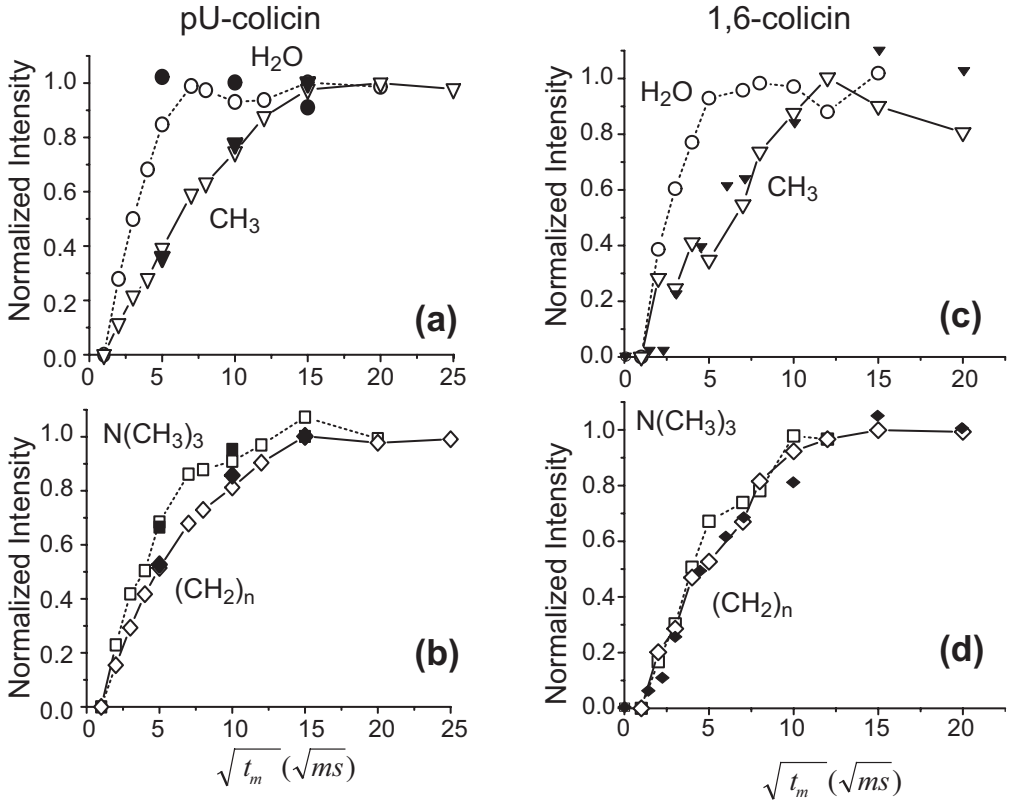


Figure 3. Comparison between the ^1H spin diffusion buildup curves from the 1D CHH experiment (open symbols) and from the 2D HHC experiment (filled symbols) at selected

mixing times. (a, b) pU-colicin. (c, d) 1,6-colicin. The magnetization sources are H_2O (circles), $N(CH_3)_3$ (squares), $(CH_2)_n$ (diamonds), and CH_3 (triangles). The 2D data in (c, d) are those of [2-^{13}C] colicin Ia channel domain reproduced from ref. (3). For the pU-colicin data in (a, b), the error bar (including both random and systematic errors) is smaller than or comparable to the symbol sizes. For the 1,6-colicin data in (c, d), the random error is negligible but the systematic uncertainty is up to 30% for the CH_2 and CH_3 intensities, as indicated in Table 3.

The 1D CHH buildup curves of the more sparsely ^{13}C labeled 1,6-colicin reproduce the trend of the pU-colicin data as well as the buildup curves of [2-^{13}C] colicin reported previously (3) (Figure 3c, d). However, the CH_3 buildup curve, obtained from the lowest intensity in the spectra, shows noticeable uncertainty. This results from the lower ^{13}C labeling level, as we analyze below.

Sensitivity of the 1D CHH experiment

Since the 1D CHH experiment requires that all detected ^{1}H magnetization originates from the labeled ^{13}C sites in the protein, the sensitivity of the experiment depends critically on the ^{13}C labeling level. The higher the ^{13}C labeling level, the higher the CHH sensitivity. At the same time, the sensitivity of the experiment also depends on the fraction of the mobile protons in the entire proton reservoir, since only the ^{1}H magnetization of the mobile lipid and water, not the rigid protein, is detected. The larger the fraction of the mobile protons, the more sensitive the CHH experiment.

The numbers of ^{1}H spins in the protein (H_P), lipids (H_L) and water (H_W) for the two membrane-bound colicin Ia channel domain samples are tabulated in Table 1. These are estimated based on an experimental protein : lipid molar ratio of 1:100 and 35% (by mass) of water in the sample. The total number of the ^{1}H spins in the system (H_{tot}) is the sum of the three. The number of labeled ^{13}C spins in the protein (C_P) was calculated based on the labeling schemes used. C_P of pU-colicin is about three times that of 1,6-colicin.

During the reverse CP step, the ^{13}C magnetization is transferred to the protein protons. Assuming that the spin diffusion within the protein is rapid due to the rigidity of the molecule, we can estimate the fraction of the ^{13}C magnetization that ends up in the proton reservoir during CP as $H_P/(H_P+C_P)$. This magnetization is then transferred to the mobile lipid and water molecules during the mixing time. Thus, at equilibrium, the effective number of detected protons in the CHH experiment is:

$$H_{CHH} = C_P \times \frac{H_P}{H_P + C_P} \times \frac{H_L + H_W}{H_P + H_L + H_W} \quad (1)$$

We define the ideal CHH sensitivity as the number of detected ^1H spins, H_{CHH}, relative to the total number of lipid and water protons in the system:

$$\text{Sensitivity} = \frac{H_{CHH}}{H_L + H_W} \quad (2)$$

Based on the ^{13}C labeling levels and the distribution of protons in the membrane samples, we find this ideal sensitivity to be 2.5% for pU-colicin and 1.1% for 1,6-colicin (Table 1), which represent the maximum sensitivity of the CHH spectrum relative to the equilibrium mobile ^1H magnetization.

Table 1. Spin numbers and ideal CHH sensitivities for pU-colicin and 1,6-colicin.

	H_P	H_L	H_W	H_{tot}	C_P	n.a. $^{13}C^a$	H_{CHH}	Sensitivity
pU-colicin	1544	7780	6090	15414	513	50	346	2.5%
1,6-colicin	1544	7780	6090	15414	189	53	152	1.1%

[a] the number of natural abundance lipid and protein ^{13}C spins in each sample.

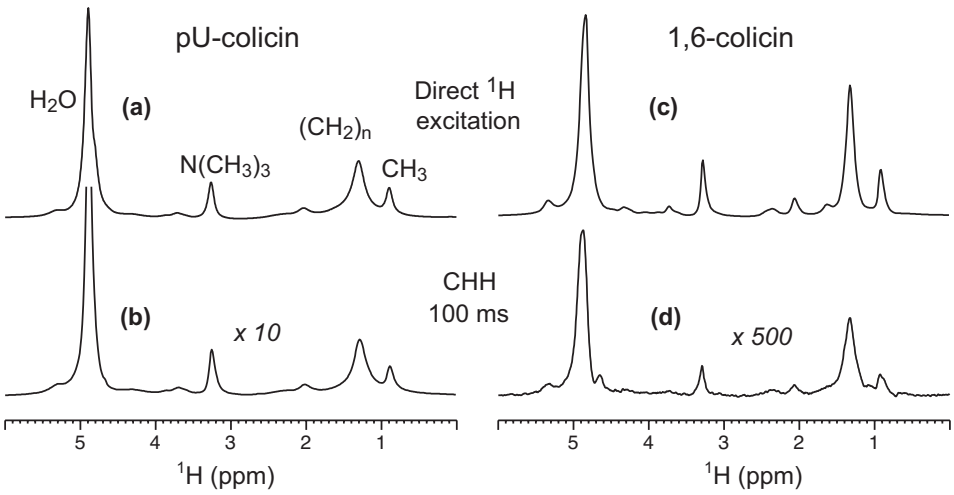


Figure 4. Direct ^{1}H excitation spectra (a, c) and 1D CHH spectra (b, d) with a mixing time of 100 ms. (a, b) pU-colicin. (c, d) 1,6-colicin. Spectra (b, d) were scaled with respect to spectra (a, c) by the ratios shown. Numbers of scans (NS) are: (a) 4, (b) 64, (c) 32, and (d) 64. The intensities of the 100 ms CHH spectra relative to the direct excitation spectra after taking into account the NS difference are listed in Table 2.

Figure 4 compares the ^{1}H direct excitation spectra with the 1D CHH spectra of pU-colicin and 1,6-colicin at a mixing time of 100 ms. The overall intensities of the 1D CHH spectra is two to three orders of magnitude lower than the direct excitation spectra. The experimental CHH sensitivities, obtained from the ratios of the individual CHH peaks (Figure 4b, d) relative to the direct excitation peaks (Figure 4a, c), are listed in Table 2. Although the CHH intensities are less than 1% of the direct excitation spectra, due to the high detection sensitivity of ^{1}H, the random noise of the spectra is still negligible compared to the signal, at less than 1% of the signal. Indeed, the experimentally achieved CHH sensitivities are reasonable compared to the ideal sensitivity calculated for each sample. For pU-colicin, these are 25-35% of the ideal value, while for 1,6-colicin, the experimental CHH sensitivities is 5-10% of the ideal value.

Suppression of the direct lipid and water ^{1}H polarization

Since the 1D CHH experiment detects only a small fraction (< 3% in our systems) of the total mobile proton magnetization, the clean suppression of the equilibrium lipid and water ^{1}H magnetization is crucial to the accuracy of the ^{1}H buildup curve. The suppression is mostly accomplished by the ^{1}H T_2 filter, during which ^{1}H-^{1}H dipolar coupling and transverse relaxation destroy the directly excited ^{1}H magnetization. The mechanism works well for the rigid protein protons; however, the suppression of the protons in the anisotropically mobile lipids and the isotropic water molecules is more difficult due to their motionally averaged dipolar couplings and long T_2 relaxation times. As an example, Figure 5 shows the ^{1}H CHH spectra of the two membrane-bound colicin samples after a mixing time of 1 μs (b, d) and 100 ms (a, c). The former correspond to the no spin diffusion case, so any residual signals represent the unsuppressed and non-spin-diffused ^{1}H magnetization. In the more highly ^{13}C-labeled pU-colicin, only the water peak is visible, while in the more sparsely labeled 1,6-colicin, the water and the CH_2 peaks both show noticeable residual intensities. The fractions of the residual ^{1}H intensities at 1 μs over the ^{1}H intensities at 100 ms are 3-5% for pU-colicin, but much higher values of 10 - 30% for 1,6-colicin (Table 3).

Table 2. Experimental CHH sensitivities for pU-colicin and 1,6-colicin, obtained from the intensity ratio of the 100 ms CHH spectra with the ^{1}H direct excitation spectra.

	H_2O	$N(CH_3)_3$	CH_2	CH_3
pU-colicin	0.90%	0.85%	0.63%	0.64%
1,6-colicin	0.08%	0.05%	0.06%	0.05%

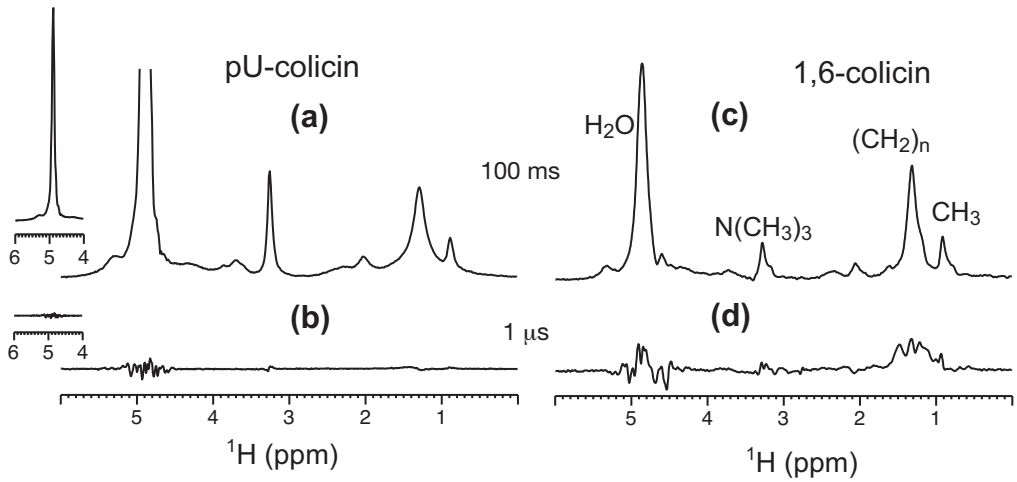


Figure 5. CHH ^{1}H spectra of (a, b) pU-colicin and (c, d) 1,6-colicin, with spin diffusion mixing times of (a, c) 100 ms and (b, d) 1 μs. Insets of spectra (a, b) show the full heights of the water peak. The bottom spectra are plotted on the same scale as the top spectra in each column. NS are (a) 64, (b) 32, (c) 96, and (d) 96. The intensities of the 1 μs spectra relative to the 100 ms spectra after taking into account the NS difference are tabulated in Table 3.

The residual equilibrium ^{1}H magnetization, when present, decreases the magnetization difference between the source protons after the reverse CP and the sink protons, thus should cause a smaller slope in the ^{1}H buildup curve. This in turn would result in an apparently longer protein-lipid distance. Indeed, this was found to be the case when incomplete suppression occurred in the absence of the 90˚ purge pulse in the T_2 filter period. This 90˚ pulse serves to rotate back to the transverse plane the ^{1}H magnetization that has relaxed due to T_1 relaxation. The ^{1}H T_1 values of the water and lipid protons in POPC/POPG membranes range from 250 ms to 450 ms. Thus, for a τ_1 of 20 ms, 4-7% of the mobile proton magnetization would have relaxed to the z-axis. Although phase cycling removes most of this magnetization, it is still not negligible, considering that the percent of detected protons in the CHH experiment is only 1-3% (Table 1). By using this 90˚ purge pulse followed by another short dephasing period, we achieve more complete removal of the direct ^{1}H magnetization. The effect of the 90˚ purge pulse and τ_2 on the spin diffusion buildup curve is shown in Figure 6. The addition of the purge pulse and τ_2 resulted in faster buildup for all resolved

proton species. The increased buildup rate is particularly pronounced for the CH_2 and CH_3 signals from the hydrophobic interior of the membrane, probably due to the fact that the majority of colicin Ia channel domain is bound immediately at the membrane surface that the initial rates for water and headgroup γ protons are already high even in the absence of the purge pulse, while the buildup rates of the CH_2 and CH_3 are generally slower than the surface protons.

Table 3. CHH intensity ratios between spectra acquired with mixing times of 1 μs and 100 ms.

	H_2O	$N(CH_3)_3$	CH_2	CH_3
pU-colicin	3%	4%	4%	5%
1,6-colicin	10%	20%	30%	30%

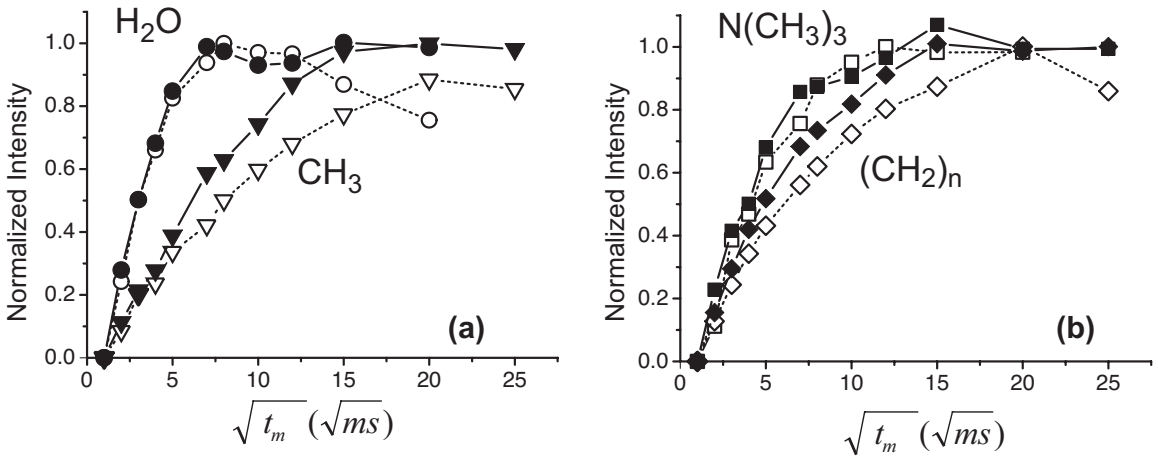


Figure 6. Effect of the 90° purge pulse on the ^1H spin diffusion buildup curves obtained from the CHH experiment. pU-colicin was used. The data were collected with the purge pulse on (filled symbols, solid lines) and off (open symbols, dotted lines). The magnetization sources are H_2O (circles), CH_3 (triangles), $N(CH_3)_3$ (squares) and $(CH_2)_n$ (diamonds).

From this analysis, it can be seen that the main source of uncertainty of the 1D CHH experiment is the systematic error introduced by the incomplete suppression of the non-spin

diffused ^{1}H magnetization of lipid and water. For pU-colicin, which has a protein ^{13}C labeling level of 53%, the signal-to-noise ratio of the water peak in the 25 ms CHH spectrum is ~1700 (Figure 2d). This translates into a random noise of 0.06%. In comparison, the error from the unsuppressed initial ^{1}H magnetization is 3% (Table 3), significantly higher than the random noise. Thus, better suppression techniques will be desirable for improving the applicability of this technique. The current 1D CHH experiment used extended phase cycles of 32 or 64 steps (Figure 1). Potentially, gradient pulses may be used to achieve better suppression in fewer scans. Cleaner suppression of the initial ^{1}H magnetization will make this 1D CHH technique more robust for proteins with lower labeling levels such as site-specifically labeled membrane peptides.

The 1,6-colicin sample gives experimental CHH sensitivities that are lower than pU-colicin by more than the ^{13}C labeling level difference. We attribute this to the fact that 1,6-colicin is predominantly labeled in the mobile sidechains, thus have less efficient CP transfer from ^{13}C to ^{1}H. 1,6-colicin also shows much larger residual ^{1}H magnetization than pU-colicin in the limit of no spin diffusion mixing (Table 3). This partly results from the increased fraction of natural abundance lipid ^{13}C sites (Table 1), which, during the reverse CP step, can give rise to ^{1}H intensities that do not originate from the protein. Indeed, a control experiment using a membrane sample without the protein showed non-negligible CH_2 peak intensities at short mixing times (data not shown), consistent with the magnetization transfer from natural abundance lipid ^{13}C to lipid ^{1}H. However, there is no systematic intensity buildup for any of the ^{1}H peaks on this control sample, confirming that the smooth buildup of the 1,6-colicin data (Figure 3c, d) contains real distance information.

In theory, the fraction of such lipid background ^{1}H intensities is only about twice higher for 1,6-colicin than for pU-colicin due to the two-fold lower H_{CHH} for 1,6-colicin (Table 1). In practice, however, the amount of the lipid-originated ^{1}H intensity depends sensitively on the reverse ^{13}C-^{1}H CP efficiency. The pU-colicin membrane sample contains many rigid backbone protein ^{13}C sites that transfer the polarization to the ^{1}H spins in 150 μs much more efficiently than the mobile natural-abundance lipid ^{13}C sites. Moreover, the optimum Hartman-Hahn match condition for the rigid protein sites can be tailored to differ from the optimum match condition for the mobile lipids. For 1,6-colicin, however, most

labeled ^{13}C sites occur in the mobile sidechains, whose CP efficiency and match condition are more comparable to the lipid. Thus, it is more difficult to select the protein-originated ^{1}H signals against the undesirable lipid-originated ^{1}H intensities. Therefore, the CHH experiment is best applied to backbone labeled proteins with high labeling levels. An analogous NHH experiment using a ^{15}N-labeled protein should also have less natural abundance contribution.

The 1D CHH experiment bears resemblance to the CHHC experiment introduced recently for determining distances within a protein and between proteins (*11*). There, the result of ^{1}H spin diffusion is detected on ^{13}C (or other rare spins) due to the difficulty of direct detection of rigid protons. In our case, this last ^{1}H–>^{13}C step is avoided because of the mobility of the lipid and water of interest. Thus, the sensitivity gain from straightforward ^{1}H detection, without the need for multiple-pulse decoupling, is an appealing aspect of this CHH technique.

This 1D CHH technique can also be applied to the low-temperature spin diffusion version, where quantitative distance and depth information can be obtained. There, the sensitivity gain will be due to γ-enhancement only, without the benefit of the reduction of the dimensionality.

Acknowledgement: The authors thank Prof. K. Schmidt-Rohr for helpful discussions. This work is supported by a CAREER grant from the National Science Foundation (MCB-93387) and a Sloan Research Fellowship to M.H.

References

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

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

3. D. Huster, X. L. Yao and M. Hong, Membrane Protein Topology Probed by 1H Spin Diffusion from Lipids Using Solid-State NMR Spectroscopy, J. Am. Chem. Soc. 124, 874 (2002).

4. J. J. Buffy, A. J. Waring, R. I. Lehrer and M. Hong, Immobilization and Aggregation of Antimicrobial Peptide Protegrin in Lipid Bilayers Investigated by Solid-State NMR, Biochemistry 42, 13725 (2003).

5. D. Huster, X. Yao, K. Jakes and M. Hong, Conformational changes of colicin Ia channel-forming domain upon membrane binding: a solid-state NMR study, Biochim Biophys Acta 1561, 159 (2002).

6. W. Luo, X. L. Yao and M. Hong, Large Structure Rearrangement of Colicin Ia Channel Domain After Membrane Binding from 2D 13C Spin Diffusion NMR, J. Am. Chem. Soc. 127, 6402 (2005).

7. M. Hong and K. Jakes, Selective and Extensive 13C Labeling of a Membrane Protein for Solid-State NMR Investigation, J. Biomol. NMR 14, 71 (1999).

8. M. Hong, Determination of Multiple Phi Torsion Angles in Solid Proteins by Selective and Extensive 13C Labeling and Two-Dimensional Solid-State NMR, J. Magn. Reson. 139, 389 (1999).

9. A. E. Bennett, C. M. Rienstra, M. Auger, K. V. Lakshmi and R. G. Griffin, Heteronuclear decoupling in rotating solids, J. Chem. Phys. 103, 6951 (1995).

10. D. J. States, R. A. Haberkorn and D. J. Ruben, A 2D NOE experiment with pure absorption phase in four quadrants, J. Magn. Reson. 48, 286 (1982).

11. A. Lange, S. Luca and M. Baldus, Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids., J. Am. Chem. Soc. 124, 9704 (2002).

Chapter 5

Determination of the Oligomeric Number and Intermolecular Distances of Membrane Protein Assemblies by Anisotropic ^{1}H-Driven Spin Diffusion NMR Spectroscopy

A paper published in the Journal of the American Chemical Society

2006, vol. 128 pp. 7242-7251

Wenbin Luo and Mei Hong

Abstract

The determination of the high-resolution quaternary structure of oligomeric membrane proteins requires both the oligomeric number and intermolecular distances. The centerband-only detection of exchange (CODEX) technique has been shown to enable the extraction of the oligomeric number through the equilibrium exchange intensity at long mixing times. To obtain quantitative distances, we now provide an analysis of the mixing-time-dependent CODEX intensities using the ^{1}H-driven spin diffusion theory. The exchange curve is fit to a rate equation, where the rate constants are proportional to the square of the dipolar coupling and the spectral overlap integral between the exchanging spins. Using a number of ^{13}C and ^{19}F labeled crystalline model compounds with known intermolecular distances, we empirically determined the overlap integrals of ^{13}C and ^{19}F CODEX for specific spinning speeds and chemical shift anisotropies. These consensus overlap integral values can be applied to structurally unknown systems to determine distances. Applying the ^{19}F CODEX experiment and analysis, we studied the transmembrane peptide of the M2 protein (M2TM) of influenza A virus bound to DMPC bilayers. The experiment proved for the first time that the M2TM peptide associates as tetramers in lipid bilayers, similar to its oligomeric state in detergent micelles. Moreover, the nearest-neighbor interhelical F-F distance between (4-19F) Phe30 is 7.9 – 9.5 Å. This distance constraints the orientation of the M2TM helices and the packing of the helices in the tetrameric bundle, and supports the structural model derived from previous solid-state NMR orientational data. The CODEX

technique presents a general method for determining the oligomeric number and intermolecular distances in the ~10 Å range in membrane proteins and other complex biological assemblies.

Introduction

One of the most challenging questions today in structural biology is the determination of the quaternary structure of membrane proteins. Many membrane proteins such as ion channels oligomerize with well-defined stoichiometry and structure to carry out their function. Elucidating the intermolecular packing of these proteins is essential not only for deciphering the structure-function relations of specific proteins but also for understanding the general principles of membrane protein folding [1, 2]. To determine the oligomeric structures, intermolecular distance constraints are essential. Many solid-state NMR techniques utilizing dipolar recoupling under magic-angle spinning (MAS) [3-6] have been developed to measure distances. Most these techniques yield only relatively short internuclear distances (~ 5 Å), albeit with high accuracy. The determination of long-range distances still remains a challenge.

Spin diffusion NMR is a robust approach for obtaining semi-quantitative homonuclear distances. Both direct ^{1}H spin diffusion and ^{1}H-driven X-nucleus (X=^{13}C, ^{19}F, etc) spin diffusion experiments have been analyzed and employed for distance determination in small molecules [7, 8] and in isotopically labeled proteins [9-12]. Because of the lack of ^{1}H decoupling during the mixing time and the long T_1 relaxation time of many X nuclei, measuring long-range distances with spin diffusion is in principle straightforward, simply involving lengthening the mixing time. As a new member of this family, the CODEX ^{1}H-driven anisotropic spin diffusion technique has been shown to enable spin counting, i.e. the determination of the oligomeric number of proteins [13]. We now show that it can also be used to extract intermolecular distances, and is thus a particularly effective tool for elucidating the oligomeric structure of membrane proteins.

The principle of the centerband-only detection of exchange (CODEX) technique [14, 15] for spin counting is magnetization exchange between orientationally different and singly

labeled molecules. Two rotor-synchronized π-pulse trains recouple the orientation-dependent chemical shift anisotropy (CSA) of X spins under MAS. Between the two π-pulse trains, ^{1}H-driven X spin diffusion occurs between chemically equivalent but orientationally inequivalent nuclei, causing a change of the CSA frequency, which prevents the complete refocusing of a stimulated echo at the end of the second π-pulse train. At long exchange mixing times, the initial magnetization is equally distributed among *n* orientations in the cluster, reducing the CODEX echo intensity to 1/*n*. Thus, from the equilibrium value of the CODEX intensity one can determine the oligomeric number of aggregating proteins or the number of magnetically inequivalent molecules in a crystal unit cell [13]. Using ^{19}F CODEX, we showed that protegrin-1, a β-sheet antimicrobial peptide, forms dimers in POPC bilayers [13].

While the CODEX equilibrium intensity permits spin counting, the time-dependent CODEX intensity decay contains distance information, since the homonuclear X-X dipolar coupling that drives the magnetization exchange scales with distance (*r*) as $1/r^3$. Thus, it is desirable to extend the capability of the CODEX technique beyond spin counting to include distance determination. According to the ^{1}H-driven spin diffusion theory, the rate of magnetization exchange depends on the dipolar coupling and the spectral overlap integral, the precise value of which is difficult to predict from first principles. Thus, here we adopt an empirical approach of calibrating the overlap integral using model compounds with known internuclear distances, and show that consensus overlap integral values can be obtained for specific chemical functional groups and experimental conditions. These consensus overlap integral values can then be used for distance determination of structurally unknown molecular assemblies.

As a first application of the ^{19}F CODEX experiment and analysis, we investigate the transmembrane peptide of the M2 protein (M2TM) of influenza A virus. The 97-residue M2 protein forms proton-conducting channels that initiate the dissociation of the viral RNA/protein complex and the fusion of the viral membrane with the endosomal wall [16, 17]. Various studies showed that the active M2 protein is a tetramer in vivo [18] and the transmembrane domain also tetramerizes in detergent micelles [19, 20]. However, the oligomeric

state of the peptide in the lipid bilayer has not been directly determined. The structure of the M2TM helical bundle has been investigated by Cys scanning and conductivity measurements [21], and the orientation of the helices has been measured accurately using ^{15}N NMR experiments of uniaxially oriented membranes [22]. Beside these, only a single short distance has been reported on this peptide [23]. Thus, currently the highest-resolution structural information about the M2 helical bundle concerns the monomer and the backbone. Few constraints about the helix-helix packing in the bundle and the sidechain conformation are available.

In this study, we first use $^{13}C'$- and ^{19}F-labeled model compounds with known intermolecular distances to extract consensus overlap integral values for the ^{13}C and ^{19}F CODEX experiments. We then apply the ^{19}F CODEX experiment to determine the oligomeric number of the M2TM peptide directly in lipid bilayers for the first time. Moreover, we extract the interhelical distance between Phe30 residues from the ^{19}F CODEX curves, and show that this distance provides an important restraint to the orientation of the peptide and the diameter of the helical bundle. This is the first time that the oligomeric number and a long intermolecular distance are determined for this homo-oligomeric helical bundle in lipid bilayers.

Materials and Methods

Sample preparation

Isotopically labeled model compounds, including $^{13}C'$-Gly, $^{13}C'$-Leu, $^{13}C'$-Phe, 5-^{19}F-Trp, and 4-^{19}F-2'-nitroacetanilide, were purchased from Cambridge Isotope Laboratories (Andover, MA) and Sigma (St. Louis, MO). 1, 2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) was obtained from Avanti Polar Lipids (Alabaster, AL). The fluorinated M2TM peptide, A30F-M2TM, where the original Ala30 residue was replaced by 4-^{19}F-Phe, was custom-synthesized by SynPep Corp. (Dublin, CA) using standard Fmoc chemistry. The amino acid sequence is NH_2-Ser22-Ser23-Asp24-Pro25-Leu26-Val27-Val28-Ala29-(4-19F) Phe30-Ser31-Ile32-Ile33-Gly34-Ile35-Leu36-His37-Leu38-Ile39-Leu40-Trp41-Ile42-

Leu43-Asp44-Arg45-Leu46-COOH. The A30F mutation has been shown by analytical ultracentrifugation experiments to stabilize the tetramer in DPC micelles [20]. The purified A30F-M2TM peptide was repeatedly washed in dilute HCl solution to remove residual trifluoroacetate (TFA) salt, until no TFA signal was observed (-75 ppm) in the ^{19}F solution NMR spectrum.

Membrane-bound A30F-M2TM peptide was prepared using a procedure similar to that of Cross and coworkers [23]. Large unilamellar DMPC vesicles were prepared by dissolving DMPC lipids in 5 mM Na_2HPO_4/NaH_2PO_4 buffer at 30˚C (pH 7), vortexing, freeze-thawing, then extruding the solution through polycarbonate filter membranes with 1 μm pores at 27˚C. Purified A30F-M2TM peptide was dissolved in the resulting vesicle solution at a P : L molar ratio of 1 : 15. The membrane mixture was vortexed, sonicated and incubated at 30˚C for 2 days to facilitate peptide reconstitution. The mixture was then centrifuged at 150,000 g for 2 hours at 28˚C and the pellet was collected. Photometric assay of the supernatant [24] showed that 90% of the peptide was reconstituted into the membrane. The pellet was transferred to a 4 mm rotor and incubated at 30˚C for two days before the NMR experiments.

Solid-state NMR experiments

CODEX experiments were carried out on a Bruker DSX-400 spectrometer (Karlsruhe, Germany) operating at a resonance frequency of 400.49 MHz for ^{1}H, 376.8 MHz for ^{19}F and 100.71 MHz for ^{13}C, using MAS probes equipped with 4 mm spinner modules. The spinning speed was 8 kHz for the ^{19}F experiments and 5 kHz for the ^{13}C experiments. The ^{19}F CODEX experiments used a H/F/X probe that allows simultaneous tuning of the ^{1}H and ^{19}F frequencies on a single channel through a combiner/splitter assembly. Experiments on the M2TM peptide were conducted at 240 K using air cooled through a Kinetics Thermal Systems XR Air-Jet Cooler (Stone Ridge, NY). Typical radio-frequency (rf) field strengths were 50 kHz for ^{19}F and ^{13}C and 60-70 kHz for ^{1}H. Recycle delays ranged from 1.5 s to 3 s. Cross-polarization (CP) contact times were 300 μs for ^{19}F and 500 μs for ^{13}C experiments.

^{13}C and ^{19}F chemical shifts were referenced to the α-glycine $^{13}C'$ signal at 176.49 ppm on the TMS scale and the Teflon™ ^{19}F signal at -122 ppm, respectively.

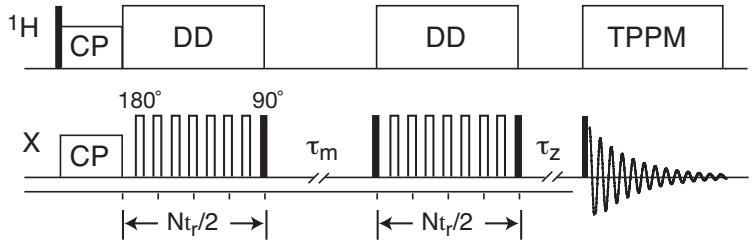


Fig. 1 CODEX pulse sequence. Filled and open rectangles indicate 90° and 180° pulses, respectively. CP: cross polarization. DD: dipolar decoupling. TPPM: two-pulse phase modulation [45].

The pulse sequence for the ^{19}F and ^{13}C CODEX experiments is shown in Fig. 1 [14]. Two rotor-synchronized π-pulse trains recouple the CSA interaction. During the mixing time (τ_m), spin diffusion changes the CSA frequency and prevents complete refocusing of the stimulated echo. To correct for spin-lattice relaxation (T_1) effects during τ_m, a z-filter (τ_z) is added at the end of the second π-pulse train. Two experiments, an exchange experiment (S) with the desired τ_m and a short τ_z (10 μs), and a reference experiment (S_0) with interchanged τ_m and τ_z values, were conducted. The normalized intensity, S/S_0, was measured as a function of the mixing time until it reaches a plateau. Error bars were propagated from the signal-to-noise ratios of the S_0 and S spectra.

Simulation of CODEX curves

The τ_m-dependent CODEX curves were calculated in MATLAB using Eqs. (2-7). *n*-dimensional rate matrices were constructed for oligomers of size *n*, where the rate constants are calculated according to Eq. (2). The mixing-time-dependent magnetization ***M(t)*** was calculated using Eq. (7). For model compounds with known internuclear distances and dipolar couplings, best-fit simulation of the experimental exchange curve yields the overlap

integral, F(0). The best fit was obtained by minimizing the root-mean-square-derivation (RMSD) between the calculated intensity I_{sim} and the experimental intensity I_{exp} as a function of F(0):

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{N}(I_{sim,i} - I_{exp,i})^2}{N}}, \quad (1)$$

where N is the number of the data points in the exchange curve. The F(0) values were incremented in 5 μs steps for ^{13}C CODEX and 1 μs steps for ^{19}F CODEX.

Structure visualization and modeling

The crystal structures of model compounds were obtained from Cambridge Structure Databank and visualized in the software Mercury. The crystal structures of L-Leucine and 5-^{19}F-L-Trptophan were also measured locally by single-crystal X-ray crystallography and confirmed to be consistent with the literature. The M2TM peptide structure was visualized and modeled in Insight II (Accelrys, San Diego). The NMR-derived tetramer model (PDB accession code: 1NYJ) [23] was used as a starting structure, with Ala30 replaced by Phe. To find tilt-angle dependent interhelical F-F distances, we separately changed the tilt angle of each helix in the tetramer while fixing the position of G34, the middle of the helix. To define the new tilt angle, we first rotated the entire helical bundle together so that the helix to be modified is parallel to the screen. This helix was then rotated around the z-axis by either 38˚ or –38˚ so that it became vertical. The 38˚ angle is the ^{15}N-NMR measured tilt angle of each M2 monomer. Finally, the vertical helix was rotated by the desired new angle around the z-axis. This process was then repeated for the other helices in the bundle. During this tilt-angle modification, the rotation angle of the helices and the interhelical G34-G34 distances remained unchanged. The interhelical F-F distances at (4-19F) Phe30 were then measured as a function of the tilt angle.

An alternative set of M2TM structural models (courtesy of DeGrado, Howard, and coworkers) based on Cys scanning data [21] and EPR constraints [25] was also examined to obtain the interhelical F-F distances. The set of models contains helix tilt angles from 10˚ to 35˚, and was used without modification.

Analysis of CODEX curves by ^{1}H-driven spin diffusion

In this section we describe the ^{1}H-driven spin diffusion theory used to fit the CODEX curve to obtain distance information. Since all samples used in our experiments are singly labeled in either ^{13}C' or ^{19}F, magnetization exchange occurs between spins with identical isotropic shifts but different shielding tensor orientations. Magnetization exchange among *n* such chemically equivalent but magnetically inequivalent sites reduces the intensity of the CSA echo to an equilibrium value of $1/n$ [15].

