# Application of molecular modeling techniques to study the structure, dynamics, and interactions of membrane proteins

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

**Lei Shi**

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
**Doctor of Philosophy**

**Gianluigi Veglia**

**August, 2011**

# Acknowledgements

I am grateful to my advisor Gianluigi Veglia for his guidance. His enthusiasm about science has inspired me and shaped my career as a scientist. Without his passion and determination, accomplishment in this thesis will not be possible. I am thankful to Prof. Jiali Gao and his lab members, especially Dr. Alessandro Cembran, for the productive collaboration. They have been another invaluable source of knowledge and inspiration. I also benefit from the collaboration with Prof. David Thomas and his group. I am grateful to work with fantastic colleagues in Veglia's lab, with whom I have built strong friendship. Specially I want to thank Dr. Larry Masterson and Kim Ha for the continuous support. Other members who I am indebted to are Prof. Nate Traaseth, Martin Gustavsson, Dr. Raffaello Verardi, Dr. Gopi Tata, Dr. Jarrod Buffy, Fa-An Chao, Kaustubh Mote. I have benefited in some way from other previous and current member in the group. I am grateful for the opportunity to work with them. Without the great resource provided by University of Minnesota, Department of Chemistry and Minnesota Supercomputing Institute, work in the thesis will not be possible. I am also grateful for the great friends in Minnesota who have made graduate school a memorable and enjoyable journey. Last, I want to specially acknowledge my wife Chu-Ting Chung's for her unconditional love and support. There are several others who have contributed to the work described here. These are Emily Metcalfe, Naomi Wash, Dr. Dan Mullen, Prof. George Barany, and Prof. Susan Taylor. Many others also contributed by providing helpful discussions and suggestions.

# Dedication

To my parents Jinlan Hua and Qingyuan Shi and my wife Chu-Ting Chung

## Abstract

Membrane proteins constitute ∼30% of all the genomes and ∼70% of the drug targets. However, less than 1% of the entries in the protein data bank are membrane proteins. The underrepresentation of membrane protein structures limits our understanding of their functions. This thesis summarizes my effects to apply theoretical methods to understand the structure and function relationships of membrane proteins. Specifically, we developed computational techniques to interpret solution and solid-state NMR data of membrane proteins and determine their high resolution structures. We further performed molecular dynamics simulations to study their dynamics, interaction with other proteins and the lipid bilayer environment. We applied these approaches to phospholamban, which is a membrane protein that is involved in cardiac muscle relaxation by regulating $Ca^{2+}$-ATPase activity. Our results provide new insights to understand how membrane proteins elicit their function.

# Contents

# List of Tables

# List of Figures

# List of Abbreviations

Care has been taken in this thesis to minimize the use of jargon and acronyms, but this cannot always be achieved. This appendix contains a table of acronyms and their meaning.

Table 1: Acronyms

| Acronym | Meaning |
|---|---|
| AKAP | A Kinase Anchoring Protein |
| ATP | Adenosine-5'-triphosphate |
| ADP | Adenosine-5'-diphosphate |
| CD | Circular Dichroism |
| COM | Center of Mass |
| CP | Cross Polarization |
| CPMG | Carr-Purcell-Meiboom-Gill |
| CSA | Chemical Shift Anisotropy |
| CMC | Critical Micelle Concentration |
| DARR | Dipolar Assisted Rotational Resonance |
| DB | Database |
| DC | Dipolar Coupling |
| DCM | Dilated Cardiomyopathy |
| DOPC | 1,2-Dioleoyl-sn-glycero-3-phosphocholine |
| DOPE | 1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine |
| DPC | dodecylphosphocholine |
| DHPC | dihexanoylphosphatidylcholine |
| GALA | Geometric Analysis of Labeled alanines |
| GBSW | Generalized Born model with Simple Switching function |
| GPCR | G-protein-coupled-receptors |
| HETCOR | heteronuclear correlation |
| HSQC | heteronuclear single quantum coherence |
| IMP | Integral Membrane Protein |
| ITC | Isothermal Titration Calorimetry |

Continued on next page

**Table 1 – continued from previous page**

| Acronym | Meaning |
|---|---|
| MAS | Magic Angle Spinnig |
| MD | Molecular Dynamics |
| MP | Membrane Protein |
| MTSSL | 1-oxyl-2,2,5,5-tetramethyl-$\eta^3$-pyrroline-3-methyl-methanethiosulfonate |
| NCS | Non-crystallographic Symmetry |
| NOE | nuclear Overhauser effect |
| NPT | constant temperature and pressure |
| NMR | Nuclear Magnetic resonance |
| OPM | Orientations of Proteins in Membranes |
| PAF | Principal Axis Frame |
| PBC | Periodic Boundary Conditions |
| PCA | Principal Component Analysis |
| PC | Principal Component |
| PDB | Protein Data Bank |
| PISEMA | Polarization Inversion Spin Exchange at the Magic Angle |
| PISA | Polar Index Slant Angle |
| PKA | Protein Kinase A |
| PKI | Protein Kinase Inhibitor |
| PLN | phospholamban |
| PME | Particle Mesh Ewald |
| PMLG | Phase Modulated Lee-Goldberg |
| PRE | Paramagnetic Resonance Enhancement |
| RMSD | Root Mean Square Deviation |
| RMSF | Root Mean Square Fluctuation |
| RMS | Root Mean Square |
| Ryr | Ryanodine Receptor |
| ssNMR | solid state NMR |
| SA | Simulated Annealing |
| SAXS | Small Angle X-ray Scattering |
| SBMF | Solomn-Boembergen Model Free |
| SDS | Sodium Dodecyle Sulfate |
| SERCA | Sarco(endo)plasmic Reticulum Calcium Pump |
| SLF | Separated Local Field |
| SLN | Sarcolipin |
| SR | Sarcoplasmic Reticulum |
| SVD | Single Value Decomposition |
| TALOS | Torsion Angle Likelihood Obtained from Shifts |

Continued on next page

**Table 1 – continued from previous page**

| Acronym | Meaning |
|---|---|
| TAMD | Torsion Angle Molecular Dynamics |
| TEV | Tobacco Etch Virus |
| TM | Transmembrane |
| TROSY | Transverse Relaxation Optimized Spectroscopy |
| vdW | van der Waals |

# Preface

The thesis is based on I to VII of the following XII publications:

I. Structural and Dynamic Basis of phospholamban and sarcolipin Inhibition of $Ca^{2+}$-ATPase. Traaseth NJ, Ha KN, Verardi R, **Shi L**, Buffy JJ, Masterson LR, Veglia G; **Biochemistry**. 2008 Jan 8;47(1):3-13

II. Tilt and azimuthal angles of a transmembrane peptide: a comparison between molecular dynamics calculations and solid-state NMR data of sarcolipin in lipid membranes. **Shi L**, Cembran A, Gao J, Veglia G; **Biophysical Journal** 2009 May 6;96(9):3648-62

III. Paramagnetic-based NMR restraints lift residual dipolar coupling degeneracy in multidomain detergent-solubilized membrane proteins. **Shi L**, Traaseth NJ, Verardi R, Gustavsson M, Gao J, Veglia G; **Journal of the American Chemical Society** 2011 Feb 23;133(7):2232-41

IV. A refinement protocol to determine structure, topology, and depth of insertion of membrane proteins using hybrid solution and solid-state NMR restraints. **Shi L**, Traaseth NJ, Verardi R, Cembran A, Gao J, Veglia G; **Journal of Biomolecular NMR** 2009 Aug;44(4):195-205

V. Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach. Traaseth NJ, **Shi L**, Verardi R, Mullen DG, Barany G, Veglia G; **Proceedings of the National Academy of Sciences U. S. A.** 2009 Jun 23;106(25):10165-70

VI. Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. Verardi R, **Shi L**, Traaseth NJ, Walsh N, Veglia G; **Proceedings of the National Academy of Sciences U. S. A.** 2011 May 31;108(22):9101-6

VII. Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy. Masterson LR [1] , **Shi L**[1], Metcalfe E, Gao J, Taylor SS, Veglia G; **Proceedings of the National Academy of Sciences U. S. A.** 2011 Apr 26;108(17):6969-74

VIII. NMR structure of the cathelicidin-derived human antimicrobial peptide LL-37 in dodecylphosphocholine micelles. Porcelli F, Verardi R, **Shi L**, Henzler-Wildman KA, Ramamoorthy A, Veglia G; **Biochemistry** 2008 May 20;47(20):5565-72.

IX. Backbone NMR resonance assignment of the catalytic subunit of cAMP-dependent protein kinase A in complex with AMP-PNP. Masterson LR, **Shi L**, Tonelli M, Mascioni A, Mueller MM, Veglia G; **Biomolecular NMR Assignment** 2009 Jun;3(1):115-7

X. What can we learn from a small regulatory membrane protein? Veglia G, Ha KN, **Shi L**, Verardi R, Traaseth NJ; **Methods in Molecular Biology** 2010;654:303-19

XI. Probing membrane topology of the antimicrobial peptide distinctin by solid-state NMR spectroscopy in zwitterionic and charged lipid bilayers. Verardi R, Traaseth NJ, **Shi L**, Porcelli F, Monfregola L, De Luca S, Amodeo P, Veglia G, Scaloni A; **Biochimica et Biophysica Acta** 2011 Jan;1808(1):34-40

XII. cAMP-dependent protein kinase A selects the excited state of the membrane substrate Phospholamban. Masterson LR, Yu T, **Shi L**, Wang Y, Gustavsson M, Mueller MM, Veglia G; **Journal of Molecular Biology** 2011 Jul 2 [Epub ahead of print]

[1] Equal contribution

# Chapter 1

# Introduction

## 1.1 Introduction to membrane proteins

### 1.1.1 Biological function of membrane proteins

Membrane proteins (MPs) are proteins that associate with biomembranes. They are an integral part of organelle cell structures, and maintain the exchange of solute and signals between and within cells. They are divided into two major categories: peripheral and integral. Peripheral MPs adhere to the surface of the lipid membrane, while integral membrane proteins (IMPs) span the entire membrane, connecting the two bilayer leaflets. In contrast to $\sim 10^3$ folds identified for soluble proteins, IMPs consist of two major ones: $\alpha$-helical, found in the inner membrane of bacterial cells and plasma membrane of eukaryotes, and $\beta$-barrel, found in outer membrane of Gram-negative bacteria, cell wall of Gram-positive bacteria and outer membrane of mitochondria and chloroplasts. Based on recent genome-wise sequence analysis[1], about 20-30% of all genomes encode for $\alpha$-helical MPs. $\sim$2-3% of the Gram-negative bacteria genome encodes for $\beta$-barrels[2].

### 1.1.2 Structures of membrane proteins

Due to their physiological importance, MPs account for more than 60% of drug targets. Among these, G-protein-coupled-receptors (GPCRs) are the largest family, the malfunction of which can result in serious diseases, such as hypertension, congestive

heart failure, stroke, and cancer. Similarly, mutations of ion-channels can lead to genetic disorder, such as cystic fibrosis, Batter syndrome, and paralysis[3]. Despite their importance, MPs remain a major challenge in structural biology. Three major obstacles hamper their structural elucidation at high-resolution: (1) MPs are difficult to over-express in both prokaryotic and eukaryotic cells due to toxicity, heterogenous post-translational modification, poor stability, and proteolysis, (2) MPs are difficult to extract from their native environment and obtain homogeneous sample preparation for biophysical characterization, and (3) detergents/lipids used to solubilize MPs can denature MPs or hinder the crystallization by phase separation. Without well diffracting crystals, it is not possible to use X-ray crystallography to obtain protein structures. The large size of MP/lipid complexes also impedes structural studies using conventional solution NMR spectroscopy. As a result, only 281 (As of Mar 14, 2011) unique MP structures have been characterized at high-resolution. Most of these structures are obtained using X-ray crystallography and typically comprise MPs that form regular arrays in their biological environments, such as bacteriorhodopsin[4], Ca-ATPase[5]. The sharp contrast of structure statistics between soluble proteins and MPs is shown in Figure 1.1.

Theoretical methods are another key technique to study MPs. Two major classes of computational techniques are structure determination and molecular simulations. Structure determination of MPs use homology modeling or *ab initio* modeling techniques to model structures of MPs or MP complexes. Molecular simulations start from a structure representation of MPs and use different physical models, such as molecular dynamics (MD), continuum electrostatic Poisson-Boltzmann theory, and Brownian dynamics to study the structure and dynamics of MPs. Computational approaches have been valuable tools to study MPs. For example, homology modeling, docking, and molecular dynamics have been widely applied to GPCRs[6].

My focus for this thesis was to develop computational techniques to determine MP structures in combination with experimental data and apply computational techniques, such as MD, to study MPs dynamics. The ultimate goal of these projects is to build molecular-level models that are capable of robustly predicting MP function upon structural perturbations. In the following section, I will briefly review high-resolution NMR

techniques used to characterize MPs, including solution and solid-state NMR (section 1.2), computational methods used in NMR structure determination (section 1.3), and MD simulations in studying MPs (section 1.4).

## 1.2 NMR study of membrane proteins

Nuclear magnetic resonance (NMR) is a phenomenon which occurs when certain nuclei with spins are immersed in a static magnetic field and then exposed to a second oscillating magnetic field. NMR spectroscopy is the use of this phenomenon to study physical, chemical, and biological properties of matter. NMR spectroscopy is routinely used by chemists to study chemical structure using simple one-dimensional techniques and by biochemist/biophysicist to determine the structure of more complicated biomolecules using high-dimensional techniques.

NMR techniques have become a major alternative (and complementary) approach to structure determination of MPs by X-ray crystallography. For solution NMR, MPs are generally reconstituted in membrane mimicking environments (micelles, bicelles, or nanodisks) that form complexes with relatively short correlation times and result in high-resolution spectra[7]. The preparation of homogenous complexes is the key to obtaining NMR spectra suitable for structural analysis. Using these preparations in concert with TROSY-based experiments, it is possible to obtain NOEs, J-couplings, and residual dipolar couplings that are converted into structural constraints for structure determination. The membrane mimics are not native environment for MPs and thus can result in irreversible aggregation and loss of native structure. Solid-state NMR (ssNMR) allows studying MPs in their native lipid environment, without the requirement of MPs to be in crystalline form or process fast tumbling properties like those found in solution. While solution NMR provides many of the structural constraints necessary to fold the structures of relatively large proteins, ssNMR probes the topology (i.e., architecture) of MPs with respect to the lipid membranes. MP topology, which is the specification of MP orientation within the membrane, is a crucial feature that influences their interaction with the environment. In addition, NMR has a unique advantage to study proteins dynamics at time scales which range from picosecond to second. Since the interplay of

structure, dynamics, and environment dictates the function of MPs[8], the combination of solution and solid-state NMR is a practical and promising approach to characterize the function of MPs.

### 1.2.1 Solution NMR study of membrane proteins

Solution NMR is well established as a high-resolution technique to study protein structures. Among the ∼74,000 structures in Protein Data Bank (PDB), ∼9,000 are solved by solution NMR (As of June 2011). The general procedure used by NMR spectroscopist is listed in Figure 1.2. Isotopically enriched MPs are over-expressed and isolated in order to acquire high resolution NMR spectra. After establishing backbone resonance assignment of the protein, conformational restraints are further collected and processed by computational modeling softwares to determine the NMR structure ensemble. The details of these steps are presented in the following sections.

#### 1.2.1.1 Sample preparation

Isotopically labeled MPs for solution NMR samples are usually obtained from E. *coli* or a cell-free expression system. MPs of interest are can be expressed together with His-affinity tags that allows purification with metal chelate columns. Small MPs are commonly fused in the N-terminus with other globular proteins, such as maltose-binding protein. This increases the expression, facilitates the correct folding and insertion into cell membranes, and aids the purification. The fusion proteins can later be cleaved at an engineered cleavage site and separated from the MPs by analytical techniques, such as size exclusion or ion exchange chromatography.

To study by NMR, purified MPs need to be solubilized in a membrane mimicking environment. Commonly used membrane mimics are detergents, which are amphiphilic molecules that have hydrophilic heads and hydrophobic tails. At the critical micelle concentration (CMC), they form micelles which cover the hydrophobic surfaces of MPs in a similar way as lipid bilayers. Different detergents have different properties, such as charge of head groups or length of hydrophobic tail. Popular detergents used in MPs study include sodium dodecyle sulfate (SDS), dodecylphosphocholine (DPC), and

dihexanoyl-phosphatidylcholine (DHPC). Due to their different properties, they work differently for different proteins.

#### 1.2.1.2 NMR experiments for sequential resonance assignment

The MP/micelle complex tumbles fast and isotropically in solution, and is amenable to study by solution NMR experiments. Standard methods to assign large proteins are triple resonance experiments that probe different correlations between three nuclei ($^1H$,$^{15}N$,$^{13}C$) in doubly labeled($^{15}N$, $^{13}C$) proteins[9]. These experiments include HNCA, HN(CO)CA, HN(CA)CB, HNCO, HN(CA)CO. They allow sequential assignment by establishing backbone connectivity using through-bond scalar couplings (Figure 1.3A). For example in the HNCA (Figure 1.3B), magnetization transfers from $H_N$ to $N$ and then to $C_\alpha$ through both intra-residue coupling ($^1J_{NC_\alpha}$=11 Hz) and inter-residue coupling ($^2J_{NC_\alpha}$=7 Hz). In the HNCA spectrum (Figure 1.3B, right), each $H_N$ has cross peaks to $C_\alpha$ of residue $i$ and residue $i-1$. In contrast, in the HN(CO)CA experiments (Figure 1.3C), the pulse sequence is designed so that magnetization transfers from $H_N$ to $N$ to $CO$ and then to $C_\alpha$, which only results inter-residue peak (residue $i$) in the spectrum. Starting from residues that have unique chemical shifts, combination of those experiments permits accurate assignment of protein backbone resonance. Due to the relative large size of the MP/micelle complex, TROSY versions of the triple-resonance experiments are used with increased resolution[10].

#### 1.2.1.3 Collection of conformational restraints in NMR

Various geometrical restraints, such as distance, angle and orientation, need to be collected in order to determine the high-resolution three-dimensional structures of proteins. These restraints are classified as short-range restraints, orientational restraints, and long-range distance restraints.

**Short-range restraints** This includes distance restraints from NOE, torsion angle restraints from J-coupling or backbone chemical shift, and hydrogen bond restraints.

- Distances

Nuclear Overhauser effect (NOE) is the major source of distance information for solution NMR spectroscopy[11]. The NOE reflects the transfer of magnetization between two (or more) dipolarly coupled nuclear spins. The secular term of the dipolar Hamiltonian is averaged to zero, while the non secular term is still present and gives rise to cross-correlations between nuclear spins in the NOESY spectra. The cross-relaxation rate constant $\sigma_{IS}^{NOE}$ is proportional to the inverse sixth power of the distance $r$ between two homonuclear spins $I$ and $S$[9].

$$\sigma_{IS}^{NOE} = \frac{\hbar^2 \mu_0^2 \gamma^4 \tau_c}{160\pi^2 r_{IS}^6}(-1 + \frac{6}{1 + 4\omega_0^2 \tau_c^2})$$

where $\hbar$ is the reduced Planck constant, $\mu_0$ is the permeability constant, $\gamma$ is the gyromagnetic ratio, $\tau_c$ is the correlation time, and $\omega_0$ is the Larmor frequency. In a NOESY experiment, the volume of a cross peak $V$ can be related to internuclear distances according to the expression:

$$V =< r^{-6} > f(\tau_c)$$

Since most of the spin pairs are relatively rigid in folded proteins, the NOE can be used to measure distances between spin pairs. The cross peaks in NOESY spectra are usually classified into strong, medium and weak NOEs based on their intensity. The classification can be calibrated using cross-peaks that arise from spins with known distance. The upper limit of the distance probed by NOE restraints is $\sim$6 Å.

- Torsion angles

NMR provides angular constraints through the measure of J-coupling constants. The values of vicinal scalar couplings $^3J$ between atoms separated by three covalent bonds from each other are correlated to the torsion angle $\theta$ using the Karplus equation:

$$^3J = A\cos^2(\theta + P) + B\cos(\theta + P) + C$$

where the parameters $A$, $B$, and $C$ are optimized for different types of $^3J$ couplings, and $P$ is the phase of the dihedral angle $\theta$ defined by the three bonds[12].

Another method uses probabilistic approaches to determine torsion angles from chemical shift, which is a sensitive reporter of the chemical environment. Although accurate interpretation of chemical shift requires high level quantum mechanical calculations, it was found statistically that backbone chemical shifts ($C_\alpha$, $C_\beta$, $CO$) relative to random coil shifts have clear correlations with the backbone torsion angle ($\phi$,$\psi$) of polypeptide[13]. The chemical shifts have been used as a qualitative way to identify secondary structure elements and as a quantitative method to infer the values for the backbone torsion angle. One widely used software, TALOS[14], uses a database of high resolution crystal structures (<2.5 Å) with known backbone chemical shift values. It allows matching chemical shift values of a query protein to the database and assign the torsion angles values found in the corresponding crystal structures.

- Hydrogen bond

  Hydrogen/Deuterium exchange (H-D exchange) NMR experiments allow one to measure the kinetics of hydrogen-deuterium exchange. The exchange is suppressed for hydrogens which participate in hydrogen bonding. The identification of hydrogen bonds, especially those for amide hydrogens, can provide important restraints in structural biology. One of the approaches[15] measure the $^1H-^{15}N$ HSQC peak volume as a function of deuterium level.

$$(yC)^{-1} = \chi\frac{1-X}{X} + 1$$

  where $y$ is the peak volume of amide peak, $C$ is a normalization factor and $X$ is the mole fraction of $H_2O$. The exchange factors $\chi$ can be converted into hydrogen bond constraints with the knowledge of secondary structure information.

**Orientational restraints** RDCs (Residual Dipolar Couplings) are another important class of restraints discovered using NMR for weakly aligned samples. The partial alignment results in incomplete averaging of dipole-dipole interactions(Figure 1.4A). The average orientation of a weakly aligned sample is described by a

symmetric second rank traceless tensor $A$, which is defined by five independent elements[16]. The elements in $A$ is:

$$A_{ij} = \frac{1}{2} < 3\cos\theta_i \cos\theta_j - \delta_{ij} >$$

where $i$,$j$ is one of $x$,$y$ and $z$; $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$, $\theta_i$ is the angle between axis $i$ and magnetic field. The brackets <> indicate time and ensemble averaging. The average dipolar couplings between two spin-1/2 nuclei $A$ and $B$ in the weak alignment is:

$$D^{AB} = -S\frac{\mu_0 \gamma_A \gamma_B h}{8\pi^3 < r_{AB}^3 >} \sum A_{ij} \cos\phi_i^{AB} \cos\phi_j^{AB}$$

where $S$ is the Lipari-Szabo generalized order parameter, which scales $r_{AB}$ for the averaging effects of bond libration,$r_{AB}$ is the distance between the two nuclei,$\gamma_A$ and $\gamma_B$ is the gyromagnetic ratios respectively,$h$ is the Planck's constant, $\mu_0$ is the magnetic permeability of vacuum, $\phi_i^{AB}$ is the angle between $r_{AB}$ vector and $i$th axis in the alignment frame. In the principal axis frame, RDC can be rewritten as:

$$D = D_a[(3\cos^2\theta - 1) + \frac{3}{2}R\sin^2\theta\cos 2\phi]$$

where $D_a$ and $R$ are the axial and rhombic tensor components of the alignment tensor, respectively. $\theta$, $\phi$ are the angles that define the orientation of the intermolecular vector in the principal axis frame (Figure 1.4B). Due to the angular dependence, RDCs have been used as a major resource of orientational restraints, that can be used to determine the relative orientation of protein domains.

**Long-range distance restraints** PRE (Paramagnetic Relaxation Enhancement) measures long-range distances between protons and spin labels that are engineered to the protein. MTSSL (1-oxyl-2,2,5,5-tetramethyl-$\eta^3$-pyrroline-3-methyl-methanethiosulfonate) is one of such spin labels that can be introduced to the protein backbone through a disulfide bridge to either a native or engineered cysteine by site directed mutagenesis. The transverse relaxation rate ($R_2$) is dramatically enhanced for protons close to an unpaired electron, such as the one present in the oxygen atom in MTSSL. The relaxation rate enhancement ($R_2^{sp}$) is related to the

distance ($r$) between the proton and the free electron[17].

$$r = \frac{K}{R_2^{sp}}(4\tau_c + \frac{3\tau_c}{1+\omega_h^2\tau_c^2})$$

where $\tau_c$ is the correlation time for electron-nuclear interaction, $\omega_h$ is the Larmor frequency of proton and $K$ is a constant ($1.23 \times 10^{-32} cm^6 s^{-2}$) based on

$$K = \frac{1}{15}S(S+1)\gamma^2 g^2 \beta^2$$

where $\gamma$ is the nuclear gyromagnetic ratio, $g$ is the electron $g$ factor, $\beta$ is the Bohr magneton. As an approximation, the global correlation time is usually used in replace of $\tau_c$. Based on the above Solomon-Bloembergen equation, PRE restraints can provide distance restraints between 15 and 25 Å using $\tau_c$=20 ns and $\omega_h$= 500-800 MHz.