During the mixing time of a ^{1}H-driven spin diffusion experiment, polarization transfer between spins X_i and X_j occurs due to X_i-X_j dipolar coupling and is facilitated by the coupling of X to the abundant protons. The rate constant k_{ij} for this process is given by the first-order perturbation theory [26, 27]:

$$k_{ij} = 0.5\pi \cdot \omega_{ij}^2 \cdot F_{ij}(0), \tag{2}$$

where ω_{ij} is the homonuclear dipolar coupling

$$\omega_{ij} = \frac{\mu_0}{4\pi}\gamma^2\hbar\frac{1}{r_{ij}^3}\frac{(1-3\cos^2\theta_{ij})}{2}, \tag{3}$$

which depends on the internuclear distance r_{ij} and the angle θ_{ij} between the internuclear vector and the external magnetic field. $F_{ij}(0)$ is the overlap integral describing the probability that single-quantum transitions occur at the same frequency for spins i and j:

$$F_{ij}(0) = \int_{-\infty}^{+\infty} f_i(\omega - \omega_i) f_j(\omega - \omega_j) d\omega, \tag{4}$$

where $f_i(\omega-\omega_i)$ is the normalized single-quantum lineshape of spin i in the absence of proton decoupling and ω_i is the center of the lineshape. The overlap integral is related to the normalized zero-quantum lineshape at zero frequency [26].

For spin diffusion among *n* X spins, the time-evolution of the *n*-dimensional vector of the z magnetization, $\vec{M}(t)$, is given by

$$\frac{d\vec{M}(t)}{dt} = -\hat{\mathbf{K}}\vec{M}(t), \tag{5}$$

where $\hat{\mathbf{K}}$ is the *n*-dimensional exchange matrix containing the rate constants k_{ij}. T_1 relaxation is not included in Eq. (5) since it is removed in the CODEX experiment by normalization of the exchange intensity S with the reference intensity S_0. Detailed balance of equilibrium magnetization requires that the sum of each column of the $\hat{\mathbf{K}}$ matrix be zero and that the rate constants satisfy $k_{ij} = k_{ji}$ for equal populations of equilibrium magnetization [28]. Thus, for example, the exchange matrix of a four-spin system ABCD is

$$\hat{\mathbf{K}} = \begin{pmatrix} k_{AB}+k_{AC}+k_{AD} & -k_{BA} & -k_{CA} & -k_{DA} \\ -k_{AB} & k_{BA}+k_{BC}+k_{BD} & -k_{CB} & -k_{DB} \\ -k_{AC} & -k_{BC} & k_{CA}+k_{CB}+k_{CD} & -k_{DC} \\ -k_{AD} & -k_{BD} & -k_{CD} & k_{DA}+k_{DB}+k_{DC} \end{pmatrix}, \tag{6}$$

The rate matrix includes not only direct but also relayed magnetization transfer effects. For example, the magnetization transfer from spin A to spin C is mediated not only by the matrix element $-k_{AC}$, which reflect the direct transfer, but also by $-k_{AB}$ and $-k_{BC}$, which reflect the relayed transfer pathway $A \rightarrow B \rightarrow C$.

The formal solution to Eq. (5) for a given initial magnetization distribution $\vec{M}(0)$ is

$$\vec{M}(t) = e^{-\hat{\mathbf{K}}t}\vec{M}(0). \tag{7}$$

The exponential operator can be treated by diagonalization of the $\hat{\mathbf{K}}$ matrix or calculated directly in a matrix-based software such as MATLAB. Expressed in terms of the diagonal exchange matrix $\hat{\mathbf{D}}$, $\vec{M}(t) = \hat{R}e^{-\hat{\mathbf{D}}t}\hat{R}^{-1}\vec{M}(0)$, where $\hat{\mathbf{R}}$ is the eigenvector matrix of $\hat{\mathbf{K}}$. For an *n*-dimensional exchange matrix with zero-sum columns, it can be shown that one eigenvalue is always zero with the associated eigenvector of $(1/\sqrt{n}, 1/\sqrt{n}, \ldots 1/\sqrt{n})$, while all other eigenvalues of $\hat{\mathbf{K}}$ are positive. As a result, at large mixing times, $\vec{M}(t)$ approaches $(1/n, 1/n, \ldots 1/n)$, corresponding to complete equilibration of the initial magnetization.

The distance-dependent dipolar coupling in Eq. (3) contains an angular term, $(1-3\cos^2\theta_{□□})$, which depends on the powder angles of individual molecules in the magnetic field. To simplify the analysis, the square of this term, which is relevant in the rate constant expression of Eq. (2), can be replaced by its powder-averaged value of 0.8 [27]. We also make the approximation that the value of the overlap integral is the same for all spin pairs. This is reasonable since most factors that influence F(0), including isotropic chemical shifts, chemical shielding principal values, X-^{1}H dipolar couplings, and ^{1}H-^{1}H dipolar couplings [27], are either identical or similar for all spin pairs in the current experiment. The main difference between spins is the CSA tensor orientation due to the random orientation distribution of molecules.

For crystalline small-molecule compounds, we consider not only the nearest-neighbor distances between magnetically inequivalent sites but also longer distances, by replacing the

ω_{ij}^2 term in Eq. (2) with its second moment, $\sum_{i,j}\omega_{ij}^2$. For amino acids, the dipolar second moment typically converges within a distance of ~15 Å, and exceeds the nearest-neighbor coupling term by about a factor of 2. In comparison, for membrane peptides, long-range dipolar couplings between different aggregates are negligible due to lipid dilution, thus simplifying the extraction of the distances within an aggregate or oligomer.

Results and Discussion

^{13}C' CODEX of model compounds

We first analyze the ^{13}C' CODEX spin diffusion data of ^{13}C'-labeled amino acids: α-Gly, γ-Gly, L-Leu and L-Phe. These amino acids have 2, 3 or 4 magnetically inequivalent molecules in the respective unit cells, and thus should give equilibrium S/S_0 values of 0.50, 0.33, and 0.25. Fig. 2a shows the τ_m-dependent S/S_0 values of α-Gly (Fig. 2a) at 5 kHz MAS, reproduced from an earlier study [13]. A single exponential decay with a time constant of 265 ms and an equilibrium value of 0.49 were found. α-Gly crystallizes in space group $P2_1/n$ and has four molecules in the unit cell divided into two pairs that are related by inversion symmetry [29, 30]. The nearest-neighbor C'-C' distance between the two inequivalent molecules is 4.22 Å. With a dipolar second moment that converges at a distance of 15 Å, the best fit yields an F(0) value of 50 μs (Fig. 2b, c).

γ-Gly crystallizes in the hexagonal space group $P3_2$ with three magnetically inequivalent molecules in the unit cell. The two nearest-neighbor C'-C' distances are 4.17 Å [31]. The γ-Gly CODEX intensities (Fig. 3a) exhibit a single exponential decay with a time constant of 121 ms and the expected equilibrium value of 0.32. Using a dipolar second moment that converges within a distance of 15 Å [31], the best-fit spin diffusion curve yields an F(0) value of 120 μs (Fig. 3b, c).

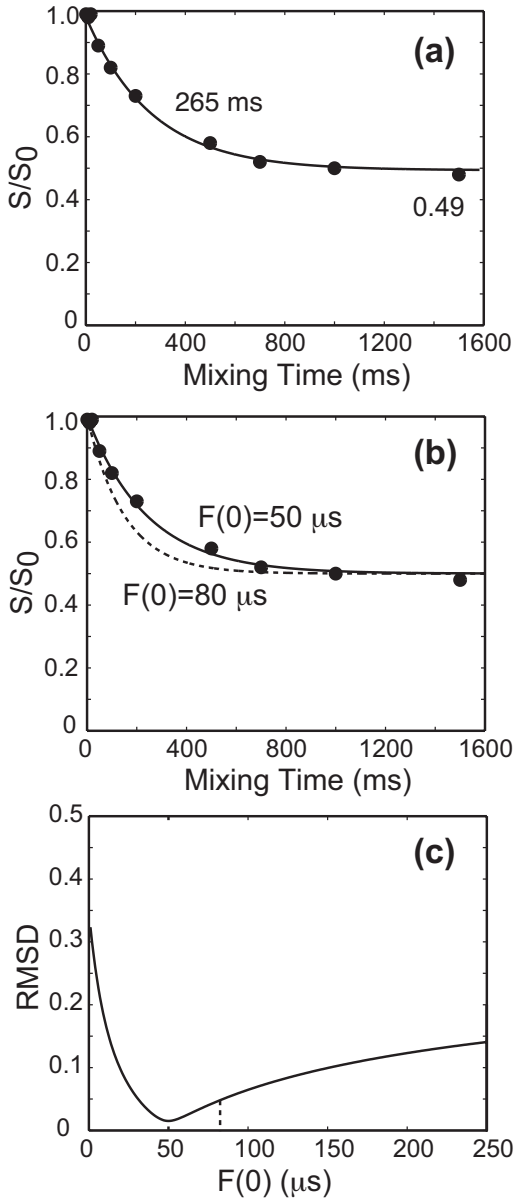


Fig. 2 ^{13}C CODEX of ^{13}C'-α-Gly. (a) Experiment data (circles) are best fit to a single exponential decay with a time constant of 265 ms. (b) Experimental data superimposed with the calculated magnetization exchange curves using the best-fit F(0) of 50 μs (solid line) and the consensus F(0) value of 80 μs (dashed line). (c) RMSD between the simulations and the experiment as a function of F(0). Vertical dashed line indicates the RMSD for the consensus F(0) value of 80 μs.

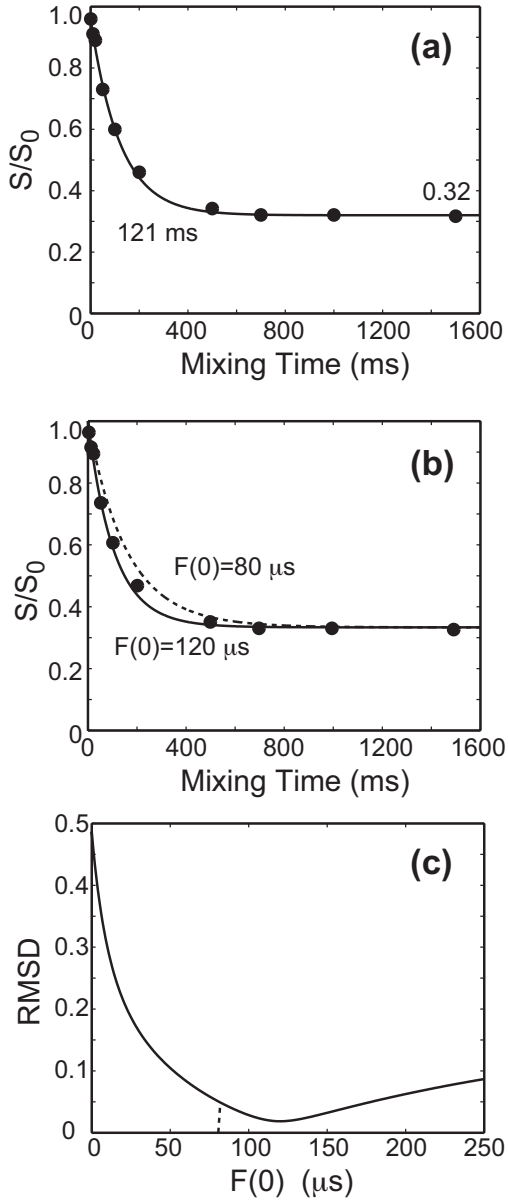


Fig. 3 ^{13}C CODEX of ^{13}C'-γ-Gly. (a) The experimental data (circles) are best fit to a single exponential decay with a time constant of 121 ms. (b) Experimental data superimposed with the calculated magnetization exchange curves using the best-fit F(0) of 120 μs (solid line) and the consensus F(0) of 80 μs (dashed line). (c) RMSD between the simulations and the experiment as a function of F(0). Vertical dashed line indicates the RMSD for the consensus F(0) value of 80 μs, which is close to that of the best-fit F(0).

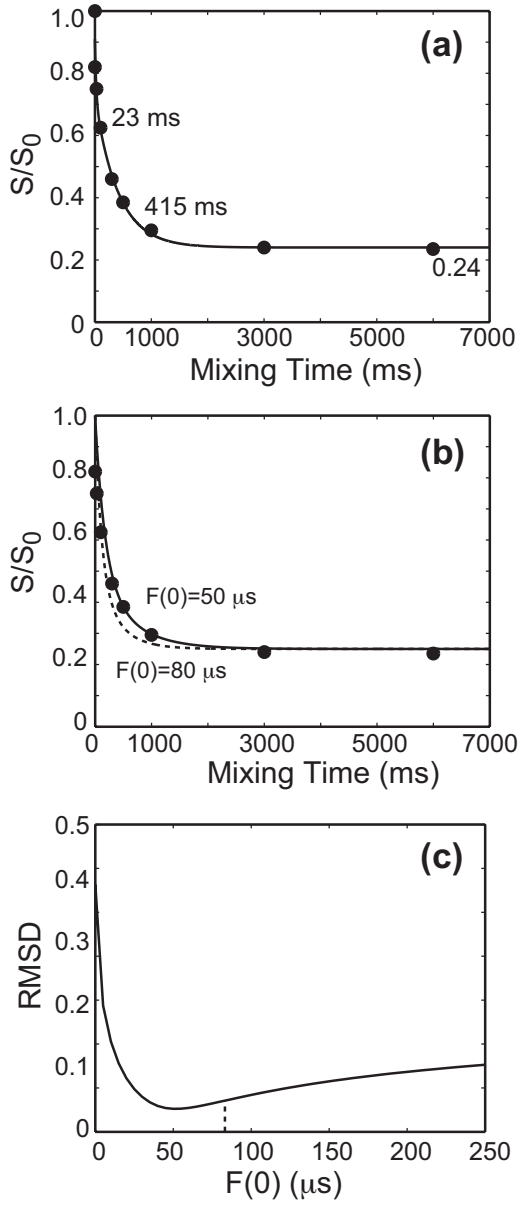


Fig. 4 ^{13}C CODEX of ^{13}C'-Leu. (a) The experimental data (circles) are best fit to a double exponential decay with time constants of 23 ms and 415 ms. (b) Experimental data superimposed with the calculated magnetization exchange curves using the best-fit F(0) of 50 μs (solid line) and the consensus F(0) value of 80 μs (dashed line). (c) RMSD between the simulations and the experiment as a function of F(0).

L-Leu and L-Phe both contain four orientationally unique molecules in the unit cell. L-Leu crystallizes in space group $P2_1$ [32, 33] while L-Phe crystallizes in space group $P2_12_12_1$ [34]. Leu has nearest-neighbor C'-C' distances of 3.84, 5.59 and 6.01 Å, while Phe has much longer nearest-neighbor distances of 5.12, 11.0, 14.0 Å. Consequently, the CODEX curves of the two compounds exhibit quite different decay rates. Leu gives rise to a bi-exponential decay with time constants of 23 ms and 415 ms (Fig. 4a), while Phe has slower decays with time constants of 347 ms and 3.0 s (Table 1). The best-fit CODEX curve of Leu yields an F(0) value of 50 μs (Fig. 4b, c), while that of Phe gives an F(0) value of 120 μs. The best-fit curve for these two compounds does not fully capture the multi-exponential nature of the decays, which may be due to imperfection of the spin diffusion theory for spin networks involving very different distances.

Table 1: Experimental decay constants and best-fit F(0) values of ^{13}C' CODEX data for different samples at various spinning speeds.

Compound	$(S/S_0)_{eq}$ [a]	2.5 kHz		5 kHz		10 kHz	
		τ_{SD} (ms)	F(0) (μs)	τ_{SD} (ms)	F(0) (μs)	τ_{SD} (ms)	F(0) (μs)
α-Gly	0.47	139	90	265	50	477	35
γ-Gly	0.32	74	180	121	120	186	85
Leu	0.24	206[b]	70	261[b]	50	523[b]	35
Phe	0.28	NA	NA	347, 3000	120	NA	NA

[a]: Equilibrium S/S_0 values were averaged from measurements at various spinning speeds.

[b]: These τ_{SD} values were obtained from stretched exponential fits to reflect the average decay constants.

Table 1 summarizes the equilibrium values $(S/S_0)_{eq}$, exchange time constants τ_{SD}, and best-fit F(0) values for the four ^{13}C'-labeled amino acids at 5 kHz MAS. The overlap integrals fall within a factor of 2.5. Since the spin-diffusion rate constant depends on $1/r^6$ but

only linearly on F(0) (Eqs. 2, 3), the choice of an average F(0) value of 80 μs introduces only a small distance uncertainty of less than 8%. The calculated $^{13}C'$ CODEX curves for the various amino acids using this consensus F(0) value are shown as dashed lines in Fig. 2-4b. In general, they fit the experimental data well, and give low RMSD values close to the best-fit results.

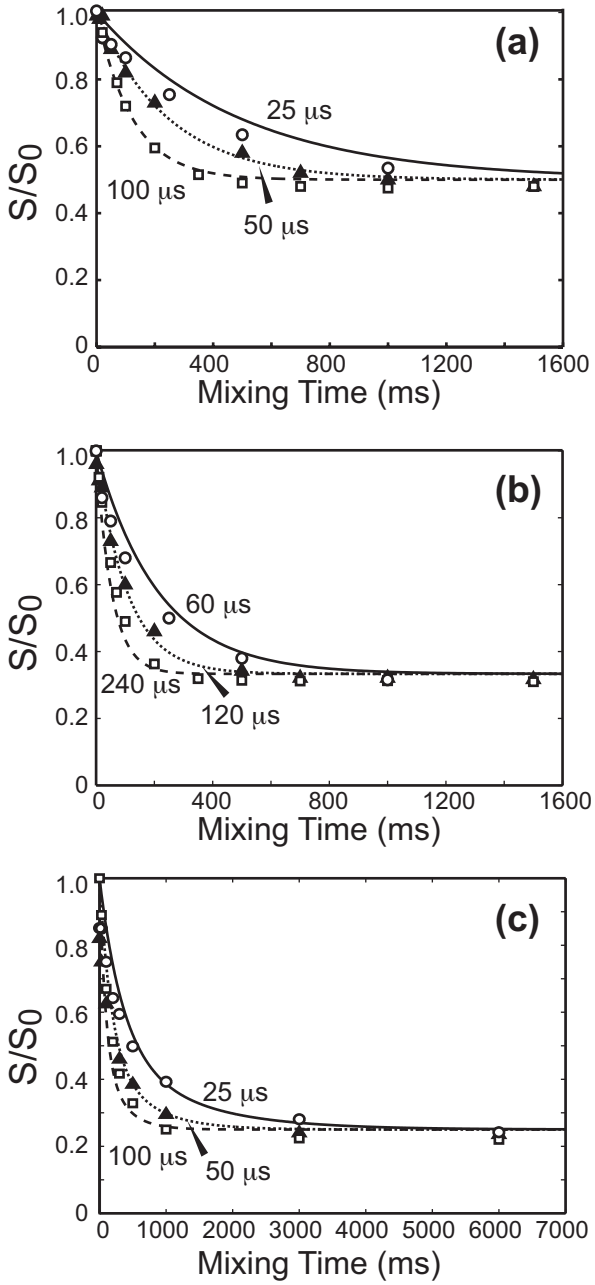


Fig. 5 ^{13}C' CODEX of (a) α-Gly, (b) γ-Gly, and (c) L-Leu at MAS rates of 10 kHz (circles, solid line), 5 kHz (triangles, dotted line), and 2.5 kHz (squares, dashed line). The best-fit F(0) for the 5 kHz data of each sample is scaled by $5000/\nu_r$ to fit the data at the other two spinning speeds.

The consensus ^{13}C' F(0) value measured here is on the same order of magnitude as that obtained by Suter and Ernst on doubly carboxylic ^{13}C labeled single-crystal malonic acid [35]. By direct measurement of the static ^{1}H-coupled zero-quantum spectra, the authors obtained zero-quantum T_2 relaxation time, T_2^{ZQ}, of 137 μs and 78.3 μs for two crystal orientations. Since $F(0) = T_2^{ZQ}/\pi$ [26], these correspond to overlap integrals of 44 μs and 25 μs, which are within a factor of 3 of our consensus F(0) value of 80 μs.

The empirical ^{13}C' F(0) value is further validated by the spinning speed dependence of the CODEX curves. The efficiency of ^{1}H-driven spin diffusion depends on the spinning speed: faster spinning attenuates spin diffusion and thus should give larger time constants and smaller F(0) values. We investigated the effect of the spinning speed (ν_r) on F(0) by measuring the ^{13}C' CODEX curves at 2.5 kHz and 10 kHz MAS. Fig. 5 shows the CODEX curves at the three spinning speeds. Using overlap integral values scaled by $5000/\nu_r$ from the F(0) value at 5 kHz MAS, we calculated the exchange curves for the two additional spinning speeds. Fig. 5 shows that these calculated curves agree with the experimental data well. More exactly, the best-fit F(0) values at the three MAS rates follow a $(1/\nu_r)^{\beta}$ dependence, where β varies between 0.5 and 0.9. This ν_r dependence is qualitatively consistent with the previously observed inverse-ν_r dependence of the spin diffusion rate [36].

^{19}F CODEX of model compounds

Compared to ^{13}C CODEX, ^{1}H-driven ^{19}F spin diffusion has two advantages: a longer distance reach due to the larger gyromagnetic ratio of ^{19}F, and higher angular sensitivity due to the larger ^{19}F CSA. We use two fluorinated compounds, 5-^{19}F-Trp and 4-^{19}F-2'-

nitroacetanilide, to measure the overlap integral of ^{19}F CODEX at 8 kHz MAS. Fig. 6(a-c) shows the Trp data: an equilibrium value of 0.50 and a τ_{SD} of 5.3 ms were found, consistent with the $P2_1$ space group of Trp (with two molecules in the unit cell) [37]. The nearest-neighbor ^{19}F-^{19}F distance is 4.62 Å, giving a large dipolar coupling of 1.1 kHz, which accounts for the short time constant observed. The best fit is obtained at $F(0) = 28$ μs (Fig. 6b, c), significantly smaller than the $^{13}C'$ F(0) values. This is reasonable: the instantaneous ^{19}F chemical shift overlap is smaller than the $^{13}C'$ chemical shift overlap due to the larger ^{19}F CSA: the anisotropy parameter (δ) of 5-^{19}F-Trp is 48 ppm, or 18 kHz, almost twice that of the carbonyl carbons. For 4-^{19}F-2'-nitroacetanilide [38], the CODEX time constant is larger by a factor of 10, since the nearest-neighbor intermolecular F-F distance is much longer, 11.5 Å. Despite this large structural difference, the best-fit F(0) value, 45 μs, is remarkably close to that of 5-^{19}F Trp. This confirms the existence of consensus F(0) values under similar experimental conditions and validates our empirical approach of extracting the F(0) value. We took the average F(0) of 37 μs as the consensus value for ^{19}F CODEX at 8 kHz MAS. Simulations using this consensus value agree with the experimental data of both compounds well (Fig. 6b, e).

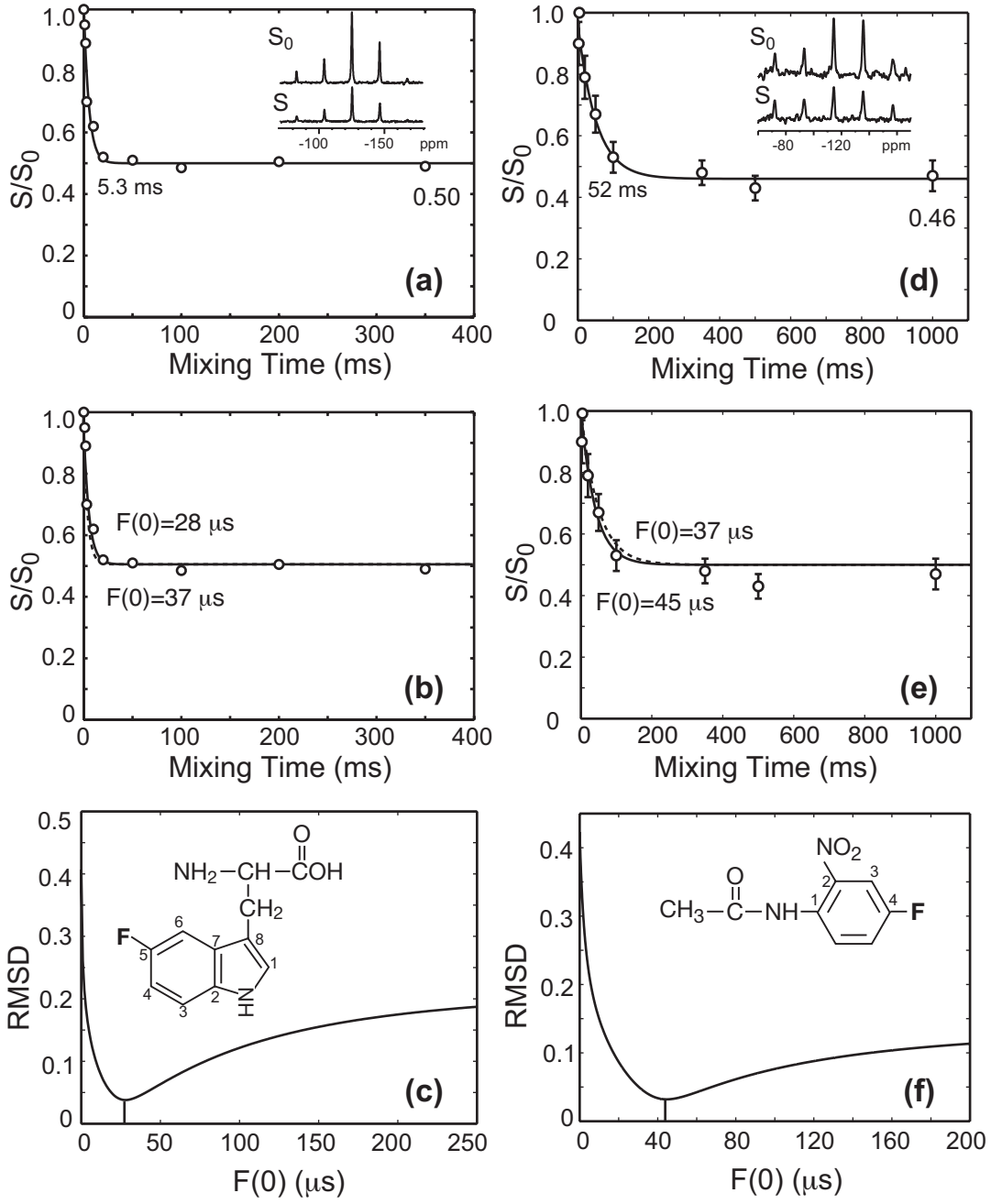


Fig. 6 ^{19}F CODEX of (a-c) 5-^{19}F-Trp and (d-f) 4-^{19}F-2'-nitroacetanilide, whose chemical structures are shown in (c) and (f). (a) The Trp data (circles) exhibits a single exponential decay with a time constant of 5.3 ms and an equilibrium value of 0.50. (b) Calculated magnetization exchange curves using the best-fit F(0) value of 28 μs (solid line) and the average F(0) value of 37 μs (dashed line) both agrees with the experimental data well. (c) Trp RMSD between the simulation and the experiment as a function of F(0). (d) ^{19}F CODEX data

of 4-^{19}F-2'-nitroacetanilide (circles), with a decay constant of 52 ms and an equilibrium value of 0.46. (e) The best-fit curve corresponds to an F(0) value of 45 μs (solid line), while the average F(0) of 37 μs still agrees with the data well. (f) RMSD as a function of F(0) for 4-^{19}F-2'-nitroacetanilide.

^{19}F CODEX of M2TM in lipid bilayers

With the F(0) value known for ^{19}F CODEX at 8 kHz MAS, we can now determine F-F distances in structurally unknown systems such as the M2TM peptide. Since the rate constant k_{ij} is proportional to $1/r^6$ but linear with F(0) (Eq. 2), small variations in the F(0) value do not affect the distance appreciably. In addition, for peptides diluted in lipid bilayers at molar concentrations of a few percent, the inter-oligomeric distances are several times longer than the intra-oligomeric distances, making the inter-oligomeric dipolar couplings more than two orders of magnitude smaller than the intra-oligomeric couplings. Thus, the second moment treatment for single-component solids is unnecessary for two-component peptide-lipid mixtures, simplifying the analysis.

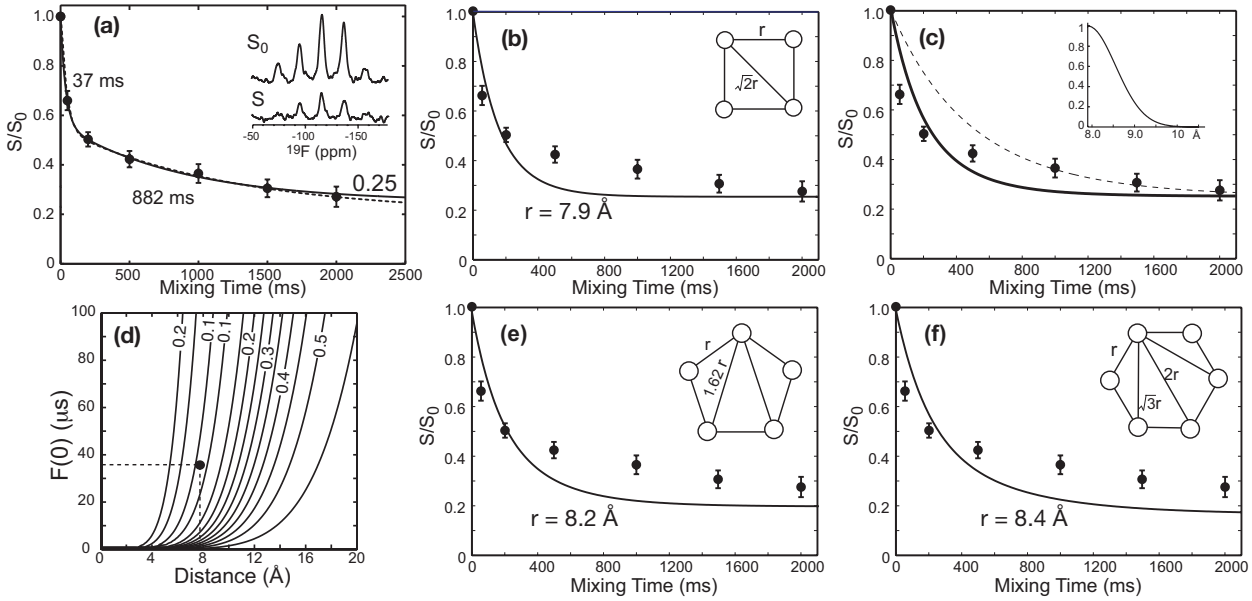


Fig. 7 (a) ^{19}F CODEX data (circles) of the M2TM peptide in DMPC bilayers at 240 K and 8 kHz MAS. The S_0 and S spectra for a mixing time of 1.0 s are shown. A bi-exponential fit

with an equilibrium value constrained at 0.25 (solid line) shows excellent agreement with the data, with $R^2 = 0.9990$. Free fitting (dashed line) yields comparable decay constants and an equilibrium value of 0.20±0.05 ($R^2 = 0.9995$). (b) Best-fit exchange curve using F(0) = 37 μs gives a nearest-neighbor F-F distance of 7.9 Å for a symmetric tetramer model, which is shown in the inset. (c) A half-Gaussian distance distribution of 7.9 – 9.5 Å (inset) using the same tetramer model fits the data well. (d) RMSD between simulated and experimental CODEX curves as a function of distance r and F(0). Contour levels vary from 0.10 to 0.55 in increments of 0.05. The minimum RMSD for F(0) = 37 μs is obtained at r = 7.9 Å (circle). (e) Best-fit CODEX exchange curve using a symmetric pentamer model deviates from the measured equilibrium S/S_0 value. (f) Best-fit CODEX exchange curve using a symmetric hexamer model also deviates from the measured equilibrium S/S_0 value.

The ^{19}F CODEX experiments on the M2TM peptide were carried out at 240 K, 56 K below the DMPC phase transition temperature (296 K) to eliminate slow peptide motion. Our previous experiments showed that at this reduced temperature, slow motion is frozen, leaving dipolar spin diffusion as the only mechanism of exchange [13]. A CSA recoupling time, $N\tau_r$, of 0.25 ms was used, giving a $2\pi\delta\tau_r$ of 13π ($\delta = 69$ ppm). This is sufficiently large to detect orientational differences as small as ~7° [39]. Fig. 7a shows the CODEX curve of the M2TM peptide and a representative pair of spectra. Uncertainties in the S/S_0 values were propagated from the spectral signal-to-noise ratios and are small because of the high sensitivity of the spectra (Fig. 7a inset). The S/S_0 value of the longest mixing time is 0.27, consistent with a tetrameric bundle. To fit the data, we first used a constrained bi-exponential function with a fixed equilibrium value of 0.25. The fit curve, $S/S_0 = 0.25 + 0.44e^{-t/37} + 0.31e^{-t/882}$ (solid line), shows excellent agreement with the experimental data ($R^2 = 0.999$), indicating that the M2TM peptide forms tetrameric bundles in lipid bilayers, similar to the situation when it binds to detergent micelles [19, 20].

Previous ^{15}N 2D dipolar-shift correlation spectra indicate that the M2TM helical bundle is highly symmetric in lipid bilayers, since each ^{15}N label gives rise to a single ^{15}N peak [22]. This symmetry indicates that the four ^{19}F spins of Phe30 are located at the corners of

a square, whose side r is the nearest-neighbor interhelical F-F distance between adjacent helices (Fig. 7b inset). Using this symmetric tetramer model and the consensus F(0) value of 37 μs, we simulated the experimental CODEX curve and found the best-fit at a F-F distance of 7.9 Å (Fig. 7b). Fig. 7d shows the 2D RMSD contour plot between the simulations and the experiment as a function of F(0) and r. The contour lines are nearly parallel to the F(0) axis, confirming that the fitting is insensitive to F(0) and primarily dependent on the distance r. The distance uncertainty, allowing for a conservative estimate of an F(0) uncertainty of a factor of 2 on each side of the consensus value (i.e. 18 – 75 μs), amounts to only 0.8 Å.

The best-fit spin diffusion curve with r = 7.9 Å falls below the experimental intensities at mixing times of 500 ms and 1 s (Fig. 7b). This discrepancy reflects the bi-exponential nature of the measured CODEX curve. Experiments at a lower temperature of 227 K verified that no residual slow motion is present at 240 K to cause the fast initial decay of the intensities. An obvious possibility for the bi-exponentiality of the CODEX curve is structural heterogeneity of the M2 helical bundle. Fig. 7c shows that the calculated CODEX curve for 9.5 Å (dashed line), which reproduces the long-time exchange intensities, is the upper bound that is still consistent with the experimental data. Thus, a conservative estimate of the interhelical F-F distance at A30F is 7.9 – 9.5 Å. Alternatively, simulation using a Gaussian distance distribution centered at 8.5 Å also gave a more balanced agreement with both the short and long time points of the experimental data (solid line, Fig. 7c). We note that although a double Gaussian distribution centered at two different distances can better capture the bi-exponentiality of the curve, there is no good physical justification for a bimodal distance distribution, and the ^{19}F spectral lineshape also does not give any sign for two distinct conformations of the peptide. Another possible reason for the imperfection of the fit is the phenomenological nature of the ^{1}H-driven spin diffusion theory used to simulate the data.

While the constrained bi-exponential fit of Fig. 7a agrees with the data well, free fitting, on the other hand, gave an equilibrium value of 0.20 ± 0.05 (dashed line, Fig. 7a). This implies that pentamers, with an expected equilibrium value of 0.20, and even hexamers, with an equilibrium value of 0.16, cannot be completely ruled out based on the equilibrium S/S_0 value. Fortunately, the ambiguity is readily resolved by calculating the CODEX curves

for a symmetric pentamer (Fig. 7e) and a symmetric hexamer (Fig. 7f). We find that even the best-fit exchange curve in each model (obtained by varying r) deviates substantially from the experimental data at long times: the calculated equilibrium intensities fall below the measured values by more than the experimental uncertainty. Thus, a combination of the CODEX equilibrium value and the exchange time course is important for determining the oligomeric size of large aggregates, whose $1/n$ values become more difficult to distinguish from $1/(n+1)$ or $1/(n-1)$ as n increases.

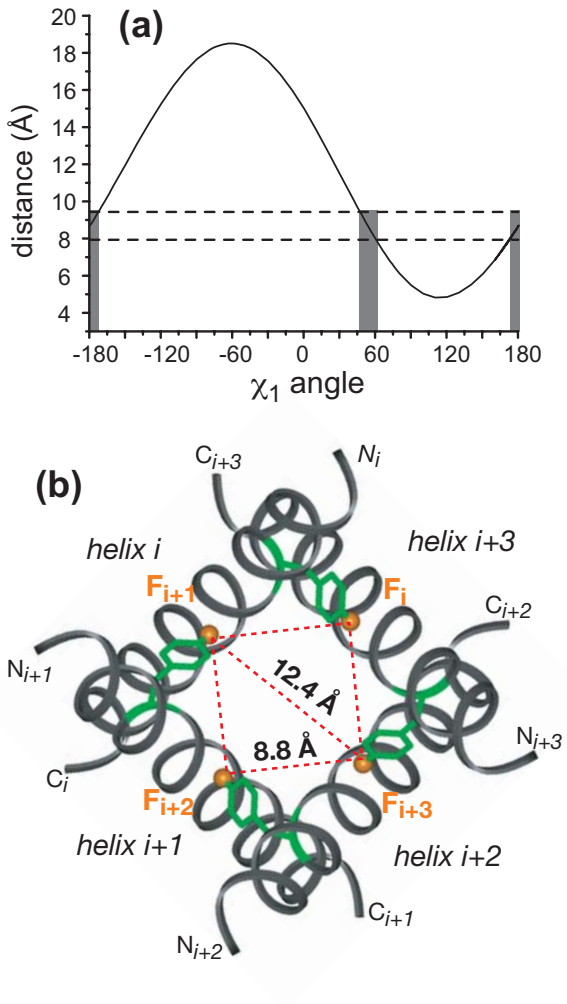


Fig. 8 (a) ^{19}F-^{19}F distance between (4-19F) Phe30 of adjacent helices as a function of Phe χ_1 angle in the NMR-derived model of M2TM tetramer (PDB: 1NYJ). A distance of 7.9 – 9.5 Å is satisfied near χ_1 angles of +60˚ and +180˚. However, the χ_1=+60˚ rotamer causes severe steric conflicts between the sidechain and the backbone and is thus ruled out. (b) Top view of

the NMR-based M2TM tetramer model with a nearest-neighbor ^{19}F-^{19}F distance of 8.8 Å at Phe30 between adjacent helices. The Phe ring of helix i + 1 protrudes from its backbone and lies close to the neighboring helix i.

What is the structural implication of the interhelical Phe30-Phe30 distance of 7.9 – 9.5 Å. The intermolecular distance depends on a number of factors: the channel diameter, the helix orientation, and the Phe χ_1 torsion angle. We first assess the χ_1 angle. In the NMR-derived tetramer model (PDB accession code: 1NYJ), the measured F-F distance can only be satisfied for χ_1=+60˚ and 180˚ (Fig. 8a). However, out of the three Phe χ_1 conformers, χ_1=+60° is energetically forbidden in α-helices due to steric conflicts between the aromatic ring and the backbone [40, 41]. Thus only the χ_1 = 180˚ rotamer agrees with the data. This is also the most favorable rotamer for Phe in α-helices [40, 41]. With χ_1 = 180˚, the F-F distance between adjacent helices in the NMR-based tetramer model is 8.8 Å, which is compatible with the ^{19}F CODEX result within experimental uncertainty. Fig. 8b shows a cross section of this tetramer model with the F-F distances highlighted. The Phe30 ring of helix i+1 protrudes from its backbone in a roughly perpendicular direction and lies close to the neighboring helix i (without causing steric conflicts). The distance between the opposing helices is 12.4 Å, which gives a measure of the diameter of the channel at this residue.

With the Phe χ_1 conformation known, the measured F-F distance provides a valuable constraint to the orientation of the M2TM helices and the diameter of the helical bundle. We first examine the tilt angle dependence. Starting from the original NMR-derived tetramer model of Cross and coworkers [22, 23], which have a helix tilt angle of 38˚, we varied the tilt angle of the helices while holding the channel diameter and the helix rotation angle unchanged. The rotation angle around the helix axis defines which residues face the channel interior versus the lipids. This information is well known qualitatively from functional studies of the M2 protein [21] and quantitatively from 2D ^{15}N NMR spectra of oriented membranes [22]. The interhelical Phe30 F-F distances are plotted in Fig. 9 for the NMR model (filled circles). It can be seen that tilt angles less than 20˚ give distances outside the measured range of 7.9 – 9.5 Å and thus can be ruled out.

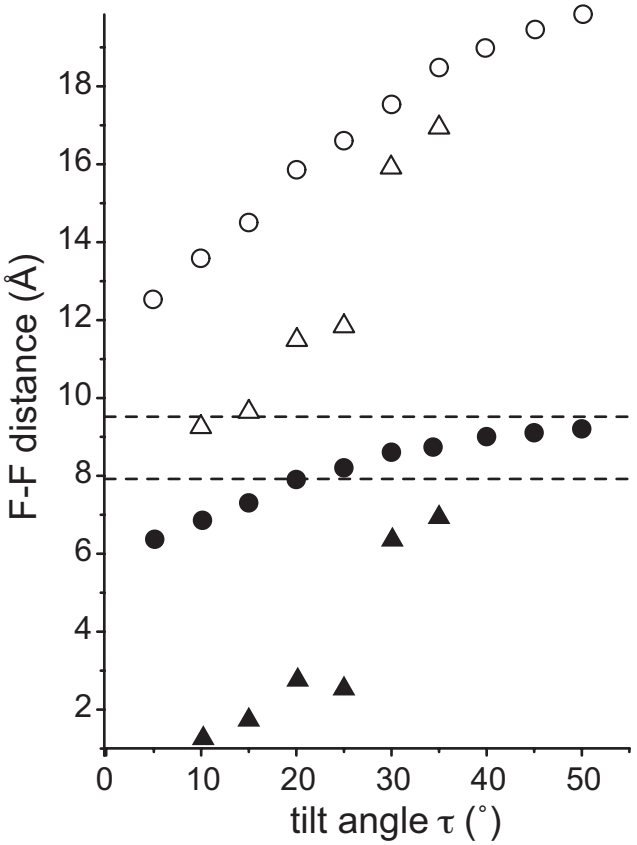


Fig. 9 Phe30 ^{19}F-^{19}F distances between adjacent helices in the M2TM tetramer as a function of the helix tilt angle for two different structural models. Circles: NMR-derived model based on ^{15}N orientational constraints. Triangles: Functional model based on Cys mutagenesis and EPR data. The χ_1=180˚ and χ_1=-60˚ rotamers are shown as filled and open symbols, respectively. The F-F distance range determined from the present study is defined by the dashed lines.

An alternative M2TM tetramer model (designated as the functional model below) was proposed by DeGrado and coworkers based on Cys scanning data [21]. The transmembrane segment of the intact protein was successively mutated with Cys and the effects of the mutation on the reverse potential, ion currents, and amantadine resistance of the channel were measured. Fourier analysis of the data showed a periodicity that is consistent with a tetrameric helical bundle. This functional model was further refined by a recent EPR study of the M2TM peptide in phosphocholine membranes of different thicknesses [25]. By observing

the dipolar line broadening of the spin label at the N-terminus of the helix, the authors found qualitative evidence that the distance between the N-termini of the various helices decrease with increasing membrane thickness, suggesting that the M2TM helices become less tilted in thicker membranes. Fig. 9 plots the F-F distances for the different tilt angles of this functional model (triangles). The distance increases with increasing tilt angles as expected. At a small tilt angle such as 15˚, a distance of 1.7 Å is found for χ_1 =180˚ (filled triangles), which is clearly unphysical. At this helix tilt, a χ_1 angle of –60˚ gives a F-F distance of 10 Å, which is more reasonable but still inconsistent with the ^{19}F CODEX result. Thus, small tilt angles can also be ruled out in this functional model. At the largest tilt angle of 35˚, the trans rotamer gives a F-F distance of 7.0 Å, which still falls short of the measured value by ~1.0 Å. While larger tilt angles would better agree with the experimental data, they would create unfavorable hydrophobic mismatch between the peptide and the lipid bilayer. An examination of the backbone structure of this functional model indicates that the channel diameter as defined by the backbone-backbone distances is comparable or even slightly larger than the channel diameter in the NMR-derived structural model. Thus, the channel diameter does not explain the short F-F distance in the model. However, the rotation angle around the helix axis is slightly different between the NMR model and this functional model. This small difference causes the Phe sidechain in the functional model to extend more into the center of the channel, and is sufficient to give rise to the 1.8 Å shorter distance compared to that of the NMR model (8.8 Å).

Clearly, it is difficult to use a single distance to uniquely define both the helix-helix diameter and the helix orientation. To determine both the interhelical separation and the peptide orientation, one needs to measure multiple intermolecular distances at multiple residues. The Phe30 F-F distance measurement shown here is only the first demonstration of the rich information content of such long-range intermolecular distance restraints. By determining distances in the 10 Å range, this anisotropic magnetization exchange technique complements other NMR distance methods such as rotational resonance that have been used to determine peptide interfacial structure [42]. The present ^{1}H-driven ^{19}F spin diffusion experiment can be compared with a ^{19}F-^{19}F radiofrequency-driven recoupling (RFDR) experiment, which yielded distances of 5–12 Å on model compounds with an accuracy of 1-2

Å [43]. The main differences between the two approaches are that the ^{19}F RFDR experiment does not allow spin counting and measures distances between chemically distinct fluorinated groups. Distance extraction from RFDR exchange curves also requires an adjustable simulation parameter, the zero-quantum T_2 relaxation time, T_2^{ZQ}, which is analogous to the spectral overlap function [26].

Compared to other spectroscopic probes such as EPR and fluorescence resonance energy transfer, the current ^{19}F CODEX NMR technique has the advantages that the label is relatively non-perturbing to the protein and does not introduce extra degrees of freedom to the distances of interest. Therefore, it promises to be a new and general probe for the oligomeric structure of membrane proteins.

Conclusion

We have shown that ^{1}H-driven spin diffusion between spins with identical isotropic shifts but different anisotropic shifts, CODEX, not only allows spin counting but also provides long-range distance constraints in oligomeric membrane peptides. Using a rate matrix analysis, in which the spin diffusion rate constant is proportional to the overlap integral and the square of the dipolar coupling, we determined consensus overlap integral values from the CODEX curves of structurally known model compounds. For $^{13}C'$-labeled amino acids at 5 kHz MAS, a consensus F(0) value of 80 μs was found. For aromatic ^{19}F-labeled compounds at 8 kHz MAS, a consensus F(0) value of 37 μs was obtained. Using ^{19}F CODEX, we proved for the first time that the M2 transmembrane peptide of influenza A virus forms tetramers in lipid bilayers. Moreover, the nearest-neighbor interhelical F-F distance between Phe30 residues is 7.9 – 9.5 Å. This supports the M2TM tetramer model obtained from NMR orientational constraints, and indicates that the helix tilt angle must be greater than 20° in DMPC bilayers. In addition, the experimental F-F distance points out a subtle difference in the helix rotation angle between the functional model of M2TM and the NMR-based model. This difference propagates to a detectably shorter sidechain F-F distance than the CODEX result. Thus, the long-range CODEX intermolecular distances are useful for refining the high-resolution oligomeric structure of membrane peptides.

The CODEX technique and analysis can be applied to both membrane proteins and protein aggregates outside the membrane [44] to determine oligomeric numbers and intermolecular distances in the 10 Å range. These intermolecular distances complement local structural parameters such as torsion angles and intramolecular distances in elucidating the full three-dimensional structure of complex biological assemblies.

Acknowledgement

This work is supported by a NSF CAREER grant to M. H. The authors thank Rajeswari Mani for various help in sample preparation and NMR experiments, and Prof. William DeGrado and Prof. Kathleen Howard for generously providing the coordinates of their M2 structure models.

References

1. Popot, J. L.; Engelman, D. M., *Biochemistry* **1990**, *29*, 4031-4037.
2. DeGrado, W. F.; Gratkowski, H.; Lear, J. D., *Protein Sci.* **2003**, *12*, 647-665.
3. Dusold, S.; Sebald, A., *Annu. Rep. NMR Spectrosc.* **2000**, *41*, 185-264.
4. Antzutkin, O. N., Solid-state NMR spectroscopy principles and applications, Duer, M. J., Ed. Blackwell Sciences, Inc.: Oxford, 2002; pp 280-390.
5. Gullion, T.; Schaefer, J., *J. Magn. Reson.* **1989**, *81*, 196-200.
6. Raleigh, D. P.; Creuzet, F.; Gupta, S. K. D.; Levitt, M. H.; Griffin, R. G., *J. Am. Chem. Soc.* **1989**, *111*, 4502-4503.
7. Bronniman, C. E.; Szeverenyi, N. M.; Maciel, G. E., *J. Chem. Phys.* **1983**, *79*, 3694-3700.
8. Tekely, P.; Potrzebowski, M. J.; Dusausoy, Y.; Luz, Z., *Chem. Phys. Lett.* **1998**, *291*, 471-479.
9. Castellani, F.; vanRossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H., *Nature* **2002**, *420*, 98-102.
10. Luo, W.; Yao, X. L.; Hong, M., *J. Am. Chem. Soc.* **2005**, *127*, 6402-6408.

11. Lange, A.; Seidel, K.; Verdier, L.; Luca, S.; Baldus, M., *J. Am. Chem. Soc.* **2003**, *125*, 12640-12648.

12. Huster, D.; Yao, X. L.; Hong, M., *J. Am. Chem. Soc.* **2002**, *124*, 874-883.

13. Buffy, J. J.; Waring, A. J.; Hong, M., *J. Am. Chem. Soc.* **2005**, *127*, 4477-4483.

14. deAzevedo, E. R.; Bonagamba, T. J.; Hu, W.; Schmidt-Rohr, K., *J. Am. Chem. Soc.* **1999**, *121*, 8411-8412.

15. Schmidt-Rohr, K.; deAzevedo, E. R.; Bonagamba, T. J., Encyclopedia of NMR, Grant, D. M.; Harris, R. K., Eds. John Wiley & Sons: Chichester, 2002.

16. Lamb, R. A.; Holsinger, K. J.; Pinto, L. H., Cellular Receptors of Animal Viruses, Wemmer, E., Ed. Cold Spring Harbor Lab Press: Plainview, NY, 1994; pp 303-321.

17. Duff, K. C.; Ashley, R. H., *Virology* **1992**, *190*, 485-489.

18. Sakaguchi, T.; Tu, Q.; Pinto, L. H.; Lamb, R. A., *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 5000-5005.

19. Salom, D.; Hill, B. R.; Lear, J. D.; DeGrado, W. F., *Biochemistry* **2000**, *39*, 14160-14170.

20. Howard, K. P.; Lear, J. D.; DeGrado, W. F., *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8568-8572.

21. Pinto, L. H.; Dieckmann, G. R.; Gandhi, C. S.; Papworth, C. G.; Braman, J.; Shaughnessy, M. A.; Lear, J. D.; Lamb, R. A.; DeGrado, W. F., *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11301-11306.

22. Wang, J.; Kim, S.; Kovacs, F.; Cross, T. A., *Protein Sci.* **2001**, *10*, 2241-2250.

23. Nishimura, K.; Kim, S.; Zhang, L.; Cross, T. A., *Biochemistry* **2002**, *41*, 13170-13177.

24. Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T., *Protein Sci.* **1995**, *4*, 2411-2423.

25. Duong-Ly, K. C.; Nanda, V.; DeGrado, W. F.; Howard, K. P., *Protein Sci.* **2005**, *14*, 856-861.

26. Meier, B. H., *Adv. Magn. Opt. Reson.* **1994**, *18*, 1-115.

27. Vanderhart, D. L., *J. Magn. Reson.* **1987**, *72*, 13-47.

28. Bain, A. D., *Prog. Nucl. Magn. Reson. Spectrosc.* **2003**, *43*, 63-103.

29. Jonsson, P.-G.; Kvick, A., *Acta Cryst.* **1972**, *B28*, 1827-1833.

30. Marsh, R. E., *Acta. Cryst.* **1958**, *11*, 654-663.

31. Iitaka, Y., *Acta. Cryst.* **1961**, *14*, 1-10.

32. Gorbitz, C. H.; Dalhus, B., *Acta. Cryst.* **1996**, *C52*, 1754-1756.

33. Harding, M. M.; Howieson, R. M., *Acta Cryst.* **1976**, *B32*, 633-634.

34. Al-karahouli, A. R.; Koetzle, T. F., *Acta. Cryst.* **1975**, *B31*, 2461-2465.

35. Suter, D.; Ernst, R. R., *Physical Review B* **1985**, *32*, 5608-5627.

36. Reichert, D.; Bonagamba, T. J.; Schmidt-Rohr, K., *J. Magn. Reson.* **2001**, *151*, 129-135.

37. Takigawa, T.; Ashida, T.; Sasada, Y.; Kakudo, M., *Bull. Chem. Soc. Jpn.* **1966**, *39*, 2369-2378.

38. Qi, J.-Y.; Zhou, Z.-Y.; Liu, D.-S.; Chan, A. S. C., *Acta Crysta.* **2001**, *E57*, o675-o676.

39. deAzevedo, E. R.; Bonagamba, T. J.; Hu, W.; Schmidt-Rohr, K., *J. Chem. Phys.* **2000**, *112*, 8988-9001.

40. Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson, D. C., *Proteins: Struct., Funct., Genet.* **2000**, *40*, 389-408.

41. Janin, J.; Wodak, S., *J. Mol. Biol.* **1978**, *125*, 375-386.

42. Smith, S. O.; Song, D.; Shekar, S.; Groesbeek, M.; Ziliox, M.; Aimoto, S., *Biochemistry* **2001**, *40*, 6553-6558.

43. Gilchrist Jr, M. L.; Monde, K.; Tomita, Y.; Iwashita, T.; Nakanishi, K.; McDermott, A. E., *J. Magn. Reson.* **2001**, *152*, 1-6.

44. Antzutkin, O. N.; Balbach, J. J.; Leapman, R. D.; Rizzo, N. W.; Reed, J.; Tycko, R., *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13045-13050.

45. Bennett, A. E.; Rienstra, C. M.; Auger, M.; Lakshmi, K. V.; Griffin, R. G., *J. Chem. Phys.* **1995**, *103*, 6951-6958.

Chapter 6

Sidechain Conformation of the M2 Transmembrane Peptide Proton Channel of Influenza A Virus from ^{19}F Solid-State NMR

A paper published in the Journal of Physical Chemistry

2007, vol. 111 pp. 10825-10832

Wenbin Luo, Rajeswari Mani and Mei Hong

Abstract

The M2 transmembrane peptide (M2TMP) of the influenza A virus forms a tetrameric helical bundle that acts as a proton-selective channel important in the viral life cycle. The sidechain conformation of the peptide is largely unknown and is important for elucidating the proton-conducting mechanism and the channel stability. Using a ^{19}F spin diffusion NMR technique called CODEX, we have measured the oligomeric states and inter-helical sidechain-sidechain ^{19}F-^{19}F distances at several residues using singly fluorinated M2TMP bound to DMPC bilayers. ^{19}F CODEX data at a key residue of the proton channel, Trp_{41}, confirms the tetrameric state of the peptide and yields a nearest-neighbor inter-helical distance of ~11 Å under both neutral and acidic pH. Since the helix orientation is precisely known from previous ^{15}N NMR experiments and the backbone channel diameter has a narrow allowed range, this ^{19}F distance constrains the Trp_{41} sidechain conformation to t90 ($\chi_1 \approx 180^\circ$, $\chi_2 \approx 90^\circ$). This Trp_{41} rotamer, combined with a previously measured ^{13}C-^{15}N distance between His_{37} and Trp_{41}[1], suggests that the His_{37} rotamer is t-160. The implication of the proposed (His_{37}, Trp_{41}) rotamers to the gating mechanism of the M2 proton channel is discussed. Binding of the antiviral drug amantadine to the peptide does not affect the F-F distance at Trp_{41}. Interhelical ^{19}F-^{19}F distances are also measured at residues 27 and 38, each mutated to 4-^{19}F-Phe. For V27F-M2TMP, the ^{19}F-^{19}F distances suggest a mixture of dimers and tetramers, whereas the L38F-M2TMP data indicate two tetramers of different sizes, suggesting sidechain conformational heterogeneity at this lipid-facing residue. This work shows that ^{19}F spin diffusion NMR is a valuable tool for determining long-range

intermolecular distances that shed light on the mechanism of action and conformational heterogeneity of membrane protein oligomers.

Introduction

The determination of the three-dimensional structure of membrane proteins and their assemblies requires long-range intermolecular distances > 10 Å. While NMR spectroscopy can readily measure short distances of ~5 Å accurately, long-range distance measurement remains a challenge [2]. Recently, we have shown ^{19}F spin diffusion to be a robust strategy for obtaining site-specific long-range distances. This approach, termed CODEX, involves measuring dipolar exchange between chemically identical but orientationally different ^{19}F spins on different molecules through a chemical-shift-anisotropy-based stimulated echo [3-5]. Intermolecular ^{19}F-^{19}F dipolar couplings in membrane peptide assemblies manifest as a decrease in the stimulated-echo intensity as a function of the exchange mixing time. The intensity decay is then fit to yield distances up to 15 Å with an uncertainty of 0.5 - 1.0 Å [6,7]. Both the number of spins in the cluster and internuclear distances can be obtained from each time-dependent echo decay curve.

The M2 protein of influenza A virus is a tetrameric proton channel essential in the life cycle of the virus. The channel is closed at neutral pH but opens in the acidic environment (low pH_{out}) of the endosome after viral entry into the host cell [8,9]. The acidification of the viral interior initiates the release of the viral RNA into the host cell, causing infection. Binding of amantadine blocks the proton channel and prevents infection. Elucidating the structure of the M2 proton channel in the closed and open states is thus important for understanding the mechanism of proton transfer and for designing antiviral drugs. Electrophysiological measurements on the M2 protein and analytical ultracentrifugation (AUC) experiments on the transmembrane domain of the M2 protein, M2TMP, have unambiguously established that His_{37} is the key residue for proton conductance [10-12].

In addition to His_{37}, Trp_{41} is another residue known to be important for proton conductance [9,13]. However, there are two different models about its specific role in M2 proton conductance. The first model was derived from proton conductance measurements showing

that the replacement of Trp_{41} by Phe causes an outward proton current when pH_{out} is high and pH_{in} is low, a condition under which the wild-type channel shows no outward proton flow [13]. This observation led to the conclusion that Trp_{41} is the gate of the channel, such that replacement by amino acids with smaller sidechains causes channel leakage. The proposed mechanism for channel gating by Trp_{41} involves conformational changes of the indole ring between the closed and open states: in the closed state it may obstruct the pore while in the open state it may become parallel to the pore axis, allowing the protons to flow.

The second model for the role of Trp_{41} in M2TMP results from pH-dependent UV resonance Raman spectra [14]: analysis of the wavenumbers and relative peak intensities indicated that there is no change in hydrogen bonding, environmental hydrophobicity, and sidechain torsion angles of the indole ring between the closed and open states [14]. Instead, based on spectral intensity changes, it was suggested that the only change is the addition of weak cation-π interactions between protonated His_{37} imidazole rings and Trp_{41} in the open state.

Because of the importance of His_{37} and Trp_{41} for proton conductance of the M2TMP channel, distance experiments and molecular dynamics (MD) simulations have been carried out to probe the sidechain conformation of these two residues. Solid-state NMR measurements of the His_{37} Nπ – Trp_{41} Cγ distance suggested (t-160, t-105) rotamers for the (χ_1, χ_2) angles of (His_{37}, Trp_{41}) [1] (PDB: 1NYJ). The t-105 rotamer of Trp_{41} points the indole ring towards the pore lumen, which was interpreted as supporting the first, Trp gating, model. However, MD simulations scanning the full conformational space of the two residues subject to this distance restraint proposed an alternative possibility, (t60, t90), that is energetically more stable than the (t-160, t-105) rotamers [15].

In this work, we probe the sidechain conformation of Trp_{41} in the closed and open states by directly measuring the inter-helical sidechain-sidechain distances between 5-^{19}F-Trp_{41} using ^{19}F spin diffusion NMR. We find that Trp_{41} has an inter-helical nearest-neighbor F-F distance of 11 ± 1 Å at both neutral (closed) and low (open) pH, constraining the Trp rotamer unambiguously to t90. Amantadine binding does not change the interhelical distance at this site. We also measured the F-F distances at residues 27 and 38, which are mutated to

4-^{19}F-Phe. We find that both sites show distance heterogeneity, which we attribute to oligomeric mixtures in one case and sidechain conformational distribution in the other. The ^{19}F spin diffusion NMR method is thus sensitive to the conformational heterogeneity of this transmembrane proton channel.

Experimental Methods

NMR samples

5-^{19}F-Trp was purchased from BioCatalysts (Pasadena, CA), Fmoc-protected by SynPep and purified using silica-gel column chromatography. 1, 2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) was obtained from Avanti Polar Lipids (Alabaster, AL). The ^{19}F-labeled wild type and mutant M2TMP samples were custom-synthesized by PrimmBiotech (Cambridge, MA) using standard Fmoc chemistry. The amino acid sequence of the Udorn strain of the influenza A virus is used in most experiments: NH_2-Ser_{22}-Ser_{23}-Asp_{24}-Pro_{25}-Leu_{26}-Val_{27}-Val_{28}-Ala_{29}-Ala_{30}-Ser_{31}-Ile_{32}-Ile_{33}-Gly_{34}-Ile_{35}-Leu_{36}-His_{37}-Leu_{38}-Ile_{39}-Leu_{40}-Trp_{41}-Ile_{42}-Leu_{43}-Asp_{44}-Arg_{45}-Leu_{46}-COOH. The purity of all peptide was checked by HPLC and mass spectrometry to be greater than 95%. The purified peptide was washed in 5 mM HCl solution to remove residual trifluoroacetate (TFA) ions from the synthesis, and checked by solution ^{19}F NMR. The ^{19}F spectra of the membrane-bound peptides contain no detectable TFA signal, indicating complete removal of TFA.

Large unilamellar DMPC vesicles were prepared by dissolving DMPC lipids either in a phosphate buffer (10 mM Na_2HPO_4/NaH_2PO_4, pH 7.5, 30˚C) or in a 10 mM citrate buffer (pH 4.5, 30˚C). The DMPC solution was freeze-thawed 8 times. Purified wild-type and mutant M2TMP peptides were dissolved in the DMPC solution at a peptide-lipid molar ratio (P/L) of 1: 15. This P/L value is sufficiently high to promote complete tetramerization of the M2TMP, as shown by analytical ultracentrifugation data [16]. The peptide-lipid mixture was vortexed for 30 minutes and incubated at 30˚C for 2 days. The solution was ultracentrifuged at 150,000 g for 3 hours at 28˚C. This yielded 90% reconstitution of the peptide, as measured by the BCA assay [17]. The pH of the membrane samples was confirmed through the supernatant to be 7.5 and 4.5 for the phosphate buffer sample and the citrate buffer sample,

respectively. The wet membrane pellet was transferred to a 4 mm rotor with borosilicate glass spacers and incubated at 30˚C for 2 days before the NMR experiments. To calculate the peptide molar concentrations of the NMR samples for comparison with the AUC data acquired on DPC micelles, the incubation and ultracentrifugation solution volume of 4 mL was used, giving a M2TMP concentration of ~0.6 mM.

To assess whether the oligomeric structure and sidechain conformation of the peptide are affected by the membrane reconstitution protocol of the peptide, we also used two other methods to mix the peptide with the DMPC lipids and measured the ^{19}F CODEX spectra of the corresponding samples to compare with the aqueous-mixed samples. One method involved cosolubilizing M2TMP with lipids in chloroform to obtain a well-mixed and clear solution. The mixture was dried under a stream of N_2 gas, lyophilized, then rehydrated to 35% water. A Trp_{41}-M2TMP sample was prepared in this way. The third method codissolves M2TMP in the detergent octyl-β-glucoside (OG) in aqueous buffer (10 mM Na_2HPO_4/NaH_2PO_4, 1 mM EDTA) with the desired pH [18,19]. The clear solution was then mixed with a DMPC vesicle solution to reach an OG concentration of 10% and a P/L of 1:15. The detergent was then removed by dialysis against the 10 mM Na_2HPO_4/NaH_2PO_4 buffer at 4 ºC for 3 days. The absence of detergent in the remaining peptide-lipid mixture was confirmed by ^{1}H solution NMR on the supernatant from ultracentrifugation. The dialyzed peptide-lipid solution was centrifuged to give the pellet for solid-state NMR experiments. A V27F-M2TMP sample was prepared in this way to compare with the aqueous mixed sample. An amantadine-bound Trp_{41}-M2TMP sample was also prepared by the detergent dialysis method, where 10 mM amantadine was added to the DMPC vesicle solution and the dialysis buffer solution.

Solid-state NMR experiments

The ^{19}F CODEX experiments were carried out on a Bruker DSX-400 spectrometer (Karlsruhe, Germany) operating at a resonance frequency of 400.49 MHz for ^{1}H and 376.8 MHz for ^{19}F, using an H/F/X probe equipped with a 4 mm MAS spinner module. The probe tunes ^{1}H and ^{19}F frequencies on a single channel. Experiments on M2TMP were conducted at

8 kHz MAS and 233 K using air cooled by a Kinetics Thermal Systems XR Air-Jet cooler (Stone Ridge, NY). Typical radio-frequency (rf) field strengths were 50 kHz for ^{19}F and ^{1}H. Recycle delay was 1.5 s. ^{1}H-^{19}F cross-polarization (CP) contact times were 200 μs. ^{19}F chemical shifts were externally referenced to the Teflon ^{19}F signal at -122 ppm.

The ^{19}F CODEX experiment uses two rotor-synchronized π-pulse trains to recouple the chemical shift anisotropy (CSA) [4]. The mixing time (τ_m) between the two trains allows spin diffusion to occur, which changes the CSA frequency and prevents complete refocusing of the stimulated echo. To correct for ^{19}F spin-lattice relaxation (T_1) effects during the mixing time τ_m, a z-filter (τ_z) is added at the end of the second π-pulse train. Two experiments were conducted: a dephasing experiment (S) with the desired τ_m and a short τ_z of 10 μs, and a reference experiment (S_0) with interchanged τ_m and τ_z values. The normalized echo intensity, S/S_0, decays to $1/n$ at long mixing times for an *n*–spin cluster. All CODEX experiments were conducted with two rotor periods (τ_r) of CSA recoupling. This resulted in $2\pi\delta N\tau_r$ values from 9π to 12π, where δ is the chemical shift anisotropy of the ^{19}F label. These values were sufficiently large to detect small orientational differences between different helices in the tetramer [5].

CODEX data analysis

The CODEX magnetization exchange curves were calculated in MATLAB using an exchange-matrix formalism [6]. Four-dimensional exchange rate matrices were constructed for the tetrameric bundle, where the rate constants are proportional to an overlap integral, F(0), and to the square of the dipolar couplings, ω_{FF}, which depend on the F-F distances. Thus all distance contacts in the four-spin cluster, including the nearest-neighbor distances and the diagonal distance, are included in the matrix. Our recent model compound studies yielded an F(0) of 37 μs for aromatic ^{19}F sites under identical MAS conditions as used here [6]. Thus, we fix this value and vary ω_{FF} to find the best-fit spin diffusion curve. The best-fit curve is evaluated by minimizing the root-mean-square-derivation (RMSD) between the calculated

intensity I_{sim} and the experimental intensity I_{exp}. The inter-helical nearest-neighbor F-F distances determining the coupling strengths were incremented at 0.1 Å steps.

Structure modeling

M2TMP structure was modeled in Insight II (Accelrys, San Diego). The structural model 1NYJ [1] of Cross and coworkers was used as the starting point and was modified to satisfy the ^{19}F CODEX data. We specify the sidechain conformations of His_{37} and Trp_{41} using the notation of the penultimate rotamer library [20]. The χ_1 angle is specified as t, p, or m, corresponding to 180˚, +60˚, and –60˚, respectively. These letters are followed by the approximate numerical value of the χ_2 angle. Thus, for example, the His_{37} and Trp_{41} conformations in the 1NYJ model are (t-160, t-105) [1,15].

Results

Inter-helical F-F distance at Trp_{41} at neutral and acidic pH

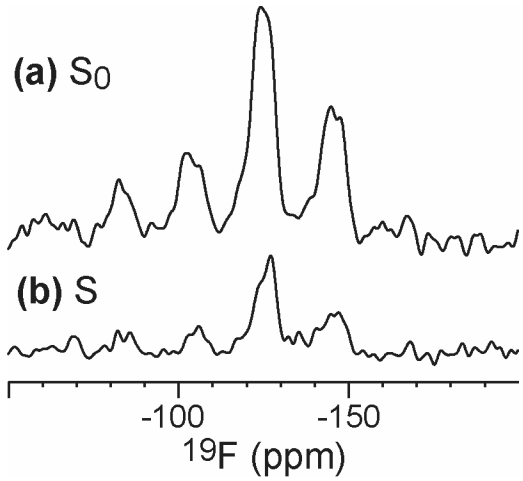


Figure 1. ^{19}F CODEX spectra of [5-^{19}F-Trp_{41}] M2TMP in DMPC bilayers at pH 7.5. The mixing time is 3.0 s. Data were collected at 233 K under 8 kHz MAS. (a) Reference spectrum S_0 to correct for T_1 relaxation effects. (b) Dephasing spectrum S. The normalized intensity S/S_0 is 0.36 ± 0.04.

Trp_{41} is recognized to be an important residue in M2TMP function, and has been proposed to act as the gate of the channel [9,13]. Thus, we used 5-^{19}F-Trp_{41} to measure the inter-helical sidechain-sidechain F-F distances in this part of the peptide assembly. Fig. 1 shows a representative pair of CODEX spectra of 5-^{19}F-Trp_{41} M2TMP in DMPC bilayers at pH 7.5, acquired with a mixing time of 3.0 s under 8 kHz MAS and at 233 K. Significant dephasing, with $S/S_0 = 0.36$, of the S spectrum is observed compared to the reference spectrum S_0, indicating the presence of multiple peptides with orientationally different ^{19}F chemical shift tensors in close proximity. The M2TMP helical bundle is known to be pseudo-symmetric [21], but the symmetry is rotational rather than translational, thus the ^{19}F chemical shift tensor orientations differ significantly between the helices in the assembly, allowing the detection of this CODEX effect.

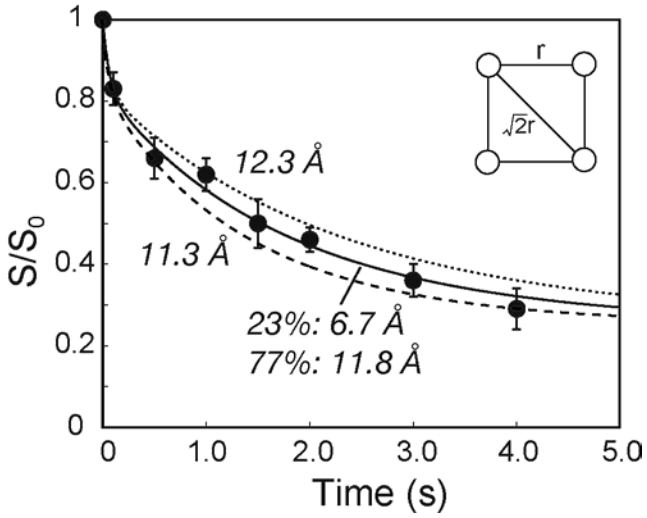


Figure 2. Normalized ^{19}F CODEX intensity of 5-^{19}F-Trp_{41} M2TMP in DMPC bilayers at pH 7.5. Error bars were propagated from the spectral sensitivity. In the CODEX simulation, an overlap integral F(0) of 37 μs was used. Bimodal fitting using a 23% component of a 6.7 Å distance and 77% of an 11.8 Å distance best fits the double exponential decay of the experimental intensities.

The complete mixing-time dependent CODEX curve of 5-^{19}F-Trp_{41} is shown in Fig. 2. The final normalized intensity is 0.29 ± 0.05, observed at a mixing time of 4.0 s,

confirming the tetrameric state of the peptide in DMPC bilayers. Longer mixing times were not measured because spectral intensities become prohibitively low due to ^{19}F spin-lattice relaxation (T_1 = 3.2 s). A separate sample prepared using organic-phase mixing of the lipid and peptide gave the same CODEX decay curve as the aqueous-mixed sample (Supporting information Fig. S1), indicating that the aqueous-mixed sample gives structurally identical tetramers as the organic samples. The only difference is that aqueous mixing produces more immobilized peptides than organic mixing [22], thus facilitating the CODEX experiment, which requires that reorientation motion be frozen during the mixing time.

The equilibrium value of 0.29 for Trp_{41} gives an upper bound to the fraction of possible monomer in the sample: 10% monomer would give an equilibrium value of 0.33 (=10% x 1 + 90% x 0.25), which is distinguishable from the experimental data. Thus any monomer component, if present, is no more than 10%, and the Trp_{41}-M2TMP sample is predominantly tetrameric in DMPC bilayers.

To fit the Trp_{41} CODEX data, we use an exchange-matrix formalism, where the magnetization decay with time is dictated by an exponential rate matrix in which the rate constants depend on the distance-dependent dipolar couplings and a spectral overlap integral F(0). Using an overlap integral value of 37 μs [6], which was calibrated based on experiments on model compounds under the same MAS conditions, we find a nearest-neighbor distance of 11.2 ± 0.5 Å using a single symmetric 4-spin model (Supporting information Fig. S2a). The presence of the diagonal distance between non-adjacent helices is automatically taken into account in the 4 x 4 exchange matrix. However, the single-tetramer fit does not capture well a minor component of fast initial decay in the experimental data. Thus we simulated the data with a double-tetramer model where the two tetramers have different side lengths. The fraction of each tetramer was obtained from the height of the turning point between the fast and slow decays. Fig. 2 shows that a 23% component with a short nearest-neighbor F-F distance of 6.7 Å and a 77% component with a longer distance of 11.8 ± 0.5 Å best-fit the experimental data. The RMSD analysis is shown in Supporting information Fig. S2b. Since the fraction of the short-distance component is low, below we analyze only the major conformer.

The intermolecular sidechain-sidechain ^{19}F-^{19}F distances between 5-^{19}F-Trp_{41} residues depend on the orientation of the helices, the pore diameter, and the sidechain conformation of Trp_{41}. The tilt angle of M2TMP in DMPC bilayers have been determined to be 35°±3° by ^{15}N NMR [21]. The rotation angle is also known with high precision. Thus, the peptide orientation is fixed. The pore diameter, indicated by the diagonal Cα-Cα backbone-backbone distance at the central residue Gly_{34}, has not been experimentally determined. However, statistical analysis of four-helix-bundle proteins indicates that tetrameric pores have remarkably consistent diameters of 10.0 -10.5 Å [23]. For M2TMP, various models derived from experimental data [12] and converged MD simulations [24,25] also put the pore diameter to 10.0 - 10.5 Å. This small variability likely reflects the requirement of tetramer stability. With this consideration, we used the NMR-derived 1NYJ structure as the starting point of our modeling since it has both the correct orientation and a suitable pore diameter of 10.2 Å for the Gly_{34} Cα-Cα diagonal distance.