#### 1.2.1.4 Structural calculation

Structural determination in NMR is accomplished by minimizing a target function composed of experimental restraints, covalent geometry and non-bonded interactions. The minimum is located by combined minimization and molecular dynamics in both Cartesian and torsion angle space. The details for the computational approach and algorithm are presented in section 1.3.

#### 1.2.1.5 Structural refinement and analysis

Structures from computational modeling are usually ranked according to their "energy" (target function values), which correlates with structure quality and agreement with experimental data. Precision of the structures is measured as the average pair-wise root mean square deviation (RMSD) across the ensemble, or the mean RMSD relative to the average structure over the whole molecule or selected segment of the molecule. The selection of structures into the final NMR ensemble can be arbitrary due to the lack of general criteria. Several procedures are usually followed in literature to improve the accuracy of NMR structure ensembles. These include:

- Rank structures by the number of violations against experimental restraints.

- Check stereochemical quality of proteins, such as bond lengths, bond angles and dihedral angles.
- Select representative structures.

The first two criteria are usually correlated with the "energy" of a protein structure (target function value). One of the widely used methods is to check the backbone torsion angle ($\phi$,$\psi$) against the Ramachandran plot and group the angles in the structure ensemble into favorable, unfavorable, or disallowed regions. The standard values for these geometrical parameters of proteins structures are inferred from high-resolution X-ray crystallographic structures. The final structure analysis can be done with software such as AQUA and Procheck-NMR[18]. Without adequate restraints from each residue, there might be structures with similar "energy" but with large pair-wise RMSD. The third criteria ensures to select representative structure with different RMSD.

### 1.2.2 Solid-state NMR analysis of membrane proteins

For solid-state NMR (ssNMR) spectroscopy, MPs are reconstituted in synthetic membranes of various lipid compositions to mimic cell membranes. The effective correlation time of the molecule is infinite on the time scale of experiments. Radio frequency irradiations, magic angle spinning and sample alignment serve as the mechanism for line-narrowing[19]. The following sections discuss two major solid-state NMR methods, magic angle spinning and oriented ssNMR.

#### 1.2.2.1 Magic angle spinning NMR

Magic angle spinning (MAS) combines cross-polarization techniques with mechanical spinning of the samples at the magic angle (54.7°). When the sample spins at high frequency (10 kHz-25 kHz), the angle between any two atoms with respect to the external field is on average at the magic angle, at which their magnetic interactions vanish. A schematic of the experimental setup is reported in Figure 1.5A. While cross polarization enhances the sensitivity of nuclear spins with low gyromagnetic ratios (i.e. heteronuclear such as $^{15}N$ and $^{13}C$), the mechanical spinning of the sample at the magic angle averages both chemical shift anisotropy (CSA) and dipolar coupling (DC), resulting in high-resolution NMR spectra similar to those obtained using solution NMR techniques[20],

providing atomic information on both MP structure and dynamics.

#### 1.2.2.2 Oriented solid-state NMR

Oriented ssNMR method studies MP samples reconstituted in oriented lipid membranes (Figure 1.5B). Pioneered by Opella and co-workers[21], the oriented ssNMR method gives orientationally dependent conformational restraints for MPs and is the method of choice to obtain MP architecture in a native-like environment. This is achieved by measuring orientationally dependent chemical shift and heteronuclear dipolar interactions. PISEMA (Polarization Inversion Spin Exchange at Magic Angle)[19] is a two-dimensional ssNMR experiment that correlates $^{15}N$ CSA and $^{1}H$-$^{15}N$ DC for proteins with uniformly or selectively $^{15}N$ labeled backbone amides, providing information on the membrane topology of the molecule.

Mathematically, the PISEMA function is represented as $\pi(\vec{B_0}) = (\sigma, \nu)$, where $\sigma$ and $\nu$ are the CSA and DC tensors, respectively[22]. The CSA tensor is asymmetric and its principal axis frame (PAF) can be described in terms of its principal components $\vec{\sigma}_{33}$, $\vec{\sigma}_{22}$, $\vec{\sigma}_{11}$ as follows:

$$PAF = (\vec{\sigma}_{33}, \vec{\sigma}_{22}, \vec{\sigma}_{11})$$

The CSA tensor can be rewritten as a function of its principal components:

$$\sigma = \sigma_{11}(\vec{B_0} \cdot \vec{\sigma}_{11})^2 + \sigma_{22}(\vec{B_0} \cdot \vec{\sigma}_{22})^2 + \sigma_{33}(\vec{B_0} \cdot \vec{\sigma}_{33})^2$$

where $\sigma_{33} > \sigma_{22} > \sigma_{11}$. The DC tensor $\nu$ is traceless and axially symmetric with a unique rotation axis ($\vec{u}$) around the covalent NH bond and can be expressed as:

$$\nu = \frac{\nu_{||}}{2}(3(\vec{B_0} \cdot \vec{u})^2 - 1)$$

where $\nu_{||}$ is the value of the DC tensor when $\vec{B_0} = \vec{u}$, which is when the NH bond vector is in line with the direction of the magnetic field. If $(x, y, z)$ are the coordinates for $\vec{B_0}$ in the PAF, then $\sigma$ and $\nu$ can be expressed as follows:

$$\sigma = \sigma_{11}x^2 + \sigma_{22}y^2 + \sigma_{33}z^2$$

$$\nu = \frac{\nu_{||}}{2}(3(\cos\alpha \sin\beta \cdot x + \sin\alpha \sin\beta \cdot y + \cos\beta \cdot z)^2 - 1)$$

with $\vec{\nu}_{||} = C\gamma_1\gamma_2/r^3$ and $x^2 + y^2 + z^2 = 1$, where $\gamma$ is the gyromagnetic ratio of the nucleus, $C$ is a constant that equals $-\frac{\mu_0 h}{2\pi}$ with $\mu_0$ is the vacuum permeability and $\hbar$ the Planck constant, and $r$ is the distance between the two nuclei, $(x, y, z)$ are the coordinates of $\vec{B_0}$ in the PAF, $\vec{u}$ is the direction of the unit NH bond vector, $\alpha$ is the polar angle made by $\vec{u}$ and the peptide plane, $\beta$ is the polar angle between $\vec{u}$ and $\sigma_{33}$. Assuming a planar peptide plane, $\alpha$ is 0° and $\beta$ is 17° determined from experiment[23] (Figure 1.6A). The principal components of the CSA tensor are obtained from experimental PISEMA powder pattern and assumed to be the same for all residues ($\sigma_{11}$=64 ppm, $\sigma_{22}$= 77 ppm, $\sigma_{33}$=217 ppm)[23]. Note that the latter is a non-trivial approximation that simplifies the interpretation. The coefficient for the DC is 9.75 kHz with the N-H bond distance equal to 1.07 Å. Based on the above considerations, PISEMA can be used to obtain both structural and topological constraints for structure refinements.

#### 1.2.2.3 Interpretation of PISEMA data from helix

The PISEMA approach has been used for several helical MPs[24, 25, 26, 27]. In an ideal helix, the orientation of each peptide plane follows a periodic geometry. This periodicity is reflected into a characteristic PISEMA pattern termed a PISA (Polar Index Slant Angle) wheel. By analyzing the PISA wheels, one can obtain the orientation of the helix with respect to magnetic field and simultaneously all the resonance assignments[19, 28, 29]. A rigorous mathematical approach to the interpretation of these regular patterns was published by Mascioni and Veglia[30]. Using this new interpretation, the PISEMA approach can be extended to any secondary structure of biopolymers, including $\alpha$-sheet, $\pi$-helices, and $3-10$-helices. Figure 1.6 demonstrates the relationship between the PISA wheel and the tilt angle $\theta$ (slant angle with respect to the field, Figure 1.6B) and the rotation angle $\rho$ (polar index around the helix axis, Figure 1.6C). As the tilt angle $\theta$ increases, the PISA wheel shifts from the top left to bottom right. Since only positive DC is observed, the bottom right wheels are reflected on top with positive DC values. The radius of the PISA wheel expands and then reduces as $\theta$ increases. The position of each resonance in the PISA wheel depends on the rotation angle $\rho$. By rotating the helix around helix axis, the size and position of the PISA wheel is invariant, however, the relative positions of the resonance changes. One way to assign the spectra of helical MPs is to match PISA wheels by iterative rotations and tilt of an ideal helix[21, 28].

This method is based on several non-trivial assumptions. First, it assumes that the geometry of the helix is ideal (typically $\phi$ and $\psi$ angles are -65° and -40°, respectively). Second, the CSA tensor components used in the calculations are identical for all of the amino acids (in reality these values are amino acid specific). Third, the interpretation of the PISEMA is done with a static helix, which does not account for protein dynamics.

Another method based on PISA wheels was recently proposed by Opella and co-workers[31]. The authors define a chain propagator, $P$, as a product of two Wigner rotation matrices with the spherical harmonics defined as:

$$Y^T(\beta_{i+1}, \alpha_{i+1}) = Y^T(\beta_i, \alpha_i)P(\phi, \psi)$$

where $(\beta_{i+1}, \alpha_{i+1})$ and $(\beta_i, \alpha_i)$ describe the orientation for the peptide plane $i$ and $i+1$ with respect to the applied magnetic field $\vec{B_0}$, and $(\phi, \psi)$ are the protein backbone dihedral angles. This structural fitting method allows one to calculate the orientation of the peptide plane $i+1$ from $i$ and reconstructed the protein backbone using a recursive algorithm. According to the authors, no resonance assignments are required. However, this approach assumes a rigid helix as a model for the structural fitting and suffers the same limitations of the PISA wheel method.

## 1.3 Structural determination of membrane proteins using solution NMR data

NMR spectroscopy does not function as a microscope with atomic resolution that provides high-resolution images of biomacromolecules[32]. Rather, NMR provides indirect structural information that must be converted into three-dimensional structures using computational methods. The computational step is crucial to (a) obtaining the final structural ensemble compatible with NMR data and (b) refining the interpretation of the NMR data. The widely used computational methods, developed by Nilges and Clore, use a hybrid simulated annealing/MD method into the program XPLOR[33], originally developed by Brunger and co-workers. In this procedure, the CHARMM force field[34] developed for molecular dynamics (MD) simulations is combined with experimental restraints for structural determination of biomacromolecules (proteins, DNA,

and RNA). The incorporation of the force field allows the optimization algorithm to generate physically meaningful structural models of the macromolecules that satisfies experimental inputs. To allow efficient sampling of the conformational space, the minimum is searched in both Cartesian space and torsion angle space, where high-frequency motions such as bond and angle vibration motions are fixed. Below, the computational methods for structural determination of biomacromolecules using NMR data are presented. The following sections are brief summaries of the computational methods used in XPLOR[35].

### 1.3.1 Energy function

The target energy function ($E$) used for the structure determination is a combination of an empirical energy function ($E_{empirical}$) and a penalty function ($E_{solution}$) that incorporates the structural constraints obtained from solution NMR.

$$E = E_{empirical} + E_{solution}$$

#### 1.3.1.1 Minimal force field terms

$E_{empirical}$ includes all the geometric parameters necessary for a minimalistic description of biomacromolecules. It contains bond lengths, bond angles, and improper angles and van der Waals interactions from the terms usually found in a force field(Figure 1.7A, used in CHARMM):

$$E_{empirical} = E_{bonds} + E_{angles} + E_{improper} + E_{vdw}$$

**Bonded terms** All of the geometrical energy terms are represented using the harmonic approximation(Figure 1.7E):

$$E_x = \sum_x k_x(x - x_0)^2$$

where $x$ is bond length, angle, or improper angle, $k$ represents the force constant, and $x_0$ is the target values of $x$. $x_0$ values are usually obtained from high-resolution X-ray structures. The harmonic potential ensures the known geometries, such as bond length are respected during structure calculation. Note that this simplified force field

does not include the energy terms for dihedral angles and the Urey-Bradley potential as in the CHARMM field[36] (see section 1.4). In fact, the dihedral angle values are obtained experimentally, and the energy terms are included as restraints into $E_{CDIH}$ as a subcomponent in $E_{solution}$.

**Non-bonded interactions** For the non-bonded interactions, usually only the van der Waals (vdW) terms $E_{vdw}$ are considered since electrostatic interactions can introduce bias during the conformational search in vacuum. There are different ways to include a van der Waals potential into structure calculation protocols. Three formulations (truncation, switch, and repel) are reported below:

$$f_{vdw}(R) = \begin{cases} (\frac{A}{R^{12}} - \frac{B}{R^6})H(R - R_{cut}) = 4\epsilon(\frac{\sigma}{R^{12}} - \frac{\sigma}{R^6})H(R - R_{cut}) & truncation \\ (\frac{A}{R^{12}} - \frac{B}{R^6})SW(R, R_{on}, R_{off}) & switch \\ C_{rep}(max(0, (k_{rep}R_{min})^{irexp} - R^{irexp}))^{rexp} & repel \end{cases}$$

where the heavy side function ($H$) and the switching function ($SW$) are defined as:

$$H(x) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{if } x \geq 0 \end{cases}$$

and

$$SW(R, R_{on}, R_{off}) = \begin{cases} 0 & \text{if } R \geq R_{off} \\ \frac{(R^2 - R_{off}^2)^2 * (2R^2 + R_{off}^2 - 3R_{on}^2)}{(R_{off}^2 - R_{on}^2)^3} & \text{if } R_{off} > R > R_{on} \\ 1 & \text{if } R \leq R_{on} \end{cases}$$

where $\epsilon$, $\sigma$ are the Lennard-Jones (L-J) parameters ($A = B = 4\epsilon\sigma$), $R_{cut}$ is the cutoff distance, $R_{on}$, $R_{off}$ are the distances used to switch on and off the non-bonded interactions. $R_{min}$ is the van der Waals radius for a given nucleus. Both the truncation and switch functions use the L-J potential form(Figure 1.7D). In the truncation method, L-J potential is cut off abruptly at $R_{cut}$. Alternatively, $SW$ function allows for L-J potential to gradually decrease from 1 to 0 in a range from $R_{on}$ to $R_{off}$. The advantage is to maintain a continuous, smooth function up to the cutoff limit. The repel function (Figure 1.7C) does not have an attractive part and does not approach infinity when the nuclear radii overlap. At very high temperatures, this allows the atoms to go through each other, sampling a larger portion of the conformational space. The repel potential

is the commonly used in the initial annealing step while full L-J potential is used in the final refinement.

#### 1.3.1.2 Solution NMR restraints

All of the NMR constraints are coded into $E_{solution}$:

$$E_{solution} = E_{NOE} + E_{CDIH} + E_{HBON} + E_{RDC} + E_{PRE} + E_{DB}$$

where $E_{NOE}$ is the energy term for short distance restraints, $E_{CDIH}$ for torsion angles, $E_{HBON}$ for hydrogen bonds, $E_{RDC}$ for orientational restraints RDCs, $E_{PRE}$ for long-range PRE distances and $E_{DB}$ for biasing potential for torsion angles derived from protein structure database (DB).

**NOE** To account for internal dynamics, NOEs are interpreted in a semi-quantitative manner with an average distance $d$ and upper and lower distances that reflect the deviations from the average value. Different forms of potential are being used to incorporate these distance restraints. The most common potential is the soft-square,$E_{NOE}^{soft}$, where the potential is zero within a specified interval ($d-d_{minus}$,$d+d_{plus}$) and a soft asymptote is used outside this range:

$$E_{NOE}^{soft} = min(ceil, SC) \begin{cases} a + \frac{b}{\Delta^{softexp}} + c\Delta & \text{if } R > d + d_{plus} - d_{off} + r_{sw} \\ \Delta^{exp} & \text{if } R \leq d + d_{plus} - d_{off} + r_{sw} \end{cases}$$

and

$$\Delta = \begin{cases} R - (d + d_{plus} - d_{off}) & d + d_{plus} - d_{off} \leq R \\ 0 & d - d_{minus} < R < d + d_{plus} - d_{off} \\ d - d_{minus} - R & R \leq d - d_{minus} \end{cases}$$

where $d$ is the average distance obtained from NOE experiments, $d + d_{plus}$ is the upper limit, $d - d_{minus}$ is the lower limit, $d_{off}$ is the cutoff distance, $R$ is the distance between the two nuclei that needs to be optimized, $r_{sw}$ is the switching distance for the asymptotic function. The constants $a$ and $b$ are used to smooth the transition at $r_{sw}$, while $c$ determines the slope of asymptote, $S$ is a scaling factor, $C$ is the force constant, *ceil* is a constant which set the upper limit for NOE scaling (Figure 1.7E). Other potential forms for NOE potential include biharmonic and square-well function. The soft

asymptote allows fine-tuning of the $E_{NOE}$ potential and allows better sampling during annealing.

Calculation of $R$ between two nuclei is less straightforward when the two nuclei involved in the NOE distance restraints are not unambiguously defined. This can occur when signal overlaps due to poor spectrum resolution or due to chemical equivalence, such as three hydrogens from a single methyl group. The distance $R$ between groups of nuclei can be calculated using different ways of averaging. For example, in the "R-6" averaging, the distance between two selected sets of atoms are calculated according to

$$R =< R_{ij}^{-6} >^{-\frac{1}{6}}$$

where $R_{ij}$ runs through all pair-wise combinations between the two sets. In case of "centering" averaging, the distance $R$ is calculated between the geometrical center between two sets of selection.

**Torsion angles** Torsion angle restraints derived from J-couplings or TALOS are modeled using a similar harmonic potential:

$$E_{CDIH} = S \sum C * well * (modulo_{2\pi}(\phi - \phi_0), \Delta\phi)^{ed}$$

where $well(a, b)$ is defined as:

$$well(a,b) = \begin{cases} a-b & \text{if } a \geq b \\ 0 & \text{if } -b < a < b \\ a+b & \text{if } a \leq -b \end{cases}$$

$C$ is the force constant, $S$ is the overall weight factor for dihedral angle, $modulo_x(y)$ is a function that returns the reminder of $y$ over $x$. Note that it is also possible to directly refine against the J couplings using Karplus equation by using similar approaches.

**Torsion angle biasing potential** Another type of energy term $E_{DB}$ with dihedral angles was introduced by Clore and co-workers[37]. They used the Protein Data Bank to select dihedral angles occurring at high probability and discard those located outside the favorable Ramachandran region. This procedure reduces the sampling space during the simulated annealing protocols and minimizes the artifacts derived from the approximate force field used by biasing torsion angle rotamer populations. The additional

conformational database potential improves the packing of structures without violating the experimental NMR restraints. In their approach, the dihedral angle space is divided into $8^\circ \times 8^\circ$ bins. By querying the protein database, the probability $P_i$ of angle $\phi$ is estimated by counting number of occurrences in bin $i$, and then normalized by the total sample points in the database. According to statistical thermodynamics, the energy $E_{DB}(\phi_i)$ is related to the probability of finding angle $\phi$ in bin $i$ by:

$$E_{DB}(\phi_i) = -k_{DB} log(P_i)$$

The derivatives of the discontinuous potential with respect to the atomic coordinates is calculated by:

$$\frac{\partial E_{DB}(\phi_i)}{\partial \phi} = -\frac{k_{DB}(E_{DB}(\phi_{i-1}) - E_{DB}(\phi_{i+1}))}{2}$$

**Hydrogen bond** Hydrogen bond, one of the most important interactions in stabilizing protein structures, is modeled in a similar manner. One method is to use NOE-type distance restraints as shown in the earlier section. In addition, it is also possible to model the hydrogen bonding explicitly by considering both distance and angular constraints on the hydrogen bonding geometry. For a hydrogen bond $D - H \cdots A - AA$, the $E_{HBON}$ is calculated using:

$$\begin{aligned} E_{HBON} \quad = \quad & \sum_{i<j} \left(\frac{A}{r_{AD}^i} - \frac{B}{r_{AD}^j}\right) * \\ & \cos^m(\theta_{A-H-D}) \cos^n(\theta_{AA-A-H}) * \\ & SW(r_{AD}^2, r_{h_{on}}^2, r_{h_{off}}^2) SW(\cos^2(\theta_{A-H-D}), \cos^2(\theta_{h_{on}}), \cos^2(\theta_{h_{off}})) \end{aligned}$$

where $AA$, $A$, $H$ and $D$ are the acceptor antecedent, acceptor, hydrogen and donor heavy atoms, $i$, $j$, $m$, $n$, $A$ and $B$ are values that allows fine tuning the potential shape. SW is a switching function defined before, and $r_{AD}$ is the acceptor and donor distance. $\theta_{A-H-D}$ is the angle between bond $A - H$ and $H - D$. A similar definition applies for $\theta_{AA-A-H}$. Distance cutoffs and angle cutoffs are specified by $r_{h_{on}}$,$r_{h_{off}}$,$\theta_{h_{on}}$ and $\theta_{h_{off}}$. Both methods are currently used, although the latter opens up the possibility to the presence of multiple acceptors that are common in protein turns.

**RDC** As with the other NMR parameters, RDCs are incorporated in the structure calculations using a flat-well potential:

$$E_{RDC} = \begin{cases} \sum_{m=1}^{N_{restraints}} k_{RDC}(|D_{c,i} - D_{o,i}| - D_{Error,i})^2 & \text{if } |D_{c,i} - D_{o,i}| > D_{Error,i} \\ 0 & \text{if } |D_{c,i} - D_{o,i}| < D_{Error,i} \end{cases}$$

where $D_{o,i}$ is observed RDC from experiment for residue $i$, $D_{c,i}$ is from calculation, $D_{Error,i}$ is the uncertainty from experiment, $k_{RDC}$ is the force constant and the energy term sums over all restraints $N_{restraints}$. This flat-well potential considers experimental errors when penalizing deviations from experimentally observed values. In order to use the RDC data for protein structure refinement, the alignment tensor must be defined in order to utilize the angular dependence from the RDC. The alignment tensor can be determined using least-square minimization or single value decomposition (SVD)[38]. Another approach used in Xplor-NIH[39], four pseudo atoms are used to define the tensor orientation, while two additional atoms are used to represent the magnitude $D_a$ and rhombicity $R$, respectively. The pseudo atoms do not interact with protein atoms. The alignment tensors can be fixed or allowed to change during structure calculation.

**PRE** PRE restraints can be implemented the same way as distance restraints. Since paramagnetic centers are usually introduced to protein backbone through linkers that have several rotational degrees of freedom, the paramagnetic centers are rather flexible. To account for the dynamics and flexibility of spin labels, the distance uncertainties are set to be ±4Å. The error bars can reduce the coordinate accuracy. Clore and coworkers[40] included a model free formalism into the Solomn-Boembergen equation (SBMF) to account for internal motions when interpreting $^1H$-PRE rate $\Gamma_m$:

$$\Gamma_m = S^2 < r^{-6} > f_{SB,m}(\tau_c) + (1 - S^2) < r^{-6} > f_{SB,m}(\tau_t)$$

where $S$ is the Lipari-Szabo generalized order parameter, $< r^{-6} >$ is the ensemble averaged distance paramagnetic center and nucleus, $f_{SB,m}$ is a function of correlation time $\tau$ from the original Solomn-Boembergen equation, $\tau_c$ is defined as $(\tau_r^{-1}+\tau_s^{-1})^{-1}$ and $\tau_t = (\tau_r^{-1} + \tau_s^{-1} + \tau_i^{-1})^{-1}$, with the $\tau_r$ denotes the macromolecule rotational correlation time, $\tau_s$ the effective electron relaxation time, and $\tau_i$ the internal motion correlation time. Based on the SBMF formulation, they used an ensemble (multiple-structure)

representation of the paramagnetic group and refined $^1H$-PRE against observed values using:

$$E_{PRE} = k_{PRE} \sum \omega_i (\Gamma_2^{obs}(i) - \Gamma_2^{calc}(i))^2$$

where $k_{PRE}$ is the force constant, $\omega_i$ is the weighting factor, $\Gamma_2^{obs}(i)$ and $\Gamma_2^{calc}(i)$ is the observed and calculated PRE rate. Their approach is able to consider the spin label flexibility explicitly and allows quantitative use of the highly valuable $^1H$-PRE distance restraints.

### 1.3.2 Structure calculation algorithm

Many algorithms have been developed to obtain NMR structure ensemble, including distance geometry that use metric matrix approach[41], and variable target functions[42] in dihedral angle space. Torsion angle molecular dynamics (TAMD)[43, 44, 45, 46, 47, 48] is currently the most efficient method to calculate biomolecular structures by NMR. TAMD is a MD algorithm that uses torsion angles as degrees of freedom rather than Cartesian coordinates. For a biomolecule of $N$ atoms, we need to search a conformation space of $3N$ dimensions. This requires a remarkable amount of computational time. In TAMD, bonds and angles are kept fixed at their equilibrium values determined from X-ray crystal structures. Also, the molecules are grouped into several rigid bodies, referred to as clusters, in which the atoms have fixed distances relative each other. Those clusters are connected via $n$ covalent bonds (or hinges), which are able to rotate freely. Compared with $3N$ Cartesian coordinates, there are only $n$ torsion angles allowed to evolve during dynamics (Figure 1.8). This greatly reduces the size of conformational space, allowing for faster and more efficient sampling, while keeping fixed the covalent geometry. The classical mechanical system is described using the Lagrange equation:

$$\frac{d(\frac{\partial L}{\partial \dot{\theta_k}})}{dt} - \frac{\partial L}{\partial \theta_k} = 0$$

with the Lagrangian $L = E_{kin} - E_{pot}$, $\theta_k$ is the torsion angle ($k = 1, \ldots, n$), $E_{kin} = \sum_{kj}^{n} M(\theta)_{kj} \dot{\theta_k} \dot{\theta_j}$ is the kinetic energy with $M$ the $n \times n$ mass matrix, $E_{pot}$ is the potential energy function which is the sum of energy terms from structural calculation. The equations of motions are:

$$M(\theta)\ddot{\theta} + C(\theta, \ddot{\theta}) = 0$$

where $M(\theta)$ is the $n \times n$ mass matrix, $C(\theta, \ddot{\theta})$ is the $n$-dimension vector which can be calculated explicitly[46, 47]. In order to integrate this equations, Jain and co-workers developed a recursive algorithm based on the tree structure of the molecules[43]. In this method, biomolecules are dissected into a tree structure, with one base cluster and $n$ children clusters[46, 49]. This method simplifies the computational complexity, reducing the complexity from $n^3$ to $n$. A variant of the Verlet algorithm is used as an integrator[44]. Temperature control is done by velocity scaling[50]. The timestep for integration is adjusted by itself to preserve the total energy during dynamics. Compared with classical MD simulations, larger timesteps can be used, allowing fast exploration of the conformation space.

Simulated annealing (SA) is a useful optimization algorithm for high dimensional problems. It is widely used with molecular dynamic or Monte Carlo simulations in structural determination. The name comes from an analogy with thermodynamics, where liquid metal freezes naturally to a state of minimum energy (defect free) when it is cooled from a high temperature. In order to use SA algorithm, it is necessary to define the energy of the system (target function), temperature of the system $T$ and a cooling schedule. The cooling schedule defines the initial temperature $T_i$, final temperature $T_f$ , and the speed to decrease from $T_i$ to $T_f$ and the time to spend at each $T$. When combined with Metropolis algorithm, SA allows the system to explore the potential energy surface in both downhill and uphill directions. At higher temperatures, it has a higher probability of accepting large uphill Monte Carlo moves. In case of MD simulations, SA can cross high barriers without getting trapped in local minima and the cross event happens more frequently at higher temperature.

During a typical structural calculation, the target function is first defined by combining different potential energy terms. A cooling schedule is defined so that SA can be performed in either Cartesian or torsion angle space. Since SA does not guarantee to find the global minimum, the calculations are performed many times to generate a collection of structures for further process. Different weighting factors in the target function and different cooling schedules can result in different results. For each different system, these need to be optimized to avoid bias and achieve convergence.

## 1.4 Molecular dynamics simulations of membrane proteins

### 1.4.1 Force field

Molecular dynamics (MD) simulation is a computational technique for generating time-dependent trajectories of condensed phases and macromolecular systems to obtain equilibrium and dynamic properties of the system. Phase space is sampled by numerically integrating the Newton's equations of motion. MD simulations have been employed to understand macroscopic phenomena (experimental observables) from microscopic point of view (atomic positions and velocities). Furthermore, MD simulations can be used to investigate systems that are not experimentally accessible, or to make predictions for properties that have not been measured experimentally. With the advances in theory, numerical algorithms and computer power, MD simulations have now evolved to be a valuable tool for studying biological systems.

### 1.4.2 Molecular dynamics algorithms

**Newton's equation of motion** For a system of $N$ atoms, each atom $i$ is associated with a mass $m_i$ and a position denoted by the 3D vector $\vec{r}_i$, the potential energy for atom $i$ is indicated by $V_i$. The dynamics of the atom $i$ is determined by solving Newton's equations of motion:

$$m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{F}_i$$

The net force ($\vec{F}_i$) on each atom $i$ is

$$\vec{F}_i = - \bigtriangledown_i V_i$$

The potential energy function $V_i$ depends on the $3N$ atomic positions in the system, and it is described by a set of analytical functions, called a force field. The potential energy function contains empirical parameters that have been optimized to reproduce properties derived from experiments or quantum mechanical calculations on a set of model systems representing common functional groups. These parameters are assumed to be transferable for simulations of the entire proteins. Presently, a number of force fields have been developed for biomacromolecules, including CHARMM[36], AMBER[51], and GROMACS[52]. Using CHARMM27 all-atom force field as an example, the empirical

potential energy ($V$) is a function of atom positions and consists of bonded and non-bonded terms (Figure 1.7A).

$$\begin{aligned} V &= \sum_{bonds} k_b(b-b_0)^2 + \sum_{angles} k_\theta(\theta-\theta_0)^2 \\ &+ \sum_{torsionangles} k_\phi[1+\cos(n\phi-\delta)] + \sum_{impropers} k_w(w-w_0)^2 \\ &+ \sum_{Urey-Bradley} k_u(u-u_0)^2 + \sum_{nonbonded} \epsilon[(\frac{R_{min}}{r_{ij}})^{12} - (\frac{R_{min}}{r_{ij}})^6] + \frac{q_i q_j}{\epsilon_0 r_{ij}} \end{aligned}$$

The bonded term describes the empirical energy function for bond, bond angle, torsion angle, improper angle and Urey-Bradley energy terms. The first two terms in the equation are energy functions for bond stretching and angle vibration (bond bending), which both are represented by harmonic potential. A harmonic potential has the general forms of $k(x-x_0)^2$, where $k$ is the force constant, and $x_0$ is the target value for $x$. The third term is for the torsion angle term that is represent by a sinusoid function, where $k_\theta$ is the torsion angle force constant, $\theta$ is the torsion angle, $\delta$ is the phase shift and $n$ is the multiplicity. The fourth term is for the improper angles used to maintain chirality and planarity, and fifth term is the Urey-Bradley component used to restraint atoms separated by two bonds. The non-bonded terms account for the interactions separated by more than 3 bonds and include the van der Waals and the electrostatic interactions. The VDW term is described with a 12-6 Lennard-Jones potential, in which $\epsilon$ is the well depth for Lennard-Jones potential and $R_{min}$ is the distance where potential approaches to zero. Electrostatic interactions are calculated using Coulomb's law, where $q_i$ and $q_j$ are the partial charges for nucleus $i$ and $j$, $r_{ij}$ is the distance between $i$ and $j$, and $\epsilon_0$ is the vacuum permittivity. The parameters for bonded, non-bonded interactions and the connectivity of the atoms in the molecule are described in a topology file and used throughout the simulations.

**Propagating simulation system** Since there is no analytical solution to the equations of motion, numerical techniques have been developed to obtain the trajectory defined by Newton's equation of motion. One of the most common integration schemes is the Verlet algorithm[53]. The original Verlet algorithm to update the positions and

velocities is:

$$\vec{r}(t+\Delta t) = 2\vec{r}(t) - \vec{r}(t-\Delta t) + \frac{d^2\vec{r}_i}{dt^2}\Delta t^2 + o(\Delta t^4)$$

$$\vec{v}(t) = \frac{\vec{r}(t+\Delta t) - \vec{r}(t-\Delta t)}{2\Delta t}$$

Note that the subscript $i$, indexing atoms, is omitted for clarity. The original Verlet algorithm uses the current position $\vec{r}(t)$, the acceleration $\frac{d^2\vec{r}_i}{dt^2}$ at a given time $t$ and the position $\vec{r}(t-\Delta t)$ from the previous step $t-\Delta t$ to calculate the next atomic position, where $\Delta t$ is the time increment. The velocity $\vec{v}(t)$ at time $t$ is calculated using the atomic positions at times $t+\Delta t$ and $t-\Delta t$. When evolving the system by $\Delta t$, the truncation error for position is on the order of $\Delta t^4$ and the velocity $\Delta t^2$. There are several variants of the Verlet algorithm, including the velocity Verlet and the leap-frog integrator.

In order to model a condensed-phase system more realistically, the system used in the numerical simulations is typically replicated in space using periodic boundary conditions (PBC). This avoids the edge effects on the surface of the simulation box. The minimum image convention ensures that pair-wise interaction between atoms $i$ and $j$ or atom $i$ and images of atom $j$ is only calculated once despite the duplication of simulating box. The evaluation of the potential energy and forces requires enumeration of all pair atomic interactions at every time step. This imposes a significant taxing on the computational costs required for long-time simulations. One way to reduce the computational costs is to truncate the long-range, non-bonded interactions at a cut-off distance (usually 10-20 Å), beyond which the interaction energies are smoothed to zero and discarded. The smoothing functions are used to avoid discontinuity in forces and potential energies[54, 55]. This method is usually applied to van der Waals interactions, which decays quickly over distance. Serious artifacts may result by applying standard cutoff schemes to evaluating electrostatic interactions[56]. Ewald summation method and its variations, such as PME(Particle Mesh Ewald)[57] and $P^3M$(particle-particle particle-mesh)[58], are standard methods to calculate electrostatic interaction for periodic systems during simulations.

**Simulating thermodynamic ensembles** Different MD techniques have been developed to simulate a macroscopic sample, which contains an enormous number of

molecules in different microscopic states. Under the Ergodic hypothesis, time averages of the microstates in these MD simulations are consistent with macroscopic parameters in the corresponding thermodynamic state[59]. The Ergodic hypothesis establishes that a microstate can explore all possible states if it is evolved in time indefinitely. We can define several different ensembles with different characteristics: canonical ensemble (NVT), isobaric-isothermal ensemble (NPT), grand canonical ensemble ($\mu$VT), and microcanonical ensemble (NVE), where $N$ is the number of atoms, $P$ is the pressure and $T$ is the temperature, $\mu$ is the chemical potential, and $V$ is the volume.

Since it corresponds to experimental conditions and can be directly validated, the NPT ensemble is the most commonly used for most of the molecular dynamics simulations. In simulations carried out under the NPT conditions, both volume and kinetic energies are fluctuating quantities and are subjected to contact with an external piston and a thermal bath. The extended system approach is one of the advanced techniques for temperature and pressure control[60, 61]. In this approach, an additional variable is introduced to the dynamics of the system (friction coefficient, for thermostat, and piston, for barostat) and integrated in the equations of the motions. In the Nosé-Hoover thermostat algorithm, after integrating the additional variable, the microcanonical distribution for the extended system results in an isothermal-isobaric distribution for the particles[62, 63]. In this algorithm, the effects of the heat reservoir are introduced via an additional degree of freedom, s, in the Lagrangian equation for a classical $N$-body system:

$$L_{Nose} = \sum_{i=1}^{N} \frac{m_i}{2} s^2 \dot{(r)}_i^2 - U(r^N) + \frac{Q}{2}\dot{s}^2 - \frac{\ln s}{\beta} L$$

where $L_{Nose}$ is the final Lagrangian, $L$ equals to $3N+1$, $Q$ is the effective mass associated with the coordinate s, $\beta$ equals $\frac{1}{k_b T}$ and $m_i$ and $p_i$ are the mass and momentum of particle $i$, respectively. The Hamiltonian of the extended system becomes:

$$H_{Nose} = \sum_{i=1}^{N} \frac{P_i^2}{2m_i s^2} + U(r^N) + \frac{P_s^2}{2Q} + \frac{\ln s}{\beta} L$$

In the Nosé-Hoover algorithm, two sets of variables are introduced: real $(p', r')$ and virtual $(p, r)$, which are related: $r' = r$, $p' = p$, $s' = s$ and $\Delta t' = \frac{\Delta t}{s}$. The variable $s$

can be interpreted as a scale factor for the time step. The equations of motions derived from the Hamiltonian are:

$$\frac{dr'_i}{dt'} = \frac{p'_i}{m_i}$$

$$\frac{dp'_i}{dt'} = -\frac{U(r'^N)}{\partial r'_i} - \frac{s'p'_s}{Q}p'_i$$

$$\frac{1}{s}\frac{ds'}{dt'} = \frac{s'p'_s}{Q}$$

$$\frac{d(s'p'_s/Q)}{dt'} = (\sum_i p_i'^2/m_i - \frac{L}{\beta})/Q$$

Integration over $s'$ and $p'_s$ give the canonical distribution for both $r'$ and $p'$. This idea can be extended to controlling the pressure of the system; in fact, the Andersen extended Lagrangian dynamics approach for controlling system pressure was the first one introduced in the field and it is still used in conjunction with the Nosé-Hoover thermostat in NPT simulations. This is achieved by coupling the system to a fictitious piston, which allows variations of the unit cell to take place in order to control the pressure.

### 1.4.3 Molecular dynamics simulation of membrane proteins

MD simulations have been successfully applied to investigating MPs and related systems, such as ion channels, aquaporins, passive and active transporters and bioenergetic proteins[64].

A general scheme to set up and run MD simulations for MP systems includes three major steps: system setup (including proper orientation of the protein and preparation of solvated lipid bilayer), production dynamics, and trajectory analyses.

#### 1.4.3.1 System setup

**Initial protein structure** During the system setup (Figure 1.9), an initial orientation of the MPs is needed before setting up the bilayer. Experimentally, oriented ssNMR is the only technique that directly measures the orientations of peptide planes with respect to the magnetic field, which coincides with the bilayer normal. Computationally,

this can be done by embedding the protein in a continuum electrostatics models as a membrane mimicking environment and running a relatively short MD simulation to equilibrate the system. This method is quite efficient and computationally inexpensive. Other approaches,, such as the one from OPM (Orientations of Proteins in Membranes) database[65], determines the depth of insertion of the transmembrane residues and the orientation of the protein by minimizing the free energy of transfer from water to lipids.

**Build protein/lipid/water system** In this step, we need to prepare the lipid bilayer in the presence of the membrane protein. Although this seems to be straightforward, in practice, there are several issues that must be carefully considered. They include the number of lipids on each side of the bilayer, the size of the system and the initial packing of the lipids. Specifically, three important considerations need to be kept in mind for selecting different bilayer-building methods: (a) the time required to equilibrate the system, (b) the manual effort required to build the membrane, and (c) the potential for artifacts due to the wrong number of lipids at each lipid leaflet. The current methods used to build bilayer have been recently reviewed by Kandt and co-workers[66]. One of the widely used and less biased methods constructs of a lipid bilayer around the protein. A collar of vdW spheres is built around the proteins to represent the lipid headgroups. Then the spheres are substituted using pre-equilibrated lipids from a pre-hydrated lipid library. In the next stage, the protein/lipid system is solvated by overlaying the lipid head group region with a pre-equilibrated water box and let the system equilibrate over time. Counter ions are further added to reach neutrality and physiological ionic strength ($\sim$150 mM).

#### 1.4.3.2 Production dynamics

The initial setup requires extensive minimization and equilibration of the system to remove bad contacts. In practice, harmonic restraints are applied and then further decreased during the equilibration. During the same process, the temperature can also gradually increase to the desired target from a lower value.

During production dynamics, constant temperature and constant pressure ensemble (NPT) are preferred for MPs simulations. The methods have been well-established for homogeneous systems. In MPs simulations, however, there is strong disagreement

regarding the methods for pressure coupling. There are three different types of pressure coupling methods have been devised: isotropic, semi-isotropic and anisotropic. In the isotropic pressure coupling, pressure along $x$, $y$ and $z$ are coupled together no surface tension is present in the simulations. In semi-isotropic coupling, $x$ and $y$ are coupled aside from $z$; in anisotropic coupling, no coupling between pressure along $x$, $y$ and $z$ is present. In both the semi-isotropic and anisotropic coupling, there is significant surface tension. The surface tension $\gamma$ is related to the pressures $(P_{xx}, P_{yy}, P_{zz})$ along each dimension by the following:

$$\gamma =< L_z \times (P_{zz} - \frac{1}{2}P_{xx} - \frac{1}{2}P_{yy}) >$$

where $L_z$ is the length of the unit cell, $P_{zz}$ is the pressure normal to the surface, while $P_{xx}$ and $P_{yy}$ is the tangential pressures. The existence of surface tension at the membrane-water interface is controversial[67, 68]. The proponents of the lack of surface tension argue that in a fully hydrated lipid bilayer the effects of curvature that would generate surface tension are absent[67]. On the contrary, the supporters of the existence of a surface tension argue that the curvature is significant and when simulating a membrane system it should be taken into account[68].

#### 1.4.3.3 Trajectory analysis

The final stage is to analyze the trajectories and to compute the desired statistical averages. It is important to first assess the MD simulation quality by checking their agreement with available experimental data. One reason is the concern over the accuracy of the force field used to represent the protein and lipids. Although significant progress has been made in protein force field development based on high resolution soluble protein structures, there is a lack of accurate experimental data to further improve the lipid models. MP structures used to parameterize the force field for complex of protein/lipid complexes are scarce and at low resolution. As a result, the current force field used is the mere combination of force fields specific for protein and lipids. In a recent paper, White and co-workers[69] evaluated different force fields: CHARMM22, CHARMM27 and GROMACS by comparing bilayer simulation with X-ray and neutron scattering density profiles. It was shown that these force fields can reproduce experimental data reasonably well. However, they still cannot model certain physical chemical

properties of the bilayers. In particular, the spacing between headgroups in the DOPC bilayer (or bilayer thickness) obtained using these force fields is larger than that determined experimentally[69]. Until recently[70], the newly developed lipid parameters (CHARMM36) allows MD simulations in tensionless NPT ensemble to reproduce several experimental findings, including bilayer thickness, surface area for lipids, density profiles from X-ray and neutron diffraction and NMR deuterium order parameters. After checking MD simulations quality, atomic details, such as interactions between proteins or protein/lipids, can be gleaned from MD simulations. These interactions provide important insights to understand the structural/functional relationship of MPs.

## 1.5 Organization of the thesis

Advised by Prof. Gianluigi Veglia, I have been working with other colleagues in his group on studying several MPs involved in dilated cardiomyopathy (DCM). I have focused the computational aspects in studying MP structure, dynamics, and interactions. Specifically, we developed new approaches to include both solution and solid-state NMR constraints into a hybrid approach to determine MP structures. We also applied docking and molecular dynamics simulation techniques to MPs to study their dynamics and interaction with other proteins and lipids. These work are summarized in the following chapters (Chapters 2-9), which are reprints (with permission from publisher) of peer-reviewed articles.

- Chapter 1: Introduction to membrane proteins, structures determination of membrane proteins by NMR techniques and molecular modeling methods to study membrane protein structure and dynamics.

- Chapter 2: Background of SERCA regulation by phospholamban and sarcolipin
  **Reprint from**: Structural and Dynamic Basis of phospholamban and sarcolipin Inhibition of $Ca^{2+}$-ATPase. Traaseth NJ, Ha KN, Verardi R, **Shi L**, Buffy JJ, Masterson LR, Veglia G; **Biochemistry**. 2008 Jan 8;47(1):3-13

- Chapter 3: Molecular dynamics simulations of sarcolipin in lipid bilayer.
  **Reprint from**: Tilt and azimuthal angles of a transmembrane peptide: a comparison between molecular dynamics calculations and solid-state NMR data of

sarcolipin in lipid membranes. **Shi L**, Cembran A, Gao J, Veglia G; **Biophysical Journal** 2009 May 6;96(9):3648-62.

- Chapter 4: Structure determination of membrane protein phospholamban in detergent micelle.
  **Reprint from**: Paramagnetic-based NMR restraints lift residual dipolar coupling degeneracy in multidomain detergent-solubilized membrane proteins. **Shi L**, Traaseth NJ, Verardi R, Gustavsson M, Gao J, Veglia G; **Journal of the American Chemical Society** 2011 Feb 23;133(7):2232-41

- Chapter 5: Hybrid solution and solid-state NMR method to determine membrane protein structure and topology.
  **Reprint from**: A refinement protocol to determine structure, topology, and depth of insertion of membrane proteins using hybrid solution and solid-state NMR restraints. **Shi L**, Traaseth NJ, Verardi R, Cembran A, Gao J, Veglia G; **Journal of Biomolecular NMR** 2009 Aug;44(4):195-205.

- Chapter 6: Application of hybrid method to determine phospholamban monomer structure and topology in lipid bilayers.
  **Reprint from**: Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach. Traaseth NJ, **Shi L**, Verardi R, Mullen DG, Barany G, Veglia G; **Proceedings of the National Academy of Sciences U. S. A.** 2009 Jun 23;106(25):10165-70.

- Chapter 7: Application of hybrid method to determine of phospholamban pentamer structure and topology in lipid bilayers.
  **Reprint from**: Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. Verardi R, **Shi L**, Traaseth NJ, Walsh N, Veglia G; **Proceedings of the National Academy of Sciences U. S. A.** 2011 May 31;108(22):9101-6

- Chapter 8: Molecular dynamics simulations of PKA complexes.
  **Reprint from**: Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy. Masterson LR [1] ,

[1] Equal contribution

**Shi L**[1], Metcalfe E, Gao J, Taylor SS, Veglia G; **Proceedings of the National Academy of Sciences U. S. A.** 2011 Apr 26;108(17):6969-74.

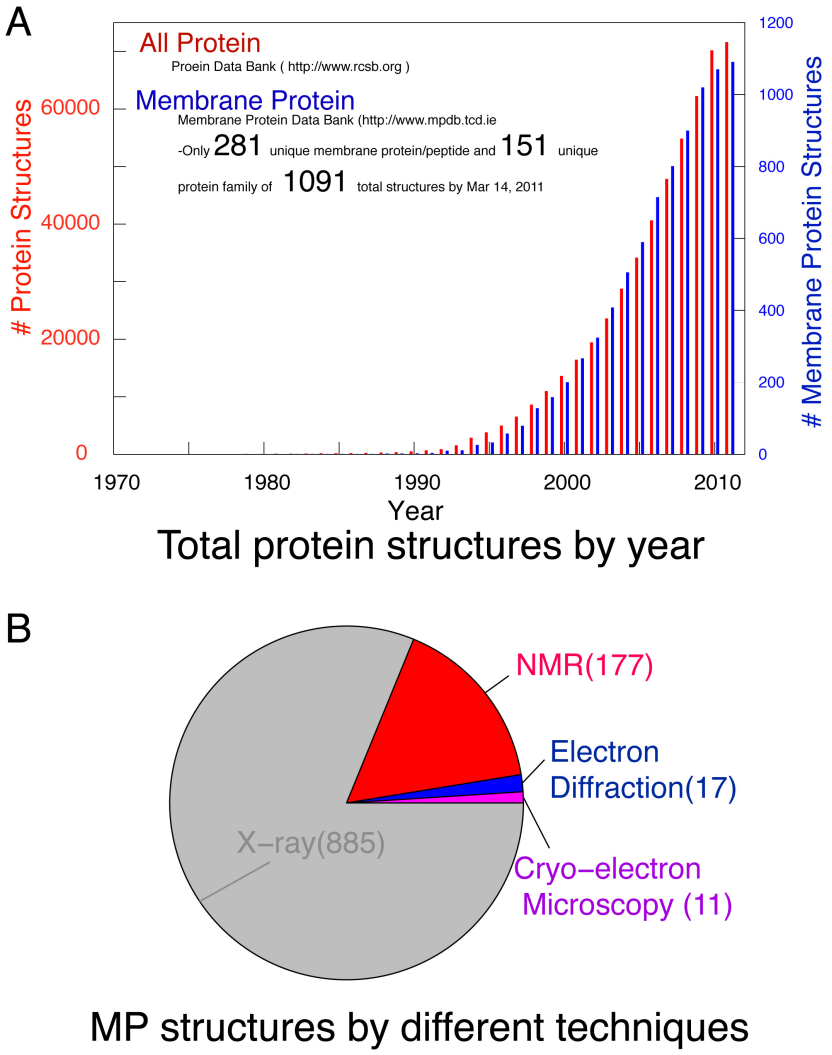


Figure 1.1: Structure statistics of MPs. (A) Yearly growth of protein structures: soluble in red, membrane proteins in blue. (B) The proportion of MP structures by different techniques.