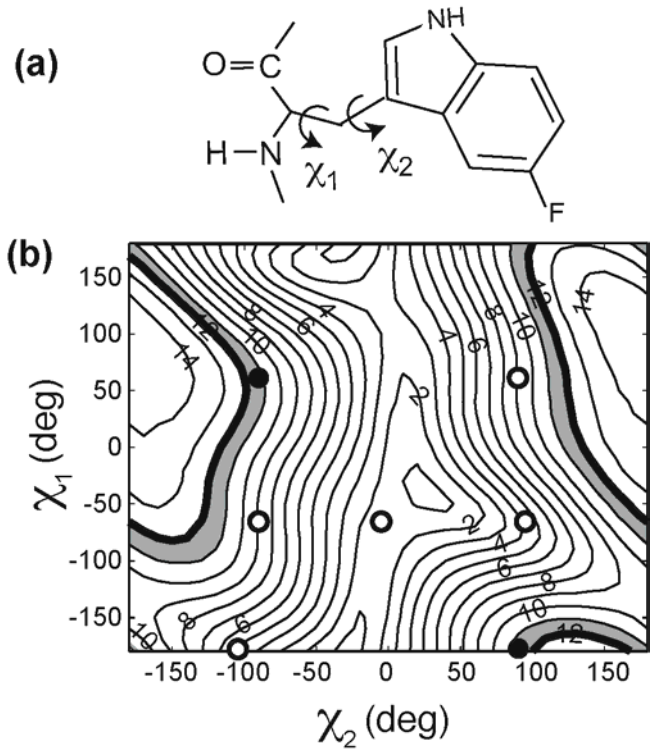


Figure 3. Nearest-neighbor intermolecular ^{19}F-^{19}F distance between 5-^{19}F-Trp_{41} as a function of torsion angles χ_1 and χ_2. (a) Definition of the χ_1 and χ_2 torsion angles. (b) Contour plot of

the nearest-neighbor F-F distance as a function of (χ_1, χ_2). Circles: rotameric states of Trp residue in α-helical proteins [20]. Open circles: rotamers that are ruled out based on the measured F-F distance. Filled circles: rotameric states allowed by the measured F-F distance. Shaded region: the experimentally measured F-F distance including experimental uncertainty.

These considerations leave the Trp_{41} sidechain conformation as the main parameter influencing the inter-helical sidechain-sidechain ^{19}F-^{19}F distance. Indeed, the χ_1 and χ_2 angles, defined in Fig. 3a, change the F-F distance dramatically from 2 Å to 14 Å (Fig. 3b). The (χ_1, χ_2) combinations that give a nearest-neighbor F-F distance of 11-12 Å are highlighted as gray areas in Fig. 3b. The seven rotamers of Trp populated in α-helical proteins are superimposed as circles in this distance plot [20,26]. As can be seen, only two out of seven rotamers, t90 and p-90, agree with the experimental data (solid circles), thus constraining the conformation of Trp_{41} in the closed state to one of these two possibilities. Even when the variability of pore diameter is taken into account, the F-F distance cannot change by more than ± 0.5 Å, which is already included in the experimental uncertainty.

To determine whether low pH corresponding to the open state of the channel changes the sidechain conformation of Trp_{41}, we measured the ^{19}F CODEX curve of 5-^{19}F-Trp_{41} labeled M2TMP in DMPC bilayers at pH 4.5. Fig. 4a shows that the low pH data (solid circles) overlaps with the pH 7.5 data (open squares) within experimental uncertainty. The best-fit distance using a single tetramer model is 10.8 ± 1.0 Å (solid line) based on RMSD analysis (Supporting information Fig. S3). Although the best-fit distance is 0.4 Å shorter than the single-tetramer fit for the neutral pH data, the distance error bar and the CODEX intensity error bars both overlap with those of the pH 7.5 data, therefore within experimental uncertainty there is no detectable difference in the Trp_{41} F-F distances between the closed and open states.

To assess the effect of the antiviral drug amantadine on the Trp_{41} sidechain conformation, we also measured the ^{19}F CODEX intensities of amantadine-bound M2TMP in DMPC bilayers at pH 7.5. Amantadine binding is known to prevent proton conductance and

channel opening [9,27]. The CODEX curve of the amantadine-bound Trp_{41} sample is shown in Fig. 4b. The data is best fit with a distance of 11.5 Å using a single-tetramer model. This is again similar to the 11.2 Å distance of the amantadine-free peptide within experimental uncertainty. The lack of distance change is consistent with the fact that the amantadine binding site is known to be on the N-terminus of the channel [28], on the other side of Gly_{34}, which is the narrowest point of the tetramer. Thus amantadine binding is not expected to affect the local sidechain conformation of Trp_{41} but at most only indirectly affects the interhelical distances through small changes in the helix tilt angle. Moreover, since the amantadine-bound membrane sample is prepared with the detergent dialysis method, the similar distance with the aqueous-mixed samples further indicates the independence of the tetramer structure on the sample preparation protocol.

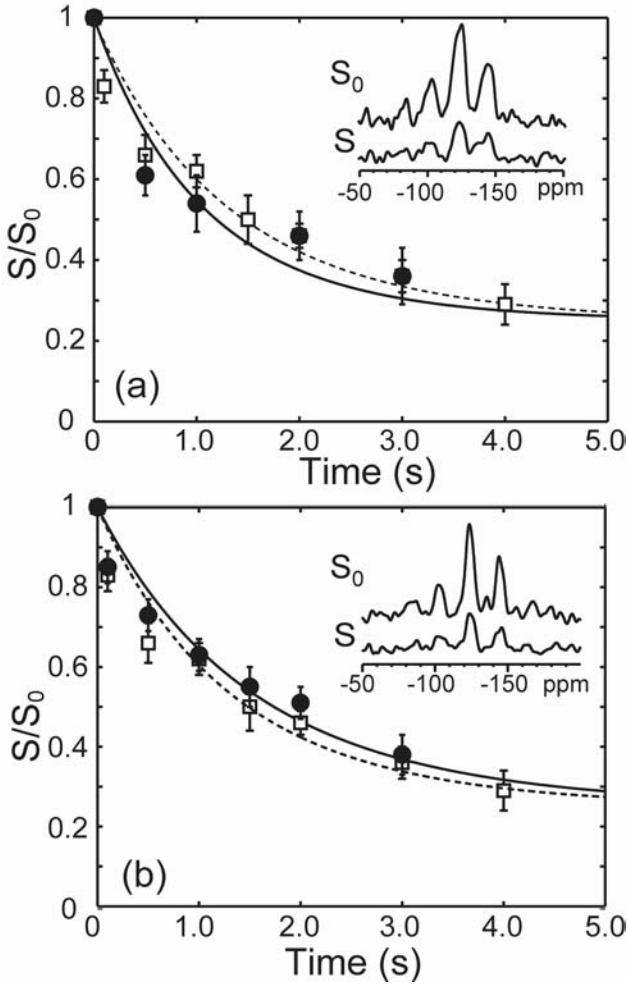


Figure 4. ^{19}F CODEX data of 5-^{19}F-Trp$_{41}$ M2TMP in DMPC bilayers at pH 4.5 (solid circles) compared with the pH 7.5 data (open squares). Inset shows the S_0 and S spectra at a mixing time of 3.0 s. The best fit of the pH 4.5 data using a single tetramer model gives a nearest-neighbor distance of 10.8 Å (solid line), which is the same as the pH 7.5 result (dashed line) within experimental uncertainty.

Inter-helical F-F distances at V27F and L38F at neutral pH

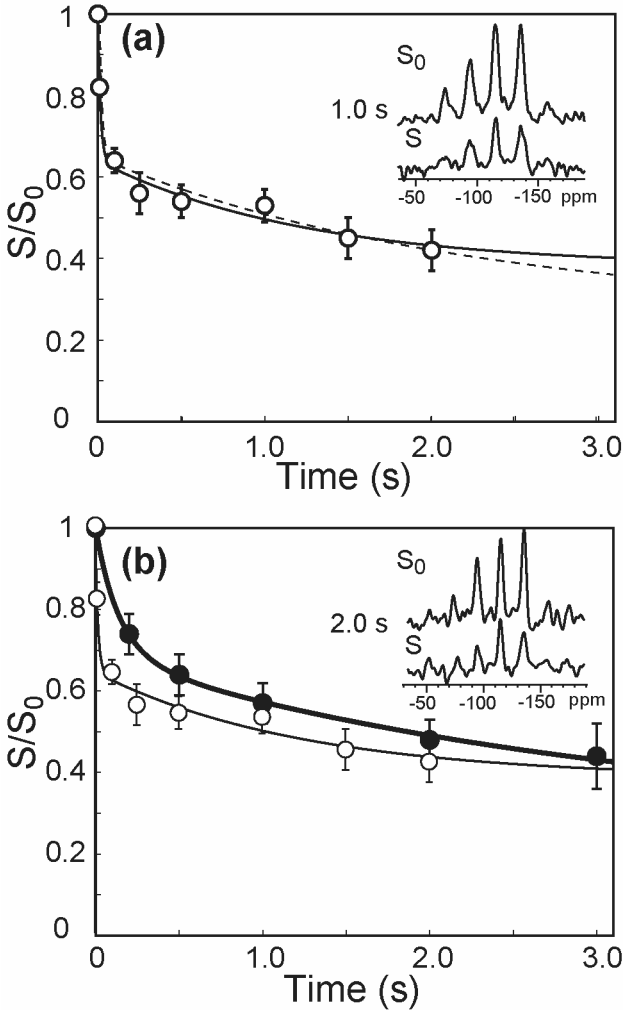


Figure 5. ^{19}F CODEX data of V27F- and L38F-M2TMP in DMPC bilayers at pH 7.5. (a) 4-^{19}F-Phe labeled V27F-M2TMP. The best fit using a dimer-tetramer mixture (solid line) gives a distance of 10.5 Å and 5.3 Å (nearest-neighbor distance) for the dimer and tetramer at a

weight ratio 52 : 48. Best fit using a tetramer-tetramer mixture (dashed line) gives distances of 12.3 Å and 5.3 Å at a ratio of 52 : 48. (b) 4-^{19}F-Phe labeled L38F-M2TMP (filled circles). The best fit gives two tetramers with side lengths of 13.1 Å and 7.7 Å at a ratio of 58 : 42. The V27F data is reproduced in (b) for comparison. Data were obtained at 233 K under 8 kHz MAS. Representative of S_0 and S spectra are shown in the insets.

To investigate the intermolecular packing of the M2TMP helical bundle at other residues, we mutated V27 and L38 to 4-^{19}F-Phe and measured their CODEX curves. These positions were chosen based on previous AUC data showing that the helical bundle stability is not significantly altered by the mutation to Phe. V27 is located at the *a* position of the heptad repeat and faces the pore lumen, thus is expected to have a short inter-helical distance. In comparison, L38 lies at the *e* position of the heptad repeat and is expected to point to the lipid-peptide interface. The L38F mutant is also found in nature in the Weybridge strain of influenza A virus: this proton channel has amantadine sensitivity [27] similar to the Udorn strain studied in the rest of this paper. Thus, the L38F mutation is expected to be particularly non-perturbing to the tetramer structure.

Fig. 5 shows the CODEX data of V27F-M2TMP and L38F-M2TMP in DMPC bilayers at pH 7.5. The V27F curve decays quickly in the first 500 ms, then appears to plateau to ~0.40 by 2.0 s. The same decay trend is reproduced in a detergent-dialysis sample (Supporting information Fig. S1). In comparison, the L38F intensity decays more slowly and at the longest mixing time used (3.0 s) the intensity continues to decrease. The different decay rates between the initial and final parts of the V27F data suggest a distribution of distances. We first used a model of two tetramers with different distances to fit the V27F data. Fig. 5a shows that a mixture of 52% of a tetramer with a side length of 12.3 Å and 48% of a tetramer with a side length of 5.3 Å fits the data well (Fig. 5a, dashed line). Since the pore diameter is unlikely to change by more than 1 Å, the very different distances would mean that Phe_{27} adopts different χ_1 angles, with –100˚ and –160˚ for the long and short distances, respectively (Supporting information Fig. S6a).

However, the relatively high final value of ~0.4 and the apparent plateau at long mixing times suggests an alternative interpretation where the V27F mutant forms a mixture of dimers and tetramers in DMPC bilayers. Thus we simulated the CODEX data using varying fractions of dimers and tetramers. A model where 52% of the peptide is in a dimer state with an intermolecular distance of 10.5 Å and 48% of the peptide is in a tetramer state with a nearest-neighbor inter-helical distance of 5.3 Å captures the experimental data best (Fig. 5a, solid line). The RMSD plot for the simulation is shown in Supporting information Fig. S4a. Increasing the dimer fraction to as much as 70% still gave acceptable fits while using less than 50% dimer fraction disagrees with the data (Supporting information Fig. S5). This suggests that V27F-M2TMP may be at least half populated as dimers while the rest are well-defined tight tetramers. The tetramer distance of 5.3 Å would be satisfied by the t80 conformer, which is the most populated rotamer of Phe in α-helices [20,26] (Supporting information Fig. S6a).

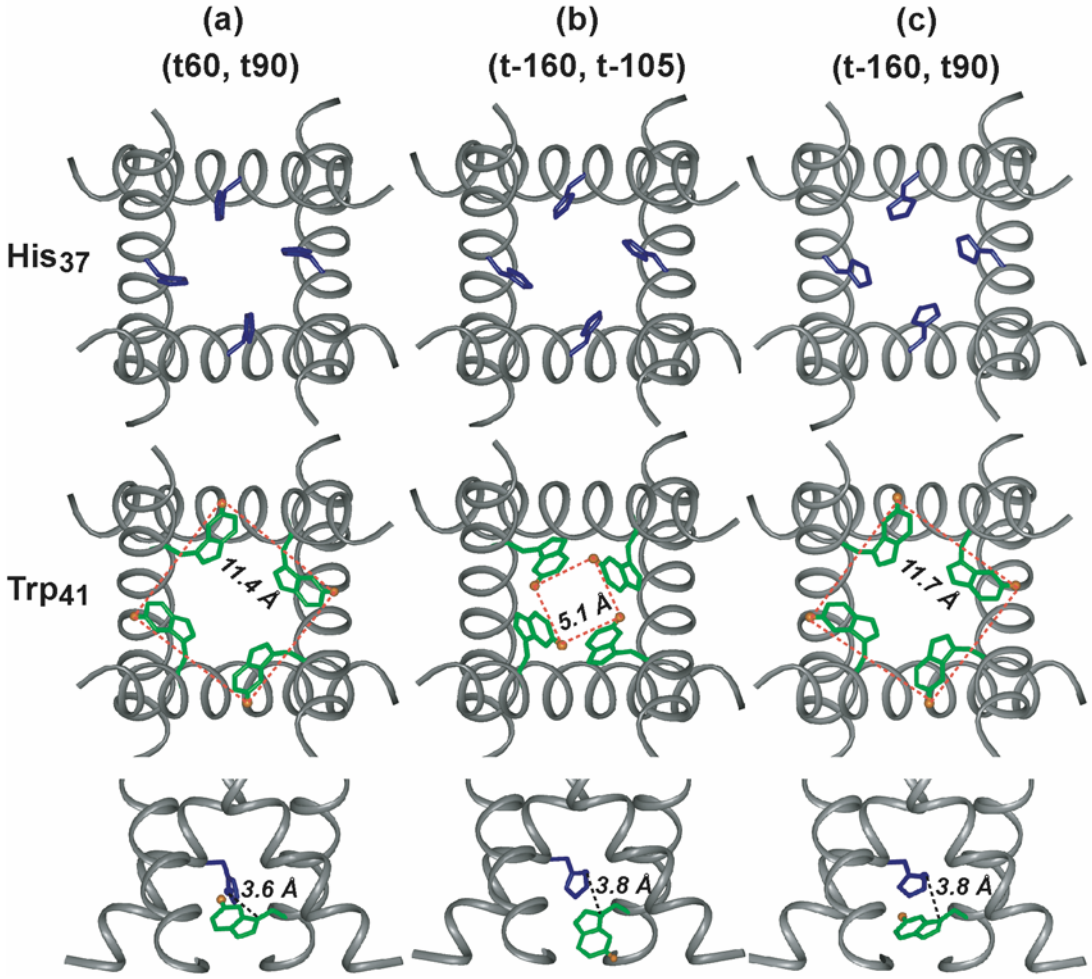


Figure 6. ^{19}F-NMR restrained conformation of Trp_{41} and His_{37} in M2TMP in DMPC bilayers. Top row: His_{37} (blue) in the top view of the channel from the C-terminus to the N-terminus. Middle row: Trp_{41} (green) and the F-F distance (orange). Bottom row: side view of the channel showing both Trp_{41} and His_{37} and the Cγ-Nδ distance between them. The (His_{37}, Trp_{41}) sidechain conformation is: (a) (t60, t90), (b) (t-160, t-105), and (c) (t-160, t90). Structure (a) disagrees with the C-N distance data [1] due to unfavorable orientation of the C-N vector, whereas structure (b) disagrees with the F-F distance data. Only structure (c) satisfies both constraints.

The L38F CODEX curve shows a clearer decay trend through the 3.0 s mixing time, suggesting that the data can be fit straightforwardly with tetramers only. But similar to the

V27F sample, a single distance does not fit the data well, but two distances of 13.1 Å (58%) and 7.7 Å (42%) for the tetramer side lengths give excellent fit to the data (Fig. 5b, with the RMSD plot in Supporting information Fig. S4b). Weight fractions differing by more than 15% from the best-fit fractions can be ruled out on physical grounds (Supporting information Fig. S5c). The longer distance of 13.1 Å translates to a Phe_{38} rotamer of t80 (58%) [20,26], whereas the shorter distance indicates a χ_1 angle of about -130˚ (42%) (Supporting information Fig. S6). The latter rotamer, although non-canonical, is still found in Phe residues in proteins [26].

Discussion

Trp_{41} sidechain conformation

Inter-helical F-F distances between 5-^{19}F-Trp_{41} labeled M2TMP have been measured in three states: the neutral pH state, the low pH state, and the amantadine-bound neutral pH state. All three samples exhibit nearest-neighbor distances between 10.8 Å and 11.8 Å, with uncertainties of ± 0.5 Å to ± 1.0 Å. We first consider the neutral pH amantadine-free state. The 11.8±1.0 Å distance obtained for the major conformer is a strong constraint of the Trp sidechain conformation. The existing model (1NYJ) put the Trp_{41} rotamer as t-105, based on a measured His_{37} Nδ – Trp_{41} Cγ distance of < 3.9 Å [1]. However, this would give rise to a F-F distance of 5.1 Å (Fig. 3b and Fig. 6b), in clear disagreement with the current ^{19}F data. Instead, the measured F-F distance restrains the Trp_{41} conformation to either t90 ($\chi_1 \approx 180^\circ$, $\chi_2 \approx 90^\circ$) or p-90 ($\chi_1 \approx +60^\circ$, $\chi_2 \approx -90^\circ$) (solid circles in Fig. 3b). The χ_1=60˚ (p) rotamer causes steric conflicts between the indole ring and the Cα of residue Leu_{38} and thus is rarely populated in α-helices (2%) [20,26]. In contrast, the t90 rotamer is one of the most populated rotamers of Trp in α-helices [20,26]. Thus, we propose the t90 rotamer for Trp_{41}.

The t90 rotamer of Trp_{41} for both the closed and open states is in excellent agreement with a number of observations from the UV resonance Raman spectra of Takeuchi and coworkers [14]. First, the wavenumber of the W3 peak in the Raman spectra predicted a χ_2 angle of around 100˚, in agreement with the current result. Second, the W7 peak intensity,

which reflects the environmental hydrophobicity of the indole ring, and the W3 wavenumber, a marker of the absolute value of χ_2, both showed little change between the open and closed states, suggesting that Trp_{41} sidechain conformation is similar between the open and closed states. This is consistent with the observed lack of change in the F-F distance between high and low pH (Fig. 4). Therefore, even though the importance of Trp_{41} in proton conductance is unambiguous, the detail of Trp's role in gating is more subtle and does not appear to involve the large conformational change hypothesized by Tang et al [13].

Implications of the Trp_{41} conformation to His_{37} conformation

Fixing the Trp_{41} conformation to t90 and varying the His_{37} conformation, we find two His_{37} rotamers, t60 and t-160, that give a His_{37} Nδ – Trp_{41} Cγ distance within the experimental range of < 3.9 Å measured by Cross and coworkers. However, the t60 rotamer of His_{37} in combination with a Trp_{41} t90 rotamer is unlikely because it gives an Nδ-Cγ bond orientation close to the magic angle relative to the helical bundle axis (Fig. 6a), thus the fast uniaxial rotation of the tetrameric helical bundle [22] would significantly average the Nδ-Cγ dipolar coupling to much less than the experimental value of 63±12 Hz [1]. In contrast, the t-160 conformation of His_{37} in combination with the Trp_{41} t90 rotamer gives an Nδ-Cγ vector that is roughly parallel to the bundle axis (Fig. 6c), thus is consistent with the measured motionally averaged C-N dipolar coupling. Thus, we propose the (t–160, 90) conformation for (His_{37}, Trp_{41}) (Fig. 6c).

Wu and Voth carried out a MD simulation that examined the Nδ-Cγ distances in four possible (His_{37}, Trp_{41}) adducts and proposed the (t60, t90) rotamer pair [15]. The t90 rotamer for Trp_{41} agrees with the current ^{19}F distance data, however the t60 rotamer for His_{37} is suspect. The MD simulations used a helical tilt angle of 30˚, which is significantly lower than the experimental value, and the His37 χ_2 angle that reproduced the Nδ-Cγ distance constraint in the simulation is 110˚, which is 50º away from the ideal value of 60˚. When the ^{15}N-NMR extracted tilt angle of 38˚ is used and the His χ_1 and χ_2 angles are kept within 20º of the ideal values, we find the t60 rotamer of His_{37} to no longer give a suitable Nδ-Cγ dipolar coupling of 63 ± 12 Hz.

The (t-160, t90) rotamer pair we propose was previously overlooked by Cross and coworkers for steric reasons. However, within ±20˚ of the ideal torsion angle values, the two residues in this rotamer combination maintain a minimum separation of 3.3 Å between the imidazole ring and indole ring while still satisfying the F-F and C-N distance constraints (Fig. 6c). Thus, steric conflict is not a problem for this rotamer combination.

To prove beyond doubt the sidechain conformation of His_{37}, rigid-limit dipolar couplings in the absence of complicating motions need to be measured. The previous C-N distance measurement was conducted at 38˚C, well into the liquid-crystalline phase of the DMPC bilayer. At this temperature, the His_{37} Nδ position is influenced by (χ_1, χ_2) torsional motions in addition to the uniaxial whole-body rotation of the helical bundle. These multiple degrees of freedom average the dipolar coupling in a complicated fashion and cannot be adequately taken into account in the distance extraction. The mobility of the peptide in the liquid-crystalline phase does not, however, affect the helix orientation measurement on glass-plate samples, since these samples used backbone ^{15}N-labeled peptides, whose uniaxial rotation around the bilayer normal or the magnetic field is invisible in the spectra [21].

The implication of the Trp_{41} conformation to the gating mechanism

How do the Trp_{41} sidechain conformation and the proposed (His_{37}, Trp_{41}) rotamer pair explain the observed proton blockage in the closed state and proton conduction in the open state, given the fact the Trp_{41} 5-^{19}F inter-helical distance remains largely unchanged at ~11 Å between neutral and acidic pH? We propose two models. In the first model, the (t-160, t90) rotamer pair places the His_{37} imidazole rings close to the Trp_{41} indole rings (Fig. 6c). Thus when pH_{in} is low and pH_{out} is high, the C-terminal indole rings prevent the intracellular protons from protonating His. The His rings occlude the pore, either by formation of imidazole-imidazolium dimers [29] or sterically. When pH_{out} is low and pH_{in} is high, the extracellular protons from the N-terminus are able to protonate all four imidazole rings. This either results in electrostatic repulsion that opens the constriction at His_{37}, or causes the excess proton on the imidazolium ion to be relayed as His returns to its neutral state [9,29]. When Trp_{41} is mutated to Phe, the phenylene ring, in its most populated rotamer of t80, is

further away from the imidazoles (Fig. 7), thus allowing protons to protonate His_{37} from either direction, causing a leaky channel [13]. Thus, in this model, close proximity and interaction between Trp_{41} and His_{37} combined with the constriction at His_{37} gate the channel.

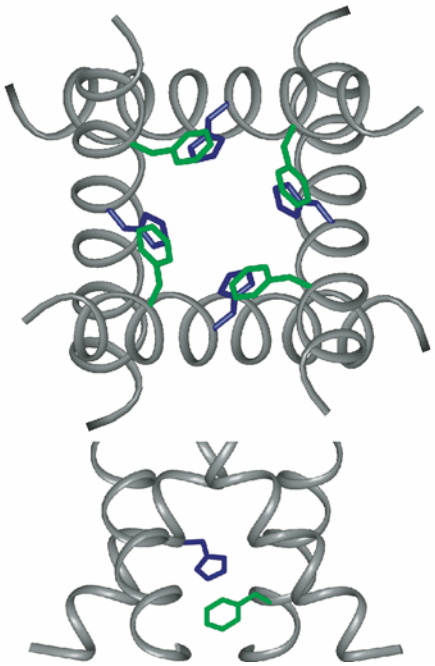


Figure 7. Proposed conformation of W41F mutant of M2TMP and the His_{37} contact with Phe_{41}. Only the neutral pH_{out} state is considered. Top row: top view of M2TMP from the C-terminus to the N-terminus. Bottom row: side view. His_{37} is in the t-160 state and Phe_{41} is in the most populated t80 state. The two rings are further away from each other than in Fig. 6c.

Examination of the pore constrictions at Trp_{41} and His_{37} raises a second possibility for channel gating. At the t90 Trp_{41}, the shortest diagonal distance between protons of the indole rings is 2.2 – 6.5 Å, taking into account the hydrogen atom van der Waals radius of 1.2 Å. This significant distance variation reflects χ_1 and χ_2 uncertainties of ±20° that still satisfy the measured F-F distance. Varying the backbone tilt angle within the ^{15}N NMR experimental uncertainty does not affect this Trp_{41} constriction. Thus, it is possible that under high pH_{out} and low pH_{in}, the Trp_{41} t90 conformation is such that the pore constriction is at its lower limit of 2.2 Å, which is sufficient to block protons, whereas under low pH_{out}, protonation of the His_{37} imidazole ring may change the Trp_{41} conformation slightly through cation-π

interactions [30,31] so that the Trp_{41} constriction opens up to 4–7 Å, allowing protons to pass. The lower-bound constriction may be achieved by Trp_{41} (χ_1, χ_2) ≈ (-157˚, 110˚) (Fig. 8a), whereas the larger constrictions may be achieved by (χ_1, χ_2) = (163˚, 80˚-90˚) or (χ_1, χ_2) = (-177˚, 70˚-90˚) (Fig. 8b). Further experiments would be required to ascertain if this conformational change model is correct.

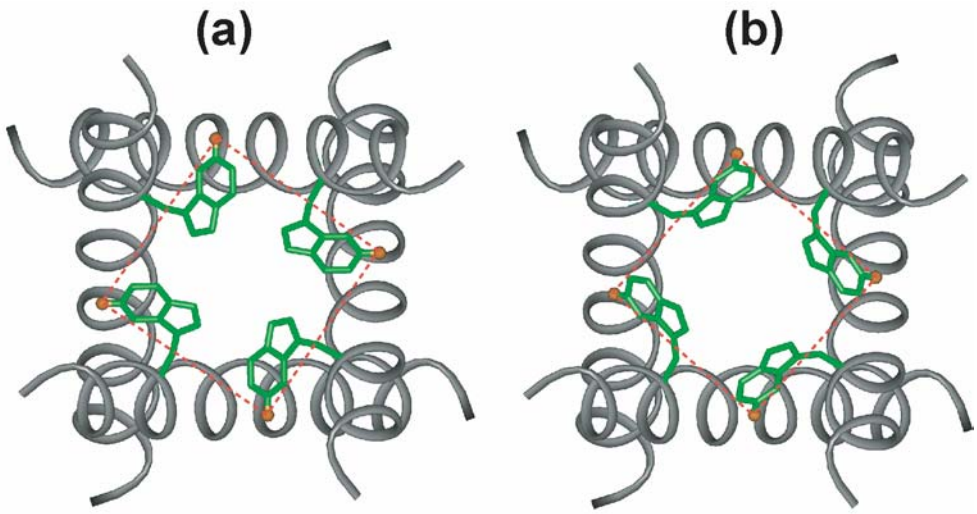


Figure 8. Two Trp_{41} t90 conformations that satisfy the ^{19}F CODEX distance constraint while changing the pore constriction. (a) χ_1=-157˚ and χ_2=110˚. The pore constriction, indicated by the shortest diagonal proton-proton van der Waals distance, is 2.2 Å, and the F-F nearest-neighbor inter-helical distance is 11.4 Å. (b) χ_1=-177˚ and χ_2=80˚. The pore constriction is 4.8 Å while the F-F nearest-neighbor interhelical distance is 10.7 Å.

Tetramer stability and conformational heterogeneity in V27F- and L38F-M2TMP

Unlike the Trp_{41} ^{19}F CODEX data, which conclusively shows a tetrameric state, the V27F and L38F CODEX intensities both decay to intermediate values of ~0.4 within the mixing times allowed by the ^{19}F T_1 relaxation time. The L38F data show a clearer decaying trend whereas the V27F intensities appeared to have plateaued by 2.0 s. Nevertheless, the V27F data can be fit either to a tetramer-tetramer mixture with different side lengths or a dimer-tetramer mixture. However, the tetramer-tetramer fit would require Phe_{27} to adopt χ_1 angles differing by 60˚. Since residue 27 lies at the *a* position of the heptad repeat and is

known to face the pore lumen [12], we believe this significant sidechain conformation heterogeneity is unlikely. Instead, the V27F mutant may destabilize the helical bundle to partially form dimers or a dimer of dimers that makes up a loose tetramer. In the latter case, the outer dimer may have a ^{19}F-^{19}F distance larger than ~15 Å that cannot be detected in the CODEX experiment. This hypothesis would be consistent with the AUC data [11] showing that the Gibbs free energy of tetramer formation of the V27F mutant is 0.7 kcal/mol higher or less stable than that of the wild-type M2TMP in DPC micelles [11]. The fact that a significant fraction of dimers (50-70%) is observed in lipid bilayers despite the small Gibbs free energy difference may be partly due to intrinsic environmental differences between DPC micelles and lipid bilayers.

In comparison, the CODEX data of the L38F mutant, which is naturally present in the Weybridge virus, is better fit to a tetramer-tetramer mixture. This oligomeric mixture and distance heterogeneity are reasonable because residue 38, unlike residues 27 and 41, faces the lipid molecules [32], whose thermal disorder can readily affect the sidechain conformation of Phe_{38}. Moreover, residue 38 is close to the center of the helix, thus the phenylene ring lies at a depth near the middle of the bilayer, where the lipid chain disorder is maximal. Thus, a distribution of the Phe χ_1 angle is reasonable. Modeling shows that the F-F distances of 13.1 Å (58%) and 7.7 Å (42%) are achieved using χ_1 angles of 180˚ and –130˚, respectively (Supporting information Fig. S6b). The former is the most populated rotamer of Phe in α-helices while the latter is less populated but still found in protein structures at non-negligible percentages [26].

Compared to L38F, the Trp_{41} sidechain resides largely in the pore lumen except for the end of the six-membered ring (Fig. 6c), which points close to the helical interface [12], thus it is less subject to lipid-induced thermal motion. Trp_{41} is also near the C-terminus of the peptide and thus should be embedded at the membrane-water interface, where the lipid is the most rigid. Thus, conformational heterogeneity should be much reduced at Trp_{41}, as observed by the much lower fraction of a second component (23%) in the best-fit simulation.

To conclusively determine the oligomeric states and fractions of mixtures, additional experiments such as the four-time CODEX experiment are desirable [5]. The main challenge

will be to increase the sensitivity of such experiments to make them applicable to membrane-bound peptides and proteins.

The tetramer stability of M2TMP depends not only on the amino acid sequence and site-specific mutations, but also on the membrane environment. DeGrado and coworkers have shown that increasing the lipid chain length, adding cholesterol and amantadine increase the tetramer stability [16]. The presently observed conformational heterogeneity for the V27F and L38F mutants thus may very well change in different membranes. However, this does not change the conclusion that the L38F mutant is relatively stable compared to the V27F mutant in the same membrane.

Conclusion

Inter-helical sidechain ^{19}F-^{19}F distances have been measured for 5-^{19}F-Trp$_{41}$, 4-^{19}F-V27F, and 4-^{19}F-L38F positions of M2TMP bound to DMPC bilayers under various conditions. At neutral and acidic pH, the peptide shows the same nearest-neighbor distance of ~ 11 Å, which is unchanged upon amantadine binding. This distance constrains the Trp$_{41}$ rotamer to t90, and no significant conformational change occurs between the closed and open states. Combined with a previously measured ^{13}C-^{15}N distance between His$_{37}$ and Trp$_{41}$, this suggests that the His$_{37}$ rotamer is t-160 at neutral pH. Gating of the proton channel may be explained either by a cooperative interaction between His$_{37}$ and Trp$_{41}$ that changes the protonation state of the His rings, thus closing or opening the constriction at His$_{37}$, or by a subtle conformational change of Trp$_{41}$ that changes the pore constriction without affecting the ^{19}F-^{19}F distance.

The inter-helical ^{19}F-^{19}F distances at V27F and L38F are heterogeneous. The V27F mutant data is attributed to a mixture of dimers and tetramers due to the known destabilization of the helical bundle by mutation at this pore-facing site, whereas the data of the naturally occurring L38F mutant is best explained by sidechain conformational heterogeneity of this lipid-facing residue.

Acknowledgement

This work is supported by the National Science Foundation grant MCB-0543473.

References

(1) Nishimura, K.; Kim, S.; Zhang, L.; Cross, T. A. *Biochemistry* **2002**, *41*, 13170.

(2) Hong, M. *Structure* **2006**, *14*, 1731.

(3) Schmidt-Rohr, K.; deAzevedo, E. R.; Bonagamba, T. J. Centerband-Only Detection of Exchange (CODEX): Efficient NMR Analysis of Slow Motions in Solids. In *Encyclopedia of NMR*; Grant, D. M., Harris, R. K., Eds.; John Wiley & Sons: Chichester, 2002.

(4) deAzevedo, E. R.; Bonagamba, T. J.; Hu, W.; Schmidt-Rohr, K. *J. Am. Chem. Soc.* **1999**, *121*, 8411.

(5) deAzevedo, E. R.; Bonagamba, T. J.; Hu, W.; Schmidt-Rohr, K. *J. Chem. Phys.* **2000**, *112*, 8988.

(6) Luo, W.; Hong, M. *J. Am. Chem. Soc.* **2006**, *128*, 7242.

(7) Buffy, J. J.; Waring, A. J.; Hong, M. *J. Am. Chem. Soc.* **2005**, *127*, 4477.

(8) Lamb, R. A.; Holsinger, K. J.; Pinto, L. H. The Influenza A virus M2 ion channel protein and its role in the influenza virus life cycle. In *Cellular Receptors of Animal Viruses*; Wemmer, E., Ed.; Cold Spring Harbor Lab Press: Plainview, NY, 1994; pp 303.

(9) Pinto, L. H.; Lamb, R. A. *J. Biol. Chem.* **2006**, *281*, 8997.

(10) Wang, C.; Lamb, R. A.; Pinto, L. H. *Biophys. J.* **1995**, *69*, 1363.

(11) Howard, K. P.; Lear, J. D.; DeGrado, W. F. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8568.

(12) Pinto, L. H.; Dieckmann, G. R.; Gandhi, C. S.; Papworth, C. G.; Braman, J.; Shaughnessy, M. A.; Lear, J. D.; Lamb, R. A.; DeGrado, W. F. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11301.

(13) Tang, Y.; Zaitseva, F.; Lamb, R. A.; Pinto, L. H. *J. Biol. Chem.* **2002**, *277*, 39880.

(14) Okada, A.; Miura, T.; Takeuchi, H. *Biochemistry* **2001**, *40*, 6053.

(15) Wu, Y.; Voth, G. A. *Biophys. J.* **2005**, *89*, 2402.

(16) Cristian, L.; Lear, J. D.; DeGrado, W. F. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 14772.

(17) Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. *Prot. Sci.* **1995**, *4*, 2411.

(18) Liu, W.; Crocker, E.; Siminovitch, D. J.; Smith, S. O. *Biophys. J.* **2003**, *84*, 1263.

(19) Smith, S. O.; Eilers, M.; Song, D.; Crocker, E.; Ying, W.; Groesbeek, M.; Metz, G.; Ziliox, M.; Aimoto, S. *Biophys. J.* **2002**, *82*, 2476.

(20) Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson, D. C. *Proteins: Structure, Function, and Genetics* **2000**, *40*, 389.

(21) Wang, J.; Kim, S.; Kovacs, F.; Cross, T. A. *Prot. Sci.* **2001**, *10*, 2241.

(22) Cady, S. D.; Goodman, C.; Tatko, C. D.; DeGrado, W. F.; Hong, M. *J. Am. Chem. Soc.* **2007**, *129*, 5719.

(23) Harris, N. L.; Presnell, S. R.; Cohen, F. E. *J. Mol. Biol.* **1994**, *236*, 1356.

(24) Sansom, M. S. P.; Kerr, I. D.; Smith, G. R.; Son, H. S. *Virology* **1997**, *233*, 163.

(25) Torres, J.; Kukol, A.; I.T., A. *Biophys. J.* **2001**, *81*, 2681.

(26) Janin, J.; Wodak, S. *J. Mol. Biol.* **1978**, *125*, 375.