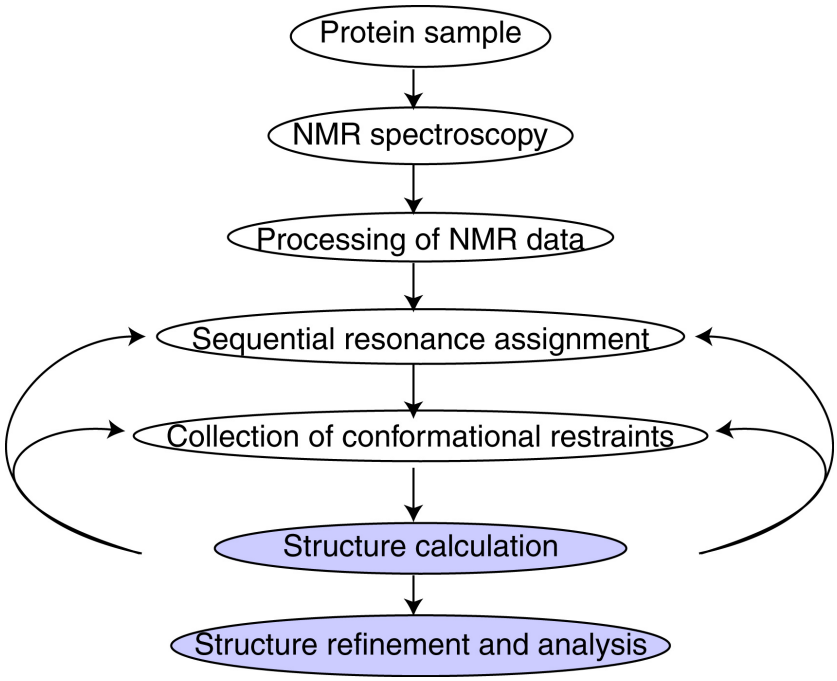


Figure 1.2: General procedures for structure determination by NMR

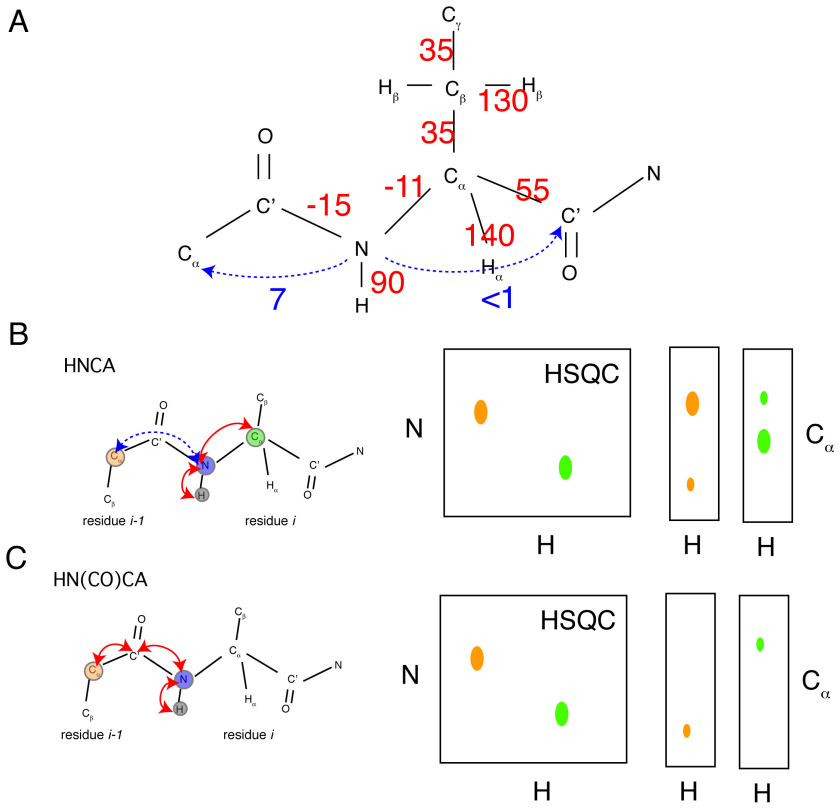


Figure 1.3: Sequential backbone assignment. (A) Strength of through-bond J-couplings in protein structures (Hz), $^1J$ is in red and $^2J$ is in blue. (B) Magnetization transfer and idealized spectrum in HNCA experiments. (C) Magnetization transfer and idealized spectrum in HN(CO)CA experiments.

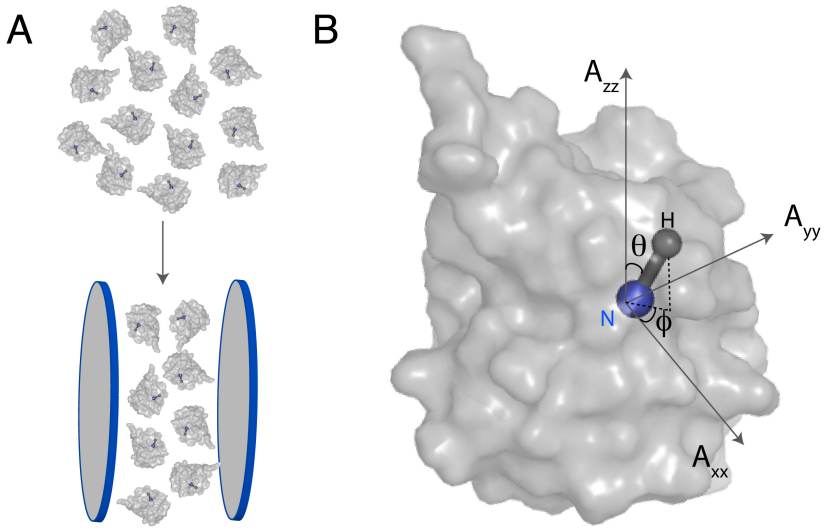


Figure 1.4: Residual dipolar couplings. (A) Isotropical tumbling of proteins in solution and restricted tumbling under partial alignment. (B) Orientation of a bond vector in the alignment tensor.

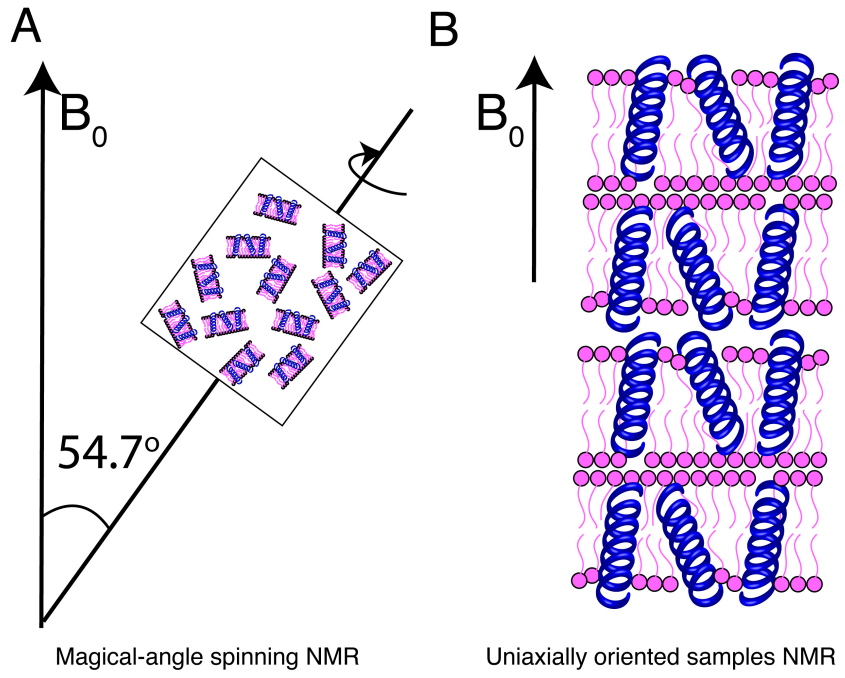


Figure 1.5: Magic angle spinning ssNMR and oriented ssNMR. (A) In MAS, samples are rotated along the magic angle (54.7°) with respect to the magnetic field $\vec{B}_0$. (B) In oriented ssNMR, uniaxially stacked lipid bilayers are aligned, typically with normal along $\vec{B}_0$.

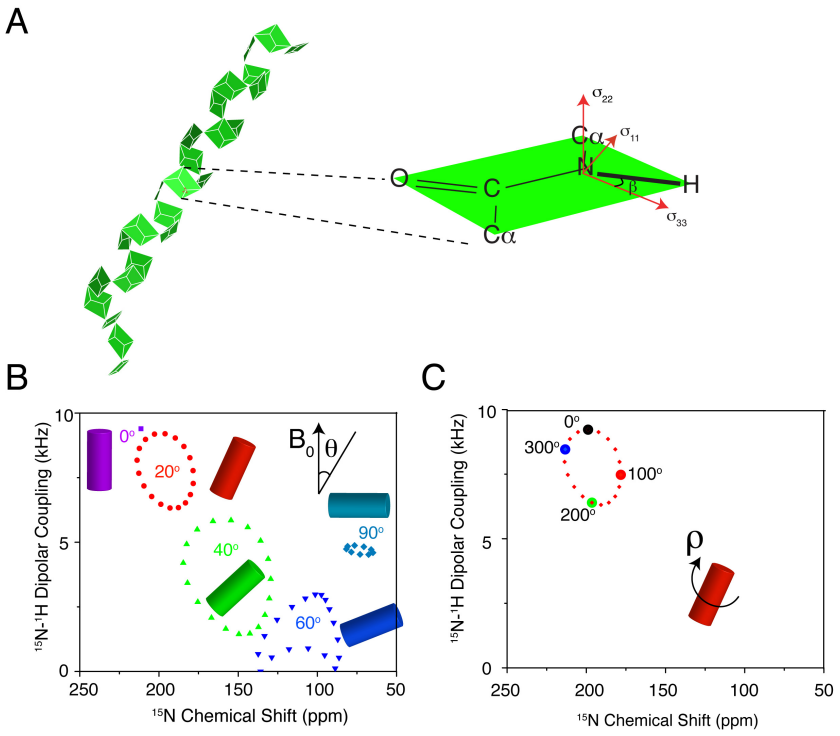


Figure 1.6: Peptide plane geometry and chemical shift tensor. (A) A protein structure connected by planar peptide plane, where each peptide plane can define the CSA tensor orientation. (B) PISA wheel patterns as a function of rotation angle $\theta$. (C) PISA wheel as a function of rotation angle $\rho$ at $\theta = 20°$.

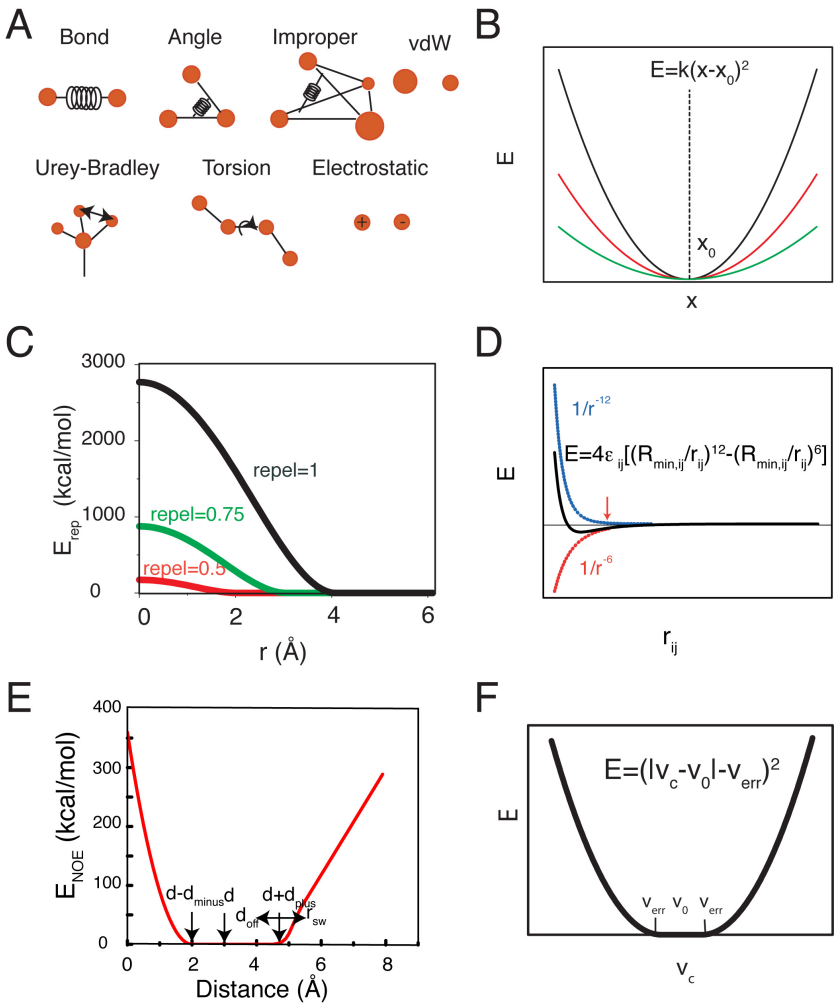


Figure 1.7: Energy terms used in structure calculation. (A) Energy terms used from the force field, including bond stretching, bond-angle vibration, improper angle, Urey-Bradley term, non-bonded interactions and electrostatic interactions. (B) The harmonic potential used to model terms in bond, bond-angle, improper angle and Urey-Bradley. (C) The repel function used in structural calculation to model non-bonded vdw interactions. (D) The L-J potential used in most MD simulations. (E) The soft square well function used to model NOE restraints. $E_{NOE}$ is plotted as a function of $r$ with $r_{sw}$=0.5 Å, $softexp$=1, $ceil$=100 kcal/mol, $a$=0,$b$=0.015,$c$=1,$S$=30,$C$=30 kcal/mol, $exp$=2, $d$=3 Å, $d_{plus}$=2.2 Å,$d_{minus}$=1 Å,$d_{off}$=0.5 Å. (F) The square well function used to model other experimental restraints, $v_c$ is the calculated value for observable $v$, $v_o$ is the experimental value, $v_{err}$ is the experimental error.

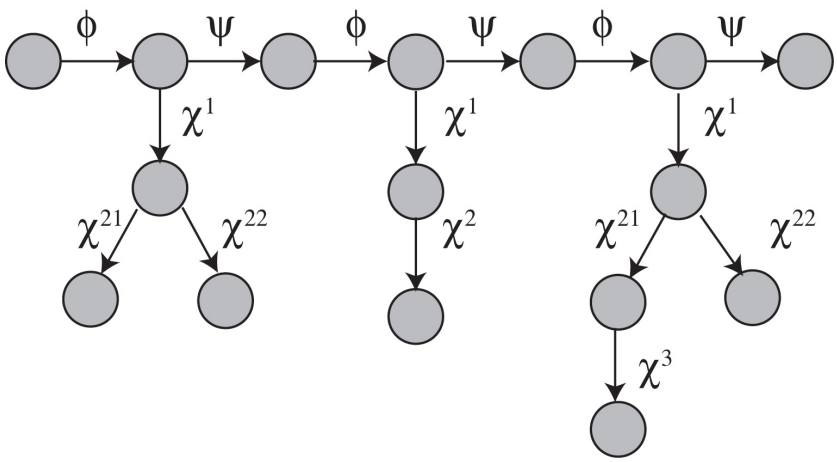


Figure 1.8: Tree structures in torsion angle dynamics. Only backbone torsion angles $\phi$, $\psi$ and sidechain torsion angles $\chi$ are allowed to change during torsion angle dynamics

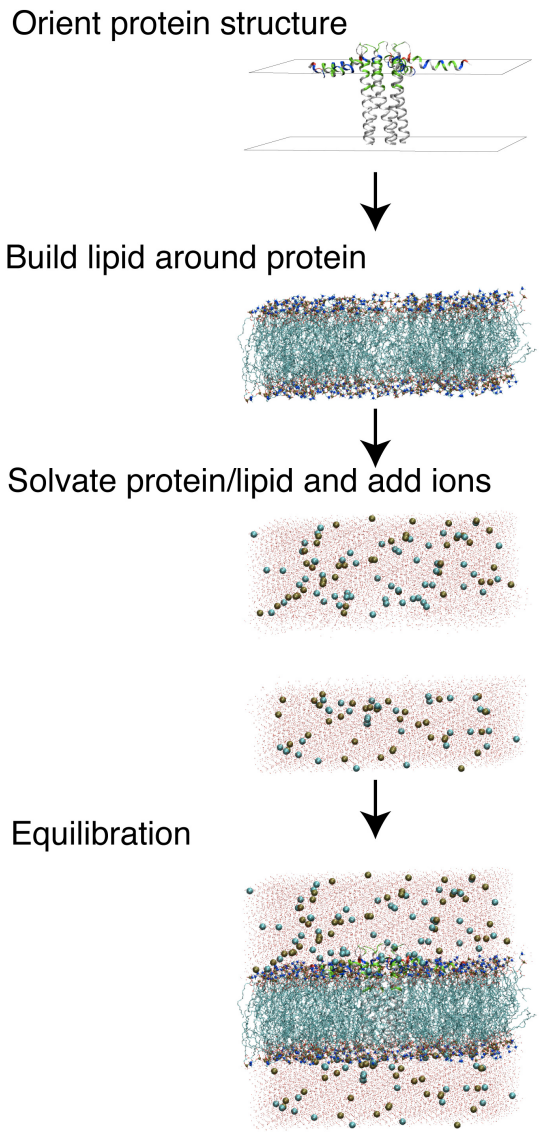


Figure 1.9: Procedures used to setupMD simulations of MPs in lipids. A proper orientation of the MP is determined from experiment or computational methods. Lipid molecules are constructed around the protein. A water box is built to solvate the protein/lipid complex and counter ions are added to reach neutrality and desired ionic strength. The assembly of protein, lipid, water and ions are subjected to equilibration before production dynamics.

# Chapter 2

# Structural and dynamic basis of phospholamban and sarcolipin Inhibition of $Ca^{2+}$-ATPase



## 2.1 Introduction

Two membrane proteins, sarco(endo)plasmic reticulum calcium ATPase (SERCA)1 and the ryanodine receptors (Ryr), play major roles in calcium regulation within muscle cells. Ryr1 and Ryr2 are responsible for releasing $Ca^{2+}$ from the sarcoplasmic reticulum (SR) of skeletal and cardiac muscle, respectively, resulting in muscle contraction. SERCA1a and SERCA2a pump calcium from the cytosol into the SR in skeletal and cardiac muscle, respectively, initiating muscle relaxation. Phospholamban (PLN), a 52-residue protein spanning the SR membrane[72], is an endogenous inhibitor of SERCA, lowering the apparent calcium affinity of the ATPase. The relief of SERCA inhibition is achieved by phosphorylation of PLN at Ser16 by protein kinase A and/or Thr17 by $Ca^{2+}$/calmodulin-dependent protein kinase[73]. In vivo studies demonstrate that PLN phosphorylation at Ser16 and Thr17 has different physiological effects, suggesting that

these mechanisms act independently[74, 75].

Sarcolipin (SLN) has a primary sequence homologous to that of the transmembrane domain of PLN[72, 76]. Initially, SLN was thought to be the counterpart of PLN within skeletal muscle, playing only an ancillary role in cardiac muscle. Recently, however, significant expression levels of SLN have been detected in cardiac atrial muscle, with lower levels in ventricular muscle[77, 78, 79], suggesting that it may play an important role in regulation of the heart. When SLN was initially copurified with fast-twitch skeletal SERCA1a[80], no post-translational modifications were identified, which led to the conclusion that the regulation of SLN depended on its variable expression levels[81, 82]. Recent evidence shows that SLN inhibition can be fully reversed by isoproterenol, a $\beta$-adrenergic receptor agonist in PLN knockout mice[83]. These results led to the hypothesis that the inhibitory effect of SLN can be reversed via phosphorylation in a manner similar to that of PLN. In vitro experiments have shown that SLN can be phosphorylated at Thr5 when cotransfected with serine/threonine kinase 16 (STK16)[83]. From these biological data, it is clear that phosphorylation of PLN and SLN constitute driving forces for calcium re-uptake into cardiac SR.

Several crystal structures of SERCA in different conformations within the enzymatic cycle have revealed important atomic details regarding SERCA's mechanism (Figure 2.1)[84, 85, 86, 87, 88, 89]. PLN is thought to bind and inhibit the low-affinity calcium form of SERCA (E2) and detach from the enzyme (either partially or totally) upon phosphorylation at Ser16, reversing its inhibitory effect and restoring the affinity of SERCA for $Ca^{2+}$ ions. The only experimental structure of the SERCA-PLN complex is a low-resolution cryoelectron microscopy image (8-10 Å) obtained by Stokes and co-workers[90]. Hampering the formation of large, highly diffracting SERCA-PLN and SERCA-SLN cocrystals is the dynamic interplay between the proteins and lipids. For this reason, MacLennan and co-workers have used a plethora of biological data (mutagenesis studies, coimmunoprecipitation assays, and cross-linking experiments) in concert with molecular dynamics simulations to model SERCA-PLN[91], SERCA-SLN[92], and SERCA-PLN-SLN complexes[92] (Figure 2.2). Hutter et al.[93] have also modeled the solution structure of C41F PLN determined in a chloroform/methanol mixture with the

E2 form of SERCA using molecular mechanics. While these models shed light on the interaction between SERCA and PLN and SLN, there are inconsistencies concerning the topology and structure of PLN and SLN within the complexes.

This review reports on our recent progress involving the structure determination of PLN and SLN and toward elucidation of the interaction with SERCA using solution and solid-state NMR. While many contributions to the structural analysis of PLN and SLN from other laboratories are cited and related to our work, this review is not intended to be an exhaustive overview of the large amount of structural and biological information about PLN and SLN within the literature.

## 2.2 Choice of membrane mimicking environments for spectroscopic studies

For both solution and solid-state NMR, we use two major criteria for sample preparation: (1) functionality of proteins under NMR conditions and (2) the ability to acquire high-quality NMR spectra. Being membrane-embedded proteins, SLN and PLN need lipid environments to elicit their biological function. After scanning several different conditions, we chose dodecylphosphocholine (DPC) detergent micelles for solution NMR studies and a 4/1 mixture of dioleoyl-sn-glycero-3-phosphocholine and dioleoyl-sn-glycero-3-phosphoethanolamine (DOPC/DOPE) lipids for solid-state NMR studies. Panels A and B of Figure 2.3 show the calcium dependence of SERCA activity in the presence and absence of PLN (or SLN) under these conditions. The hydrolysis of ATP by SERCA in DPC micelles is followed directly using $^{31}P$ NMR spectroscopy and indirectly in lipids using the coupled enzyme assay[94, 95, 96]. An example of the kinetics from the $^{31}P$ spectroscopic assay is shown in Figure 2.3C. The kinetics of ATP hydrolysis was determined by following the formation (build-up) of inorganic phosphate (Figure 2.3C). In both lipid bilayers and detergent micelles, SERCA samples are fully functional.

## 2.3 PLN structure and dynamics in DPC micelles

In the SR, PLN is thought to exist as an inactive pentamer (storage form), which depolymerizes into functional monomers prior to interaction with SERCA (Figure 2.2)[97, 98]. Accordingly, we have focused our attention on a fully functional monomer of PLN (AFA-PLN) obtained by mutating the three cysteine residues (Cys36, 41, and -46) of the transmembrane domain to Ala, Phe, and Ala, respectively. Using NOE restraints in structural calculations, we have determined that PLN adopts an L-shaped conformation in DPC micelles comprised of three distinct structural domains: cytoplasmic domain Ia helix (residue 1 to 16), loop (residue 17 to 22), and transmembrane domain helix (residue 23 to 52) (Figure 2.4A)[99]. From simulated annealing calculations, we obtained a structural ensemble with very good convergence for each single structural domain (see ref[99]), but with the PLN conformers displaying an interhelical angle (angle between cytoplasmic and transmembrane domains) of 80±22° (Figure 2.4A). Indeed, the limited number of NOEs detected in the loop did not allow us to restrain the two helical domains and obtain an ensemble of structures with a low RMSD over the entire protein backbone. Because of the lack of convergence on the orientation of helical domain Ia with respect to the surface of the micelle, we used $Mn^{2+}$ and 5'- and 16'-doxyl stearic acids as paramagnetic probes of the topological arrangement of PLN in the micelle[99]. Our results show that $Mn^{2+}$ ions cause a reduction in the resonance intensities of residues located at both termini, the loop, and Ser16 and Thr17, showing that both phosphorylation sites are solvent-exposed. Both 5'- and 16'-doxyl stearic acids caused a reduction in the resonance intensities of residues located in the micellar region. 16'-Doxyl stearic acid affected residues in the core of the micelle (residues 35-45), as well as Leu7, located in the middle of domain Ia, a residue that is likely buried in the hydrophobic region of the micelle. At higher 16'-doxyl stearic acid/PLN ratios, domain Ia is also considerably affected by the paramagnetic center, with Ala11 displaying a reduced resonance intensity.

Our structural ensemble of the PLN monomer is in qualitative agreement with studies carried out in organic solvents[100, 101], where the authors found an overall L-shaped structure of PLN with the intervening loop in either a short flexible turn or a type III $\beta$-turn conformation. However, we found an orientation of the cytoplasmic helix, which

is in better agreement with the amphipathicity of the PLN sequence[99]. The most relevant PLN conformations in PDB entry 1N7L are those in which the hydrophobic side chains in domain Ia (Val4, Leu7, Ala11, and Ala15) are oriented toward the interior of the micelle and the hydrophilic residues point toward the bulk solvent, rendering the two phosphorylation sites exposed for interaction with protein kinases.

While our NMR structural studies identified three structural domains, nuclear spin relaxation measurements and solvent accessibility experiments carried out on uniformly $^{15}N$-labeled PLN[102, 103] further subdivided the helical segment from residue 23 to 52 into two domains: domain Ib (residues 23-30), which is more dynamic and constituted by hydrophilic residues, and domain II (residues 31-52), more hydrophobic and motionally restricted. Therefore, we can divide PLN into four dynamic regions, which are characterized by different order parameters (see Figure 2.4A). Carr-Purcell-Meiboom-Gill (CPMG)-based relaxation measurements also indicated the presence of slow dynamics (microsecond to millisecond motion) in domain Ia, the loop, and domain Ib[102, 103]. While the slow dynamics of domain Ia and the loop were predicted on the basis of the H-D exchange factors, the flexibility of residues within domain Ib was unexpected. The plasticity of this region supports biological evidence for the great importance of domain Ib and the molding necessary to fit into the binding groove of SERCA[91, 104].

Recently, a structure of wild-type PLN (wt-PLN) was reported showing the pentamer to be in a bellflower arrangement, where the cytoplasmic domain helix of each monomer makes an $\sim 20^{\circ}$ angle with respect to the bilayer normal[105], substantially different from the topology we reported for monomeric AFA-PLN[99]. Our recent studies using solution NMR and EPR in DPC detergent micelles also point toward a dominating L-shaped conformation (or pinwheel model) for the pentamer[106]. We also looked at the topology of pentameric wt-PLN in lipid bilayers using solid-state NMR, finding unambiguous evidence for the pinwheel model (see below).

## 2.4 SLN structure and dynamics in micelles

The initial structure of SLN was determined in sodium dodecyl sulfate (SDS) micelles using a synthetic polypeptide[107]. Under these experimental conditions, SLN has an

$\alpha$-helical conformation from residue Phe9 to Arg27 with RMSDs of 0.65 and 1.66 Å for backbone and side chain atoms, respectively. Both the N-terminus (M1-L8) and the C-terminus (S28-Y31) were found to be unstructured. Subsequently, we expressed uniformly $^{15}N$-labeled recombinant SLN in Escherichia coli, which enabled the use of higher-resolution $^{1}H$-$^{15}N$ NMR experiments to reduce several ambiguities due to resonance overlap. With recombinant SLN, we determined the structure in DPC micelles, conditions that ensured the activity of SERCA, and found it to remain a single transmembrane helix with approximately five unstructured residues at either terminus[95]. The superposition of the $C_\alpha$, NH, and C' backbone atoms from residue Arg6 to Arg27 gave an RMSD of 0.4±0.2 Å with an RMSD of 1.7±0.3 Å for side chain atoms.

In addition to structure determination, we also measured spin relaxation rates and found that the backbone dynamics, similar to that of PLN, is more complex than the structure. Relaxation measurements reveal four dynamic domains: a short unstructured N terminus (residues 1-6), a short dynamic helix (residues 7-14), a more rigid helix (residues 15-26), and an unstructured C-terminus (residues 27-31)[95]. H-D exchange factors also support the existence of four dynamic domains[95]. The similarity of SLN structure and dynamics with those of PLN domains Ib and II shows that sequence conservation is reflected in the conservation of both structure and dynamics (Figure 2.4B).

## 2.5 PLN and SLN topologies and dynamics in lipid bilayers

Solid-state NMR in lipid bilayers has emerged as a complement to solution NMR studies in detergent micelles for elucidating structure, dynamics, and interactions between membrane proteins[108, 109]. For our first solid-state NMR studies, we synthesized AFA-PLN labeled with $^{15}N$ at Ala11, Ala15, and Ala36 and measured the $^{15}N$ chemical shift anisotropy in 4/1 DOPC/DOPE mechanically oriented lipid bilayers on glass plate supports[110]. Using these solid-state NMR measurements in concert with rigid body molecular mechanics, we found that domains Ib and II are oriented approximately perpendicular to the plane of the bilayer with the interhelical (i.e., interdomain) angle

ranging between 60 and 100°, ruling out the possibility of a continuous $\alpha$-helix and also suggesting that the cytoplasmic domain of PLN interacts with the membrane surface.

Similar solid-state NMR measurements on SLN oriented in mechanically aligned DOPC/DOPE bilayers revealed the approximate parallel orientation of the SLN helix with respect to the membrane bilayer normal[107]. Since the limited number of labeled sites did not allow us to give quantitative topological angles for SLN and PLN within the bilayer, we then proceeded to use two-dimensional (2D) $^{1}H$-$^{15}N$ PISEMA (polarization inversion spin exchange at the magic angle) experiments[111]. This separated-local-field experiment correlates the $^{15}N$ chemical shift anisotropy (CSA) with the $^{1}H$-$^{15}N$ dipolar coupling (DC). Since the values of both CSA and DC depend on the orientation of the peptide plane with respect to the direction of the magnetic field, the assignment of the amide resonances allows for the determination of the structure and topology in aligned lipid bilayers.

We determined that in mixed 4/1 DOPC/DOPE lipid bilayers, AFA-PLN has an overall L-shaped conformation where the helix comprising domains Ib and II makes a tilt angle of $\sim 21^{\circ}$ with respect to the bilayer normal[112]. As expected from the one-dimensional (1D) solid-state NMR studies on synthetically $^{15}N$-labeled AFA-PLN[107], PISEMA NMR spectroscopy clearly shows that domain Ia interacts with the membrane surface, making an angle of $\sim 90^{\circ}$ with respect to the bilayer normal[112]. A current model of the PLN monomer is reported in Figure 2.5E.

In addition to the structural and topological information obtained from oriented alignments, tilting the aligned samples to different angles with respect to the direction of the static field makes it possible to investigate the rotational dynamics of the protein within the bilayer[113, 104, 114]. Tilting the AFA-PLN sample by 90° revealed that domains Ib and II undergo fast long-axial rotational diffusion about the bilayer normal with the cytoplasmic domain undergoing this motion and other complex dynamics, scaling both the values of CSA and DC[112]. The dynamics detected in both our solution and solid-state NMR experiments may explain variability within the literature regarding the topology of the cytoplasmic domain of PLN. For example, a magic-angle-spinning (MAS) solid-state NMR study carried out by Baldus and co-workers found

that while cross-polarization (CP)-based pulse sequences were adequate for detection of the transmembrane domain of AFA-PLN, showing the existence of a well-defined helix, the cytoplasmic domain residues were too dynamic to be detected[115]. Instead, J coupling coherence transfers, similar to those in solution NMR experiments, were used to detect the dynamic cytoplasmic domain, leading to the conclusion that the cytoplasmic domain was completely unstructured. While this study represents an advancement in MAS methodology, the structure most likely represents a minor conformational state and is inconsistent with a wealth of data, including those from our laboratory, which consistently show a predominant helical cytoplasmic domain with an overall L-shaped monomeric structure in lipid bilayers and detergent micelles.

A close inspection of PISEMA spectra from selectively labeled samples reveals the presence of two peak populations that exemplify two slightly different topologies for AFA-PLN domains Ib and II[112]. The two topologies have the same tilt angle ($\theta$) for domains Ib and II with respect to the membrane normal, but slightly different rotational angles around the helix axis ($\rho$). Multiple populations of PLN have also been observed by the Lorigan and Middleton groups using MAS NMR experiments in lipid vesicles[116, 117]. The detection of multiple conformers underscores the plasticity of PLN and might be an important recognition mechanism for SERCA, protein kinase A, $Ca^{2+}$/calmodulin-dependent protein kinase, and protein phosphatase 1, previously shown to be necessary for physiological processes[104].

As with those of AFA-PLN, the PISEMA spectra of SLN obtained on uniformly $^{15}N$-labeled and selectively labeled [$^{15}N$-Leu], [$^{15}N$-Ile], and [$^{15}N$-Val] samples also revealed the existence of two distinct topologies[118]. Both the major and the minor populations of the resonances corresponding to domains Ib and II are oriented $\sim 23^{\circ}$ with respect to the lipid bilayer normal but vary in the rotation angle about the helical axis by $\sim 5^{\circ}$ (in remarkable agreement with AFA-PLN). The primary sequence homology between SLN and PLN results in nearly identical structural and dynamic properties of these two regulatory proteins.

## 2.6 Pentameric wt-PLN topology in lipid bilayers

More recently, our group has embarked on the validation of the pentameric structure of wt-PLN in lipid bilayers and detergent micelles[106]. While monomeric PLN has previously been shown to bind and inhibit SERCA, recently it was hypothesized that the pentamer could also bind and inhibit SERCA[119]. While there is broad consensus regarding the secondary structure of pentameric wt-PLN, there is disagreement in the literature about the orientation of the cytoplasmic helix. In particular, there are four proposed models for pentameric wt-PLN (Figure 2.6). The first model (extended helix/sheet) shows wt-PLN to be comprised of two $\alpha$-helices connected by an antiparallel $\beta$-sheet (residues 22-32), where the cytoplasmic domain is oriented 50-60° relative to the bilayer normal[120]. The second model depicts wt-PLN as a continuous $\alpha$-helix with a tilt angle of 28±6° with respect to the bilayer normal[121, 122]. The third model (pinwheel) shows that the most stable pentamer has a pinwheel geometry in which the cytoplasmic domain helices are oriented $\sim 90°$ with respect to the membrane bilayer normal[123]. The fourth and most recent model (bellflower) shows the structure of the pentamer to be in a bellflower assembly with the cytoplasmic domain helices oriented $\sim 20°$ with respect to the bilayer normal[105].

To study the topology of pentameric wt-PLN, we reconstituted the protein in mechanically aligned 4/1 DOPC/DOPE lipid bilayers and analyzed the protein's architecture using $^{1}H$-$^{15}N$ PISEMA spectroscopy. As with the AFA-PLN monomer[112], we found that the wt-PLN PISEMA spectrum is composed of three different populations of resonances (see the 1D spectrum in Figure 2.7A) corresponding to domains Ib and II (with the resonances located between 170 and 220 ppm), an in-plane cytoplasmic domain (with resonances located between 50 and 100 ppm), and a more flexible region (loop and termini) with resonances clustered around ∼110 ppm (isotropic portion of the spectrum). Since our 1D spectrum shows three distinct regions, indicating three unique wt-PLN domain alignments with respect to the membrane bilayer normal, this eliminates the possibility of the continuous helix model.

To distinguish among the other models depicted in Figure 2.6, we performed PISEMA experiments using selectively labeled [$^{15}N$-Ala], [$^{15}N$-Thr], [$^{15}N$-Leu], [$^{15}N$-Ile], [$^{15}N$-Cys], and [$^{15}N$-Asn] wt-PLN samples. Our experimental PISEMA spectra, reported as an overlay in Figure 2.7, show a remarkable similarity with the AFA-PLN monomer. In fact, simulations for domain Ia resonances (Leu7, Thr8, Ala11, Ile12, Ala15, Thr17, and Ile18) correspond to a helix with a tilt angle of $\sim 90^{\circ}$ with respect to the bilayer normal (Figure 2.7E)[106].

Figure 2.8 shows PISEMA spectra simulated from the pinwheel and bellflower PDB coordinates for those selectively labeled sites shown in Figure 2.7. If the pentamer topology corresponded to the pinwheel model, the cytoplasmic domain residues would resonate in the upfield region of the spectrum (50-100 ppm) (Figure 2.8A). On the other hand, if the architecture of wt-PLN were consistent with the bellflower model, domain Ia resonances in the PISEMA pattern would occupy the downfield portion of the spectrum (170-220 ppm), as represented in Figure 2.8B. If the experimental domain Ia spectra are compared, it is clear that in lipid bilayers the cytoplasmic domain Ia is oriented perpendicular with the bilayer normal forming an overall pinwheel geometry.

Structural fitting with an ideal helix in Figure 2.7D revealed that the helix corresponding to domains Ib and II of pentameric wt-PLN has a tilt angle ($\theta$) of $\sim 15^{\circ}$ with respect to the bilayer normal[106]. Monomeric AFA-PLN has a tilt angle of $\sim 21^{\circ}$[112], which requires pentamer formation to tilt by $\sim 6^{\circ}$ to accommodate the leucine/isoleucine zipper holding the pentamer together[124, 125].

## 2.7 Allosteric activation model

The functionality of SERCA under NMR conditions and the quality of AFA-PLN spectra upon addition of SERCA enabled the unprecedented atomic mapping of the interactions between these two integral membrane proteins in detergent micelles[126]. In its free form, AFA-PLN exists in a dynamic equilibrium between two conformations, T and R states, where the T state or L-shaped conformation is thermodynamically stable and the R state or extended form is identified with a more dynamic cytoplasmic domain

(Figure 2.9)[126, 127]. These two states are readily detected using EPR spectroscopy in both micelles and lipid bilayers[126, 127], but due to the time scale of the exchange, NMR can only imply the existence of these forms from relaxation dispersion measurements (i.e., conformational interconversion). However, upon addition of SERCA to AFA-PLN, chemical shift perturbation analyses reveal the appearance of a second population of peaks within domain Ia, the loop, and domain Ib, indicating a conformational switch of AFA-PLN from the T state to the R state, exemplifying an allosteric activation mechanism[126].

Resonances from the hydrophobic portion of the transmembrane region (domain II) also show chemical exchange to the R state[96]. A difference plot of the $^1H_N$ chemical shift before and after addition of SERCA for residues within domain II (residues 31-52) shows a symmetric bimodal behavior where the C-terminal part of domain II shifted upfield and the residues near the N-terminal part downfield[96]. Since upfield and downfield shifts have been correlated to the strength of hydrogen bonds[128], one possible explanation for the data is that the C-terminal end of the transmembrane domain (residues 46-52) unwinds upon binding SERCA. This hypothesis was first proposed by MacLennan and co-workers[91], who indicated an overall change in the secondary structure of the domain II, with residues 49-52 unwinding upon interaction with SERCA, a process that might facilitate binding.

These results are echoed in the binding of SLN to SERCA[95]. Overall, SLN behaves like the domains Ib and II of AFA-PLN, with each dynamic domain mimicking the behavior of the corresponding domain in AFA-PLN. Upon addition of SERCA, the transmembrane domain is in fast exchange between two free forms (T and R states). As previously indicated, spin relaxation measurements divided the transmembrane domain of SLN into two regions we named domain Ib and domain II in analogy with AFA-PLN. The chemical shift changes of these two regions follow the bimodal behavior of the transmembrane domain of AFA-PLN, indicative of a similar mechanism involving an unwinding of the C-terminal residues and a stabilization of the residues in the N-terminal portion of the protein as a result of the interactions with SERCA[95]. This supports the hypothesis that both SLN and PLN transmembrane domains bind SERCA

in the same site and with an identical mechanism.

## 2.8 Effects of PLN phosphorylation on the allosteric mechanism

The inhibition of SERCA by PLN can be reversed by phosphorylation at Ser16 by cAMP-dependent protein kinase A[73]. We determined the structure of Ser16-phosphorylated AFA-PLN (pS16-AFA-PLN) and found that residues 14-16, previously helical, became unwound upon phosphorylation, revealing an order-to-disorder transition[103]. In addition, we found that there are pronounced changes in pS16-AFA-PLN backbone dynamics on both the picosecond to nanosecond and microsecond to millisecond time scales[103]. Although small, some of the changes are propagated throughout the entire protein backbone, demonstrating that while the structural transitions following phosphorylation are localized, the changes in backbone dynamics are radiated throughout the protein.

How can this order-to-disorder transition help in understanding the interaction with SERCA? To answer this question, we proceeded with the analysis of the chemical shift perturbation of pS16-AFA-PLN induced by SERCA[96]. We found that the conformational equilibrium between the T and R states upon addition of SERCA is influenced by single phosphorylation at Ser16; specifically, phosphorylation shifts the equilibrium toward the R state in a cooperative manner[96]. Another considerable difference upon phosphorylation includes a change in both the surface and the dynamics of domain Ib. In particular, a remarkable change is observed for the side chain binding behavior. In contrast with unphosphorylated PLN, the Gln26 resonance in the phosphorylated protein is unperturbed by SERCA with other smaller changes seen for side chain residues Asn27, 30, and 34. A possible mechanism for explaining these results is the rotation or rearrangement of domain Ib upon phosphorylation that disrupts crucial intermolecular hydrogen bonds, resulting in relief of inhibition. On the basis of the molecular model of MacLennan and co-workers[91], we proposed that a crucial hydrogen bond between Arg324 and Gln26 may be broken between SERCA and PLN after phosphorylation at Ser16[96]. These findings are in agreement with mutagenesis studies, showing that Q26A

is a loss-of-function mutant[129]. In addition, cross-linking studies show that an N27C mutation in pS16-AFA-PLN is no longer able to cross-link with SERCA[91]. While there are changes in the binding interface for domain Ib, the overall binding of domain II to SERCA is not affected by phosphorylation. The dissociation constants ($K_d$) for domain II in both pS16-AFA-PLN and AFA-PLN are $\sim 60\mu$M[96]. This demonstrates that the major changes are in domain Ia, loop, and domain Ib, with domain II only marginally affected, supporting the hypothesis that phosphorylation at Ser16 does not dissociate PLN from SERCA completely. We proposed domain Ib as a bridgehead region, which transmits the dynamics induced by phosphorylation at Ser16 from domain Ia to domain Ib, thereby regulating the intramembrane protein-protein interaction[96]. A schematic of the allosteric model for phosphorylation is reported in Figure 2.9.

## 2.9 Perspective

What can we learn from the analysis of the structure and dynamics of PLN? More importantly, how are structure and dynamics of PLN correlated to SERCA's function, and can we control the extent of inhibition of SERCA by manipulating PLN structural dynamics? These are questions we have begun to address in our recent publication, showing that indeed it is possible to control SERCA's activity by tuning PLN structural dynamics[130]. We will continue to tackle these questions in the upcoming years.

Given the plethora of biochemical and molecular biology data currently available, it is a very exciting moment for structural biologists involved in research on PLN and SLN and their interactions with SERCA. While our studies to date have focused on monitoring the effects of PLN induced by SERCA, our future challenge involves detecting SERCA changes from PLN within the entire enzymatic cycle. Another important challenge is to study the SERCA-PLN complex in the complex network of interactions involving protein kinase A and protein phosphatase 1. While solution NMR will help identify some important pieces of this complex puzzle, solid-state NMR will be the method of choice for the elucidation of the structural dynamics and interactions in these large complexes.

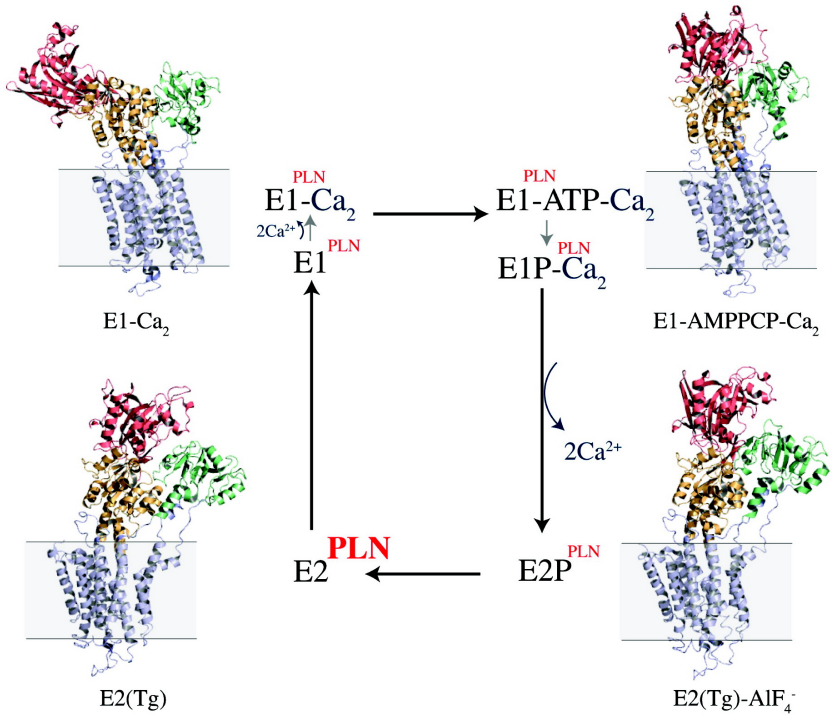


Figure 2.1: Proposed enzymatic cycle mimicking the four major conformational states of SERCA: E2 (1IWO)[85], E1-Ca2 (1SU4)[84], E1-ATP-Ca2 (1VFP)[86], and E2P (1XP5)[88]. PLN and SLN are believed to inhibit the E2 conformation as indicated within the model.

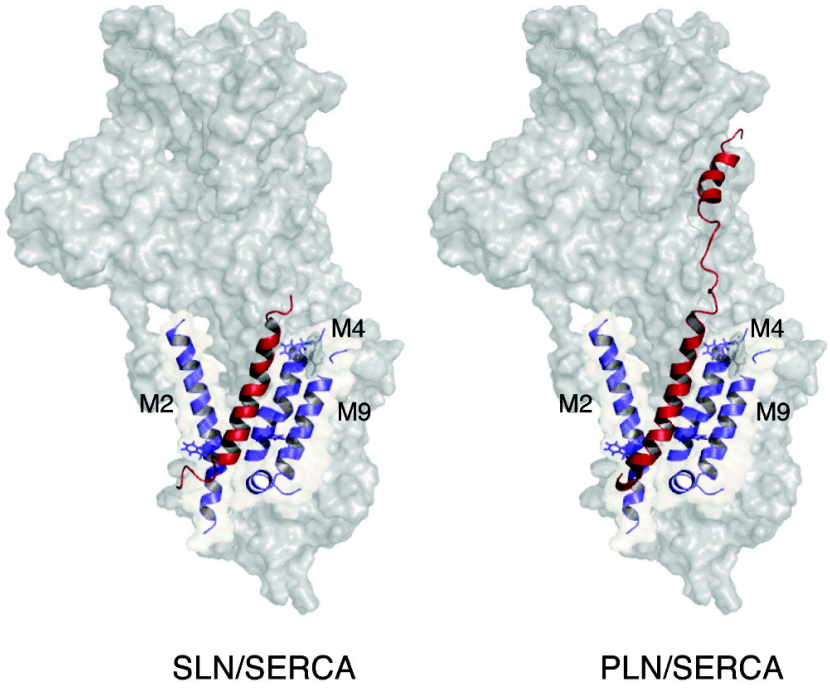


Figure 2.2: Molecular modeling of SERCA-SLN[92] and SERCA-PLN[91] complexes. Coordinates of complexes generously provided by D. H. Maclennan and C. Toyoshima.

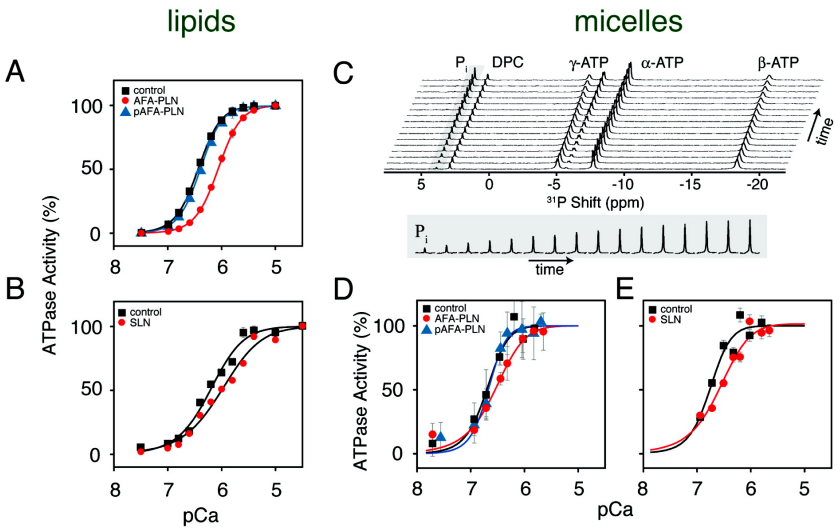


Figure 2.3: Activity assays of SERCA in the presence and absence of PLN and SLN in 4/1 DOPC/DOPE lipid bilayers (A and B) and DPC detergent micelles (D and E). SERCA activity in lipid bilayers was measured via the coupled enzyme assay and is reproduced with permission from ref[94]. Copyright 2003 Elsevier. The activity assays in DPC micelles were conducted using $^{31}P$ NMR spectroscopy (C) as previously reported[96]. Reproduced with permission from refs[131] and[132]. Copyright 2006 Elsevier.