(27) Wang, C.; Takeuchi, K.; Pinto, L. H.; Lamb, R. A. *J. Virology* **1993**, *67*, 5585.

(28) Duff, K. C.; Gilchrist, P. J.; Saxena, A. M.; Bradshaw, J. P. *Virology* **1994**, *202*, 287.

(29) Hu, J.; Fu, R.; Nishimura, K.; Zhang, L.; Zhou, H.-X.; Busath, D. D.; Vijayvergiya, V.; Cross, T. A. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 6865.

(30) Gallivan, J. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9459–9464.

(31) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303.

(32) Stouffer, A. L.; Nanda, V.; Lear, J. D.; DeGrado, W. F. *J. Mol. Biol.* **2005**, *347*, 169.

Chapter 7

Immobilization of Membrane Proteins by Eukaryote-Mimetic Lipid Membranes for Solid-State NMR Structure Determination

A paper published in Biochemistry

2009, vol. 48 pp. 6361-6368

Wenbin Luo, Sarah D. Cady and Mei Hong

Abstract

The dynamic structure of membrane proteins is intimately affected by the properties of the lipid bilayer. One property of membrane proteins is uniaxial rotational diffusion, which depends on the bilayer viscosity and thickness. This rotational diffusion is readily manifested in solid-state NMR spectra as temperature-dependent line narrowing or broadening. We show here that this whole-body uniaxial diffusion is suppressed in lipid bilayers mimicking the composition of eukaryotic cell membranes, which are rich in cholesterol and sphingomyelin. We demonstrate this membrane-induced immobilization on the transmembrane peptide of the influenza A M2 (AM2-TM) proton channel protein. At physiological temperature, AM2-TM undergoes uniaxial diffusion faster than $>1.2\times10^5\ s^{-1}$ in DLPC, DMPC and POPC bilayers, but the motion is slowed by over two orders of magnitude, to $<<10^3\ s^{-1}$, in a cholesterol-rich virus-envelope-mimetic membrane. The immobilization is manifested as rigid-limit ^{2}H, ^{13}C and ^{15}N spectra. Moreover, the protein conformation became more homogeneous in the viral membrane, as evidenced by the higher spectral resolution compared to frozen PC membranes. The immobilization and structural homogenization may be general for eukaryotic membrane proteins in eukaryote-mimetic membranes and should greatly facilitate structure determination by NMR.

Introduction

Lipid bilayers are now recognized to have significant effects on membrane protein structure and dynamics. The thickness, fluidity, and charge of lipid membranes can affect the orientation, dynamics, oligomeric state, and function of membrane proteins (1, 2). In particular, the fluid bilayer endows membrane proteins with abundant dynamics that include both internal segmental motions and whole-body rotational and translational diffusions (3-5) from picoseconds to milliseconds. These motions usually have functional importance, such as facilitating large conformational changes (6), channel formation (7), and membrane disruption (8, 9).

The rate of large-amplitude molecular motions has a large effect on NMR spectra. Motional rates much higher than the rigid-limit nuclear spin interaction of interest scale the interaction and cause spectral narrowing, while rates similar to the interaction strengths cause severe line broadening and intensity loss ("exchange broadening") (10). Saffman and Delbrück considered Brownian motions of proteins in lipid bilayers (11) and concluded that the rotational diffusion rates, D_R, of membrane proteins around the bilayer normal depend on the viscosity of the bilayer and the volume of the membrane protein. For a cylindrical protein traversing the bilayer, the rotational diffusion rate is directly proportional to temperature (T) and inversely proportional to the membrane viscosity (η), thickness (h), and the square of the radius (r) of the cylinder, $D_R = kT/4\pi\eta r^2 h$ (11). The equation predicts that in PC bilayers with η = 10 poise at 298 K, a transmembrane (TM) helical bundle with a radius of 12.5 Å has rotational diffusion rates of $1\times10^5\ s^{-1}$, which is larger than most NMR interactions. Indeed, motionally averaged ^{2}H quadrupolar couplings, ^{13}C-^{1}H and ^{15}N-^{1}H dipolar couplings, and ^{13}C chemical shift anisotropies were observed for a TM helical bundle (12).

While this uniaxial diffusion stems from general physical principles and has practical benefits such as enabling orientation determination (12), it can also complicate NMR structure determination of membrane proteins due to exchange broadening at ambient temperature. While one can overcome this exchange broadening by freezing the membrane samples, low-temperature NMR often yields lower-resolution spectra compared to the

spectra of rigid solids at ambient temperature and may change the protein structure from its physiological state. Thus there is a strong incentive to develop alternative methods to create well ordered and immobilized membrane proteins at physiological temperature.

The M2 proteins of influenza A and B viruses (AM2 and BM2) are integral membrane proteins that form pH-activated proton channels essential for virus replication (13, 14). As the smallest ion channel proteins that have complete ion selectivity and activation properties, the M2 proteins are excellent model systems for understanding the effects of lipid membranes on protein structure and dynamics. AM2 is the target of the antiviral drug amantadine (15, 16). Thus elucidating the influence of the lipid membrane and drug on M2 structure and dynamics also has public health importance. Recently, the structure of the AM2 transmembrane domain (AM2-TM) was solved in detergent matrices using X-ray crystallography (17) and solution NMR (18); however, the two structures differed in the drug binding site, helix orientation, and sidechain conformation. The AM2 structure has also been extensively studied in model PC bilayers using solid-state NMR (SSNMR). These spectra clearly indicate that AM2-TM has extensive conformational dynamics in the liquid-crystalline phase of DLPC, DMPC, and POPC bilayers. The main motion is uniaxial rotational diffusion of the helical bundle relative to the bilayer normal, as evidenced by ^{15}N NMR spectra (12, 19). As the temperature decreases, the motion slows, giving rise to exchange-broadened spectra at intermediate temperatures (~ 293 K) (20) and slow-limit high-intensity spectra at low temperatures (~243 K).

So far the model PC membranes used to characterize AM2 structure do not resemble the viral envelope lipid composition, which contains significant amounts of cholesterol (Chol), saturated-chain PC and phosphoethanolamine (PE), and sphingomyelin (SM) (21). This composition is characteristic of eukaryotic cell membranes, because the virus takes the lipids from its host cells. We show that in this virus- and eukaryote-mimetic lipid membrane, AM2-TM rotational diffusion is slowed down by over two orders of magnitude due to the high viscosity of the membrane. This immobilization yields high-resolution NMR spectra at ambient temperature and allows inter-atomic distances to be measured without freezing the sample. The conformation of various residues in this viral membrane was examined by 2D ^{13}C and ^{15}N spectra. We find that except for lipid-facing residues and known hot spots of the

protein, the average conformation of AM2-TM is not changed from that in model membranes, but the conformational homogeneity is enhanced. Thus, this eukaryote-mimetic membrane presents great spectroscopic and biological advantages for NMR structure determination.

Results

Calculated rotational diffusion rates of membrane proteins in viral membranes

The lipid composition of the influenza virus envelope resembles the host cell membrane it buds from and contains SM, PC and PE lipids (21). Together they account for 70-80% of the total lipid mass (22). The lipid chains are largely saturated with 16 or 18 carbons (23). Cholesterol (Chol) is abundant in the virus envelope. The lipid : cholesterol mass ratio is 2 – 3, depending on the host cell, the nature of the virus, and the presence of small molecules such as vitamin A (23). SM and Chol-rich lipid membranes have been extensively studied for their role in raft formation (24), and diffusion coefficients, membrane thickness, and headgroup area per lipid have been estimated (25). The virus-mimetic membrane is generally more viscous, thicker, and denser than one-component low-melting PC bilayers.

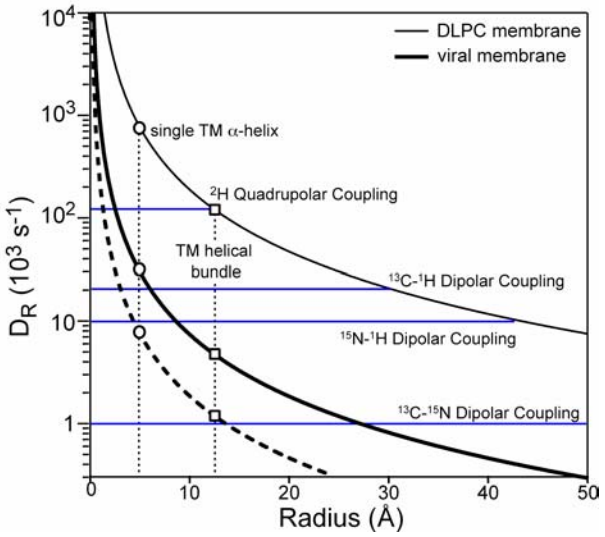


Fig. 1. Calculated rotational diffusion rates (D_R) of membrane proteins in viral membranes (thick line) and DLPC bilayers (thin line) at 298 K. The DLPC curve was calculated with η = 5 poise and *h* = 35 Å. The viral membrane curve was calculated with η = 100 poise and *h* = 45 Å. The dashed line was calculated with η = 400 poise and *h* = 45 Å. The radii of a single TM helix and a TM helical bundle are 5 Å and 12.5 Å, respectively. The sizes of several NMR spin interactions are shown as blue horizontal lines.

Figure 1 plots the calculated protein rotational diffusion rates D_R at 298 K in DLPC bilayers and viral membranes as a function of the protein radius. The main parameter affecting the rotational diffusion rate is the viscosity, η. Low-melting PC bilayers have an η of 1-10 poise (11, 26). We used η = 5 poise to calculate D_R in DLPC bilayers. For SM and Chol-rich viral membranes, the viscosities were estimated to be at least 20-fold higher than simple PC membranes based on pulse field gradient NMR (27) and molecular dynamics simulations (25). Thus we used η = 100 poise to calculate D_R in the viral membrane. Further, the bilayer thickness, d_{PP}, is ~35 Å for DLPC and ~45 Å for the viral membrane. The resulting D_R curves show the expected difference between the two membranes. A single TM helix (radius = 5 Å) has $D_R \approx 7\times10^5\ s^{-1}$ in DLPC bilayers but only $3\times10^4\ s^{-1}$ in the viral membrane. Thus, a TM helix is expected to have rigid-limit ^{2}H couplings in the viral membrane but motionally averaged ^{13}C and ^{15}N dipolar and chemical shift spectra. For a TM helical bundle (radius = 12.5 Å), D_R is about $1\times10^5\ s^{-1}$ in DLPC bilayers but only $5\times10^3\ s^{-1}$ in the viral membrane. Thus the helical bundle should exhibit rigid-limit spectra of all NMR interactions, with the possible exception of the one-bond ^{13}C-^{15}N dipolar coupling, which is 1 kHz.

Observed AM2-TM immobilization in viral membranes

Figure 2 shows the ^{2}H spectra of two Ala-CD_3 sites in AM2-TM reconstituted into viral and DLPC membranes. Ala CD_3 quadrupolar couplings reflect the dynamics of the Cα-Cβ bond (28). Motions additional to methyl three-site jumps are manifested as couplings less

than 40 kHz. For DLPC-bound samples, the ^{2}H splittings are 10.0 kHz for A30 and 15.7 kHz for A29 at 313 K (12), indicating fast rotational diffusion of the helix backbone. In contrast, in the viral membrane, both methyl groups show ^{2}H splittings of 37.6 kHz at 313 K. Thus, the M2 backbone is immobilized in the viral membrane on the 10^{-5} s timescale, or the inverse of 120 kHz. The remaining scaling factor of 0.94 reflects small-amplitude local motions.

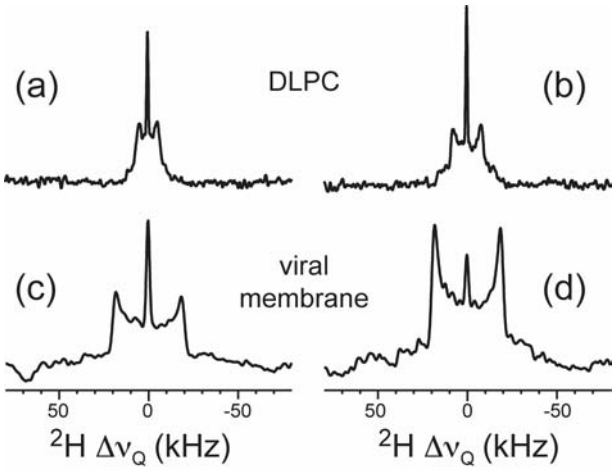


Fig. 2. ^{2}H spectra of Ala CD_3-labeled AM2-TM in different lipid membranes at pH 7.5 and 313 K. (a) DLPC bilayers, A30-CD_3. (b) DLPC bilayers, A29-CD_3. (c) Viral membrane with amantadine bound, A30-CD_3. (d) Viral membrane, A29-CD_3.

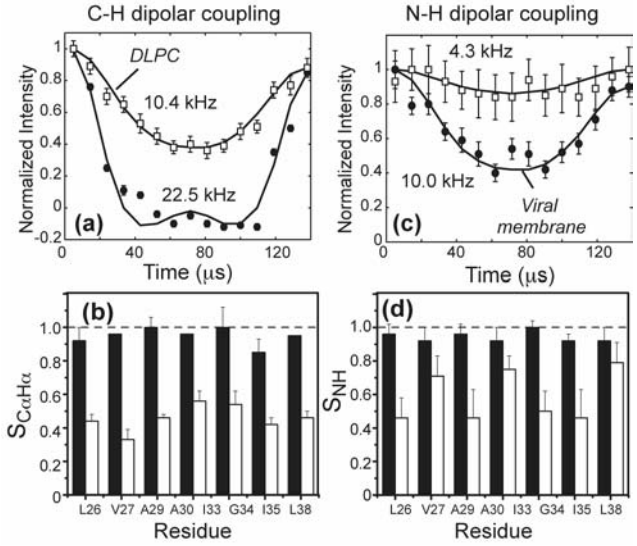


Fig. 3. C-H and N-H dipolar couplings of AM2-TM in different membranes at 313 K. Filled symbols: viral membrane data. Open symbols: DLPC data. (a) A29 Cα-Hα dipolar coupling. (b) Cα-Hα order parameters of AM2-TM in viral membranes (filled bars) and DLPC bilayers (open bars). (c) L26 N-H dipolar coupling. (d) N-H order parameters in viral membranes (filled bars) and DLPC bilayers (open bars).

To determine if AM2-TM diffuses at rates slower than $10^5\ s^{-1}$, we measured the C-H and N-H dipolar couplings of the peptide backbone. **Figure 3** shows selected Cα-Hα and N-H dipolar DIPSHIFT curves at 313 K for DLPC- and viral-membrane-bound AM2-TM. For all residues, the viral-membrane samples exhibit nearly rigid-limit couplings whereas the DLPC samples show 2-3 fold smaller couplings, indicating large-amplitude motions. The C-H and N-H order parameters range from 0.85 to 1.0 in the viral-membrane (solid bars), but only 0.30 – 0.70 in DLPC bilayers (open bars). Thus, AM2-TM is immobilized to below $10^4\ s^{-1}$ in the viral membrane.

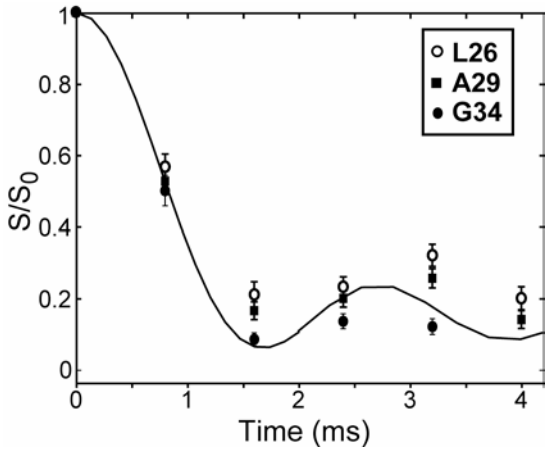


Fig. 4. ^{13}C-^{15}N dipolar couplings of the amantadine-bound LAGI-M2 in viral membranes at 303 K. The calculated REDOR curve is for a 1.0 kHz dipolar coupling.

To place a tighter upper limit to the motional rates of AM2-TM in the viral membrane, we measured an even smaller spin interaction, the one-bond ^{15}N-$^{13}C\alpha$ dipolar coupling, which has a rigid limit of 1.0 kHz. We used a ^{13}C homonuclear decoupled ^{13}C-^{15}N

REDOR technique (29) for this purpose. **Figure 4** shows the REDOR data of several Cα sites. The first intensity minimum appears at 1.6 ms, which is the time expected for the rigid-limit ^{15}N-^{13}Cα coupling. Thus, AM2-TM is immobilized to below 10^3 s^{-1}. The lack of kilohertz motion deviates somewhat from the calculation in **Figure 1**, which reflects the uncertainty of the membrane viscosity and the helical bundle radius. For example, an η of 400 poise (dashed line in **Figure 1**) or higher yields D_R consistent with the ^{15}N-^{13}C REDOR result.

Effect of M2 immobilization on ^{13}C and ^{15}N NMR spectra

The two orders of magnitude decrease of motional rates by the viral membrane gives rise to slow-limit ^{13}C and ^{15}N NMR spectra of AM2-TM with high intensities and narrow lines at ambient temperature. **Figure 5** compares the AM2-TM ^{13}C cross-polarization (CP) spectra in the viral membrane and in the DLPC bilayer. The viral-membrane samples give strong and narrow backbone Cα signals from 303 K to 243 K, and the intensities increase monotonically with decreasing temperature, consistent with the Boltzmann law. In contrast, the DLPC samples have minimum backbone intensities at intermediate temperatures of ~263 K, indicating that the motional rates are comparable to the C-H dipolar coupling.

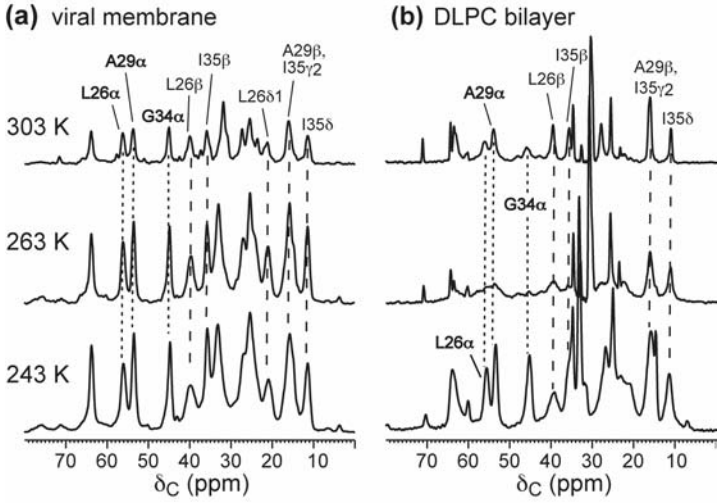


Fig. 5. ^{13}C CP-MAS spectra of LAGI AM2-TM in two different membranes. (a) Viral membrane with amantadine bound. (b) DLPC bilayers. Spectra measured at 303 K, 263 K

and 243 K were shown. The peptide shows stronger and narrow lines in the viral membrane than in DLPC bilayers at physiological temperature.

The sidechain ^{13}C signals in **Figure 5** have high intensities in both bilayers but the linewidths are slightly larger in the viral membrane. This can be explained by fast torsional motions of the sidechains in addition to the backbone motion. When the helical backbone diffuses on the intermediate timescale in DLPC bilayers, the combined motion of the sidechains is fast, giving rise to narrow lines. When the backbone is immobilized in the viral membrane, the sidechains have slightly slower motions, thus giving broader lines. But in neither bilayers do the sidechain motions fall outside the fast limit, thus high sensitivity is observed in both membranes.

To examine whether the change of membrane composition affects the peptide conformation, we compared 2D ^{13}C-^{13}C and ^{13}C-^{15}N correlation spectra of AM2-TM in the viral and DLPC membranes. Supporting information **Figure S1** shows regions of the 2D ^{13}C-^{13}C spectra, where most sites exhibit little chemical shift changes (< 0.5 ppm). The main exception is I33, whose sidechain carbons have chemical shift changes of 0.4 – 1.9 ppm. This is consistent with the lipid-facing location of I33 (30), which makes this site sensitive to the membrane compositional change. In comparison, channel-facing residues such as A30 and G34 have minimal ^{13}C chemical shift changes between the two membranes.

For the pore-facing V27, the Cα chemical shift decreased by 1.5 ppm while the C' chemical shift increased by 1.0 ppm compared to the DLPC values. V27 conformation and dynamics are known to be sensitive to drug binding (31) and pH (32). MD simulations indicate that at low external pH the V27 region adopts a closed conformation whereas at high pH the V27 region opens. This pH-dependent V27 gate, together with the NMR-detected drug-binding sensitivity and membrane sensitivity, indicate that V27 is a hot spot of conformational change in response to the environment.

^{15}N chemical shifts are even more sensitive to protein conformational homogeneity than ^{13}C chemical shifts. **Figure 6** compares 2D ^{15}N-^{13}C correlation spectra of AM2-TM in DLPC- and viral membranes without and with amantadine. All spectra were measured on

immobilized protein, at 243 K for the DLPC samples but ~303 K for the viral membrane samples. Remarkably, the viral membrane samples show much narrower ^{15}N lines than the DLPC samples, indicating that the helices adopt more homogeneous conformations in the viral membrane. In addition, G34 ^{15}N peak shows interesting dependences on the membrane composition and drug binding. In DLPC bilayers, the peak is broad (107.0 ppm) without the drug, split into two peaks, G34a (106.3 ppm) and G34b (109.7 ppm), after amantadine binding. In the viral membrane, the G34 ^{15}N signal has the same two-peak pattern in the absence of drug, but changes to a single peak at the downfield position (G34b) in the presence of amantadine, Thus, the viral membrane and amantadine both promote the downfield ^{15}N chemical shift, so that under the combined effects of the two, the G34b conformer dominates.

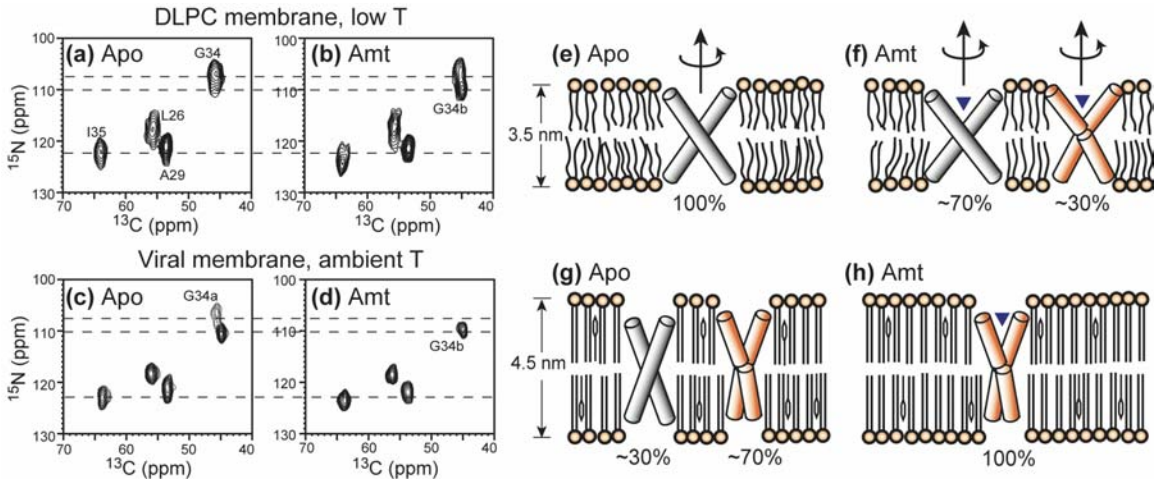


Fig. 6. 2D ^{15}N-^{13}C HETCOR spectra of LAGI AM2-TM in different lipid membranes and drug binding states. (a) Apo AM2-TM in DLPC bilayers at 243 K. (b) Amantadine-bound AM2-TM in DLPC bilayers at 243 K. (c) Apo AM2-TM in viral membranes at 294 K. (d) Amantadine-bound AM2-TM in viral membranes at 303 K. The corresponding helix orientations are shown on the right. (e) Apo peptide in DLPC bilayers. The helices are straight and tilted by 35˚ from the bilayer normal (31). (f) Amantadine-bound peptide in DLPC bilayers. ~30% of the helices exhibit a kink of 10˚ at G34. (g) Apo peptide in the viral membrane. ~70% of the helices have the kinked conformation. (h) Amantadine-bound peptide in the viral membrane. All helices have the kinked conformation.

Discussion

The ^{2}H quadrupolar spectra, C-H, N-H and ^{13}C-^{15}N dipolar couplings indicate unequivocally that the > 100 kHz rotational diffusion of AM2-TM in DLPC bilayers is slowed to < 1 kHz in the eukaryote- and virus-mimetic membrane at physiological temperature. This immobilization narrowed and increased the backbone ^{13}C and ^{15}N signals where previously no signals were observed due to intermediate-timescale motion (12). The dramatic immobilization of the AM2-TM helical bundle by over two orders of magnitude is a direct result of the increased viscosity and thickness of the viral membrane over model PC membranes. It also reflects the structural plasticity of AM2-TM, which has been well documented by the sensitivity of helix orientation (12, 33, 34), conformation (31, 35-37), and mobility to membrane thickness and drug binding. The extent of AM2-TM tetramerization is also membrane dependent: thicker bilayers and cholesterol-containing bilayers stabilize the tetramer (38) based on thiol disulfide exchange data.

The ability of the eukaryote-mimetic viral membrane to immobilize a classical TM helical bundle at physiological temperature offers significant opportunities for structure determination of membrane proteins. So far with few exceptions (39) membrane protein NMR studies that involve synthetic lipids use one or two-component lipid bilayers without cholesterol, which create unfavorable motional properties for the proteins. In cases where more native membrane extracts were used, the most common choices are *E. coli* lipids (40), asolectin (41), and purple membrane lipids (42, 43), none of which have the immobilizing properties of the viral membrane. The purple membrane is noteworthy, as it is the matrix in which bacteriorhodopsin (bR) forms immobile trimeric crystalline arrays. However, the purple membrane consists of ether-linked diphytanoyl lipids that do not have a phase transition between -120˚C and 80˚C (44). Thus bR immobilization is due to its dense packing, despite the fluidity of the purple membrane. Overall, the virus-envelope mimetic membrane is the only one identified so far to directly immobilize and homogenize proteins. By removing whole-body motion and enhancing NMR spectral resolution and sensitivity,

structure determination of eukaryotic membrane proteins is both more biological and more favorable spectroscopically.

The effects of the viral membrane on the AM2-TM conformation are interesting. ^{13}C and ^{15}N chemical shifts indicate that the viral membrane does not affect the *average* conformation of most residues except for lipid-exposed sidechains, but reduces the conformational distribution of the protein compared to frozen PC membranes. While it may be puzzling at first that a more complex membrane mixture would create a more homogeneous protein structure, it is in fact understandable because the more viscous viral membrane should promote oligomeric structures that depend more on helix-helix interactions rather than protein-lipid interactions. In contrast, a fluid PC membrane is more able to distort the helix structure and assembly, thus giving rise to larger chemical shift distributions at low temperature. Further, the viral membrane mixture used here exists in a single liquid ordered phase, because the high percentage of cholesterol (30 mol%) and the lack of unsaturated lipids are known to suppress domain formation (45). This single liquid-ordered phase also promotes more homogeneous conformation of the protein.

The observed chemical shift perturbation of G34 and V27 by the membrane composition change does not result from specific cholesterol binding, since the number of cholesterol molecules bound per tetramer is low (0.5 – 0.9) (38, 46). Instead, it reflects the intrinsic conformational flexibility of G34 and V27 to environmental changes. G34 is the site at which a helix kink of 11˚ was observed as a result of amantadine binding (34), where the C-terminal segment became less tilted than the N-terminus segment. By inference, the downfield ^{15}N isotropic peak upon amantadine binding (**Figure 6**) is also due to the kinked helices. This implies that the two G34 ^{15}N peaks in the apo viral membrane result from the coexistence of straight and kinked helices. The growth of kinked and less tilted helices is consistent with the larger thickness of the viral membrane, by reducing the hydrophobic mismatch between the protein and the membrane. Since amantadine and the viral membrane affect the helix orientation in the same direction, the drug-complexed peptide in the viral membrane shows only the kinked conformation. **Figure 6(e-h)** depicts the proposed AM2-TM orientations under the four membrane/drug conditions, where the populations of the straight and kinked helices are estimated from the relative intensities of the two G34 peaks in

the 2D spectra. Taken together, the chemical shift data is most consistent with the peptide affected by the cholesterol- and SM-induced changes in the membrane physical properties.

In virus-infected host cells, the M2 protein is known to not concentrate in raft-like microdomains, in contrast to influenza glycoproteins hemagglutinin and neuraminidase (46, 47). But this does not make studies of M2 proteins in cholesterol-rich membranes less relevant, because the level of M2 proteins is known to be low in the virus envelope, which indicates that there is little disordered phase for M2 to partition to, but that most M2 proteins reside in a liquid-ordered phase similar to that used here.

In conclusion, a membrane mixture mimicking the eukaryotic cell membrane in general and the influenza virus envelope composition in particular, is found to immobilize the whole-body uniaxial rotation of a tetrameric helical bundle and homogenizes its conformation. The use of this cholesterol-rich eukaryotic membrane mixture should greatly facilitate NMR structure determination of membrane proteins by eliminating dynamic broadening of lineshapes, thus enhancing spectral resolution and sensitivity. It also allows structural information to be extracted at physiological temperature.

Materials and Methods

Membrane sample preparation

The TM domain of the M2 protein of the Udorn strain (residues 22–46, SSDPLVVAASII GILHLILWILDRL) was synthesized by Fmoc chemistry and purified to >95% purity. Two peptide samples containing eight uniformly ^{13}C, ^{15}N-labeled residues were used. The labeled sites are L26, A29, G34 and I35 in one sample (LAGI), and V27, A30, I33 and L38 in the other (VAIL).

Egg sphingomyelin (SM), which is predominantly (84%) N-palmitoyl-D-*erythro*-sphingosylphosphorylcholine, was dissolved in chloroform/methanol (10:2) solution, then mixed with DPPC, DPPE and Chol to reach a molar ratio of SM: DPPC: DPPE : Chol = 28% : 21% : 21% : 30%. The membrane mixture was lyophilized, dissolved in pH 7.5 phosphate

buffer, vortexed, and freeze-thawed several times to form large unilamellar vesicles. The membrane has a broad phase transition around 263 K based on static ^{31}P spectra (not shown).

AM2-TM was reconstituted into the viral membrane mixture by detergent dialysis as described before (37) at peptide : lipid molar ratios of 1: 15 – 1:12. The proteoliposomes were centrifuged at 150,000 g to obtain the membrane pellet. Photometric assay showed >95% binding of the peptide. The pellet was packed into 4 mm MAS rotors for NMR experiments. For the amantadine-bound sample, amantadine hydrochloride in the pH 7.5 buffer was directly titrated to the pellet at a peptide: amantadine molar ratio of 1: 2.

Solid-state NMR

SSNMR experiments were carried out on a 400 MHz (9.4 Tesla) and a 600 MHz (14.1 Tesla) spectrometer using 4 mm MAS probes. Experiments on viral membrane samples were conducted near 303 K, well above the phase transition temperature. ^{15}N-^{1}H and ^{13}C-^{1}H dipolar couplings were measured under 7 kHz MAS using dipolar-doubled DIPSHIFT experiments (48). An FSLG sequence with an effective field of 61.2 kHz was used for ^{1}H homonuclear decoupling. 2D ^{15}N-^{13}C correlation spectra (49) were measured under 7 kHz MAS using a REDOR pulse train (50) of 0.7 ms for ^{13}C-^{15}N coherence transfer. Typical radio-frequency fields were 50 kHz for ^{13}C and ^{15}N and 60-70 kHz for ^{1}H. ^{13}C and ^{15}N chemical shifts were referenced to the α-Gly CO signal at 176.49 ppm on the TMS scale and the ^{15}N signal of N-acetyl-valine at 122 ppm on the liquid ammonia scale, respectively.

Acknowledgement: This work is funded by an NSF grant MCB-543473. The authors thank R.G. Griffin for discussions on the purple membrane composition.

Supporting Information

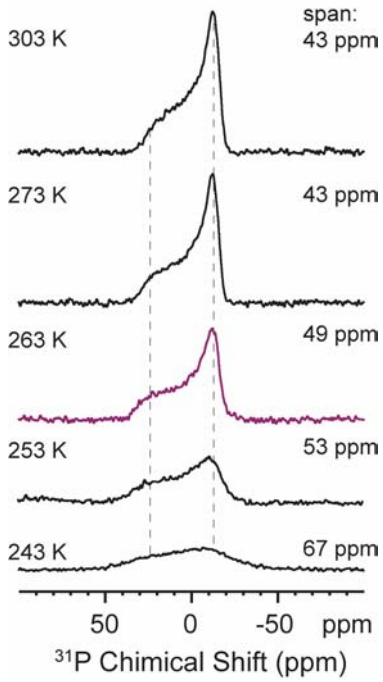


Figure 7.S1. Static ^{31}P direct-polarization spectra of AM2-TM containing viral membranes as a function of temperature. The width of the chemical shift anisotropy is indicated on the right. The membrane broadens significantly at 243 K, which correlates with the ^{15}N line broadening of the protein, indicating that the membrane phase property strongly affects the protein conformational averaging and conformational distribution.

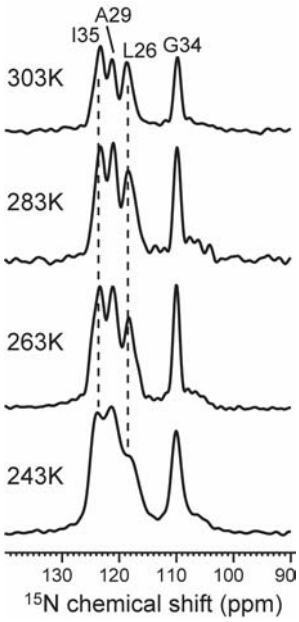


Figure 7.S2. ^{15}N CP-MAS spectra of LAGI-M2TMP in the viral membrane from 303 K to 243 K. The spectra show the same monotonic intensity increase with decreasing temperature as the ^{13}C spectra in **Fig. 5**. In addition, the ^{15}N spectra exhibit pronounced line broadening around 243 K, consistent with the large ^{31}P chemical shift anisotropy increase around 243 K. This suggests that the protein line broadening is due to the phase behavior and disorder of the viral membrane at low temperature.

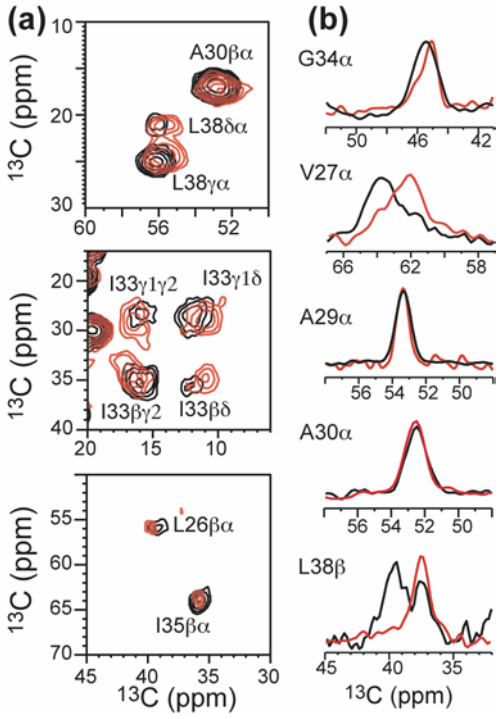


Figure 7.S3. ^{13}C chemical shifts of AM2-TM in viral membranes (red) versus DLPC bilayers (black). (a) Selected regions of the 2D spectra. The data were obtained at 303 K for the viral membrane samples and 243 K for the DLPC samples. (b) Selected 1D cross sections.

References

1. Hong M (2006) Oligomeric structure, dynamics, and orientation of membrane proteins from solid-state NMR. *Structure* 14:1731-1740.
2. Hong M (2007) Structure, topology, and dynamics of membrane peptides and proteins from solid-state NMR spectroscopy. *J. Phys. Chem. B.* 111:10340-10351.
3. Reuther G*, et al.* (2006) The lipidated membrane anchor of full length N-Ras protein shows an extensive dynamics as revealed by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* 128:13840-13846.
4. Park SH, Mrse AA, Nevzorov AA, De Angelis AA, & Opella SJ (2006) Rotational diffusion of membrane proteins in aligned phospholipid bilayers by solid-state NMR spectroscopy. *J. Magn. Reson.* 178:162-165.
5. Hong M (2006) Solid-state NMR studies of the structure, dynamics, and assembly of beta-sheet membrane peptides and alpha-helical membrane proteins with antibiotic activities. *Acc. Chem. Res.* 39:176-183.
6. Luo W, Yao XL, & Hong M (2005) Large Structure Rearrangement of Colicin Ia Channel Domain After Membrane Binding from 2D 13C Spin Diffusion NMR. *J. Am. Chem. Soc.* 127:6402-6408.
7. Huster D, Xiao LS, & Hong M (2001) Solid-State NMR Investigation of the dynamics of colicin Ia channel-forming domain. *Biochemistry* 40:7662-7674.
8. Tang M, Waring AJ, Lehrer RI, & Hong M (2008) Effects of Guanidinium-Phosphate Hydrogen Bonding on the Membrane-Bound Structure and Activity of an Arginine-Rich Membrane Peptide from Solid-State NMR. *Angew. Chem. Int. Ed. Engl.* 47:3202-3205.
9. Doherty T, Waring AJ, & Hong M (2008) Dynamic structure of disulfide-removed linear analogs of tachyplesin-I in the lipid bilayer from solid-state NMR. *Biochemistry* 47:1105-1116.
10. Bain AD (2003) Chemical exchange in NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 43:63-103.
11. Saffman PG & Delbruck M (1975) Brownian motion in biological membranes. *Proc. Natl. Acad. Sci. U.S.A.* 72(8):3111-3113.

12. Cady SD, Goodman C, C.Tatko, DeGrado WF, & Hong M (2007) Determining the orientation of uniaxially rotating membrane proteins using unoriented samples: a 2H, 13C, and 15N solid-state NMR investigation of the dynamics and orientation of a transmembrane helical bundle. *J. Am. Chem. Soc.* 129:5719-5729.
13. Pinto LH, Holsinger LJ, & Lamb RA (1992) Influenza virus M2 protein has ion channel activity. *Cell* 69:517-528.
14. Pinto LH & Lamb RA (2007) Controlling influenza virus replication by inhibiting its proton flow. *Mol. BioSyst.* 3:18-23.
15. Hay AJ, Wolstenholme AJ, Skehel JJ, & Smith MH (1985) The molecular basis of the specific anti-influenza action of amantadine. *EMBO J.* 4:3021-3024.
16. Wang C, Takeuchi K, Pinto LH, & Lamb RA (1993) Ion Channel Activity of Influenza A Virus M2 Protein: Characterization of the Amantadine Block. *J. Virol.* 67(9):5585-5594.
17. Stouffer AL*, et al.* (2008) Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* 451(7178):596-599.
18. Schnell JR & Chou JJ (2008) Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* 451:591-595.
19. Kovacs FA & Cross TA (1997) Transmembrane four-helix bundle of influenza A M2 protein channel: structural implications from helix tilt and orientation. *Biophys. J.* 73:2511-2517.
20. Li C, Qin H, Gao FP, & Cross TA (2007) Solid-state NMR characterization of conformational plasticity within the transmembrane domain of the influenza A M2 proton channel. *Biochim. Biophys. Acta* 1768:3162-3170.
21. Klenk HD, Becht H, & Rott R (1972) Structure of Influenza-Virus Envelope. *Virology* 47(3):579-591.
22. Blough HA & Merlie JP (1970) Lipids of Incomplete Influenza Virus. *Virology* 40(3):685-692.
23. Blough HA (1971) Fatty Acid Composition of Individual Phospholipids of Influenza Virus. *J. Gen. Virol.* 12:317-320.

24. Edidin M (2003) The State of Lipid Rafts: From Model Membranes to Cells. *Annu. Rev. Biophys. Biomol. Struct.* 32:257-283.
25. Niemela PS, Hyvonen MT, & Vattulainen I (2009) Atom-scale molecular interactions in lipid raft mixtures. *Biochim. Biophys. Acta* 1788:122-135.
26. Gambin Y*, et al.* (2006) Lateral mobility of proteins in lipid membranes revisited. *Proc. Natl. Acad. Sci. U.S.A* 103:2098-2102.
27. Lindblom G, Oradd G, & Filippov A (2006) Lipid lateral diffusion in bilayers with phosphatidylcholine, sphingomyelin and cholesterol - An NMR study of dynamics and lateral phase separation. *Chem. Phys. Lipids* 141(1-2):179-184.
28. Jelinski LW, Sullivan CE, & Torchia DA (1980) 2H NMR study of molecular motion in collagen fibrils. *Nature* 284:531-534.
29. Jaroniec CP, Tounge BA, Rienstra CM, Herzfeld J, & Griffin RG (1999) Measurement of 13C-15N distances in uniformly 13C labeled biomolecules: J-decoupled REDOR. *J. Am. Chem. Soc.* 121:10237-10238.
30. Pinto LH*, et al.* (1997) A functionally defined model for the M2 proton channel of influenza A virus suggests a mechanism for its ion selectivity. *Proc. Natl. Acad. Sci. U.S.A* 94:11301-11306.
31. Cady SD & Hong M (2008) Amantadine-Induced Conformational and Dynamical Changes of the Influenza M2 Transmembrane Proton Channel. *Proc. Natl. Acad. Sci. U.S.A* 105:1483-1488.
32. Khurana E*, et al.* (2009) Molecular dynamics calculations suggest a conduction mechanism for the M2 proton channel from influenza A virus. *Proc. Natl. Acad. Sci. U. S. A.* 106:1069-1074.
33. Wang J, Kim S, Kovacs F, & Cross TA (2001) Structure of the the transmembrane region of the M2 protein H+ channel. *Prot. Sci.* 10:2241-2250.
34. Hu J*, et al.* (2007) Backbone structure of the amantadine-blocked trans-membrane domain M2 proton channel from Influenza A virus. *Biophys. J.* 92:4335-4343.
35. Nishimura K, Kim S, Zhang L, & Cross TA (2002) The closed state of a H+ channel helical bundle combining precise orientational and distance restraints from solid state NMR. *Biochemistry* 41(44):13170-13177.

36. Hu J, *et al.* (2006) Histidines, heart of the hydrogen ion channel from influenza A virus: Toward an understanding of conductance and proton selectivity. *Proc. Natl. Acad. Sci. USA* 103(18):6865-6870.

37. Cady SD, Mishanina TV, & Hong M (2009) Structure of Amantadine-Bound M2 Transmembrane Peptide of Influenza A in Lipid Bilayers from Magic-Angle-Spinning Solid-State NMR: the Role of Ser31 in Amantadine Binding. *J. Mol. Biol.* 385:1127-1141.

38. Cristian L, Lear JD, & DeGrado WF (2003) Use of thiol-disulfide equilibria to measure the energetics of assembly of transmembrane helices in phospholipid bilayers. *Proc. Natl. Acad. Sci. USA* 100:14772-14777.

39. Yang J, Prorok M, Castellino FJ, & Weliky DP (2004) Oligomeric beta-structure of the membrane-bound HIV-1 fusion peptide formed from soluble monomers. *Biophys. J.* 87:1951-1963.

40. Hiller M, *et al.* (2005) Solid-state magic-angle spinning NMR of outer-membrane protein G from Escherichia coli. *ChemBioChem* 6:1679-1684.

41. Schneider R, *et al.* (2008) Solid-state NMR spectroscopy applied to a chimeric potassium channel in lipid bilayers. *J. Am. Chem. Soc.* 130:7427-7435.

42. Struts AV, *et al.* (2007) Structural analysis and dynamics of retinal chromophore in dark and meta I states of rhodopsin from 2H NMR of aligned membranes. *J. Mol. Biol.* 372:50-66.

43. Lewis BA, Harbison GS, Herzfeld J, & Griffin RG (1985) NMR structural analysis of a membrane protein: bacteriorhodopsin peptide backbone orientation and motion. *Biochemistry* 24:4671-4679.

44. Kates M, Kushwaha SC, & Sprott GD (1982) Lipids of purple membrane from extreme halophiles and of methanogenic bacteria. *Methods Enzymol.* 88:98-111.

45. Bunge A, Müller P, Stöckl M, Herrmann A, & Huster D (2008) Characterization of the ternary mixture of sphingomyelin, POPC, and cholesterol: support for an inhomogeneous lipid distribution at high temperatures. *Biophys. J.* 94:2680-2690.

46. Schroeder C, Heider H, Möncke-Buchner E, & Lin TI (2005) The influenza virus ion channel and maturation cofactor M2 is a cholesterol-binding protein. *Eur. Biophys. J.* 34:52-66.

47. Leser GP & Lamb RA (2005) Influenza virus assembly and budding in raft-derived microdomains: a quantitative analysis of the surface distribution of HA, NA and M2 proteins. *Virology* 342:215-227.

48. Hong M, *et al.* (1997) Coupling amplification in 2D MAS NMR and its application to torsion angle determination in peptides. *J. Magn. Reson.* 129(1):85-92.

49. Hong M & Griffin RG (1998) Resonance assignments for solid peptides by dipolar-mediated C-13/N-15 correlation solid-state NMR. *J. Am. Chem. Soc.* 120(28):7113-7114.

50. Gullion T & Schaefer J (1989) Rotational-Echo Double-Resonance NMR. *J. Magn. Reson.* 81(1):196-200.

Chapter 8

Water – Protein Interactions of the Influenza M2 Proton Channel in Lipid Bilayers from Solid-State NMR

A paper to be submitted to the Structure

Wenbin Luo and Mei Hong

Abstract

The influenza A virus M2 protein is a pH-gated and amantadine-inhibited proton channel important for the viral life cycle. Although the M2 channel activity is known to involve water, direct experimental evidence of M2-water interaction has been absent. Using ^{1}H spin diffusion solid-state NMR, we have now determined the water accessibility of the M2 transmembrane domain (M2-TM) in virus-envelope-mimetic lipid membranes. Site-specific water-protein magnetization transfer indicates that, in the absence of amantadine, the initial spin diffusion rate is primarily sensitive to the radial position of the residues from the water-filled pore: pore-lining residues along the helix have similar water accessibilities. However, upon drug binding, significant buildup rate differences are observed between two pore-lining residues: water-Gly_{34} spin diffusion is much slower than water-Val_{27} spin diffusion, indicating that amantadine is bound in the lumen between these two residues. Comparison of the buildup curves indicates that spin diffusion is the fastest in the low-pH state, slower in the high-pH state, and the slowest in the high-pH amantadine-bound state. These data provide the first experimental evidence of the pH-induced change of the diameter of the water-filled pore and the drug-induced dehydration of the pore. Simulations of the spin diffusion curves yielded quantitative values of the water-accessible surface area changes of the protein by pH and drug binding. This spin diffusion NMR technique is generally useful for elucidating water interactions with ion channels, water pores, and proton pumps, and for

probing membrane protein conformational changes that involve significant changes of water-accessible surfaces.

Introduction

Water underlies the folding and functions of many ion channels (1, 2), water pores (3), and proton pumps in biological membranes (4, 5) and is important for the solvation of charged residues in lipid bilayers (6-8). Elucidating water interaction with membrane proteins in the low-dielectric lipid membrane and the dynamics of water in the confined environment of membrane proteins (9) is thus of fundamental interest. The influenza A M2 protein forms a pH-gated proton channel in the virus envelope that is important for the virus lifecycle (10-12). Acidification of the virus particle triggers the release of the viral RNA into the infected cell, initiating viral replication. The M2 channel activity is mediated by water molecules and by the action of a key residue, His_{37} (13), and is inhibited by amantadine and rimantadine (14). Recent high-resolution structural studies of the M2 protein by X-ray crystallography (15) and NMR spectroscopy (16-18) have provided a wealth of information about the global conformation and the structures of specific residues important for proton conduction. However, little direct experimental evidence is yet available on how water interacts with the M2 protein under different pH and drug-binding conditions. Most proposals for the mechanism of proton conduction so far came from molecular dynamics simulations (19-21).

Solid-state NMR (SSNMR) spectroscopy provides a unique and powerful tool for studying water-protein interactions directly in native lipid bilayers (22, 23). Correlation of water-protein ^{1}H-^{13}C signals after dipolar-driven ^{1}H spin diffusion gives detailed information on the proximity of protein residues to water. The rate of this ^{1}H spin diffusion was initially used to determine the global topology of membrane proteins (24), but was recently shown to also contain specific information about the water-protein surface area (25).

In this study, we use water-to-protein ^{1}H spin diffusion experiments to investigate the water accessibilities and water dynamics of the M2 transmembrane peptide (M2-TM) in virus-envelope-mimetic lipid bilayers (26). We demonstrate that the spin diffusion buildup

rates are site specific and differ between lipid-facing and pore-lining residues in the absence of drug, thus the source of water magnetization is primarily the pore water. Drug binding significantly changes the water accessibilities of two pore-lining residues, Val_{27} and Gly_{34}, indicating that the drug is bound in the pore lumen between these two residues. Using a 3D lattice model, we simulated the water-to-protein spin diffusion buildup curves at low pH, high pH, and the amantadine-bound high-pH states, and found that the water-exposed surface area of the protein decreased by ~25% from the open state to the closed state, consistent with a smaller pore for the closed channel, but the amantadine-bound state has the smallest water-exposed surface area, indicating obstruction of the water pathway in the pore. This work provides the first direct evidence of a close correlation between the water accessibility and the function of the M2 proton channel.

Result and Discussion

Differential water accessibility of M2-TM under different pH and drug binding conditions

We investigated the M2-water contact by measuring the protein ^{13}C signals that originated from water by spin diffusion. Using a 2D ^{1}H-^{13}C correlation experiment with a ^{1}H T_2 filter (2 ms) and no ^{1}H homonuclear decoupling during the evolution period, we removed all ^{1}H magnetization of the rigid protein (26), thus the protein ^{13}C signals must have originated from the mobile water or lipids, as verified by the ^{1}H chemical shifts in the 2D spectra. The use of a lipid mixture mimicking the virus-envelope lipid composition was essential for obtaining high-sensitivity spectra of M2-TM at physiological temperature, since the peptide undergoes intermediate-timescale motion in simple phosphocholine bilayers, which severely broadens the NMR spectra (27, 28). The protein reconstituted into the lipid bilayer in this way is completely tetramerized, as shown directly by ^{19}F NMR spin counting experiments before (29) and by analytical ultracentrifugation experiments (30). The five labeled residues chosen represent different proximities to water: the N-terminal Leu_{26}, Val_{27}, Ala_{29} are close to the bilayer surface water while Gly_{34} and Ile_{35} are far. On the other hand, Val_{27} and Gly_{34} line the pore lumen while Ala_{29} faces the lipids. Thus, the labeled residues allow us to examine whether spin diffusion primarily depends on the residue proximity to the bilayer-surface water or the proximity to the pore-lumen water.

Fig. 1 compares the 2D ^{1}H-^{13}C spectra of M2-TM under different pH and drug binding conditions. Spectra measured with a short mixing time of 4 ms were compared to those of 100 ms to first deduce qualitatively the water accessibility of the residues. At pH 4.5, the water cross peak intensities at 4 ms relative to 100 ms are similar between the N-terminal residues Leu_{26} and Val_{27} and the middle residue Gly_{34}. Since Gly_{34} is significantly deeper in the membrane, the similarity indicates that there is a continuous water pathway in the pore from the N-terminus to the center of the helical bundle, which gives pore-facing Val_{27} and Gly_{34} similar water accessibilities. In comparison, the water cross peak of the lipid-facing Ala_{29} at 4 ms is weaker than the other residues, indicating that the experiment is sensitive to the water accessibility difference between lipid-facing residues and pore-lining residues, and that the radial distance to the luminal water is the main determining factor for the water cross peak intensity at short mixing times.

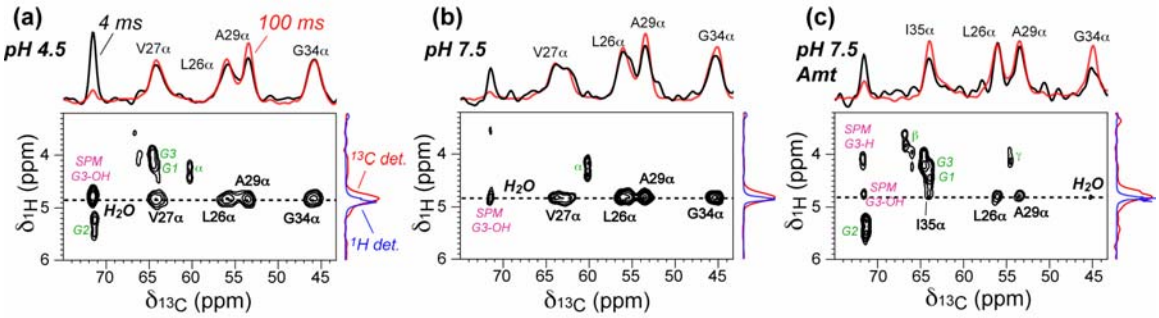


Fig. 1. 2D ^{13}C-^{1}H correlation spectra of M2-TM in virus-mimetic lipid membranes at 293 K. A ^{1}H T_2 filter time of 2 ms and a spin diffusion mixing time of 4 ms were used. (a) pH 4.5. (b) pH 7.5. (c) pH 7.5 with amantadine. Assignments for intermolecular water-protein cross peaks (black) as well as intramolecular phospholipid (green) and SPM peaks (magenta) are indicated. The water ^{1}H cross section is shown at the top (black), superimposed with the 100 ms water cross section (red). The G34 Cα cross section at 100 ms mixing is shown on the right (red) and is superimposed with the ^{1}H 1D spectra (blue) to indicate the small upfield shift of the membrane-associated water compared to the bulk water.

At pH 7.5, the main features of the pH 4.5 spectrum is preserved, but now the Gly_{34} peak is slightly lower than the N-terminal residues at 4 ms (**Fig. 1b**), suggesting that the amount of water in the middle of the pore is less than at pH 4.5.

When the protein is complexed to amantadine, the Gly_{34} and Ile_{35} cross peaks with water become significantly weaker at 4 ms compared to those of Leu_{26} and Ala_{29} (**Fig. 1c**). Thus, in the presence of drug, the N-terminus has substantially higher water accessibilities than the middle portion of the protein, suggesting that amantadine interrupts the water pathway between Ala_{29} and Gly_{34}.

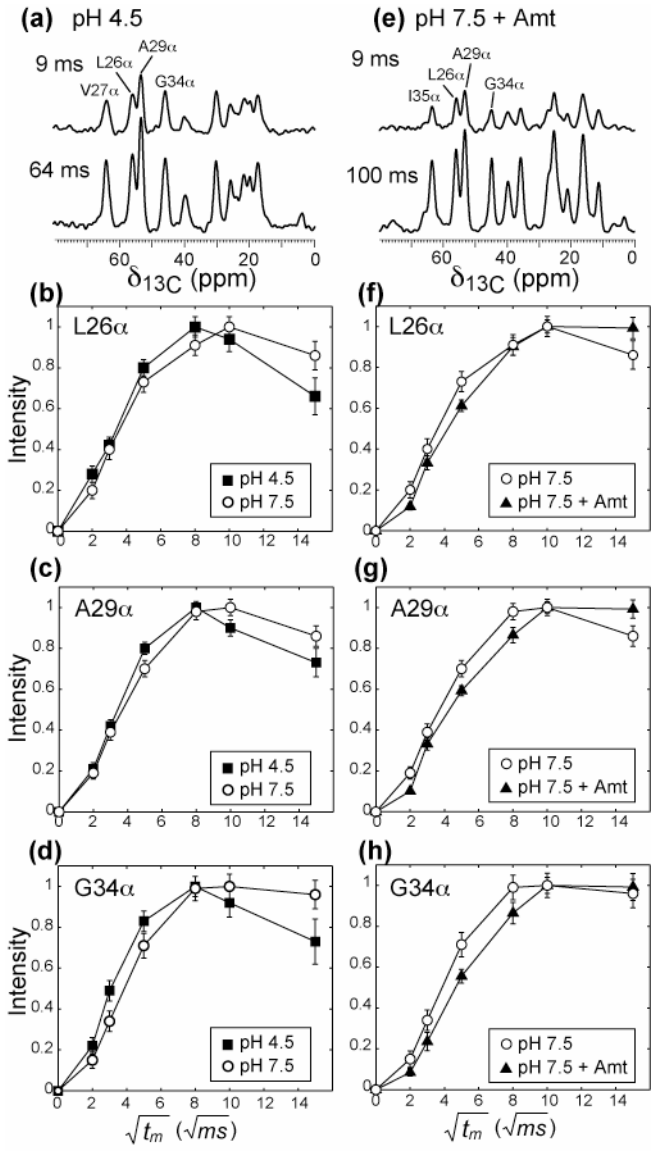


Fig. 2. Water-to-M2 ^{1}H spin diffusion buildup curves from 1D ^{13}C DQ filtered spectra. (a) Representative ^{13}C DQ filtered spectra at pH 4.5. (b-d) Buildup curves of several ^{13}C sites at pH 4.5 (filled squares) and at pH 7.5 (open circles). (e) Representative ^{13}C DQ filtered

spectra at pH 7.5 in the presence of amantadine. (f-h) Buildup curves of several ^{13}C sites at pH 7.5 without (open circles) and with amantadine (filled triangles).

The full 2D ^{1}H-^{13}C spectra indicate a lack of lipid ^{1}H – protein ^{13}C cross peaks within the mixing times of interest (< 225 ms) (**Fig. S1**). Thus, a 1D version of the 2D experiment is sufficient for extracting the water-protein buildup rates. The shorter experiments allow more mixing times to be measured so that quantitative buildup curves can be obtained. To suppress the ^{13}C signals of unlabeled lipids and cholesterol in the 1D spectra, we also added a ^{13}C double-quantum (DQ) filter. **Fig. 2** and **Fig. S2** shows representative 1D ^{13}C DQ spectra and the resulting buildup curves for the three states of M2-TM. For all sites studied, the intensity buildup is the fastest at pH 4.5, moderately slower at pH 7.5, and significantly slower upon amantadine binding. Moreover, while this trend is true for all labeled sites, amantadine decreases the Gly_{34} buildup rate more significantly than it does the N-terminal residues, confirming that water accessibility is lower in the middle of the TM helices than at the N-terminus in the drug-bound state. Quantifying the buildup rates using the initial slope yielded t_m^S values (**Fig. S3** and **Table S1**), which are inversely related to the water-accessible surface area. Increasing the pH from 4.5 to 7.5 increased the average t_m^S by 20% while amantadine binding increased t_m^S by 56% compared to the open state.

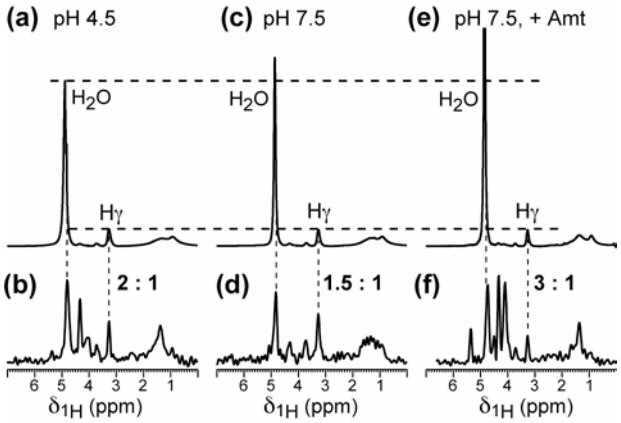


Fig. 3. 1D ^{1}H direct-polarization (DP) spectra and ^{31}P-detected ^{1}H spectra extracted from 2D ^{31}P-^{1}H correlation spectra of membrane-bound M2-TM. (a, c, e): 1D ^{1}H spectra showing the full water peak. (b, d, f): Projection of the ^{1}H cross sections of the 2D ^{31}P-^{1}H spectra with 64

ms spin diffusion (full spectra in **Fig. S4**). (a, b) pH 4.5. (c, d) pH 7.5. (e, f) pH 7.5 with amantadine.

To test whether the substantially slower spin diffusion of the amantadine-bound sample is due to lower water content of the sample instead of true obstruction of the pore, we measured the 1D ^{1}H spectra and 2D ^{1}H-^{31}P correlation spectra of the three samples. The 1D ^{1}H spectra report the total water content of each sample, including both membrane-associated water and bulk water, while the 2D ^{1}H-^{31}P correlation spectra indicate the amount of inter-bilayer water in the multilamellar vesicles. **Fig. 3** shows that normalized to the lipid Hγ intensity, which reflects the lipid amount, the water intensity increases in the order of pH 4.5 ≈ pH 7.5 < pH 7.5 with amantadine. 2D ^{1}H-^{31}P correlation spectra measured confirm that the drug-bound sample has the highest water intensity relative to the lipid Hγ intensity (**Fig. S4**). Thus, the amantadine-bound sample has ample membrane-associated water, and the slow Gly_{34} intensity buildup must be attributed to obstruction of the water pathway in the channel.

Comparison of the 2D correlation spectra and the 1D ^{1}H spectra indicates that the protein- and lipid-correlated water resonates at a ^{1}H chemical shift of 4.73-4.83 ppm, which is about 0.1 ppm lower than the bulk water chemical shift of 4.80-4.93 ppm (**Fig. 1, 3**). Thus, the inter-bilayer (correlated with ^{31}P) and pore (correlated with protein) water together has detectably different physical properties from the bulk water outside the multilamellar vesicles, as expected due to the confinement of the membrane-associated water. On the other hand, the inter-bilayer water and pore water, which are within nanometers of each other, are expected to be fully averaged in a few milliseconds due to the fast water translational diffusion, thus their chemical shifts should be indistinguishable. Indeed, the ^{13}C-detected and ^{31}P-detected water chemical shifts are identical within the experimental resolution (**Fig. S1** and **S4**).

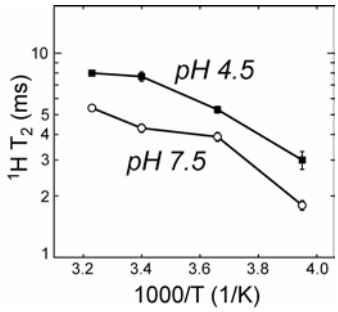


Fig. 4. ^{1}H T_2 of the inter-bilayer and channel water as a function of temperature. The water T_2 was measured by protein ^{13}C detection after 100 ms spin diffusion. Filled squares: pH 4.5. Open circles: pH 7.5.

The ability to selectively measure the pore- and inter-bilayer water but not bulk water allowed us to probe the dynamics of protein-associated water. The ^{1}H T_2 relaxation times reflect the reorientational rates of water. Water molecules in the fast motional limit should exhibit long T_2 relaxation times that increase with increasing temperature. **Fig. 4** shows the protein-^{13}C detected water ^{1}H T_2 at low and high pH without amantadine. Under both conditions the water ^{1}H T_2 increases with temperature between 253 K and 313 K, indicating fast reorientational motion, but the low pH state has a longer water T_2 than the high pH state. Thus, water is more dynamic at low pH, again consistent with a wider water-filled pore in the open state.

Water-M2 surface areas from 3D lattice calculations

To obtain more quantitative information about how pH and amantadine change the protein-water surface area, we simulated the integrated water-protein ^{1}H spin diffusion buildup curves using a 3D lattice model. In the simulation, a 44-Å thick lipid bilayer (31) was constructed from 2-Å sized cubes, in which the four-helix bundle M2-TM was represented by appropriate numbers of cubes in each plane (**Fig. S5-S7**) so that the helices were tilted by ~25˚ from the bilayer normal. This tilt angle was extrapolated from the measured M2-TM orientations in bilayers of varying thickness, specifically, DLPC, DMPC, and POPC bilayers (27, 32-34). Additionally, amantadine causes a kink of the helix at Gly_{34} (33) with a smaller tilt angle for the C-terminal segment, thus we adjusted the protein cube

positions for the amantadine-bound state to create a less tilted C-terminal segment. The total volume of the protein was kept constant at ~12.7 nm^3 (**Table S2**) based on an average protein density of 1.43 g/cm^3 (35, 36). These low-resolution models do not attempt to delineate the shape and volume of the sidechains, but they are sufficient for determining the *change* of the protein-water surface area by pH and drug binding. The center of the helical bundle was filled with water cubes, and one layer of interface cubes was used between the protein and water cubes. For the amantadine-bound state, the drug, whose approximate volume is 0.2 nm^3, was centered near Ser_{31} to be consistent with the observed maximal chemical shift perturbation at this site (37). The amide group of the drug was assumed to point down based on the recent crystal structure (15).

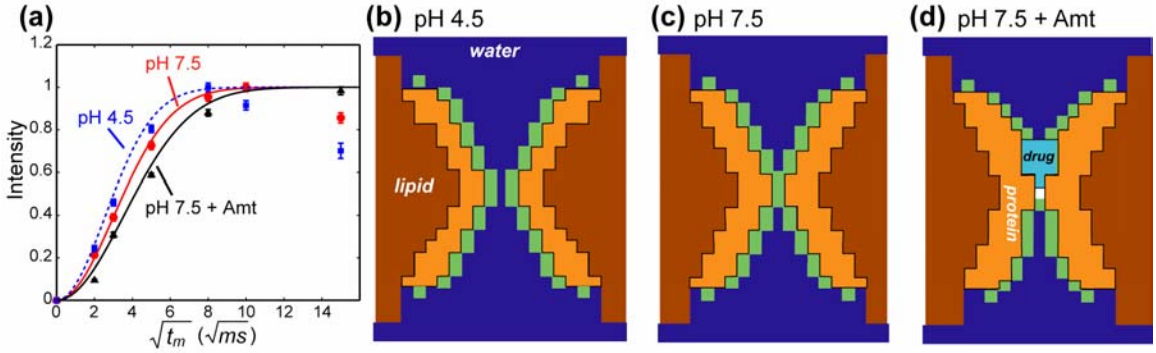


Fig. 5 Quantification of the water-accessible surface area of M2-TM from ^1H spin diffusion buildup curves. (a) Normalized water-to-M2 spin diffusion buildup curves from the integrated intensities (64-16 ppm) of the 1D ^{13}C DQ filtered spectra. Error bars are 1-2% on the normalized scale and are thus smaller than the symbols. Best-fit buildup curves (lines) were obtained as described in the text. (b-d) Low-resolution structural models of the M2-TM channels used to obtain the best fits. (b) pH 4.5. (c) pH 7.5. (d) pH 7.5 with amantadine bound. Water: blue; protein: orange; lipid: brown; water-protein interface: green; amantadine: cyan.

The number of water cubes and protein-water interface cubes were varied to simulate the experimental buildup curves, which were the integrated intensities from 64 to 16 ppm in the ^{13}C DQ filtered spectra. **Fig. 5a** shows the best-fit buildup curves for the three states. Side views of the structural models used to obtain the best fits are given in **Fig. 5b-d** and the molecular distributions in all the planes across the bilayer are given in **Fig. S5-S7**. The low

pH sample exhibits the fastest buildup and was best fit by a protein-water surface S_{WP} of 18.9 nm^2. Increasing the pH to 7.5 reduced S_{WP} by 25%, to 14.2 nm^2 (**Table S2**). Correspondingly, the minimum pore diameter is 0.6 nm at pH 4.5 and decreases to 0.2 nm at pH 7.5. The requirement of keeping the protein volume constant while reducing the water-exposed area resulted in a tighter helical bundle with thicker cross sections at high pH (**Fig. 5c**), and a thinner and more expanded helical bundle at low pH to accommodate the large pore. This simple change already reproduces certain features of molecular dynamics simulations, such as the significantly reduced water amount in the vicinity of Val_{27} at high pH (20, 21).

Amantadine binding decreased S_{WP} further to 10.0 nm^2, representing a 47% reduction of the water accessibility from the low pH state (**Table S2**). The channel is now devoid of water for about 6 planes or 12 Å along the pore axis (**Fig. S7**). Thus the slow buildup of Gly_{34} and Ile_{35} is the direct result of amantadine-induced dehydration of the channel lumen and the interruption of the water wire between Val_{27} and Gly_{34}. It is worth noting that the spin diffusion experiment detects only *mobile* water sources and filters out the magnetization of potential rigid water molecules. The crystal structure of M2-TM suggests that there may be rigid water molecules near Gly_{34} (15), which would not be detectable by the current technique.

In our simulations we assumed the water reservoir to be infinitely large, which was achieved by keeping the magnetization of each water cube at 100% throughout the spin diffusion process. Alternate simulations that allowed the water magnetization to decrease indicate that a water layer of about 10 nm is necessary to reproduce the infinite-reservoir buildup curves. The actual water amount in our membrane samples corresponded to a water layer thickness of about 4.5 nm for each lipid bilayer. Thus, while this water content is sufficient to fully hydrate the membrane, it departs significantly from the $f_W \approx 1$ approximation in Eq. 6. Thus, the normalized intensity buildup with time is modified to

$$\frac{I_P(t_m)}{I_P(\infty)} \approx \sqrt{\frac{Dt_m}{\pi}} \frac{S'_{WP}}{V_P}, \text{ where } S'_{WP} = \frac{S_{WP}}{f_W} \text{ with } f_W < 1. \quad (1)$$

Therefore, the true protein-water surface areas are larger than the simulated values by a factor of 2-3. In addition, the less complexity of the low-resolution 3D lattice models compared to the true protein structure also reduces S_{WP}. A recent study comparing S_{WP} obtained from spin diffusion data and from the VADAR web server (38) found that the spin diffusion analysis underestimates the water-protein surface by 3-fold (25), consistent with the estimate here. However, these systematic factors do not affect the *relative change* of the protein-water surface area, thus the high-pH induced pore tightening and the drug induced obstruction of the water pathway remain quantitatively valid.

In summary, using ^{1}H spin diffusion we have obtained the first experimental evidence of the pH- and drug-induced change of the water accessibilities of the influenza M2 transmembrane protein. The water-protein surface area decreased by about 25% from pH 4.5 to pH 7.5. This change is much smaller than the chimeric potassium channel KcsA-Kv1.3, which decreased its S_{WP} by ~40% from the open to the closed state (25). Thus, the conformational changes associated with M2 channel activation is more modest, which is consistent with the fact that all key functions of this channel, including the selective filter and gating, are contained within a single TM helix, in contrast to the multi-spanning topology of the larger potassium channels. Moreover, the current spin diffusion data indicate that amantadine binds to the pore lumen rather than the surface of the helical bundle, as suggested by a recent solution NMR study of M2(18-60) (18). The significant slowing down of spin diffusion to Gly_{34} and Ile_{35} can only be attributed to occlusion of the pore between Val_{27} and Gly_{34}. At short mixing times, the experiment is sensitive to differences between lipid-facing residues and pore-lining residues, as evidenced by the slower spin diffusion to Ala_{29} than to Val_{27}. This spin diffusion SSNMR approach is thus promising for investigating water-protein interactions and their changes on both a site-specific level and globally.

Materials and Methods

Membrane protein samples. M2(22-46) of the Udorn strain (SSDPLVVAASIIGILHLIL WILDRL) was synthesized and purified by PrimmBiotech (Cambridge, MA). Several samples containing uniformly ^{13}C, ^{15}N-labeled residues at Leu_{26}, Val_{27}, Ala_{29}, Gly_{34} and Ile_{35} were synthesized. The peptide was reconstituted by detergent dialysis into a lipid mixture

mimicking the virus envelope lipid composition (26, 37). The mixture contains egg sphingomyelin (SPM), DPPC, DPPE and cholesterol at molar ratios of 28% : 21% : 21% : 30%. SPM was dissolved in a chloroform/methanol (5: 1) solution before mixing with the other lipids. The lipid mixture was lyophilized, suspended in a buffer of desired pH, vortexed, and freeze-thawed several times to form large unilamellar vesicles. A phosphate buffer containing 10 mM Na_2HPO_4/NaH_2PO_4, 1 mM EDTA, and 0.1 mM NaN_3 was used for the pH 7.5 samples, and a citrate buffer with 10 mM citric acid/sodium citrate, 1 mM EDTA, and 0.1 mM NaN_3 was used for the pH 4.5 sample. The protein/lipid molar ratio was 1:15. The proteoliposome suspensions were centrifuged at 150,000 g to obtain ~40% hydrated pellets. Photometric assays showed >95% binding of M2-TM to the lipid membrane. For the amantadine-bound sample, amantadine hydrochloride in the pH 7.5 buffer was directly titrated to the pellet to reach a M2: amantadine (Amt) molar ratio of 1 : 2.

Solid-state NMR experiments. NMR experiments were carried out on wide-bore Bruker NMR spectrometers at 14.1 and 9.4 Tesla using 4 mm magic-angle spinning (MAS) probes. Typical radio-frequency field strengths were 50 kHz for ^{13}C and ^{31}P and 60-70 kHz for ^{1}H. ^{13}C and ^{31}P chemical shifts were referenced to the α-Gly ^{13}CO signal at 176.49 ppm on the TMS scale and the hydroxyapatite ^{31}P signal at +2.73 ppm on the phosphoric acid scale, respectively. ^{1}H chemical shifts were internally referenced to the lipid Cγ signal at 3.26 ppm on the TMS scale.

All 1D and 2D spin diffusion experiments with ^{13}C or ^{31}P detection (39) were conducted at 293 K, where water is mobile but the protein is immobilized (26). The 2D ^{1}H-^{13}C and ^{1}H-^{31}P correlation experiments used a ^{1}H T_2 filter of 2 ms and 0.8 ms, respectively, to suppress the ^{1}H magnetization of the rigid components. Spin diffusion mixing times (t_m) were 64 ms for the 2D ^{1}H-^{31}P experiments and 4 to 100 ms for the 2D ^{1}H-^{13}C experiments. ^{13}C double-quantum (DQ) filtered 1D spin diffusion experiments, which removed the lipid ^{13}C signals, used the SPC-5 sequence (40) to create the DQ coherence and a ^{1}H T_2 filter of 2 ms. Most spin diffusion spectra were measured under 5 kHz MAS.

Water-to-protein spin diffusion intensity as a function of the square root of t_m (Eq. 6) was plotted after correcting for water T_1 relaxation by multiplying the intensity with e^{t_m/T_1}.

The water ^{1}H T_1 was measured using the standard inversion recovery sequence. Water ^{1}H T_2 relaxation times were measured using a Hahn-echo experiment and detected through the protein ^{13}C signals.