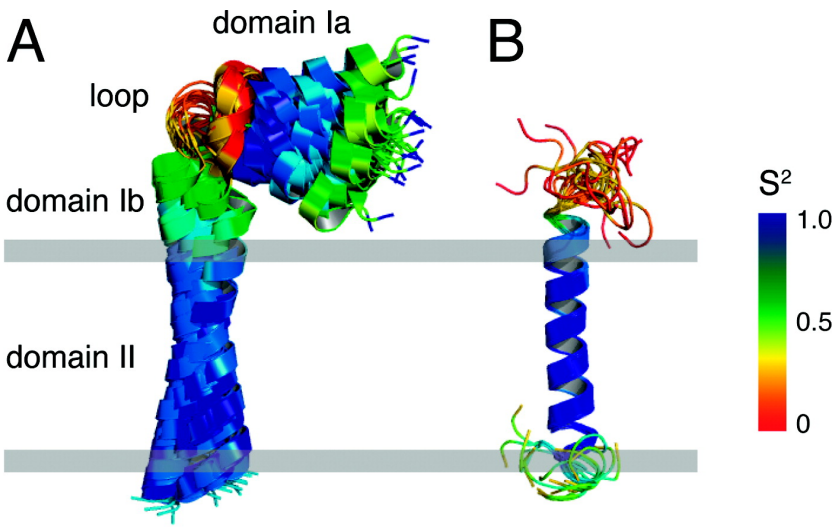


Figure 2.4: Solution NMR structure of PLN and SLN. (A) Backbone overlay of PLN from PDB entry 1N7L[99]. (B) Backbone overlay of SLN as reported by Buffy et al.[95]. The color coding on the structures corresponds to the order parameters ($S^2$) as previously determined[102, 103].

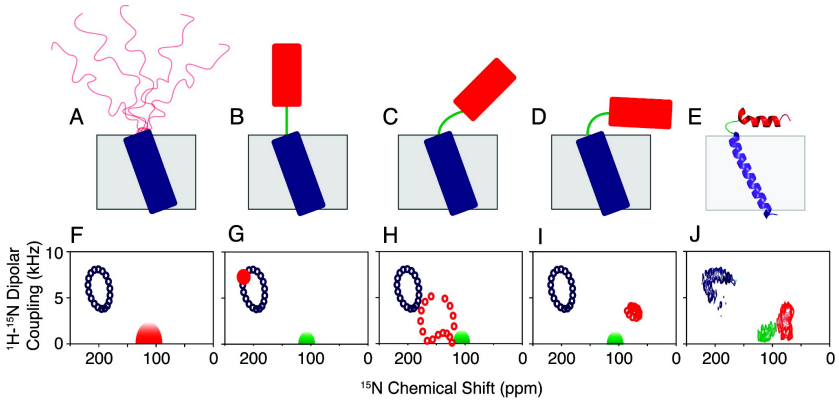


Figure 2.5: Structural models (A-D) proposed for PLN in lipid bilayers with PISEMA simulations (F-I). The ensemble of conformers reported by our laboratory (PDB entry 1N7L) is shown in panel E[99]. The colors are coded with the regions of the protein: red for cytoplasmic, blue for transmembrane, and green for loop regions. The L-shaped topology (D and E) agrees best with the experimental PISEMA (J). Reproduced with permission from ref[112]. Copyright 2006 American Chemical Society.

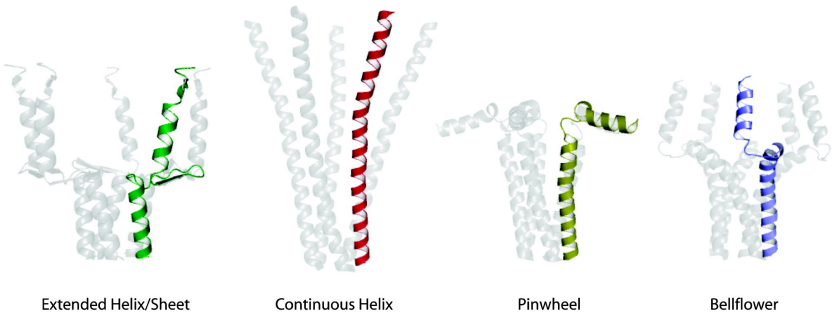


Figure 2.6: Structural models of wt-PLN. The pinwheel (1XNU) and bellflower (1ZLL) pentamer models were taken directly from PDB coordinates. Reproduced with permission from ref[106]. Copyright 2007 National Academy of Sciences of the United States of America.

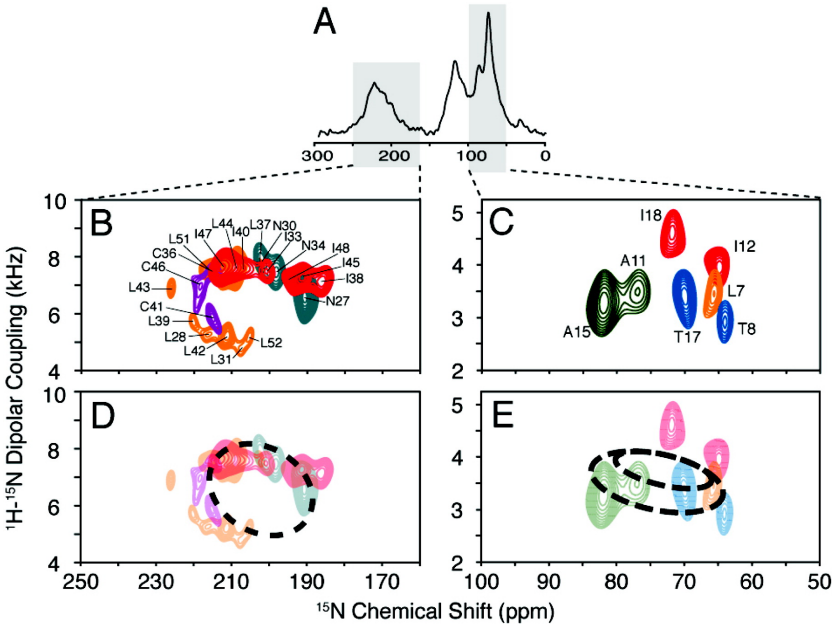


Figure 2.7: PISEMA spectra of the PLN pentamer in lipid bilayers. (A) 1D cross-polarization spectrum of [U-$^{15}N$]wt-PLN in DOPC/DOPE oriented lipid bilayers. (B and C) Selectively labeled wt-PLN PISEMA spectra for the transmembrane and cytoplasmic helices, respectively. The residues are color coded with the PISEMA spectra: green for [$^{15}N$-Ala], purple for [$^{15}N$-Cys], orange for [$^{15}N$-Leu], red for [$^{15}N$-Ile], gray for [$^{15}N$-Asn], and blue for [$^{15}N$-Thr]. (D and E) Simulated PISA wheels for both transmembrane ($\theta = 15°$) and cytoplasmic ($\theta = 92°$) domains. Reproduced with permission from ref[106]. Copyright 2007 National Academy of Sciences of the United States of America.

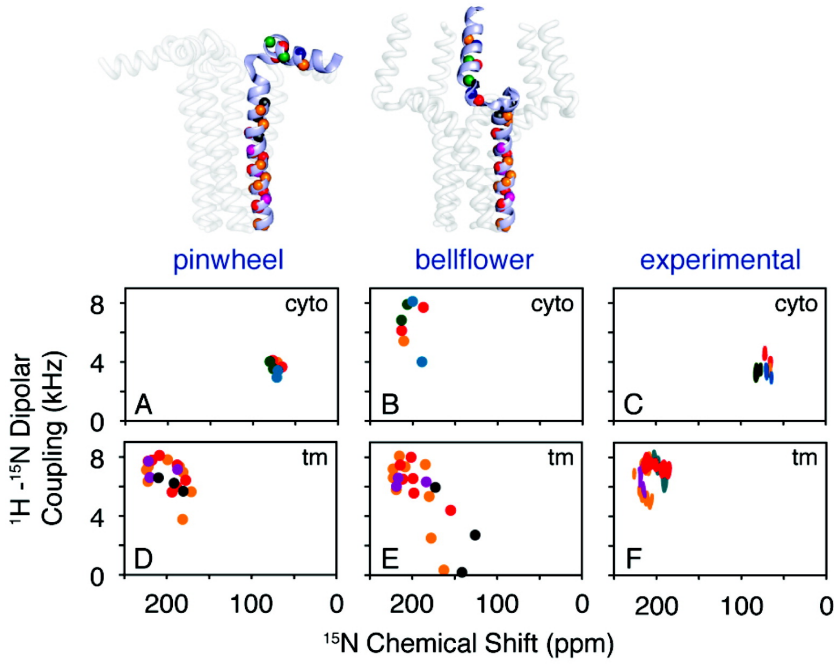


Figure 2.8: Simulated and experimental PISEMA spectra for the pinwheel and bellflower models. Unlike the pinwheel model, the bellflower model shows no high-field resonances. Experimental PISEMA spectra (C and F) show the remarkable agreement with the pinwheel model (A and D). Reproduced with permission from ref[106]. Copyright 2007 National Academy of Sciences of the United States of America.

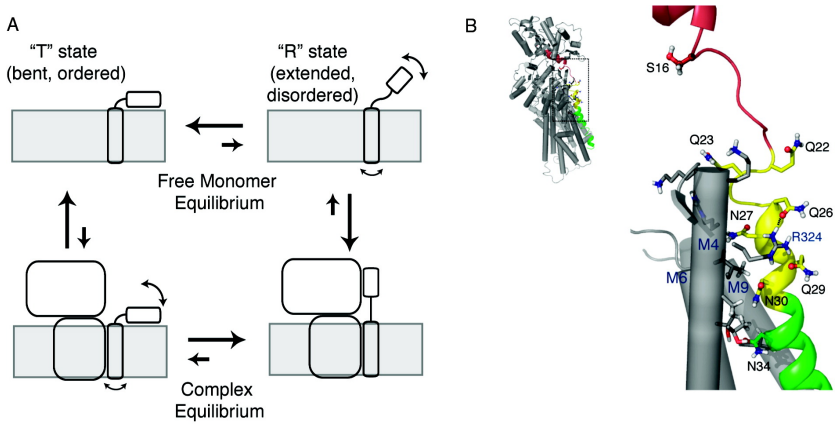


Figure 2.9: (A) Allosteric model of interaction of PLN with SERCA. The PLN monomer interconverts between the L-shaped form (T state) and the less stable (more dynamically disordered) extended form (R state). (B) PLN-SERCA model developed by Toyoshima et al.[91] highlighting the long-range allosteric control phosphorylation at Ser16 has on domain Ib. Panel A reproduced with permission from ref[126]. Copyright 2005 National Academy of Sciences of the United States of America. Panel B reproduced with permission from ref[96]. Copyright 2006 Elsevier.

# Chapter 3

# Tilt and azimuthal angles of a transmembrane peptide: a comparison between molecular dynamics calculations and solid-state NMR data of sarcolipin in lipid membranes



## 3.1 Introduction

Membrane proteins constitute ∼15%-30% of all genomes[1] and play an important role in many processes such as signal transduction, ion conduction, and transport of small molecules and proteins. To date, the Protein Data Bank (PDB) contains <200 unique membrane protein structures, compared with >36,000 structures of soluble proteins[134]. The disproportionally small number of membrane protein structures in the PDB is a result of several experimental challenges that hinder membrane protein

structural characterization using x-ray crystallography and conventional solution-NMR techniques. These challenges include 1), poor recombinant expression systems; 2), difficulties in obtaining well-diffracting crystals for x-ray; and 3), the large size of protein/lipid complexes. Careful reconstitution of membrane proteins in detergent micelles often results in well-behaving samples suitable for solution-NMR analysis using TROSY-based techniques[10]. Several outstanding examples have been reported in the literature and reviewed[7]. Many detergents can interfere with membrane protein function, but there are examples where membrane proteins and enzymes are still functional under detergent solubilization, allowing for the characterization of protein native states and protein-protein interactions[135, 132, 131, 136]. However, detergent micelles represent only a rough approximation of membranes and fall short in reproducing characteristics of membrane bilayers such as lipid-water interface, fluidity, dynamics, and curvature.

Together with x-ray crystallography, solution NMR, and cryo-electron microscopy, solid-state NMR is becoming another high-resolution method for structure determination of membrane proteins[19, 137, 138, 139, 140, 141]. This technique can be applied to membrane proteins reconstituted in lipid membranes, as the spectral resolution does not suffer as a result of the large size of membrane protein/lipid complexes. There are two major solid-state NMR approaches for determining the structure of membrane proteins: 1), magic-angle spinning NMR, using membrane proteins reconstituted in lipid vesicles; and 2), oriented solid-state NMR, using mechanically or magnetically oriented bilayer systems. Both methods can give site-specific resolution of protein resonances. Although the former can give high-resolution information also on the side-chain conformations, the latter has the advantage of simultaneously providing structure and topology of membrane proteins reconstituted in lipid bilayers[19, 142].

For oriented solid-state NMR, membrane proteins are reconstituted either on mechanically aligned glass plates or in magnetically aligned bicelle samples. $^{15}N$ and $^{13}C$ labeled sites in the protein backbone are used as reporters of orientationally dependent chemical shift anisotropy (CSA) and dipolar couplings (DC). Recent successes in the application of these approaches include membrane active peptides[143], single-pass membrane proteins[27, 26, 144, 145], membrane protein oligomers[146, 147], and

multispan membrane proteins[148, 149]. A similar approach has been pioneered by Killian and co-workers using orientationally dependent quadrupolar couplings measured on selectively $^2H$-labeled Ala residues along the peptide sequence (the so-called GALA (geometric analysis of labeled alanines) method)[150], whereas Ulrich and co-workers utilize nonperturbing $^{19}F$ substitutions along peptides or protein side chains to obtain $^{19}F$ CSA[151, 152].

A unifying element for all of these anisotropic properties is the wavelike patterns that they possess as a function of the residue number. These waves are two-dimensional projections of the protein's three-dimensional structure and orientation with respect to the lipid bilayers and can be converted into structural and topological restraints for structure determination[153, 154, 30, 155, 156]. As with all NMR parameters, these anisotropic chemical shifts and dipolar couplings are time-averaged properties and their values can be affected by protein internal dynamics. Although the dynamic effects are being analyzed and interpreted for residual CSA and DC measured in weakly aligned biomolecules[157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167], they have not been fully rationalized for membrane proteins in strongly aligned systems, with only a handful of studies carried out[168, 169, 170].

According to two recent studies[171, 172], scaling of quadrupolar couplings due to side-chain dynamics results in underestimation of the tilt angle of helical membrane peptides embedded in lipid membranes. Specifically, after analyzing molecular dynamics (MD) trajectories, it was concluded that there was a 30° discrepancy between the calculated and the experimental tilt angles. Both groups suggested that $^{15}N$ data obtained from separated-local-field experiments would give a more faithful estimation of both the tilt and azimuthal rotation angles of membrane proteins in lipid bilayers[171, 172]. A recent article by Koeppe and co-workers reconciles the experimental topology derived from both $^{15}N$ CSA and DC with that of the GALA method[173], highlighting the precision of both approaches in determining the membrane protein topology. This contrasts with work carried out by Watts and co-workers[169, 170], who found that fast structural fluctuations (rotation, librations, etc.) that occur in transmembrane peptides affect the appearance of the wavelike patterns (or polarization index slant angle (PISA) wheels)

derived from the separated-local-field experiments, jeopardizing both the assignment and interpretation of experimental data and raising more questions regarding the role of the dynamics[169]. To what extent do protein dynamics affect CSA, DC, and the wavelike pattern on aligned systems? Is this effect common to all proteins? Can MD calculations assist the resonance assignment process? To answer these questions, we have carried out molecular dynamics simulations on monomeric sarcolipin (SLN), which we have characterized extensively using both solution NMR in micelles and solid-state NMR in lipid bilayers[131, 27, 174, 71, 175]. SLN is a 31-amino-acid, single-pass membrane protein that regulates the sarcoplasmic reticulum Ca2+-ATPase (SERCA) in both fast- and slow-twitch muscle fibers by lowering the apparent affinity of SERCA for Ca2+ in a manner similar to phospholamban[176, 177]. The structures determined by solution NMR data in sodium dodecyl sulfate and dodecylphosphocholine micelles[131, 175] show that SLN comprises four different domains: a short unstructured N-terminus (residues 1-6), a short helical domain Ib (residues 7-15), a more dynamic helical domain II (residues 16-26) and the unstructured C-terminus (residues 27-31) (Figure 3.1). Solid-state NMR polarization inversion spin exchange at magic angle (PISEMA) experiments[178] carried out on selective labeled SLN reconstituted in mechanically aligned lipid bilayers show a well defined PISA wheel. Using a static model with an ideal helix to fit the experimental PISA wheel pattern, it was determined that monomeric SLN adopts a tilt angle of $\sim 23 \pm 2^{\circ}$ with respect to the membrane normal[27]. Starting from the structure determined in detergent micelles, we carried out MD simulations of monomeric SLN in pure 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) lipid bilayers and compared the calculated topology with that derived from solid-state NMR experiments (PISA wheel). Specifically, we analyzed the time evolution of the tilt and azimuthal rotation angle of SLN during 150 ns of molecular dynamics simulations The NMR parameters back-calculated from molecular dynamics trajectories are in agreement with those measured using solid-state NMR spectroscopy (separated-local-field experiments).

## 3.2 Theory and methods

### 3.2.1 Initial setup

The starting structure of SLN was selected from the conformational ensemble derived from a solution NMR study (PDB code: 1jdm)[175]. The protein was treated as a rigid body and equilibrated into a hydrophobic slab of adjustable thickness. The initial orientation of SLN was obtained from the OPM (orientations of proteins in membrane) database[179]. The depth of insertion of the transmembrane residues and the orientation of the protein were determined by minimizing the free energy of transfer from water to lipids[179]. The initial tilt angle obtained with this method was $1\pm7^{\circ}$. Subsequently, SLN was embedded in an implicit membrane using the generalized Born model with simple switching function (GBSW)[180], where the anisotropic membrane environment was modeled by different dielectric constants. The bilayer hydrocarbon core was represented by a low dielectric region ($\epsilon = 2$, thickness $\sim 30$Å). The bulk water region was approximated by $\epsilon = 80$ and the interface region (a 2.5Å slab) had a dielectric constant between those of the water and the hydrocarbon core. This was accomplished by using a simple switching function. The total membrane thickness was set to 35 Å, a value experimentally determined for the $L_{\alpha}$ phase[181, 182]. Molecular dynamics simulations were performed for 2 ns at a temperature of 298 K. The resulting tilt angle of SLN was $17\pm6^{\circ}$. This configuration was saved for subsequent MD simulations in explicit DOPC lipid bilayers. The calculations with explicit bilayers were performed using the program CHARMM, Version 33a2[34], utilizing the PARAM27 force field with the CMAP correction[183, 184, 185].

To build the DOPC bilayer, we used a protocol developed by Woolf and Roux[186, 187]. The lipid headgroups were initially represented by effective Lennard Jones spheres and were placed at +17.5 Å and -17.5 Å along the z direction[188, 189] with respect to the center of the bilayer to mimic a membrane thickness of 35 Å. After energy minimization, the Lennard-Jones spheres were replaced by prehydrated DOPC molecules taken from the CHARMM library. Lipid rotation and translation were successively applied during energy minimization steps to relax the system and to avoid close contacts.

Then, two preequilibrated hexagonal water boxes were placed at the top and the bottom of the lipid membrane containing SLN to obtain the desired level of hydration of the lipids. Successive minimization using steepest-descent and Newton-Raphson methods was applied to minimize the initial configuration of the SLN/DOPC system. The cross-sectional area of one DOPC phospholipid was 74 Å2. The entire bilayer system was constructed using 27 DOPC molecules for the top and the bottom layer, resulting in a total of 54 DOPC molecules in the xy (membrane) plane. Hexagonal periodic boundary conditions were used with a unit cell of 50 Å × 50 Å × 80 Å. The final system consists of a total of 18,090 atoms, including 549 atoms from SLN, 7452 atoms from the 54 DOPC lipids, 10,080 atoms from 3360 waters using the three-point-charge TIP3P model[190] and counter anions, which were added to neutralize the system and achieve an ionic strength near physiological conditions. To gradually relax the protein, harmonic restraints on the backbone and side chains of SLN were added and then slowly removed during the initial energy minimization and the early equilibration period using Langevin molecular dynamics. The SHAKE algorithm was applied to constrain the distances of all covalent bonds involving hydrogen atoms[191]. This allowed us to use an integration time step of 2 fs. To compute the nonbonded interactions, the van der Waals terms were evaluated using an atom-based cutoff scheme feathered to zero between 10 and 16 Å with a switching function. The particle-mesh Ewald (PME) method was used for the Coulomb terms, avoiding truncation of the long-range electrostatic interactions[192, 193, 194]. We used a charge mesh of 48*48*81 grid points, at a spacing of ∼1 Å, along with the Ewald parameter, $\kappa = 0.34$ Å$^{-1}$, and a direct interaction distance of 16 Å. A constant temperature of 310 K was maintained using the Nosé-Hoover chain[195, 196]. Pressure was kept constant at 1 atm along the z direction by Langevin piston Nosé-Hoover[197, 198] with the xy area fixed. To prevent translational drift of the protein, a cylindrical potential function was applied on SLN heavy atoms to remove the center-of-mass dynamics along the z axis for the first 4 ns. In addition, a planar potential function was applied on DOPC lipid headgroups to prevent the translational drift of the bilayer and at the same time keep the center of mass of lipids close to the xy plane. After 4 ns of MD simulation, all the constraints were removed and the system was further simulated up to a total of 150 ns. The first 10 ns were discarded and the last 140 ns were selected for the final analysis. The simulations and analyses were

carried out on the SGI (Sunnyvale, CA) Altix Cluster at the Minnesota Supercomputing Institute.

### 3.2.2 Analysis of fast motion (picoseconds to nanoseconds)

The generalized order parameter $S^2$ was calculated using the long-time tail of the second-order Legendre function of the correlation function for each NH vector ($\vec{v}$) of the protein backbone,

$$S^2 = \lim_{x \to \infty} C_2(t) \tag{3.1}$$

where

$$C_2(t) =< P_2(\vec{t}(\tau) \cdot \vec{v}(\tau + t) >_\tau \tag{3.2}$$

$P_2(x) = \frac{1}{2}(3x^2 - 1)$is the second-order Legendre polynomial. The angular brackets represent the ensemble average over the MD trajectories. The unit vectors $\vec{t}(\tau)$ and $\vec{t}(\tau + t)$ describe the orientation of NH bond vectors at time $\tau$ and $\tau + t$ with respect to a fixed reference frame. The reference frame is constructed by removing the translational and rotational motion in the MD trajectory by least root-mean-square (RMS) fitting of the backbone atoms onto the starting structure, which was performed using the ANALYSIS module of the CHARMM package[34].

Experimental spin relaxation data (R1, R2, and heteronuclear nuclear Overhauser effects) were taken from our previous work on SLN reconstituted in dodecylphosphocholine micelles[131]. The data were analyzed using the model-free approach, according to the Lipari-Szabo formalism[199]. The simulations were carried out using MODELFREE software package interfaced with FastModelFREE[200]. Model selection was performed by the fastModelFree software according to the method proposed by Mandel et al[201]. An axially symmetric diffusion model was able to fit the relaxation data for all of the protein amide sites. There was no difference between the order parameters obtained from the axially symmetric diffusion model and those from the isotropic overall diffusion model (data not shown).

### 3.2.3 Essential dynamics using principal component analysis

To extract the large amplitude motions of SLN during the MD simulations and project their effects on the PISEMA spectra, the molecular dynamics trajectories were analyzed using principal component analysis (PCA). PCA is a powerful method for extracting large-scale, low-frequency motions from MD trajectories[202, 203]. PCA is based on the diagonalization of a covariance matrix $C$, which contains the atomic fluctuation around the average atomic positions.

$$C_{ij} =< (r_i(t) - r_i^m)(r_j(t) - r_j^m) > \tag{3.3}$$

Here, $C_{ij}$ represents an element in covariance matrix $C$, $r_i$ is the position for atom $i$, and $r_i^m$ is the mean position. To focus on the intrinsic dynamics of protein SLN, translational and rotational motions are removed by superimposing each structure to the reference configuration by means of least-RMS deviation (RMSD) before analysis. We obtain a set of eigenvalues and the corresponding eigenvectors by diagonalizing $C$. The eigenvectors correspond to the quasiharmonic motions of individual modes and the eigenvalues represent the mean square fluctuations along the eigenvectors. The eigenvalues are usually selected based on their amplitude and referred to as principal components (PCs). The first (lowest) eigenvalue corresponds to the largest fluctuation of the system along the first eigenvector. In fact, the overall dynamic fluctuations of the protein can be described by just a few essential eigenvalues with large amplitudes.

It is possible to project the structure on to an eigenvector using the relationship.

$$q_l = (\vec{r} - < \vec{r} >) \cdot \vec{\eta_l} \tag{3.4}$$

where $q_l$ is the displacement of the structure along the $l$th eigenvector with respect to the average structure, $\vec{r}$ is the coordinate vector of a given structure, $< \vec{r} >$ is the vector of the average structure, and $\vec{\eta_l}$ is the eigenvector. The 3-D structure corresponding to a displacement along a single eigenvector can be calculated using

$$\vec{r} = q_l \vec{\eta_l} + < \vec{r} > \tag{3.5}$$

This provides a way to visualize the quasiharmonic motions of a specific mode, represented by the average structure. For PCA analysis of SLN, we used the software Bio3d, which was developed for comparative analysis of protein structures[204].

### 3.2.4 Back calculations of PISEMA spectra from MD trajectories

The formalism developed by Cross and co-workers was used to calculate PISEMA spectra[22]. In PISEMA experiments, both chemical shift anisotropy (CSA) and dipolar coupling (DC) tensors are measured at a fixed orientation with respect to the direction of the static magnetic field $\vec{B_0}$. The CSA tensor is asymmetric and its principal axis frame (PAF) can be described in terms of its principal components $\vec{\sigma}_{11}$,$\vec{\sigma}_{22}$,$\vec{\sigma}_{33}$ as follows:

$$PAF = (\vec{\sigma}_{11}, \vec{\sigma}_{22}, \vec{\sigma}_{33}) \tag{3.6}$$

where diagonal elements are arranged such that $\sigma_{33} > \sigma_{22} > \sigma_{11}$. The scalar value of the CSA tensor can be rewritten as a function of its principal components as

$$\sigma = \sigma_{11}(\vec{B_0} \cdot \vec{\sigma}_{11})^2 + \sigma_{22}(\vec{B_0} \cdot \vec{\sigma}_{22})^2 + \sigma_{33}(\vec{B_0} \cdot \vec{\sigma}_{33})^2 \tag{3.7}$$

The DC tensor $\nu$ is traceless and axially symmetric, with a unique rotation axis ($\vec{u}$) around the covalent NH bond, and can be expressed as

$$\nu = \frac{\nu_{||}}{2}(3(\vec{B_0} \cdot \vec{u})^2 - 1) \tag{3.8}$$

where $\nu_{||}$ is the value of the DC tensor when $\vec{B_0} = \vec{u}$, i.e., it occurs when the NH bond vector is in line with the direction of the magnetic field. In the MD trajectory, $\vec{B_0}$ is set to a unit vector along the $z$ axis that is parallel to the membrane normal. If $(x, y, z)$ are the coordinates for the unit vector $\vec{B_0}$ in the PAF, then DC and CSA can be expressed as:

$$\sigma = \sigma_{11}x^2 + \sigma_{22}y^2 + \sigma_{33}z^2 \tag{3.9}$$

$$\nu = \frac{\nu_{||}}{2}(3(\cos\alpha \sin\beta \cdot x + \sin\alpha \sin\beta \cdot y + \cos\beta \cdot z)^2 - 1) \tag{3.10}$$

where $\vec{u}$ is the direction of the unit NH bond vector, $\alpha$ is the polar angle made by $\vec{u}$ and the peptide plane, and $\beta$ is the polar angle between $\vec{u}$ and $\sigma_{33}$. In the above equations, we have the relationships $x^2 + y^2 + z^2 = 1$ and $\vec{\nu}_{||} = A\gamma_1\gamma_2/r^3$, where $\gamma_1$ and $\gamma_2$ are the gyromagnetic ratios of nuclei 1 and 2, $r$ is the distance between the two atoms, and $A$ is a constant equal to $\mu_0 h/8\pi^3$, with $\mu_0$ being the vacuum permeability and $h$ being Planck's constant. Assuming a planar peptide structure, we have $\alpha = 0°$ and $\beta = 17°$, which is determined experimentally[23]. The principal components of the CSA

tensor are obtained from the experimental PISEMA powder pattern and assumed to be the same for all residues ($\sigma_{11}$ = 57.3 ppm, $\sigma_{22}$ = 81.2 ppm, $\sigma_{33}$ = 228 ppm)[137]. The coefficient for the DC is 10.74 kHz, with the NH bond distance equal to 1.042 Å[137]. These numerical values are nontrivial approximations employed to simplify the calculations (different residues usually have different values of CSA tensors). The NH bond length used in CHARMM force field is 0.99 Å[34].

## 3.3 Results

### 3.3.1 Interactions between SLN and its environment

Molecular dynamics simulations of SLN embedded in DOPC bilayer in water were performed for a total of 150 ns. Figure 3.2 A shows the density profile of the lipid bilayer principal structural groups along the z direction after the initial 4 ns of MD calculations. The time-average distribution profile for DOPC bilayer in the presence of SLN corresponds to a typical liquid crystalline ($L_\alpha$) phase[205]. Under these conditions, the lipid bilayer is divided into interfacial and hydrocarbon regions, with the hydrocarbon center populated by methyl, methylene, and unsaturated groups of the oleoyl moieties. The interfacial region, on the other hand, is dominated by the presence of glycerol, phosphate, and choline groups, with the water molecule profile intercalating into the bilayer core up to the unsaturated carbon of the oleoyl moieties[205].

Based on analysis of the primary sequence, SLN displays the characteristic properties of transmembrane proteins, including a cluster of hydrophobic residues, charged amino acids flanking the transmembrane domain, and an aromatic belt that anchors the protein to the bilayer. To analyze the distributions of these residues during the MD simulations, the amino acids were classified as hydrophobic and hydrophilic based on the hydrophobicity profile by Wimley and White[206]. The distributions of the amino acids within the DOPC bilayer are reported in Figure 3.2, B-D. As expected, the hydrophobic amino acids populate the inner core of the membrane bilayer, whereas the charged residues located at both termini populate the water/headgroup interfaces. R6 and R27 are distributed around the interfacial region, forming salt bridges with the lipid headgroups. T5 (the putative phosphorylation site of SLN[207]) forms hydrogen bonds

with a lipid headgroup near the glycerol/choline group interface. N11 leans toward the lipid headgroups, making a hydrogen bond with a water molecule (Figure 3.3 A). It is interesting that T13 and T18 are distributed in the center of the hydrocarbon region of the lipid bilayer and their polar hydroxyl groups form hydrogen bonds with the backbone carbonyl groups of F9 and I14, respectively (Figure 3.3 B). The propensity of threonine residues to form hydrogen bonds with the protein backbone has been pointed out already by Johansson and Lindahl[208]. The aromatic side chains of F9 and F12 located toward the N-terminus intercalate in the lipid apolar region, keeping the average orientation of the aromatic rings parallel with respect to the lipid chains. This configuration is common for phenylalanine residues of membrane-embedded polypeptides and maximizes their packing efficiency within the lipid membranes[208, 209, 210]. The tryptophan side chain is located in the other leaflet of the membrane bilayer, with its polar indole group directed toward the interfacial region (Figure 3.3 C), though it did not form any interactions with the polar headgroups throughout the simulations. It is interesting that R27 and W23 seem to be stabilized by cation-$\pi$ interactions. This is indicated by the close distances of the arginine side chain and the centroid of the tryptophan ring (Figure 3.3 D). To calculate these distances, we utilized the convention described by Nielsen and co-workers[211]. Y29 can also form cation-$\pi$ interactions with the positively charged nitrogen atom in the choline group (Figure 3.3 C), anchoring the C-terminus of SLN to the lipid bilayer. Note that these aromatic residues are believed to be necessary for membrane localization of SLN in the sarcoplasmic reticulum and for interaction with SERCA[212].

### 3.3.2 Structural dynamics and topology of SLN in lipid bilayers

According to NMR analysis in detergent micelles, SLN can be partitioned into four dynamic domains: two mobile termini (N-terminus, 1-6, and C-terminus, 27-31), a short dynamic helix domain Ib (residues 7-14), and a more rigid helix domain II (residues 15-26)[131]. The structural changes of SLN throughout the MD simulation were analyzed by examining the histogram of the RMSDs of each domain as functions of time (Figure 3.4 A). It is apparent that the system has reached a stable conformation for an extended period of time during MD simulations. After 10 ns, the SLN backbone showed RMSD

values fluctuating ∼1.6 Å, indicating significant structural changes. This is primarily due to the flexible N-terminus that undergoes significant migrations near the lipid headgroups and converges to a set of stable configurations. In contrast, the two helical domains, Ib and II, and the C-terminal portion of the protein are stable throughout the simulations, with domain Ib slightly more flexible than domain II. This is consistent with both relaxation data and solvent-accessibility measurements carried out using solution NMR in DPC micelles[131]. The time-averaged RMS fluctuations (RMSF) for each residue give a more complete picture of the local protein flexibility (Figure 3.4 D). Overall, the RMSF values of the side chains are greater than those for the backbone atoms, indicating higher flexibility. The side chains for both termini are more flexible than those in the central portion of SLN, comprising residues 13-20. It is of interest that the side chains for residues in domain Ib show periodic oscillation, with R6, F9, and F12 displaying greater mobility. In general, it is expected that longer side chains would display higher RMSF values than shorter side chains (Ala, Gly). However, residues E7, L8, L10, and N11, which possess long side chains, show substantially lower RMSF values relative to those of R6, F9, and F12. Since these residues are positioned in the same face of the domain Ib helix, it is possible that one face of the helix is more dynamic than the other.

To analyze the topology during MD trajectories, the helical domain of SLN was first considered as a continuous helix and the topology was analyzed by reporting the azimuthal angle, $\rho$, and the overall tilt angle, $\theta$. The SLN helical axis was defined using the convention described by Aqvist and co-workers[213]. The overall SLN tilt and azimuthal angles are reported in Figure 3.5. As mentioned in the Methods section, the molecular dynamics carried out in the implicit solvent resulted in an overall tilt angle of 17±6°. After an abrupt initial excursion (during the first 5 ns), the SLN tilt angle oscillates around 28±6°. The calculated tilt is in good agreement with the experimental data, derived from PISEMA experiments. In fact, from the structural fitting of the experimental PISA wheel with a static ideal helix model, it was found that $\theta$ is 23±2°. Note that the MD simulations were carried out in pure DOPC, whereas the solid-state NMR experiments were performed with a mixture of DOPC/DOPE. It has been reported that the closer packing of the PE causes an increase of bilayer thickness

that in turn would cause SLN to tilt$< 25°$[214]. In agreement with previous studies[150, 215, 216], this could cause an increase in the computed tilt angle of SLN. Although the MD simulations described here were performed for 150 ns, this is still short with respect to the timescales probed by the PISEMA experiments. Improved statistical results may be obtained by running multiple 150-ns simulations or by significantly extending the overall computational time. Finally, the experimental tilt angle was determined using a limited number of amide sites and an ideal helix to fit the overall spectral pattern[27].

A parameter important for defining the topology within lipid bilayers is the azimuthal (or rotation, $\rho$) angle of SLN around its helical axis. The azimuthal angle of a helical protein in lipid bilayers is an index of the specific interactions occurring between the protein side chains and the lipid membranes that can be obtained from the PISA wheel derived from PISEMA experiments[203, 28]. To calculate $\rho$, we choose the starting configuration of SLN as a reference frame and use a method similar to that of Killian and Delgado[171, 172], as depicted in Figure 3.5 A. Figure 3.5 A shows the $\sim$150-ns time evolution for the azimuthal angle. The initial value of the azimuthal angle is $\sim 20°$. After 25 ns, there is a sharp transition and the angle continues to change monotonically up to an average value of 66°. In the last 100 ns of MD simulations, the azimuthal angle fluctuates around 66°. It is noteworthy that the motions of the N-terminal and C-terminal residues appear to drive the final topological arrangement of the protein. The change in the azimuthal angle is more clearly represented in Figure 3.6, where the starting structure, the ending snapshot and the ideal helix model derived from PISEMA experiments are compared. The starting configuration of the MD calculations rotates $\sim 66°$ to reach a configuration almost superimposable on that of the solid-state NMR model[27]. In the initial configuration, two of the N-terminal hydrophobic residues (M1 and I3) are unfavorably placed near the polar headgroups, whereas the two C-terminal tyrosine residues are located in the bulk water phase. After 60 ns, these residues are embedded into the headgroup/hydrocarbon interface of the bilayer and remain in this environment throughout the rest of the trajectory. These residues are known to be responsible for the SLN location in the sarcoplasmic reticulum membrane[212]. The cation-$\pi$ interaction between R27 and W23 occurs at $\sim$50 ns, which is concomitant with the transition of the azimuthal angle from $\sim 20°$ to $\sim 50°$. This suggests that together

with the hydrogen bonds formed by the N-terminus with the lipids, these cation-$\pi$ interactions may constitute a driving force for SLN helix rotation. At the same time, the helical axis of SLN rotates to optimize the exposure of the most hydrophilic residue side chains (N4, T5, and N11) toward the lipid polar headgroups. In the last 100 ns of MD simulations, most of the polar residues of the N-terminus are anchored to the lipid headgroup region, forming electrostatic interactions to stabilize the SLN topology. The resulting azimuthal angle is in good agreement with the experimental results, with the face containing I3, L8, V15, L16, V19, and I26 pointing toward the sarcoplasmic reticulum lumen[27].

### 3.3.3 Analysis of fast backbone dynamics

The fast dynamics of the SLN backbone (on the picosecond to nanosecond timescale) was analyzed by calculating the generalized order parameter ($S^2$). Since the experimental data are not available for SLN in lipid bilayers, we compared the computed $S^2$ value with experimental values obtained in detergent micelles. The computed and experimental $S^2$ order parameters are compared in Figure 3.7. Although there is good agreement among experimental and theoretical values for the helical region of SLN, our simulations show that the terminal residues display higher $S^2$ values than the experimental data. Spin-relaxation data in detergent micelles revealed four different dynamics regions of SLN, mimicking the motions of the transmembrane domain of phospholamban[217], the SLN homolog in the cardiac muscle[176, 177]. The $S^2$ values obtained from MD simulations are more uniform, with an average value of 0.89$\pm$0.04 from residues 5-30. For comparison, the experimental values show uniformity from residues 7-27 (0.91$\pm$0.05). The slight difference between the calculated and experimental values may be attributed to the higher fluidity of the micellar environment causing higher fluctuations in the experiment compared with the lipid bilayers used in dynamics simulations. A similar observation was made by Sansom and co-workers, who compared the dynamics of secondary-structure domains in both micellar systems and lipid bilayers[218]. These researchers showed that the dynamic flexibility is greater in a micellar environment than in lipid bilayers, attributing the differences in mobility to the fast diffusion of detergent molecules and their reduced packing in the micellar core with respect to the lipid

bilayers. Another important factor that may cause these discrepancies is the approximations used to calculate the electrostatic energy in the MD simulation and the lack of polarizable force field.

### 3.3.4 Analysis of the large-scale dynamics by PCA

The essential dynamics of SLN was determined using the protein configurations from the MD trajectories from 10 to 150 ns. These structures have been translated and superimposed with the initial (reference) structure to remove lateral diffusional motions. The backbone atoms from these configurations were rewritten into a new trajectory file in every 40-ps interval, with which the principal component analysis was performed using Bio3d[204]. The first and second components (PC1 and PC2) represent the rotation of SLN around the helix axis and the dynamics of the tilt angle with respect to the lipid bilayer normal. The third component (PC3) features a smaller amplitude of rotation of SLN along the helix axis. PC1 has a larger N-terminal fluctuation than PC3. The first three principal components represent ∼86% of the overall structural fluctuations of the entire protein during the MD simulations. In Figure 3.8 A, we have projected the structures of the backbone atoms (N, $H_N$, $C_\alpha$, C', and O) in the MD trajectories onto the essential space (planes) defined by PC1/PC2, PC1/PC3, and PC2/PC3. This allows one to visualize the conformational space sampled during MD calculations. In the plots of Figure 3.8 A, each point represents one conformation of SLN saved during the MD simulations, and the density of points is an indication of the population of conformations sampled in the molecular dynamics trajectory. The proportion of variance versus the eigenvalue rank (Figure 3.8 B) clearly shows the relative importance of the different motions, with PC1 accounting for ∼59%, PC2 for ∼19%, and PC3 for only ∼8% of the overall fluctuations. Since these results are based on 150 ns of MD simulations they only indicate the essential motions in the nanosecond timescale. Nevertheless, the PCA analysis provides a glimpse of the key features of the low-frequency large-amplitude motions of SLN in lipid bilayers.

### 3.3.5 Computation of the PISEMA spectra and the effect of dynamic averaging

To calculate the PISEMA spectra, we used the method described by Strauss and co-workers[169], where the CSA and DC values for each residue were back-calculated from SLN snapshots from the MD trajectories. When the PISEMA spectra are calculated from instantaneous (snapshot) conformations of SLN sampled during every 1-ps interval during the MD simulation, each amide site gives rise to a wide distribution of both CSA and DC, as illustrated in Figure 3.9 A. This has also been found in the simulation carried out by Straus and co-workers[169, 170], who suggested that the fast motions of the amide vectors (libration, etc.) affect the appearance of the computed spectrum. The spread of the CSA and DC values around the average values for each site approaches an anisotropic Gaussian distribution with a distinct maximum (Figure 3.9 B). To illustrate the effects of fast motions involving the amide hydrogen, we performed a series of short-time averages of the out-of-plane torsion angle of the peptide plane at different time intervals, and we found that using this approach, it is possible to reduce the thermal noise caused by the fast dynamics of the amide vector during the MD simulations. Depicted in Figure 3.9, C and D, are the results for residue 25 that were obtained using the average amide vector at 10- and 20-ps intervals, respectively. This short-time average of fast motions reduces the original data points (Figure 3.9 B) from 140,000 to 14,000 (Figure 3.9 C) and 7000 (Figure 3.9 D) for 10- and 20-ps intervals, respectively. We found that conformation averaging over time periods>20 ps does not further narrow the distribution significantly.

The calculated PISEMA spectrum and the agreement between experimental and theoretical DC and CSA are shown in Figure 3.10. The agreement between the experimental and calculated CSA and DC is good. However, the computed average CSA and DC values are somewhat smaller than those experimentally determined, corresponding to a larger tilt angle of SLN. The calculated PISEMA spectrum shows a distinct PISA wheel pattern, although the distribution of DC and CSA data around the average value is greater than the line widths obtained experimentally. The average standard deviations (error bars) between experiment and theory are 15.8±4.2 ppm for CSA and 1.6±0.4 kHz for DC. We also tested the agreement between experimental and calculated data using

different CSA tensor values taken from the work of Wu et al[23] (tensor 1), Straus et al[169] (tensor 2), and Page et al[137] (tensor 3). Figure 3.11 shows the calculation of the PISEMA patterns using the three tensors. The different values of the CSA tensors have little effect on the average CSA values calculated from the MD trajectories. As previously mentioned, the systematic deviations observed for both CSA and DC values from the experimental values could be due to the use of different lipids, hydration levels, etc. (i.e., bilayer thickness) in the experiments and in the MD simulations[214].

To understand the effects of slow motions on the calculated PISEMA spectra, we used PCA analysis to probe the low-frequency motions during the MD trajectories. Using the principal-component vectors, we projected the essential motions in the PISEMA plane. Figure 3.13, A-D, in the Supporting Material, shows the effects of the three main mode of motion as identified by the PCA analysis on the back-calculated PISEMA spectra. For clarity, we eliminated the residues corresponding to the two termini, which are the most mobile, and focused on the analysis of the residues in the helical portion of the PISEMA spectra. The combination of the PCA components is reported in Figure 3.12. We have also investigated the effects of mosaic spread (i.e., spatial distribution of conformations) on the PISEMA experiments. The mosaic spread was calculated using a uniform distribution of angles according to the method suggested by Cross and co-workers[168]. We modeled the effects of spatial distribution around 4, 10, and 20°. It was interesting to find that the 10° mosaic spread (Figure 3.12 B) would have an effect on the linewidths similar to that generated by the combined essential motions (PC1, PC2, and PC3 (Figure 3.12 A)). This suggests that a spatial distribution of the polypeptide conformations>10° would be the dominating factor in the observable linewidths.

## 3.4 Discussion

Molecular dynamics simulations at the molecular level have reached a state of sophistication such that it is possible to reproduce the physical chemical properties of proteins and membranes and, at the same time, to describe the details of protein-lipid interactions[219]. In particular, MD calculations have been utilized to interpret

and predict solid-state NMR parameters derived from membrane proteins reconstituted in oriented lipid bilayers such as $^{2}H$ quadrupolar couplings, $^{15}N$ and $^{13}C$ CSA and DC[185, 220, 221, 222, 223]. For instance, Cross and co-workers applied a combined MD/NMR approach to obtain the first high-resolution structure of a membrane protein in lipid bilayers[224].

Subsequently, Roux and co-workers applied this combined MD/NMR approach to understand the structural dynamics of several other membrane proteins and peptides in various lipid environments[220, 223]. More recently, MD simulations have been used with success to study the structural dynamics and membrane interactions of antimicrobial peptides[225, 226]. From these studies emerged excellent agreements between experimental and calculated NMR parameters (CSA and DC), which made it possible to reproduce the topologies and lipid interactions of these membrane-embedded polypeptides.

Recent reports[171, 172], however, have pointed out that dynamic averaging of $^{2}H$ quadrupolar couplings can cause substantial errors in the calculation of the tilt and azimuthal angles of small transmembrane peptides, suggesting that $^{15}N$ CSA and DC obtained from separated-local-field experiments would provide more accurate values of these angles. This view contrasts with a recent article from Koeppe and co-workers showing that the interpretation of $^{15}N$ CSA and DC and the quadrupolar couplings obtained from the GALA approach produce the same structural topology[173]. However, protein dynamics affects the appearance of the NMR spectra. In a recent MD simulation study, Straus and co-workers concluded that dynamics affects the appearance and the interpretation of CSA and DC on both helical and $\beta$-barrel membrane proteins[169, 170].

To analyze the effects of dynamic averaging on helical polypeptides, the dynamics of SLN were characterized using 150 ns of MD simulations in explicit DOPC lipid bilayer. This small membrane protein has been extensively studied in both SDS and DPC micelles[131, 175]. In addition, its topological orientation (tilt and azimuthal angles) was determined in DOPC/DOPE lipid bilayers by solid-state NMR[27]. From the analysis of the MD trajectories, we found that SLN maintains its helical conformation throughout the simulations, with both termini adopting a more helical structure than

experimentally observed in both DPC and SDS[131, 175]. This can be attributed to the membrane environment that stabilizes the SLN secondary structure more than the micellar environment does. This prediction awaits for an experimental validation.

It was interesting to find that T13 and T18 are positioned in the hydrocarbon region of the membrane bilayers. In the monomeric form of SLN that we have studied by MD simulations, these residues are hydrogen-bonded with the polypeptide backbone for SLN. However, they might be crucial for SLN oligomerization at high protein/lipid ratios[227] and for ion-conducting activity we recently measured in mercury-supported lipid bilayers[228]. Taken with previous studies on micelles and liposomes[227], these data support the hypothesis that SLN is able to form ion-conducting pores permeable to small inorganic anions such as chloride, phosphate, or sulfate and impermeable to cations such as sodium and potassium[228].

The topology of SLN converges and oscillates around average values that are in close agreement with those determined using solid-state NMR. If the main driving force for SLN solubilization by the lipid bilayer is the hydrophobic mismatch, these simulations clearly show that both the N- and C-terminal residues determine the rotation (azimuthal) angle. These conclusions agree with previous work carried out by Di Nola and colleagues with model membrane proteins embedded in implicit membrane environments[229]. However, simulations in explicit membrane environment converge into a slightly different topology for SLN. Previous studies using Monte Carlo folding and an implicit generalized Born membrane reported that SLN is less helical with a tilt angle of $16\pm8^{\circ}$[230]. Indeed, using CHARMM and implicit model membranes, we obtained approximately the same result ($\theta = 17\pm6^{\circ}$). In a similar way, Brooks and co-workers found for the fd coat protein only a semiquantitative agreement between solid-state NMR data and MD simulations when using replica-exchange molecular dynamics and the implicit-membrane generalized Born model[231]. In that work, the simulations of SLN in explicit DOPC bilayer resulted in tilt and azimuthal angles in closer accord with those determined experimentally. This emphasizes the importance of specific interactions between the explicit lipid membrane and SLN.

In agreement with a previous MD study by Straus and co-workers[169], our simulations show that fast dynamics (on the picosecond to nanosecond timescale) of the amide backbone sites dramatically affects the appearance of back-calculated PISEMA spectra, resulting in broad distributions of peaks, which makes it difficult to observe regular PISA wheels[169]. These fast fluctuations involving the amide hydrogen (such as amide libration motion) are more rapid then the time resolution of the PISEMA experiments. We found that the fast fluctuations can be largely filtered out by averaging the torsion angle of the amide out-of-plane mode over 10- to 20-ps intervals along the MD trajectory. As a result, it is possible to match the experimental results and obtain reasonably regular oscillatory patterns for both CSA and DC (i.e., the PISA wheel).