Theory for determining the water-protein surface area from ^{1}H spin diffusion. The theory for determining the water-protein surface areas from spin diffusion NMR has been well developed for heterogeneous polymers (41). The ^{1}H z-magnetization $M(\vec{r},t_m)$ diffuses spatially according to:

$$\frac{\partial M(\vec{r},t_m)}{\partial t_m} = \vec{\nabla} \bullet \left\{ D(\vec{r}) \vec{\nabla} M(\vec{r},t_m) \right\}, \tag{2}$$

where $D(\vec{r})$ is the spin diffusion coefficient. Initially, the magnetization is homogeneously distributed in water while the protein magnetization is 0 due to the ^{1}H T_2 filter. The integrated protein magnetization at time t_m due to magnetization transfer from water is the product of the water-protein surface area, S_{WP}, and the magnetization density perpendicular to the surface, $M(x_\perp,t_m)$,

$$I_P(t_m) \equiv \int M(r,t_m) d^3r = S_{WP} \cdot \int_0^\infty M(x_\perp,t_m) dx_\perp . \tag{3}$$

For spatially constant diffusivity D, $M(x_\perp,t_m)$ can be solved (41), leading to:

$$I_P(t_m) = S_{WP} \cdot M_0 \sqrt{\frac{Dt_m}{\pi}}, \tag{4}$$

where M_0 is the initial water magnetization per unit volume. The equilibrium protein magnetization $I_P(\infty)$ is proportional to the volume fractions of water (f_W) and protein (f_P) and the total volume of the sample V_{tot}:

$$I_P(\infty) = f_P I_{tot} = f_P M_0 f_W V_{tot} \tag{5}$$

Thus, the normalized spin diffusion buildup is

$$\frac{I_P(t_m)}{I_P(\infty)} \approx \sqrt{\frac{Dt_m}{\pi}} \frac{1}{f_W f_P} \frac{S_{WP}}{V_{tot}}, \tag{6}$$

In the limit of $f_W \approx 1$, the buildup curve is thus proportional to the surface-to-volume ratio of the protein,

$$\frac{I_P(t_m)}{I_P(\infty)} \approx \sqrt{\frac{D}{\pi}} \frac{S_{WP}}{V_P} \sqrt{t_m}\,. \tag{7}$$

Therefore, the spin diffusion intensity buildup as a function of $\sqrt{t_m}$ reports the water-protein surface area S_{WP}. However, proteoliposomes usually do not fulfill the assumption of homogeneous diffusivity, thus the D in Eq. 7 should be replaced by an effective diffusion coefficient, D_{eff}. The time (t_m^s) for the protein to reach $I_P(\infty)$, extracted from the initial slope of the buildup curve, is thus inversely related to S_{WP}:

$$\sqrt{t_m^s} = \sqrt{\frac{\pi}{D_{eff}}} \frac{V_P}{S_{WP}} \tag{8}$$

3D lattice calculation of spin diffusion. To obtain the water accessibilities of M2-TM under the realistic condition of heterogeneous diffusion coefficients, we calculated the ^{1}H spin diffusion buildup curves numerically using a 3D lattice model (25, 41). The lattice is a low-resolution model of an M2-TM tetramer in a 44-Å thick lipid bilayer. Cubes with a 2-Å edge were used to define water, lipids, the protein, the water-protein interface, and amantadine. The time-dependent ^{1}H magnetization at a lattice point, $M_{x,y,z}(t_m)$, was calculated in MATLAB as

$$M_{x,y,z}(t_m + \Delta t_m) = M_{x,y,z}(t_m) + \sum_{i=W,\ P} \frac{D_{ij}\Delta t_m}{d^2}\left[M_i(t_m) - M_{x,y,z}(t_m)\right]. \quad (9)$$

$M_{x,y,z}(t_m)$ exchanges with its six neighbors at a rate determined by the diffusion coefficient D_{ij}. Previous measurements have established a water D_{WW} of 3 nm^2/ms and a protein D_{PP} of 0.3 nm^2/ms (24, 25, 39, 42, 43). The high water D_{WW} reflects the fast spin diffusion and fast physical diffusion of water molecules. The protein D_{PP} was based on measurements of rigid organic polymers (44). For the water-protein interface, we used a low D_{WP} of 0.008 nm^2/ms to reflect the inefficiency of spin diffusion across the interface at the high temperature (303 K) of our experiments. The indirect pathway of water spin diffusion to lipids and then to the protein can be neglected due to the lack of any lipid ^{1}H - protein ^{13}C cross peaks in the 2D ^{13}C-^{1}H spectra within the mixing times used, which indicates extremely slow spin diffusion from the lipid to the protein (25).

The water magnetization of each cube was kept at 1 throughout the simulation, which represents the large water pool limit ($f_W \approx 1$). The protein magnetization was read out in 100 steps from time 0 to 625 ms to obtain the time-dependent intensity buildup curve.

Acknowledgement

This work is supported by NSF grants MCB-0543473 and DBI-0421374, and NIH grant R01 GM088204. The authors thank Prof. Klaus Schmidt-Rohr for helpful discussions and Fanghao Hu for experimental assistance.

References

1. Swartz KJ (2008) Sensing voltage across lipid membranes. *Nature* 456:891-897.
2. Swanson JM, *et al.* (2007) Proton solvation and transport in aqueous and biomolecular systems: insights from computer simulations. *J. Phys. Chem. B* 111:4300-4314.

3. Fu D & Lu M (2007) The structural basis of water permeation and proton exclusion in aquaporins. *Mol. Membr. Biol.* 24:366-374.
4. Brzezinski P & Gennis RB (2008) Cytochrome c oxidase: exciting progress and remaining mysteries. *J. Bioenerg. Biomembr.* 40:521-531.
5. Lanyi JK (2006) Proton transfers in the bacteriorhodopsin photocycle. *Biochim. Biophys. Acta* 1757:1012-1018.
6. Freites JA, Tobias DJ, von Heijne G, & White SH (2005) Interface connections of a transmembrane voltage sensor. *Proc. Natl. Acad. Sci. U.S.A.* 102:15059-15064.
7. Dorairaj S & Allen TW (2007) On the thermodynamic stability of a charged arginine side chain in a transmembrane helix. *Proc. Natl. Acad. Sci. U. S. A.* 104:4943-4948.
8. MacCallum JL, Bennett WF, & Tieleman DP (2008) Distribution of amino acids in a lipid bilayer from computer simulations. *Biophys. J.* 94:3393-3404.
9. Rasaiah JC, Garde S, & Hummer G (2008) Water in nonpolar confinement: from nanotubes to proteins and beyond. *Annu. Rev. Phys. Chem.* 59:713-740.
10. Pinto LH, Holsinger LJ, & Lamb RA (1992) Influenza virus M2 protein has ion channel activity. *Cell* 69:517-528.
11. Pinto LH & Lamb RA (2006) The M2 Proton Channels of Influenza A and B Viruses. *J. Biol. Chem.* 281:8997-9000.
12. Cady SD, Luo WB, Hu F, & Hong M (2009) Structure and function of the influenza M2 proton channel. *Biochemistry* 48:7356-7364.
13. Wang C, Lamb RA, & Pinto LH (1995) Activation of the M2 ion channel of influenza virus: a role for the transmembrane domain histidine residue. *Biophys. J.* 69:1363-1371.
14. Wang C, Takeuchi K, Pinto LH, & Lamb RA (1993) Ion channel activity of influenza A virus M2 protein: characterization of the amantadine block. *J. Virol.* 67:5585-5594.
15. Stouffer AL*, et al.* (2008) Structural basis for the function and inhibition of an influenza virus proton channel. *Nature* 451:596-599.
16. Hu J*, et al.* (2006) Histidines, heart of the hydrogen ion channel from influenza A virus: toward an understanding of conductance and proton selectivity. *Proc. Natl. Acad. Sci. U.S.A.* 103:6865-6870.

17. Luo W, Mani R, & Hong M (2007) Side-chain conformation of the M2 transmembrane peptide proton channel of influenza a virus from 19F solid-state NMR. *J. Phys. Chem.* 111:10825-10832.
18. Schnell JR & Chou JJ (2008) Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* 451:591-595.
19. Smondyrev AM & Voth GA (2002) Molecular dynamics simulation of proton transport through the influenza A virus M2 channel. *Biophys. J.* 83:1987-1996.
20. Yi M, Cross TA, & Zhou HX (2008) A secondary gate as a mechanism for inhibition of the M2 proton channel by amantadine. *J. Phys. Chem. B* 112:7977-7979.
21. Khurana E*, et al.* (2009) Molecular dynamics calculations suggest a conduction mechanism for the M2 proton channel from influenza A virus. *Proc. Natl. Acad. Sci. U. S. A.* 106:1069-1074.
22. Hong M (2006) Oligomeric structure, dynamics, and orientation of membrane proteins from solid-state NMR. *Structure* 14:1731-1740.
23. Hong M (2006) Solid-state NMR studies of the structure, dynamics, and assembly of beta-sheet membrane peptides and alpha-helical membrane proteins with antibiotic activities. *Acc. Chem. Res.* 39:176-183.
24. Kumashiro KK*, et al.* (1998) A novel tool for probing membrane protein structure: solid-state NMR with proton spin diffusion and X-nucleus detection. *J. Am. Chem. Soc.* 120:5043-5051.
25. Ader C*, et al.* (2009) Structural rearrangements of membrane proteins probed by water-edited solid-state NMR spectroscopy. *J. Am. Chem. Soc.* 131:170-176.
26. Luo W, Cady SD, & Hong M (2009) Immobilization of the Influenza A M2 Transmembrane Peptide in Virus-Envelope Mimetic Lipid Membranes: A Solid-State NMR Investigation. *Biochemistry* 48:6361-6368.
27. Cady SD, Goodman C, Tatko CD, DeGrado WF, & Hong M (2007) Determining the orientation of uniaxially rotating membrane proteins using unoriented samples: a 2H, 13C, and 15N solid-state NMR investigation of the dynamics and orientation of a transmembrane helical bundle. *J. Am. Chem. Soc.* 129:5719-5729.

28. Cady SD & Hong M (2009) Effects of amantadine binding on the dynamics of membrane-bound influenza A M2 transmembrane peptide studied by NMR relaxation. *J. Biomol. NMR* 45:185-196.
29. Luo W & Hong M (2006) Determination of the oligomeric number and intermolecular distances of membrane protein assemblies by anisotropic 1H-driven spin diffusion NMR spectroscopy. *J. Am. Chem. Soc.* 128:7242-7251.
30. Cristian L, Lear JD, & DeGrado WF (2003) Use of thiol-disulfide equilibria to measure the energetics of assembly of transmembrane helices in phospholipid bilayers. *Proc. Natl. Acad. Sci. USA* 100:14772-14777.
31. Duong-Ly KC, Nanda V, DeGrado WF, & Howard KP (2005) The conformation of the pore region of the M2 proton channel depends on lipid bilayer environment. *Protein Sci.* 14:856-861.
32. Cady SD & Hong M (2008) Amantadine-Induced Conformational and Dynamical Changes of the Influenza M2 Transmembrane Proton Channel. *Proc. Natl. Acad. Sci. U.S.A* 105:1483-1488.
33. Hu J*, et al.* (2007) Backbone structure of the amantadine-blocked trans-membrane domain M2 proton channel from Influenza A virus. *Biophys. J.* 92:4335-4343.
34. Wang J, Kim S, Kovacs F, & Cross TA (2001) Structure of the the transmembrane region of the M2 protein H+ channel. *Prot. Sci.* 10:2241-2250.
35. Fischer H, Polikarpov I, & Craievich AF (2004) Average protein density is a molecular-weight-dependent function. *Protein Sci.* 13:2825-2828.
36. Quillin ML & Matthews BW (2000) Accurate calculation of the density of proteins. *Acta Crystallogr. D Biol. Crystallogr.* 56:791-794.
37. Cady SD, Mishanina TV, & Hong M (2009) Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding. *J. Mol. Biol.* 385:1127-1141.
38. Willard L*, et al.* (2003) VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res.* 31:3316-3319.

39. Huster D, Yao XL, & Hong M (2002) Membrane Protein Topology Probed by 1H Spin Diffusion from Lipids Using Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* 124:874-883.

40. Hohwy M, Jakobsen HJ, Eden M, Levitt MH, & Nielsen NC (1998) Broadband dipolar recoupling in the nuclear magnetic resonance of rotating solids: a compensated C7 pulse sequence. *J. Chem. Phys.* 108:2686-2694.

41. Schmidt-Rohr K & Spiess HW (1994) *Multidimensional Solid-State NMR and Polymers* (Academic Press, San Diego) p 478.

42. Gallagher GJ, Hong M, & Thompson LK (2004) Solid-state NMR spin diffusion for measurement of membrane-bound peptide structure: gramicidin A. *Biochemistry* 43:7899-7906.

43. Mani R*, et al.* (2006) Membrane-dependent oligomeric structure and pore formation of a beta-hairpin antimicrobial peptide in lipid bilayers from solid-state NMR. *Proc. Natl. Acad. Sci. U.S.A.* 103:16242-16247.

44. Clauss J, Schmidt-Rohr K, & Spiess HW (1993) Determination of domain sizes in heterogeneous polymers by solid-state NMR. *Acta Polym.* 44:1-17.

Supporting Information

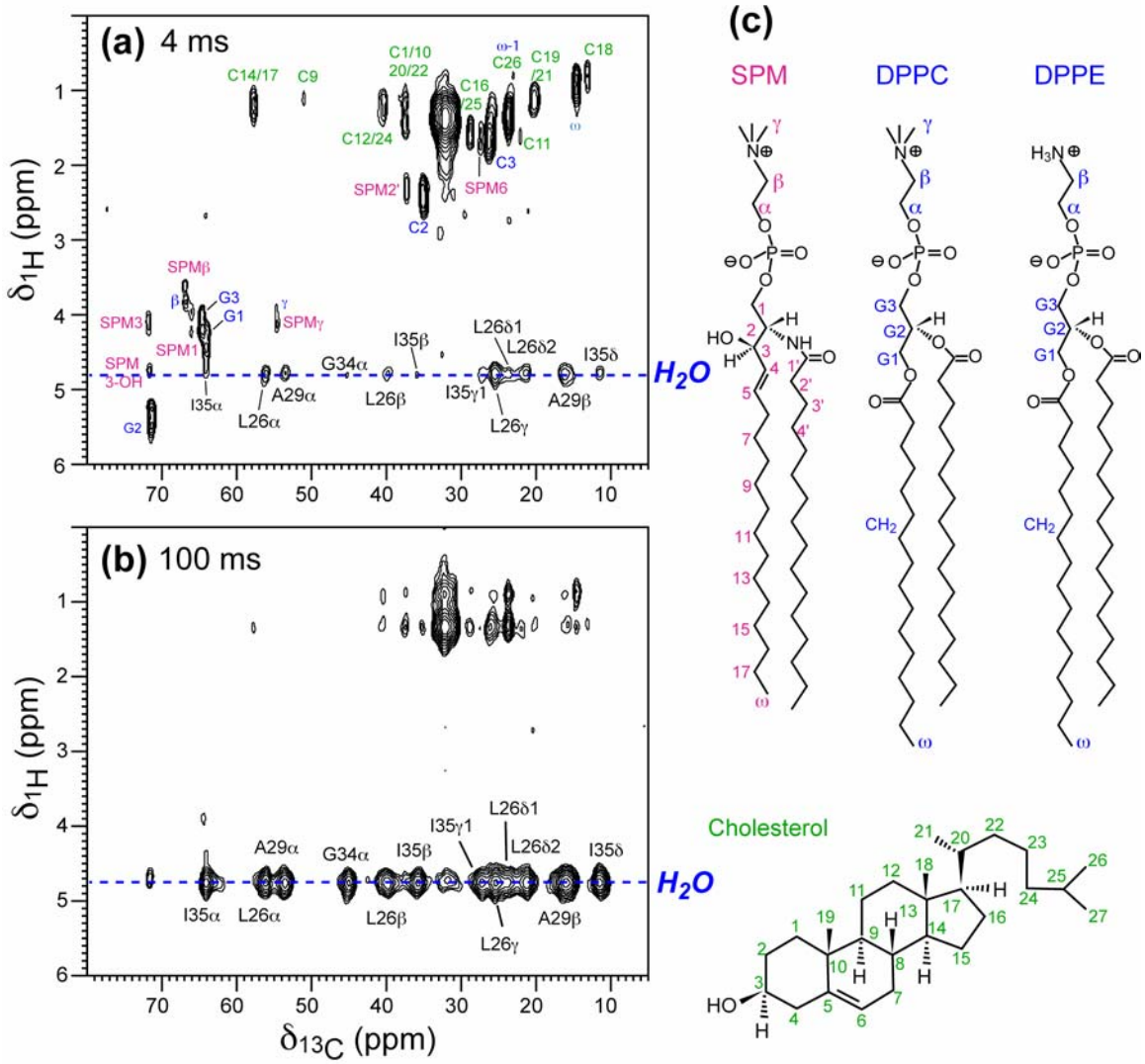


Figure 8.S1. Representative full 2D ^{13}C-^{1}H spin diffusion correlation spectra of M2-TM in viral membranes. The sample is the amantadine-bound M2-TM at pH 7.5. The spectra were measured at 293 K under 5 kHz MAS using a ^{1}H T_2 filter of 2 ms and varying spin diffusion mixing times. (a) 4 ms mixing. (b) 100 ms mixing. Intermolecular water-protein cross peaks are assigned in black. Intramolecular lipid and cholesterol ^{1}H-^{13}C cross peaks are also assigned (cholesterol: green; phospholipids: blue; SPM: magenta).

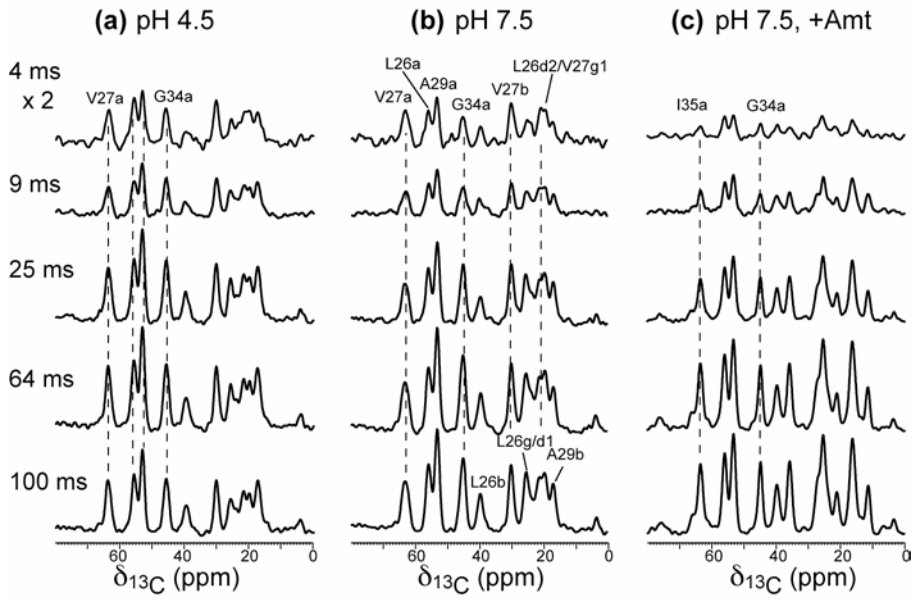


Figure 8.S2. ^{13}C DQ filtered spectra of M2-TM in viral membranes after ^{1}H spin diffusion from water. All spectra were measured using a ^{1}H T_2 filter of 2 ms at 293 K under 5 kHz MAS. Spin diffusion mixing times are indicated on the left. (a) pH 4.5. (b) pH 7.5. (c) pH 7.5 with amantadine. The spectra were plotted to scale within each sample. ^{1}H spin diffusion from water to Gly_{34} is slower than water spin diffusion to N-terminal residues in the amantadine-bound sample, but is comparable in the two amantadine-free samples. Thus, in the absence of drug, both the low and high pH pores contain significant amount of water to contact the pore-facing residues, while amantadine interrupts the water pathway between Ala_{29} and Gly_{34}. In the two apo protein samples, water to Ala_{29} spin diffusion buildup is slightly slower than to Val_{27} and Leu_{26}, consistent with the lipid-facing position of Ala_{29}.

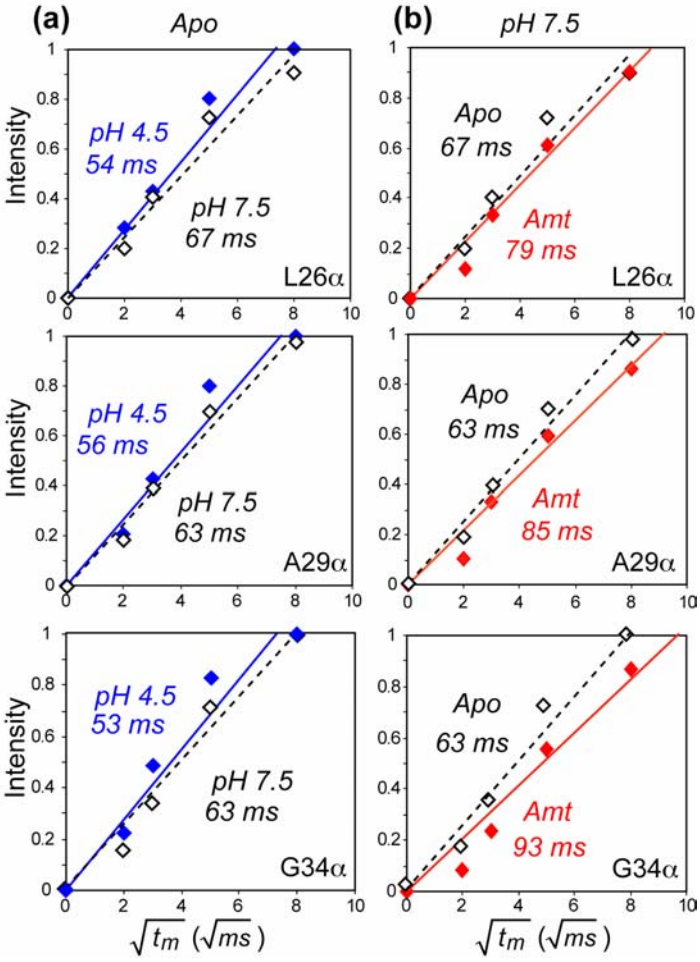


Figure 8.S3. Linear fitting of the initial rates of the water-to-M2 ^{1}H spin diffusion buildup curves to extract t_m^s values. Intensities are obtained from the 1D ^{13}C DQ filtered spectra (**Fig. S2**). (a) Comparison of the pH 4.5 (blue) and pH 7.5 (black) data without amantadine. (b) Comparison of the apo pH 7.5 data (black) and the amantadine-bound pH 7.5 data (red). For all labeled sites, the amantadine-bound protein at pH 7.5 has the longest t_m^s values, indicating that it has the lowest water accessibility, while the apo M2 at pH 4.5 has the shortest t_m^s values, indicating that the open state has the highest water accessibility.

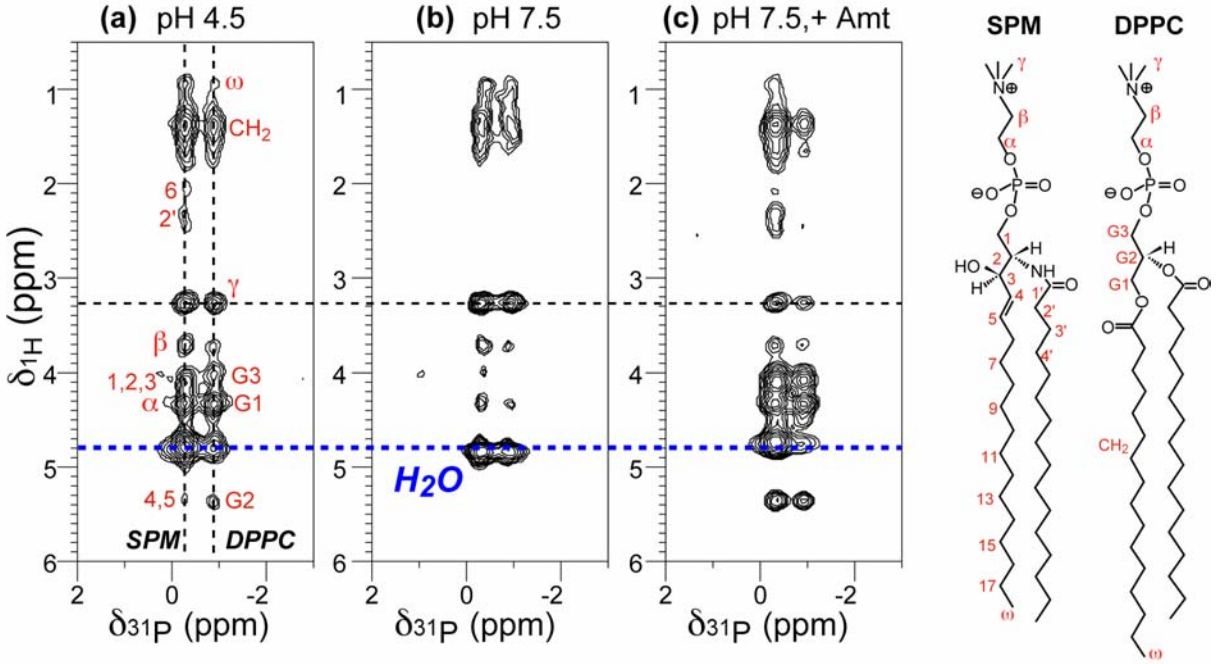


Figure 8.S4. 2D ^{1}H-^{31}P correlation spectra of membrane-bound M2-TM after ^{1}H spin diffusion. The spectra were measured with a ^{1}H T_2 filter of 0.8 ms and a spin diffusion mixing time of 64 ms at 293 K under 7 kHz MAS. (a) pH 4.5. (b) pH 7.5. (c) pH 7.5 with amantadine. Peak assignment is given in (a) along with the chemical structure and nomenclature of sphingomyelin (SPM) and DPPC on the right. The lipid Hγ signal is calibrated to be 3.26 ppm in each spectrum.

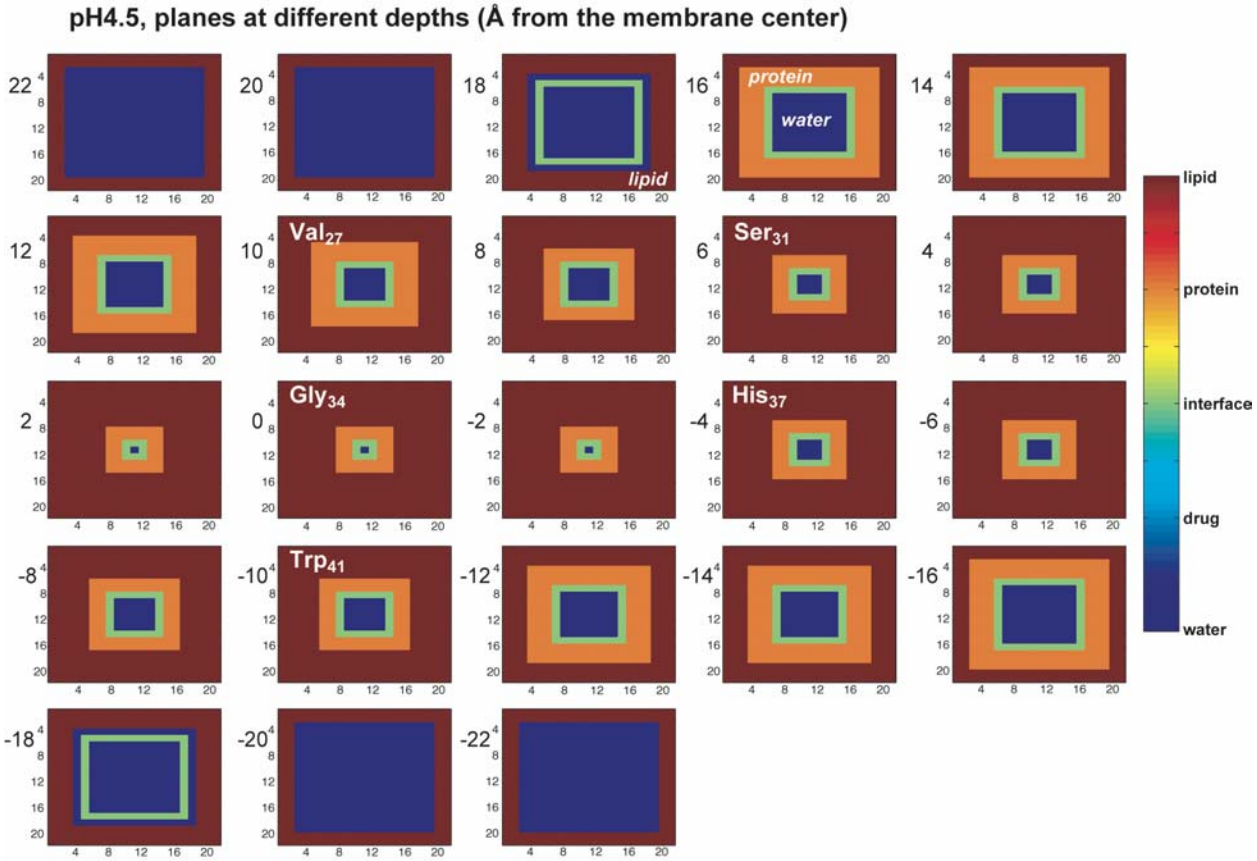


Figure 8.S5. Three-dimensional lattice used in the best-fit simulation of the water-to-protein spin diffusion buildup curve of M2-TM at pH 4.5. The spatial distributions of the water (dark blue), protein (orange), water-protein interface (green), and lipid (brown) cubes are shown for 23 planes spaced at 2 Å intervals along the bilayer normal. Approximate z-positions of key pore-lining residues on the lattice are indicated at appropriate planes.

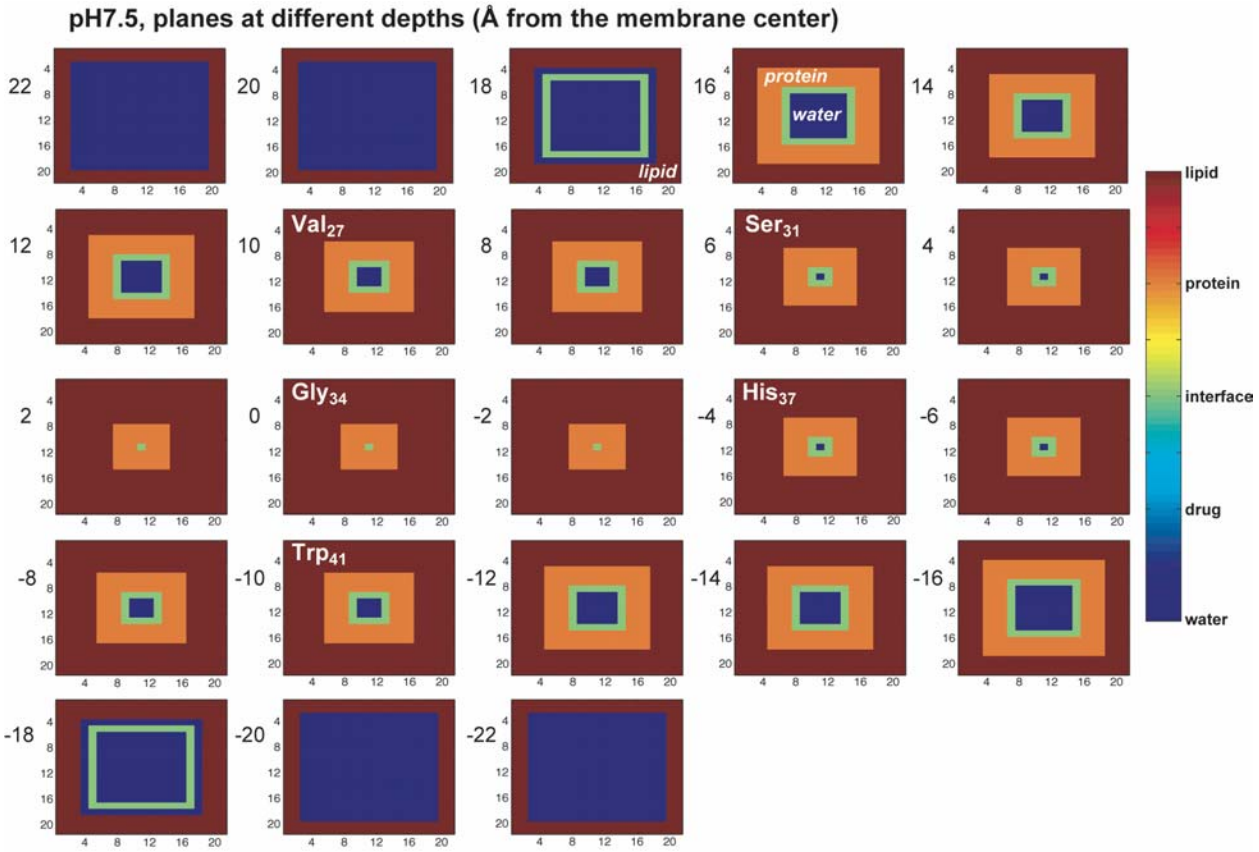


Figure 8.S6. Three-dimensional lattice used in the best-fit simulation of the water-to-protein spin diffusion buildup curve of M2-TM at pH 7.5. The spatial distributions of the water (dark blue), protein (orange), water-protein interface (green), and lipid (brown) cubes are shown for 23 planes spaced at 2 Å intervals along the bilayer normal. Approximate z-positions of key pore-lining residues on the lattice are indicated.

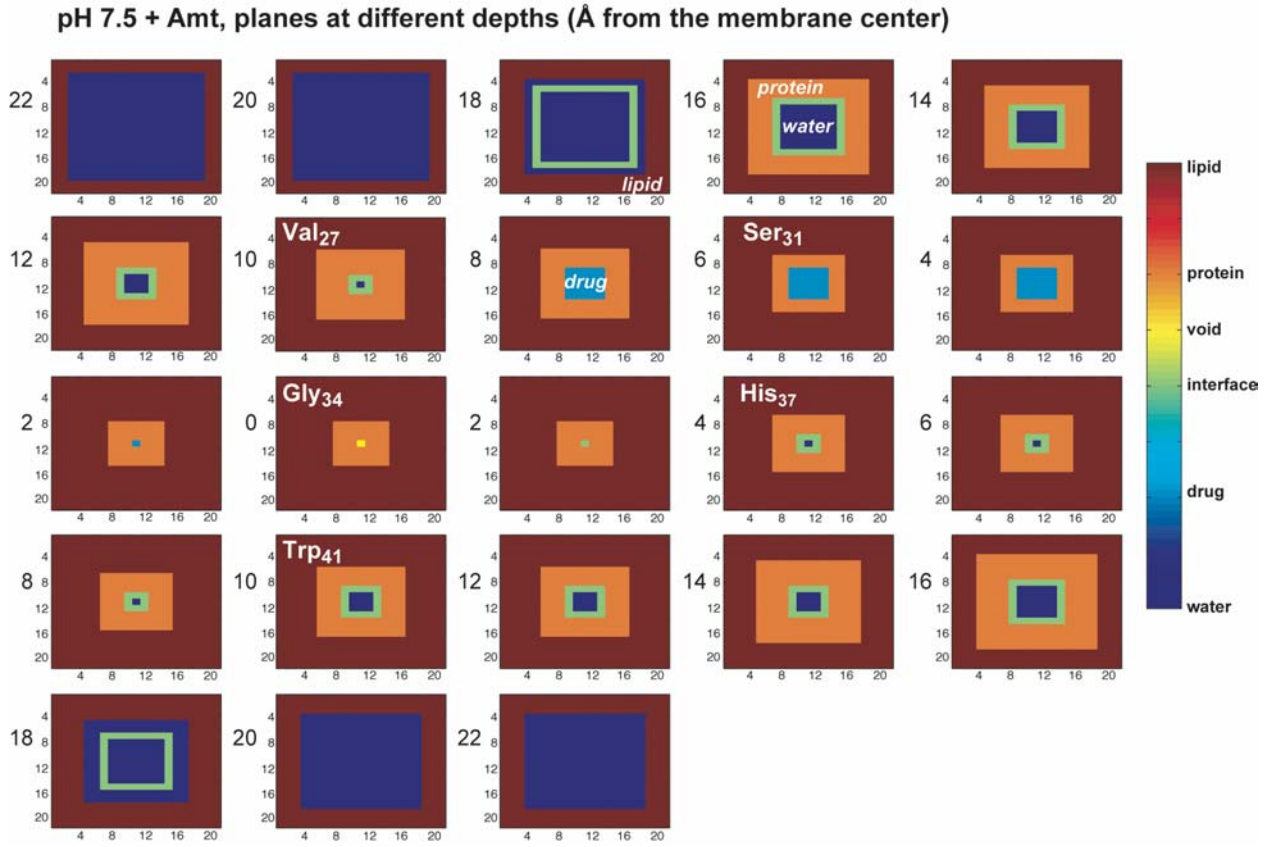


Figure 8.S7. Three-dimensional lattice used in the best-fit simulation of the water-to-protein spin diffusion buildup curve of amantadine-bound M2-TM at pH 7.5. The spatial distributions of water (dark blue), protein (orange), water-protein interface (green), amantadine (cyan), and lipid (brown) cubes are shown for 23 planes spaced at 2 Å intervals along the bilayer normal. Approximate z-positions of key pore-lining residues on the lattice are indicated. Note the exact z-position of the drug is only qualitative and is not precisely determined from the present experiments.

Table 8.S1. t_m^S values (ms) from the initial buildup rates of the ^{13}C DQ filtered ^{1}H spin diffusion spectra of M2-TM under different conditions.

Site	pH 4.5	pH 7.5	pH 7.5 + Amt
L26α	54±5	67±7	79±8
L26β	57±4	69±8	83±8
L26γ	53±7	68±8	82±9
L26 δ1	55±6	68±8	85±9

L26δ2/V27γ1	56±5	66±5	-
V27α	59±6	66±6	-
V27β	57±6	56±6	-
V27γ2	52±6	57±5	-
A29α	56±6	63±5	85±9
A29β	55±6	68±10	84±10
G34α	53±7	63±6	93±12
I35α	-	-	87±11
I35β	-	-	91±13
I35δ	-	-	88±11
Mean	55 ± 6	66 ± 7	86 ± 10

Table 8.S2 Water-accessible surfaces and pore parameters of M2-TM in viral membranes obtained from 3D lattice calculations.

Parameters	pH 4.5	pH 7.5	pH 7.5, Amt
Number of protein cubes	1592	1584	1608
Number of interface cubes	472	355	249
Number of drug cubes	0	0	28
$V_P\ \left(nm^3\right)$	12.7	12.7	12.8
$S_{WP}\ \left(nm^2\right)$	18.9	14.2	10.0
Relative S_{WP}	100%	75%	53%
S_{WP}/V_P	1.48	1.12	0.78
Minimum pore diameter (nm)	0.6	0.2	0

Appendix A

Viral Membrane Sample Preparation Protocols

Viral Membrane Vesicle Preparation

Sphingomyelin (Egg, Chicken) (SM), 1,2-Dipalmitoyl-sn-Glycero-3-Phosphocholine (DPPC), 1,2-Dipalmitoyl-sn-Glycero-3-Phosphoethanolamine (DPPE) and cholesterol should be first checked for dates on the bottle to make sure they are not more than 2 years old. SM is dissolved in chloroform/methanol (10:2) solution to obtain a clear solution first. The SM solubility in this solvent is greater than 400mg/ml as tested in our lab. Then this SM solution is mixed with DPPC, DPPE and cholesterol in chloroform at a molar ratio of SM: DPPC: DPPE=4:3:3 and a total lipid/cholesterol molar ratio of 7:3. This composition was chosen according to the influenza virus membrane content reported in the literature, therefore it is called as viral membrane here. The mixture is dried under a stream of N_2 gas and lyophilized overnight in cyclohexane. Large unilamellar lipid vesicles are formed by dissolving the mixture in 1ml buffer (10 mM Na_2HPO_4/NaH_2PO_4 or Tris, 1 mM EDTA, and 0.1 mM NaN_3) at pH 7.5, vortexing and freeze-thawing the solution six to eight times. Tris buffer is used when the ^{31}P SP lineshape of the membrane sample is closely monitored because it does not contain ^{31}P. Tris is better than HEPES, which can also be used for pH 7.5 samples, because it requires half the mass compared to HEPES. Otherwise the phosphate buffer is widely used. In all those buffers, 1 mM EDTA, and 0.1 mM NaN_3 are always added and pH 7.5 is always maintained. Viral membranes prepared in this way have a phase transition temperature around 263 K as shown by the ^{31}P spectra as a function of temperature.

M2 Peptide Reconstitution

M2TMP powder is codissolved with the detergent octyl-β-D-glucopyranoside (OG) in 1 ml of the same buffer solution as above and vortexed for 15 minutes. The OG concentration is 7 to 10 mg/ml. Then the lipid vesicles are added to the M2/OG solution at a P: L molar ratio of 1: 15 or 1:20 and vortexed for 2 hours at room temperature to facilitate

peptide reconstitution into the viral membrane. The mixture is then centrifuged at 150,000 g for 4 hours at 6˚C and the pellet is collected. Photometric assay of the supernatant using UV-Vis absorbance at 280nm normally shows that 90% of the peptide was reconstituted into the membrane. The pellet is then weighted to check the water content, by comparing with the original empty centrifugation tube weight. If the wet pellet exceeds twice the dry mass of the lipid+cholesterol+M2 mixture, which is almost always the case for viral membranes, we dry the pellet until it is less than twice the mass of our starting dry materials. This usually takes 1-2 hours, but the mass should be checked every half hour for the stopping point. After lyophilization, the pellet becomes non-sticky, and is easy to be transferred to a plastic insert for 4 mm rotor using steel spatula. The maximum mass we can load into the insert is ~40mg. Then put this insert into the rotor and cap it inside the rotor. The sample will be incubated overnight at room temperature and then be ready for the NMR experiments.

Amantadine Binding

If amantadine is to be added, we first use a desktop centrifuge to send the pellet down to the insert bottom, then measure the weight of the insert+rotor+pellet. Then we lyophilize the pellet to reduce 10mg water, with the rotor opening carefully covered by kimwipe tissue and compare with the original weight. This will take about one hour for pellets in an insert but take around 15 minutes for pellets in the rotor. Therefore the pellet mass should be closely checked every half hour for pellet in the insert but every ten minutes for pellet in the rotor. Then we prepare amantadine solution with pH 7.5 buffer. In my previous sample preparation, the molar ratio of M2/amantadine was 1:2. Therefore the mass of amantadine is calculated by (Mol of M2) x2x187x100 and then we dissolve this amount of amantadine in 1 ml buffer. We take 10 microliter of this amantadine solution to add into our semi-try pellet. If more than 10 mg water is lost through the lyophilization, like 15 mg water, we add additional 5 microliter pH 7.5 buffer into the pellet. Then we centrifuge the pellet for 20 minutes using the desktop centrifuge. The sample should be incubated overnight at room temperature, after which it will be ready for NMR experiments.

Appendix B

CODEX Simulation Code

The following program calculates a single exponential decay curve to simulate the CODEX experimental data using MATLAB. The required input parameters are explained next. The output value is the simulation curve and the root mean standard derivation (RMSD) for the simulation with the given F(0) value from the experimental data, as defined in Chapter 5.

dismatrix: the distance matrix which reflects the nearest neighbor distances between the spin pairs with different orientations.

f0: overlap integral in second.

nucleartype: 1 stands for ^{13}C and 2 for ^{19}F.

codexdata: the experimental CODEX data. The input format has two columns: the first column is the mixing time and the second column is the observed S/S_0 value.

tmax: the longest time in ms for the simulation. The default starting time is 0.

incr: the interval between adjacent time points.

fraction: the percentage of the oligomers. The rest component in the mixture is assumed to be monomer.

The simulation for Trp experiment is an example, showing the input and the simulation curve. The input of ω_2 is from the Trp second moment, as calculated in Table A1.

```
---------------------------------------------------------------------------------
function result=rmscodex(dismatrix,f0,nucleartype,codexdata,tmax,inct,spinsys,fraction)

M0=[];Mt=[];M0matrix=[];Mtmatrix=[];Mtavr=[];Mtsum=[];
F0=f0;
uplimit=tmax;
step=inct;
```

```
np=uplimit/step+1;
ncols=3;
nd = length(codexdata );

for i=1:spinsys
M0matrix(i,i)=1;
end

if nucleartype==1
   geroratio=1/4;
end
   if nucleartype==2
      geroratio=376/400;
   end
W=dismatrix.^(-3)*122*1000*geroratio^2;
Wsqu=W.^2;

 for i=1:spinsys
   Wsqu(i,i)=0;
end
for i=1:spinsys
xigma(i)=0;
end
omega=Wsqu;
    for i=1:spinsys
       for j=1:spinsys
          xigma(i)=xigma(i)-Wsqu(i,j);
       end
omega(i,i)=xigma(i);
end
Wsqu=omega;
%calculate the exchange matrix K.
```

Example: ^{19}F-Trp

$$\omega^2 = \begin{bmatrix} 0 & 2601473.98 \\ 2601473.98 & 0 \end{bmatrix} \text{Hz}^2$$

(see **Table 1)**

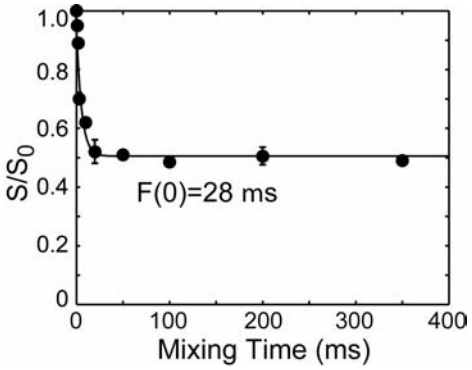
Example: ^{19}F-Trp
f(0)= 28 μs

$$K = \begin{bmatrix} 91.489 & -91.489 \\ -91.489 & 91.489 \end{bmatrix} \text{Hz}$$

```
K=0.5*3.14*Wsqu*0.8*F0*10^(-6);  % ignore (1-3*costheta^2) by replacing it with 0.8.

for count=1:spinsys
   %calculate the dipolar coupling matrix W and teh dipolar coupling square .
    M0=M0matrix(:,count);
for t=0:step:uplimit
   Mt=[Mt expm(t/1000*K)*M0];
end
   nx=(np)*(count-1)+1;
   ny=(np)*count;
   Mttra=Mt(:, nx:ny);
   Mtmatrix=[Mtmatrix; Mttra];
end
Mtsum=0;
% generate the average magnetization matrix: the first row is time points, the
% second row is the averaged magnetization.
   for i=1:spinsys
      pointer=(i-1)*spinsys+i;
for j=1:np
   Mtaver(1,j)=(j-1)*step;
end
Mtsum=Mtsum+Mtmatrix(pointer,:);
end
Mtaver(2,:)=Mtsum/spinsys;
```

Example: ^{19}F-Trp




```
%consider the several aggregate states together, the # spin system has a fraction as inputed.
Mtaver(2,:)=(1-fraction)+fraction*Mtaver(2,:);

%calculating the best fit.
sum=0;
diff=[];
for j=1:np
```

```
   for i=1:nd
      if Mtaver(1,j)==codexdata(i,1)
         diff(i)=(Mtaver(2,j)-codexdata(i,2))^2;
         sum=sum+diff(i);
      end
   end
end
sumsqrt=sqrt(sum/nd);
result=sumsqrt;
plot( Mtaver(1,:), Mtaver(2,:), codexdata(:,1), codexdata(:,2),'o'), axis([0 uplimit 0 1.1])
```

MATLAB Program to Calculate the RMSD for the CODEX Simulation

This program is to plot the RMSD map by visiting the previous program. Different distances and different F(0) values can be explored to find the best fit condition where RMSD is the least.

```
rmin = input('min r in angstrom-->');
rmax=input('max r in angstrom-->');
incr=input('increment r in angstrom-->');

f0min = input('min f0 in microsecond-->');
f0max=input('max f0 in microsecond--->');
incf0=input('increment f0 in microsecond--->');

nxx=(rmax-rmin)/incr+1;
nyy=(f0max-f0min)/incf0+1;

nucleartype=input('nuclear for CODEX expt, 1 for 13C, 2 for 19F -->');
uplimit=input('upper limit of your time point in ms -->');
step=input('increment of your time point in ms-->');
spinsys=input('# of your spin system -->');
```

```
np=uplimit/step+1;
fraction=input('fraction of this # spin system, for example: 0.8, 0.5-->');

%read CODEX expt data
codexfile=input('name of CODEX expt data file in.m,*.dat,*.txt -->', 's');
fcodexin = fopen(codexfile,'r');
if fcodexin < 0
  error(['Could not open ',codexfile,' for input']);
end
codexdata = fscanf(fcodexin,'%f');
ncols=2;
nd = length(codexdata );
nr = nd/ncols;
if nr ~= round(nd/ncols)
  fprintf(1,'\ndata: nrow = %f\tncol = %d\n',nr,ncols);
  fprintf(1,'number of data points = %d does not equal nrow*ncol\n',nd);
  error('data is not rectangular')
end
codexdata = reshape(codexdata,ncols,nr)';
M0=[];Mt=[];M0matrix=[];Mtmatrix=[];Mtavr=[];Mtsum=[]; Mttra=[];diff=[];matrix=[];
dis=rmin:incr:rmax;
f0=f0min:incf0:f0max;
for k=1:1:nyy
F0=f0(k);
for h=1:1:nxx
            leng=dis(h);
            if spinsys==2
             dismatrix=[0 leng
                   leng 0];      % for PG-1
                end
                if spinsys== 4
                      lengt=2^0.5*leng;   % for M2
```

```
                dismatrix=[0 leng lengt leng
                    leng 0  leng lengt
                    lengt leng 0 leng
                    leng  lengt leng 0];
                end
    matrix(h,k)=rmscodex(dismatrix,F0,nucleartype,codexdata,uplimit,step,spinsys,fraction);
 end
end
     matrix=matrix';
x=linspace(rmin,rmax,nxx);
y=linspace(f0min,f0max,nyy);
figure;
clevel=[0.001 0.01 0.03 0.04 0.05 0.07 0.1 0.2 0.3 0.4 0.5];
[c,h]=contourf(x,y,matrix,clevel);
```

Table A1: Distances within 15 Å between orientationally different ^{19}F spins in Trp

Atom1	Atom2	Length (Å)	□ω (Hz)	□ω^2 (Hz^2)	ω^2 sum(Hz^2)
F1U	F2U	14.16	37.99	1443.46	
F1U	F2U	11.16	77.56	6015.14	
F1U	F2U	12.01	62.23	3872.34	
F1U	F2U	8.269	190.66	36350.73	
F1U	F2U	12.01	62.23	3872.34	
F1U	F2U	8.269	190.66	36350.73	
F1U	F2U	14.16	37.99	1443.46	
F1U	F2U	11.16	77.56	6015.14	
F1U	F2U	13.78	41.21	1697.95	
F1U	F2U	13.67	42.21	1781.60	
F1U	F2U	8.804	157.97	24954.61	
F1U	F2U	11.43	72.19	5211.39	
F1U	F2U	4.619	1093.88	1196584.31	
F1U	F2U	11.43	72.19	5211.39	
F1U	F2U	4.619	1093.88	1196584.31	
F1U	F2U	13.67	42.21	1781.60	
F1U	F2U	8.804	157.97	24954.61	
F1U	F2U	13.78	41.21	1697.95	
F1U	F2U	11.89	64.07	4104.53	
F1U	F2U	9.236	136.82	18720.93	
F1U	F2U	9.236	136.82	18720.93	
F1U	F2U	11.89	64.07	4104.53	2601473.98
F2U	F1U	11.89	64.07	4104.53	
F2U	F1U	9.236	136.82	18720.93	
F2U	F1U	9.236	136.82	18720.93	
F2U	F1U	11.89	64.07	4104.53	
F2U	F1U	13.78	41.21	1697.95	
F2U	F1U	8.804	157.97	24954.61	
F2U	F1U	13.67	42.21	1781.60	
F2U	F1U	4.619	1093.88	1196584.31	
F2U	F1U	11.43	72.19	5211.39	
F2U	F1U	4.619	1093.88	1196584.31	
F2U	F1U	11.43	72.19	5211.39	
F2U	F1U	8.804	157.97	24954.61	
F2U	F1U	13.67	42.21	1781.60	
F2U	F1U	13.78	41.21	1697.95	
F2U	F1U	11.16	77.56	6015.14	
F2U	F1U	14.16	37.99	1443.46	
F2U	F1U	8.269	190.66	36350.73	
F2U	F1U	12.01	62.23	3872.34	
F2U	F1U	8.269	190.66	36350.73	
F2U	F1U	12.01	62.23	3872.34	
F2U	F1U	11.16	77.56	6015.14	
F2U	F1U	14.16	37.99	1443.46	2601473.98

Appendix C

3D ^{1}H Spin Diffusion Simulation Code

The following program calculates a single exponential build up curve to simulate the 3D ^{1}H spin diffusion process between water and M2 proton channel using MATLAB.

```
%SD fitting program 3D model with single exponential decay curves.
%M2 proton channel with pH 7.5

%input part
pai=3.1415926;
tiltN=25;  %tilt angle of helix N terminus in degree
poresize=1.4; %helix bundle diameter at G34 (nm)
step=0.2;   %site (proton) spacing a (nm) (ca. 0.2 nm)
tiltC=25;  %tilt angle of helix C terminus in degree
chanlength=3.2; %M2 channel vertical length (nm)
memthick=4.4;  %viral membrane thickness (nm)
ddrug=0;      %vertical length from G34 without water (nm)
dif=0.2;      %thickness of interface (nm)
dde=0.6;      %thickness detection region: helix thickness (nm) (ca. 0.2nm)

%------------------------define diffusion coefficient--------------------------
Diffso=3;     %D source (nm*nm/ms)
Diffdrug=0.008; %D drug (nm*nm/ms)
Diffif=0.008; %D interface (nm*nm/ms)
Diffde=0.3;   %D detection (nm*nm/ms)
tmanf=0;      %tm initial (ms)
tmend=625;    %tm final (ms)
ntm=100;
```

```
ndporewidth=round((poresize/step-1)/2);  %modified pore size. For 8A, nb=1 therefore round to 6A width.
nddrug=round(ddrug/step);
ndif=round(dif/step);
ndcl=round(chanlength/step);
ndmt=round(memthick/step);
ndmthalf=ndmt/2;
ndde=round(dde/step);

dpw=ndporewidth*step;      %modified pore size
ddrug=nddrug*step;                  %modified drug thickness
dif=ndif*step;
dclhalf=ndcl/2;
dmthalf=ndmt/2;
dde=ndde*step;

matrix=[];
i=0;
for xx=-10:10
for yy=-10:10
for zz=-ndmthalf:ndmthalf
   i=i+1;
matrix(i,1)=xx;
matrix(i,2)=yy;
matrix(i,3)=zz;
matrix(i,4)=5;   %1 for source,  2 for drug, 3 for interface, 4 for detection, 5 for lipid
matrix(i,5:10)=0;
end
end
end

nsumdet=0;
```

```
nsumif=0;
nsumdrug=0;
nsumso=0;

%-----------------------define various regions---------------------------------
for z=0: dclhalf
b=ndporewidth+round(z*sin(tiltN/180*pai)/cos(tiltN/180*pai));
for h=1:ndde
for x=-b:b
   for y=-b:b
      if or(abs(x)==b, abs(y)==b)
      for r=1:i
         if matrix(r,1)==x & matrix(r,2)==y & matrix(r,3)==z
            matrix(r,4)=4;
            nsumdet=nsumdet+1;
         end
      end
      end
end
end
b=b-1;
end

for x=-b:b
   for y=-b:b
      if or(abs(x)==b, abs(y)==b)
         for r=1:i
         if matrix(r,1)==x & matrix(r,2)==y & matrix(r,3)==z
            matrix(r,4)=3;
            nsumif=nsumif+1;
         end
         end
```

```
end
end
end

if b>0
  b=b-1;
if abs(z)<=nddrug
   for kk=0:b
      for x=-kk:kk
         for y=-kk:kk
          if or(abs(x)==kk, abs(y)==kk)
         for r=1:i
         if matrix(r,1)==x & matrix(r,2)==y & matrix(r,3)==z
            matrix(r,4)=2;
            nsumdrug=nsumdrug+1;
         end
         end
          end
         end
      end
   end
end
if abs(z)>nddrug
   for kk=0:b
      for x=-kk:kk
         for y=-kk:kk
          if or(abs(x)==kk, abs(y)==kk)
         for r=1:i
         if matrix(r,1)==x & matrix(r,2)==y & matrix(r,3)==z
            matrix(r,4)=1;
            nsumso=nsumso+1;
         end
```

```
        end
         end
        end
      end
   end
end
end
end

for z=dclhalf+1: dmthalf
   b=ndporewidth+round(z*sin(tiltN/180*pai)/cos(tiltN/180*pai));
        for r=1:i
        if  abs(matrix(r,1))<=b & abs(matrix(r,2))<=b & matrix(r,3)==z
           matrix(r,4)=1;
           nsumso=nsumso+1;
        end
      end
end

for z=0:-1:-dclhalf
b=ndporewidth+round(abs(z)*sin(tiltN/180*pai)/cos(tiltN/180*pai));
for h=1:ndde
for x=-b:b
   for y=-b:b
      if or(abs(x)==b, abs(y)==b)
      for r=1:i
        if matrix(r,1)==x & matrix(r,2)==y & matrix(r,3)==z
           matrix(r,4)=4;
           nsumdet=nsumdet+1;
        end
      end
      end
```

```
end
end
b=b-1;
end

for x=-b:b
   for y=-b:b
      if or(abs(x)==b, abs(y)==b)
         for r=1:i
         if matrix(r,1)==x & matrix(r,2)==y & matrix(r,3)==z
            matrix(r,4)=3;
            nsumif=nsumif+1;
         end
         end
end
end
end

b=b-1;
if b>=0
   for kk=0:b
      for x=-kk:kk
         for y=-kk:kk
          if or(abs(x)==kk, abs(y)==kk)
         for r=1:i
         if matrix(r,1)==x & matrix(r,2)==y & matrix(r,3)==z
            matrix(r,4)=1;
            nsumso=nsumso+1;
         end
         end
          end
         end
```

```
        end
end
end

end

for z=-dclhalf-1: -1:-dmthalf
   b=ndporewidth+round(abs(z)*sin(tiltN/180*pai)/cos(tiltN/180*pai));
         for r=1:i
         if abs(matrix(r,1))<=b & abs(matrix(r,2))<=b & matrix(r,3)==z
            matrix(r,4)=1;
            nsumso=nsumso+1;
         end
      end
end

[m,n] = size(matrix);
ratmatrix= sortrows(matrix,[3 1 2]);

for r=1:i
   b=6;
   if ratmatrix(r,3)==dclhalf+1
      for q=-b:b
      if ratmatrix(r,1)==q & abs(ratmatrix(r, 2))==b
         ratmatrix(r,4)=3;
      end
      if ratmatrix(r,2)==q & abs(ratmatrix(r, 1))==b
         ratmatrix(r,4)=3;
      end
      end
   end
      if ratmatrix(r,3)==-dclhalf-1
```

```
      for q=-b:b
      if ratmatrix(r,1)==q & abs(ratmatrix(r, 2))==b
         ratmatrix(r,4)=3;
      end
      if ratmatrix(r,2)==q & abs(ratmatrix(r, 1))==b
         ratmatrix(r,4)=3;
      end
      end
   end
end

ndrep=nsumdet+nsumif+nsumdrug+nsumso;
nxbc=nsumif+nsumdrug;

intede=1;  % integrate source (0) or non-source(=de) compon. (1)or source region (highly mobile) (2)
if and(intede~=0, intede~=2)
   intede=1;
end

ndomd2=1;                    %number of repeat domains
nxmin=((2*ndomd2+1)^2)*ndrep;       %count together 9 units
nx=nxmin;                    %total No. of grid lines for all repeats
nstep=1;

rat=0.1;      % max. transition rate per step (ca. 0.1, <0.5)
Diffmax=max(Diffde,Diffif);
Diffmax=max(Diffso,Diffmax);  %max D
deltat=rat/Diffmax*step*step;    %time steps (us),deltat*1000
imax=round(tmend/deltat);  %maximum number of steps
diffway=1;
```

```
%-----------------------relate diffusion coefficient to diffusion rates----------------------------
if Diffde==Diffmax
ratde=rat;
ratso=ratde*Diffso/Diffde;
ratdrug=ratde*Diffdrug/Diffde;
ratif=ratde*Diffif/Diffde;
end

if Diffso==Diffmax
ratso=rat;
ratde=ratso*Diffde/Diffso;
ratif=ratso*Diffif/Diffso;
ratdrug=ratso*Diffdrug/Diffso;
end

if Diffif==Diffmax
ratif=rat;
ratde=ratif*Diffde/Diffif;
ratdrug=ratif*Diffdrug/Diffif;
ratso=ratif*Diffso/Diffif;
end

if Diffdrug==Diffmax
ratdrug=rat;
ratde=ratdrug*Diffde/Diffdrug;
ratif=ratdrug*Diffif/Diffdrug;
ratso=ratdrug*Diffso/Diffdrug;
end
% assign rates to grid points

%first, paint everything with ratde
% assingh rates
```

```
for row=1:i
if ratmatrix(row,1)==10 & row>420 & ratmatrix(row,3)==ratmatrix(row-420,3)
    rowxp=row-420;
  else
    if row<(i-21) & ratmatrix(row,3)==ratmatrix(row+21,3)
    rowxp=row+21;
    end
  end

  if ratmatrix(row,1)==-10 & row<(i-420) & ratmatrix(row,3)==ratmatrix(row+420,3)
    rowxm=row+420;
  else
    if ratmatrix(row,3)==ratmatrix(row-21,3)
    rowxm=row-21;
    end
  end

  if ratmatrix(row,2)==10 & ratmatrix(row,3)==ratmatrix(row-20,3) &
ratmatrix(row,1)==ratmatrix(row-20,1)
    rowyp=row-20;
  else
    if ratmatrix(row,3)==ratmatrix(row+1,3)& ratmatrix(row,1)==ratmatrix(row+1,1)
    rowyp=row+1;
    end
  end

  if ratmatrix(row,2)==-10 & ratmatrix(row,3)==ratmatrix(row+20,3) &
ratmatrix(row,1)==ratmatrix(row+20,1)
    rowym=row+20;
  else
    if ratmatrix(row,3)==ratmatrix(row-1,3)& ratmatrix(row,1)==ratmatrix(row-1,1)
    rowym=row-1;
```

```
    end
 end

if ratmatrix(row,3)==(((i/441)-1)/2)
   rowzp=row-i+441;
else
   rowzp=row+441;
end

if ratmatrix(row,3)==-(((i/441)-1)/2)
   rowzm=row+i-441;
else
   rowzm=row-441;
end

        if ratmatrix(row, 4)==1              %assigh rates in source
        if ratmatrix(rowxp,4)==1
           ratmatrix(row, 5)=ratso;
           ratmatrix(rowxp, 6)=ratso;
        end
        if ratmatrix(rowxp,4)==2
           ratmatrix(row, 5)=ratdrug;
           ratmatrix(rowxp, 6)=ratso;
        end
         if ratmatrix(rowxp,4)==3
           ratmatrix(row, 5)=ratif;
           ratmatrix(rowxp, 6)=ratso;
         end

         if ratmatrix(rowxm,4)==1
           ratmatrix(row, 6)=ratso;
           ratmatrix(rowxm, 5)=ratso;
```

```
end
if ratmatrix(rowxm,4)==2
   ratmatrix(row, 6)=ratdrug;
   ratmatrix(rowxm, 5)=ratso;
end
 if ratmatrix(rowxm,4)==3
   ratmatrix(row, 6)=ratif;
   ratmatrix(rowxm, 5)=ratso;
 end

if ratmatrix(rowyp,4)==1
   ratmatrix(row, 7)=ratso;
   ratmatrix(rowyp, 8)=ratso;
end
if ratmatrix(rowyp,4)==2
   ratmatrix(row, 7)=ratdrug;
   ratmatrix(rowyp, 8)=ratso;
end
 if ratmatrix(rowyp,4)==3
   ratmatrix(row, 7)=ratif;
   ratmatrix(rowyp, 8)=ratso;
 end

if ratmatrix(rowym,4)==1
   ratmatrix(row, 8)=ratso;
   ratmatrix(rowym, 7)=ratso;
end
if ratmatrix(rowym,4)==2
   ratmatrix(row, 8)=ratdrug;
   ratmatrix(rowym, 7)=ratso;
end
 if ratmatrix(rowym,4)==3
```

```
    ratmatrix(row, 8)=ratif;
    ratmatrix(rowym, 7)=ratso;
 end

if ratmatrix(rowzp,4)==1
    ratmatrix(row, 9)=ratso;
    ratmatrix(rowzp, 10)=ratso;
end
if ratmatrix(rowzp,4)==2
    ratmatrix(row, 9)=ratdrug;
    ratmatrix(rowzp, 10)=ratso;
end
 if ratmatrix(rowzp,4)==3
    ratmatrix(row, 9)=ratif;
    ratmatrix(rowzp, 10)=ratso;
 end

if ratmatrix(rowzm,4)==1
    ratmatrix(row, 10)=ratso;
    ratmatrix(rowzm, 9)=ratso;
end
if ratmatrix(rowzm,4)==2
    ratmatrix(row, 10)=ratdrug;
    ratmatrix(rowzm, 9)=ratso;
end
 if ratmatrix(rowzm,4)==3
    ratmatrix(row, 10)=ratif;
    ratmatrix(rowzm, 9)=ratso;
 end
end  %sol

if ratmatrix(row, 4)==3          %assigh rates in interface
```

```
if ratmatrix(rowxp,4)==3
   ratmatrix(row, 5)=ratif;
   ratmatrix(rowxp, 6)=ratif;
end
if ratmatrix(rowxp,4)==2
   ratmatrix(row, 5)=ratdrug;
   ratmatrix(rowxp, 6)=ratif;
end
 if ratmatrix(rowxp,4)==1
   ratmatrix(row, 5)=ratso;
   ratmatrix(rowxp, 6)=ratif;
end
 if ratmatrix(rowxp,4)==4
   ratmatrix(row, 5)=ratde;
   ratmatrix(rowxp, 6)=ratif;
 end

if ratmatrix(rowxm,4)==3
   ratmatrix(row, 6)=ratif;
   ratmatrix(rowxm, 5)=ratif;
end
if ratmatrix(rowxm,4)==2
   ratmatrix(row, 6)=ratdrug;
   ratmatrix(rowxm, 5)=ratif;
end
 if ratmatrix(rowxm,4)==1
   ratmatrix(row, 6)=ratso;
   ratmatrix(rowxm, 5)=ratif;
end
 if ratmatrix(rowxm,4)==4
   ratmatrix(row, 6)=ratde;
   ratmatrix(rowxm, 5)=ratif;
```

```
 end

if ratmatrix(rowyp,4)==3
    ratmatrix(row, 7)=ratif;
    ratmatrix(rowyp, 8)=ratif;
end
if ratmatrix(rowyp,4)==2
    ratmatrix(row, 7)=ratdrug;
    ratmatrix(rowyp, 8)=ratif;
end
 if ratmatrix(rowyp,4)==1
    ratmatrix(row, 7)=ratso;
    ratmatrix(rowyp, 8)=ratif;
end
 if ratmatrix(rowyp,4)==4
    ratmatrix(row, 7)=ratde;
    ratmatrix(rowyp, 8)=ratif;
 end

 if ratmatrix(rowym,4)==3
    ratmatrix(row, 8)=ratif;
    ratmatrix(rowym, 7)=ratif;
 end
if ratmatrix(rowym,4)==2
    ratmatrix(row, 8)=ratdrug;
    ratmatrix(rowym, 7)=ratif;
 end
 if ratmatrix(rowym,4)==1
    ratmatrix(row, 8)=ratso;
    ratmatrix(rowym, 7)=ratif;
 end
 if ratmatrix(rowym,4)==4
```

```
    ratmatrix(row, 8)=ratde;
    ratmatrix(rowym, 7)=ratif;
 end

if ratmatrix(rowzp,4)==3
    ratmatrix(row, 9)=ratif;
    ratmatrix(rowzp, 10)=ratif;
end
if ratmatrix(rowzp,4)==2
    ratmatrix(row, 9)=ratdrug;
    ratmatrix(rowzp, 10)=ratif;
end
 if ratmatrix(rowzp,4)==1
    ratmatrix(row, 9)=ratso;
    ratmatrix(rowzp, 10)=ratif;
end
 if ratmatrix(rowzp,4)==4
    ratmatrix(row, 9)=ratde;
    ratmatrix(rowzp, 10)=ratif;
 end

if ratmatrix(rowzm,4)==3
    ratmatrix(row, 10)=ratif;
    ratmatrix(rowzm, 9)=ratif;
end
if ratmatrix(rowzm,4)==2
    ratmatrix(row, 10)=ratdrug;
    ratmatrix(rowzm, 9)=ratif;
end
 if ratmatrix(rowzm,4)==1
    ratmatrix(row, 10)=ratso;
    ratmatrix(rowzm, 9)=ratif;
```

```
end
 if ratmatrix(rowzm,4)==4
    ratmatrix(row, 10)=ratde;
    ratmatrix(rowzm, 9)=ratif;
 end
end  %interface

if ratmatrix(row, 4)==4            %assigh rates in detection
if ratmatrix(rowxp,4)==4
    ratmatrix(row, 5)=ratde;
    ratmatrix(rowxp, 6)=ratde;
 if ratmatrix(rowxp,4)==3
    ratmatrix(row, 5)=ratif;
    ratmatrix(rowxp, 6)=ratde;
 end
end
 if ratmatrix(rowxm,4)==4
    ratmatrix(row, 6)=ratde;
    ratmatrix(rowxm, 5)=ratde;
 if ratmatrix(rowxm,4)==3
    ratmatrix(row, 6)=ratif;
    ratmatrix(rowxm, 5)=ratde;
 end
 end
if ratmatrix(rowyp,4)==4
    ratmatrix(row, 7)=ratde;
    ratmatrix(rowyp, 8)=ratde;
 if ratmatrix(rowyp,4)==3
    ratmatrix(row, 7)=ratif;
    ratmatrix(rowyp, 8)=ratde;
 end
end
```

```
            if ratmatrix(rowym,4)==4
               ratmatrix(row, 8)=ratde;
               ratmatrix(rowym, 7)=ratde;
             if ratmatrix(rowym,4)==3
               ratmatrix(row, 8)=ratif;
               ratmatrix(rowym, 7)=ratde;
             end
            end
            if ratmatrix(rowzp,4)==4
               ratmatrix(row, 9)=ratde;
               ratmatrix(rowzp, 10)=ratde;
             if ratmatrix(rowzp,4)==3
               ratmatrix(row, 9)=ratif;
               ratmatrix(rowzp, 10)=ratde;
             end
            end
            if ratmatrix(rowzm,4)==4
               ratmatrix(row, 10)=ratde;
               ratmatrix(rowzm, 9)=ratde;
             if ratmatrix(rowzm,4)==3
               ratmatrix(row, 10)=ratif;
               ratmatrix(rowzm, 9)=ratde;
             end
            end
            end  %detection
end

%initialize distribution of magnetization at t=0
for row=1:i
   ratmatrix(row, 11)=0;
   ratmatrix(row, 12)=0;
   if ratmatrix(row, 4)==1
```

```
        ratmatrix(row,11)=1;
        ratmatrix(row,12)=1;
    end
end

%.....................run discretized diffusion---------------------------------
ijmax=0;
for j=1:ntm
ijmaxold=ijmax;
tm=tmanf+(j-1)^2*(tmend-tmanf)/(ntm-1)^2;
ijmax=round(tm/deltat);
if ge(diffway, 0.9)
ijmaxdelta=ijmax-ijmaxold;
for iii=1:ijmaxdelta
for row=1:i         %ndrep-1 points
    if ratmatrix(row,1)==10 & row>420 & ratmatrix(row,3)==ratmatrix(row-420,3)
        rowxp=row-420;
    else
        if row<(i-21) & ratmatrix(row,3)==ratmatrix(row+21,3)
        rowxp=row+21;
        end
    end

    if ratmatrix(row,1)==-10 & row<(i-420) & ratmatrix(row,3)==ratmatrix(row+420,3)
        rowxm=row+420;
    else
        if ratmatrix(row,3)==ratmatrix(row-21,3)
        rowxm=row-21;
        end
    end
```

```
  if ratmatrix(row,2)==10 & ratmatrix(row,3)==ratmatrix(row-20,3) &
ratmatrix(row,1)==ratmatrix(row-20,1)
     rowyp=row-20;
  else
     if ratmatrix(row,3)==ratmatrix(row+1,3)& ratmatrix(row,1)==ratmatrix(row+1,1)
     rowyp=row+1;
     end
  end

  if ratmatrix(row,2)==-10 & ratmatrix(row,3)==ratmatrix(row+20,3) &
ratmatrix(row,1)==ratmatrix(row+20,1)
     rowym=row+20;
  else
     if ratmatrix(row,3)==ratmatrix(row-1,3)& ratmatrix(row,1)==ratmatrix(row-1,1)
     rowym=row-1;
     end
  end

  if ratmatrix(row,3)==(((i/441)-1)/2)
    rowzp=row-i+441;
  else
    rowzp=row+441;
  end

  if ratmatrix(row,3)==-(((i/441)-1)/2)
    rowzm=row+i-441;
  else
    rowzm=row-441;
  end
ratmatrix(row,11)=ratmatrix(row,12)*(1-ratmatrix(row,5)-ratmatrix(row,6)-ratmatrix(row,7)-
ratmatrix(row,8)-ratmatrix(row,9)-
ratmatrix(row,10))+ratmatrix(rowxp,6)*ratmatrix(rowxp,12)+ratmatrix(rowxm,5)*ratmatrix(
```

```
rowxm,12)+ratmatrix(rowyp,8)*ratmatrix(rowyp,12)+ratmatrix(rowym,7)*ratmatrix(rowym,
12)+ratmatrix(rowzp,10)*ratmatrix(rowzp,12)+ratmatrix(rowzm,9)*ratmatrix(rowzm,12);
end
for row=1:i
   ratmatrix(row,12)=ratmatrix(row,11);
end
end
end

%--------------Integration over detection region--------------------
aMagncr(j)=0;
aMagnso(j)=0;
if intede==1
   ifcount=1;
for row=1:i
   %if ratmatrix(row,4)==3
   %aMagncr(j)=aMagncr(j)+ratmatrix(row,11)*ifcount/nsumif;
   %ifcount=ifcount+1;
   %end

   if ratmatrix(row,4)==4
   aMagncr(j)=aMagncr(j)+ratmatrix(row,11);
   end
end
end

if intede==0
   ifcount=1;
for row=1:i
   if ratmatrix(row,4)==3
   aMagnso(j)=aMagnso(j)+ratmatrix(row,11)*ifcount/nsumif;
   ifcount=ifcount+1;
```

```
    end
    if ratmatrix(row,4)==1
    aMagnso(j)=aMagnso(j)+ratmatrix(row,11);
    end
end
end
    aMagncr(j)=aMagncr(j)/ndrep*100;
    aMagnso(j)=aMagnso(j)/ndrep*100;

if intede==1
Mag(1,j)=tm^0.5;
Mag(2,j)=aMagncr(j);
else
Mag(1,j)=tm^0.5;
Mag(2,j)=aMagnso(j);
end
end                %j
Mag(2,1)=0;
Mag(2,:)=Mag(2,:)/Mag(2,100);
plot(Mag(1,:), Mag(2,:),'k');
```