To understand the effects of the low-frequency, large-amplitude motions characterized by essential component motions, we projected the PC1, PC2, and PC3 trajectories onto the PISEMA spectra. We found that these slow dynamical modes have a pronounced effect on the back-calculated values of DC and CSA values, exhibiting narrower spectral lines than those obtained from the full dynamics. As for the fast amide-vector dynamics, these motions are also much faster than the timescale of the PISEMA experiments, which are also averaged out under experimental conditions (i.e., slow and fast dynamics are rapid compared to the time resolution of the PISEMA experiment). We also compared the effects of protein dynamics with the mosaic spread. For SLN, we concluded that a spatial distribution of conformations of$>10^\circ$ would dominate the observable linewidths.

Although both dynamics and spatial distribution affect the linewidths, the average values computed by CSA and DC for SLN oscillate with 3.6 periodicity according to its helical conformation[30, 22, 232, 233]. It is noteworthy that the slight bend that the MD predicts for SLN causes a slight deviation in the wheel pattern (Figure 3.10) at residues 7-14, whereas residues 15-26 follow the canonical PISA wheel with a pitch of $\sim$3.6 residues per turn. This indicates that the most rigid portion of SLN on average assumes a helical conformation that is close to ideal. This is in agreement with the uniformity and ideality of strongly hydrogen-bonded transmembrane polypeptides proposed by Cross and co-workers[142, 234].

## 3.5 Conclusions

In conclusion, the MD simulations of SLN in explicit DOPC lipid membranes show that the topology of this small transmembrane protein converges to that experimentally determined by solid-state NMR in oriented lipid bilayers, with similar tilt and azimuthal (rotational) angles. From the MD trajectories, it is possible to back-calculate PISEMA spectra (CSA and DC for each amide site), which are in good agreement with the experimental results. The fast and slow motions (picosecond to millisecond) are rapid compared to the time resolution of the PISEMA experiment, and are averaged out. As a consequence, the back-calculated spectra of the helical portion of SLN from the MD trajectory reveal the classical PISA wheel pattern. This demonstrates that MD trajectories can reproduce the topology of single-pass membrane proteins and can be used to predict and help in the interpretation of solid-state NMR data.

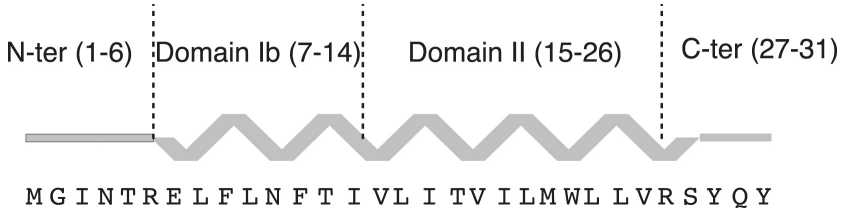


Figure 3.1: SLN primary sequence and definition of dynamic domains according to Buffy and co-workers[131]

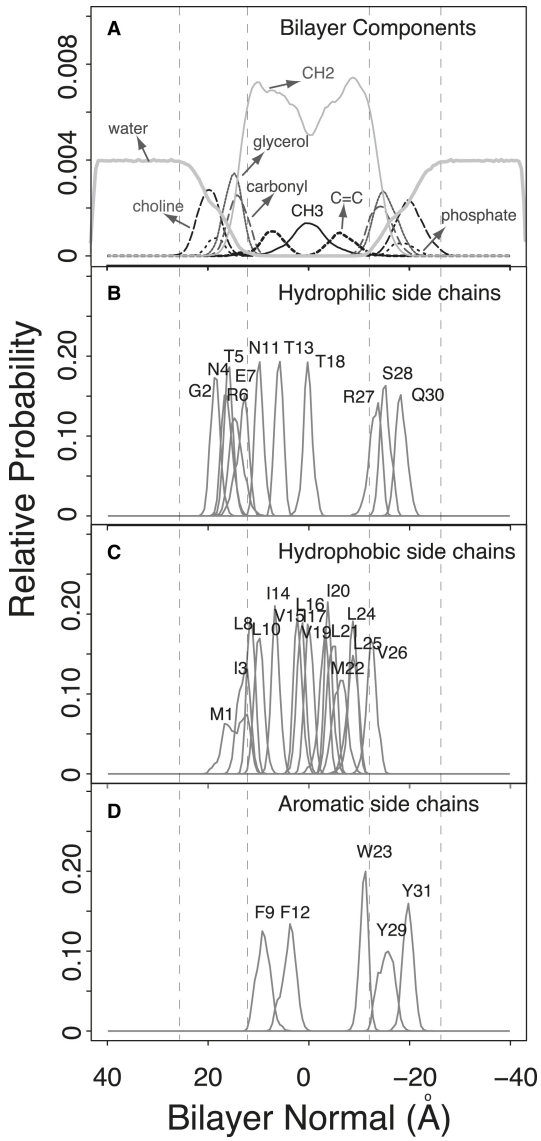


Figure 3.2: Bilayer structure distribution. (A) Atomic density profiles of the bilayer components along the bilayer normal. (B-D) Distribution of the centers of mass of the hydrophilic (B), hydrophobic (C), and aromatic (D) side chains of SLN amino acids. The partition between hydrophilic and hydrophobic was carried out according to the hydropathy scale by White and co-workers[206].

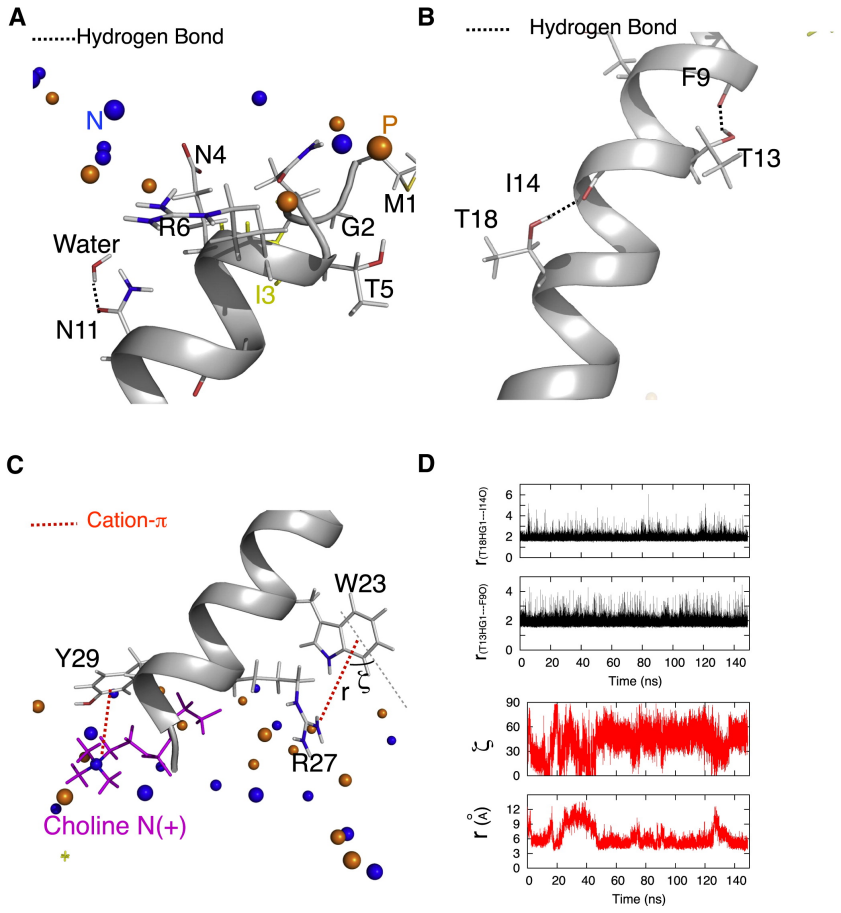


Figure 3.3: Detailed interaction of SLN during MD. (A) Snapshot (95 ns) of the SLN N-terminal portion showing the interactions between N11 with water molecules. The I3 sidechain points toward the hydrophobic core. (B) Snapshot of domain Ib showing the hydrogen bonds formed by T13 and T18 with the SLN backbone carbonyls of F9 and I14. (C) Snapshot of the C-terminal domain with cation-p interactions between R27 and W23, and of the choline group with Y29. (D) Plots of the distance versus time between g protons of the T13 and T18 side chains and F9 and I14 carbonyl groups (upper panels), and of distance r between the R27 side-chain N and the center of mass of the W23 aromatic ring, and the angle z between r and the aromatic ring.

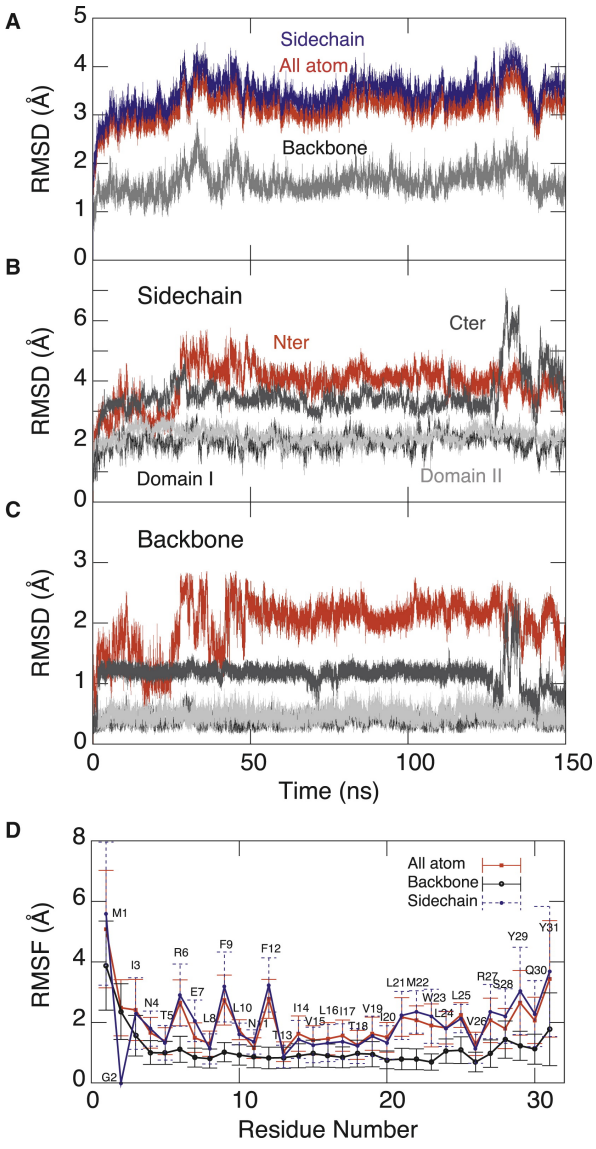


Figure 3.4: RMSD and RMSF of SLN during MD simulation. (A) Histograms of the RMSD with respect to time for the SLN backbone, side chain, and all-atom domains during the MD simulations. (B) Contribution of the SLN N-terminus, domain Ib, domain II, and the C-terminus to the RMSD of the side chain. (C) Contribution of the SLN N-terminus, domain Ib, domain II, and the C-terminus to the RMSD of the backbone. (D) RMSFs of all atoms, backbone atoms, and side-chain atoms during the MD simulations.

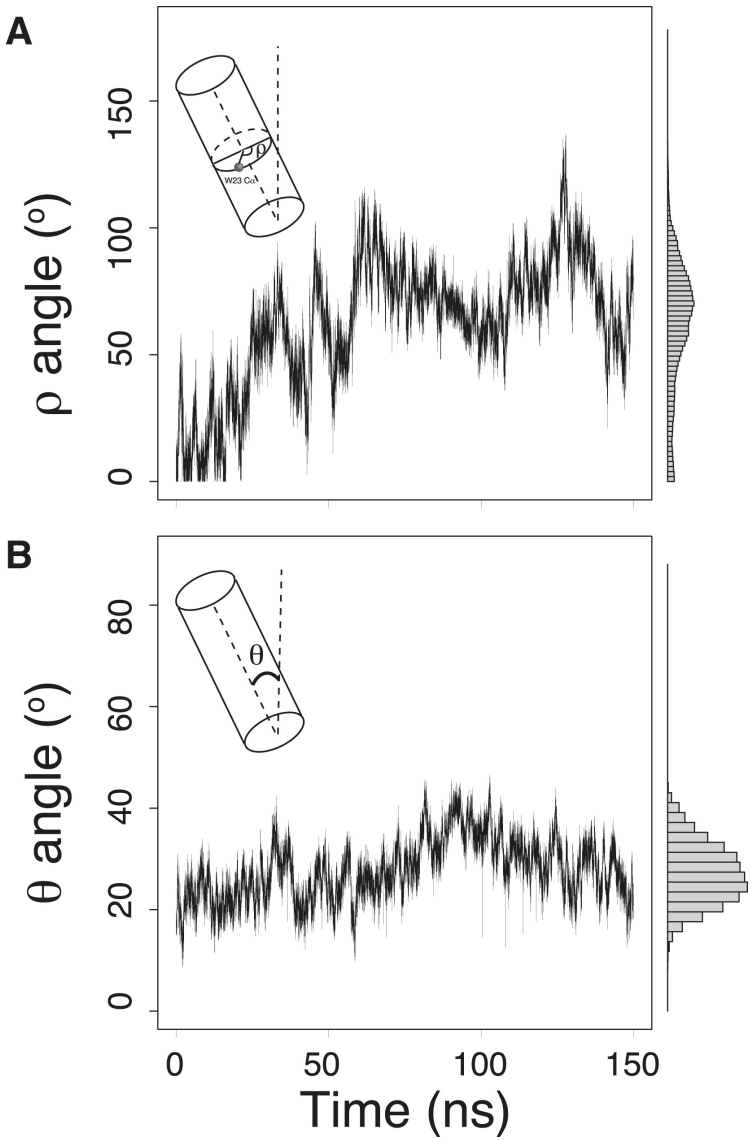


Figure 3.5: Time evolutions and distribution of the azimuthal (A) and tilt (B) angles for the entire SLN helix during MD simulations. Descriptions of the azimuthal angle, $\rho$, and tilt angle, $\theta$, used for structural analysis are shown in the upper and lower insets, respectively.

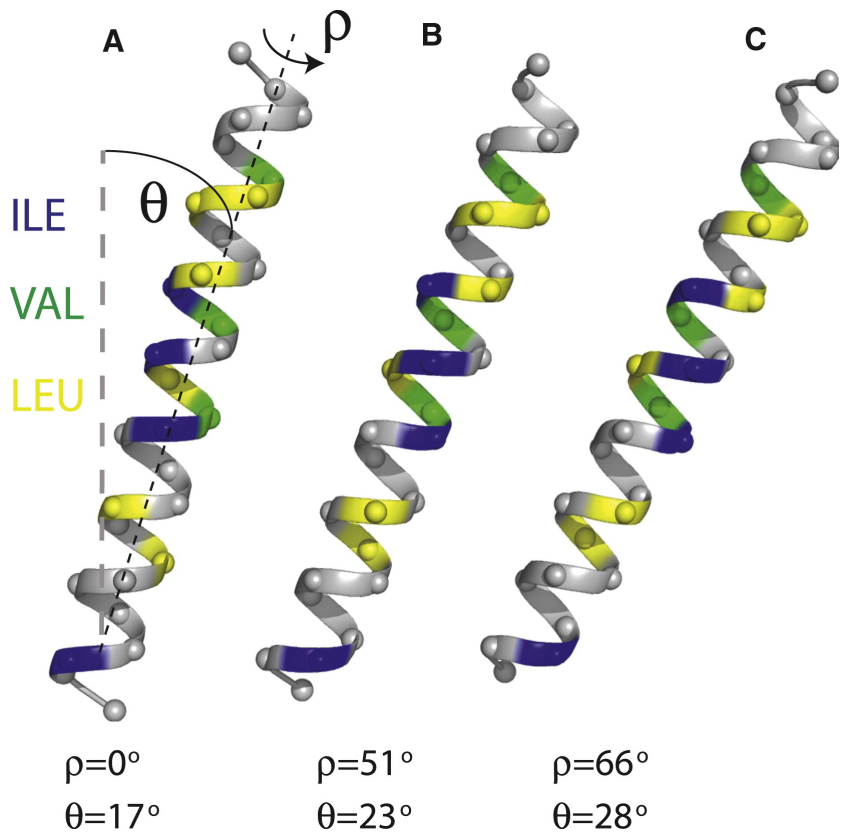


Figure 3.6: Structure comparison of the starting configuration (A), the ideal helix model best-fit PISEMA data (B), and the snapshot from the simulation having the average rotation and tilt angle (C). All Ca atoms are shown in spheres with the Ile, Val, and Leu residues colored blue, green, and yellow, respectively.

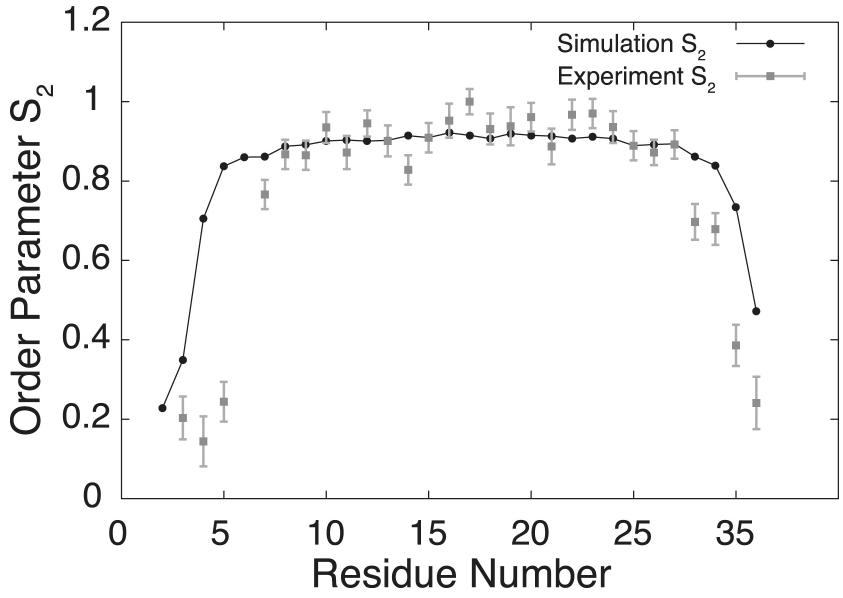


Figure 3.7: Comparison of the generalized order parameters ($S^2$) for the amide vectors of SLN obtained from MD simulations (black) and solu- tion-NMR data in DPC micelles (gray).

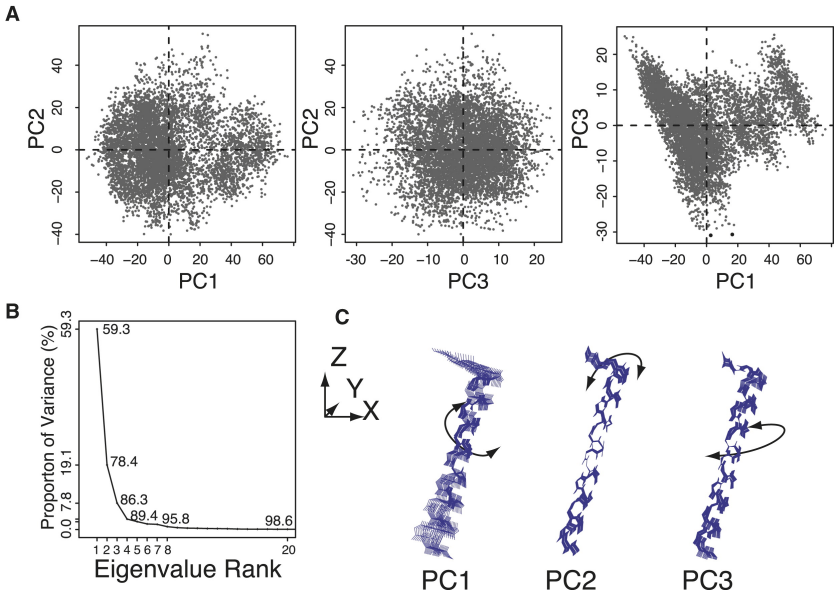


Figure 3.8: PCA of the backbone atoms of SLN for the last 40 ns of MD calculations. (A) Projection of PC1/ PC2, PC2/PC3, and PC1/PC3 planes. (B) The cumulative contribution of the first 20 principal components to the overall structural variation. (C) Graphic representation of the first three principal components. The arrows indicate the direction of the structural fluctuations.

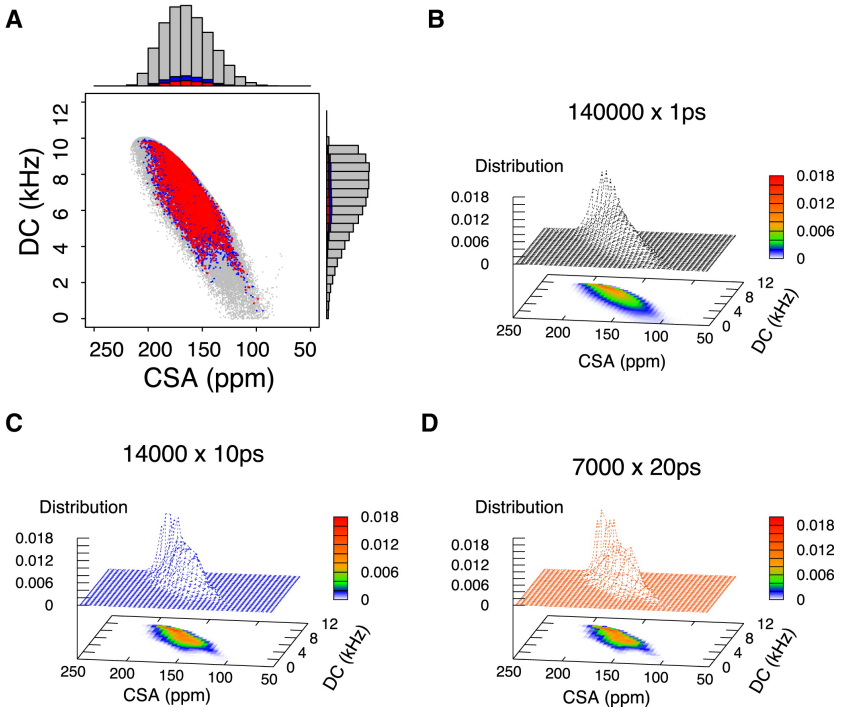


Figure 3.9: Simulated CSA and DC values obtained for residue 25 from MD trajectories of the last 140 ns of simulations. (A) Overlay of the calculated PISEMA spectra from SLN snapshots taken from MD trajectories. The distribution of the CSA and DC values is indicated on each axis. (B-D) Three-dimensional plot showing the distribution of CSA and DC for residue 25 with no torsion-angle averaging (B), with 10-*ps* torsion-angle averaging (C), and with 20-ps torsion-angle averaging (D).

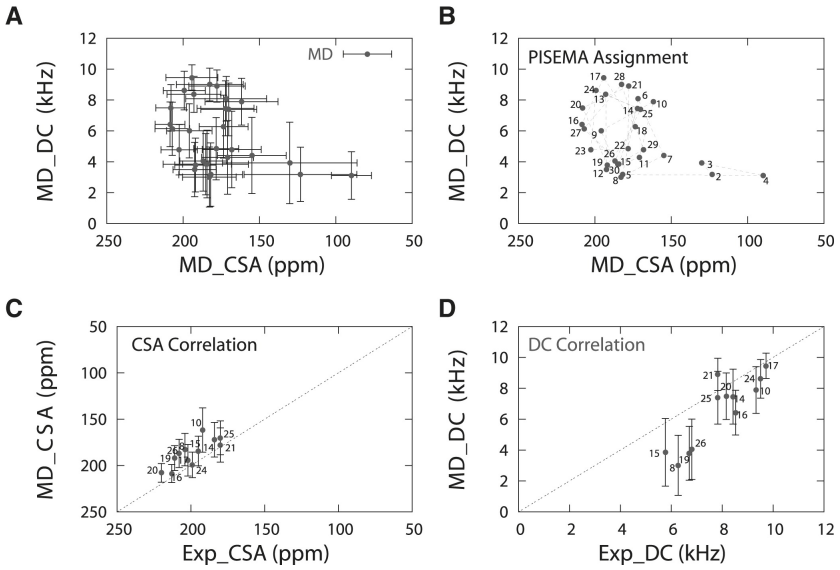


Figure 3.10: Simulated PISEMA for SLN MD simulation. (A) Full PISEMA spectrum for all of the amide sites of SLN from the last 140 ns of the simulation using 10° torsion-angle averaging. Note that the error bars indicate the standard deviations of the values. (B) Average positions from A are shown, with labels corresponding to residue number. The resonances are connected using dotted lines to show the regular PISA wheel for the SLN helix. (C and D) Correlation of experimental and calculated CSA (C) and DC (D) for Leu, Ile, and Val only.

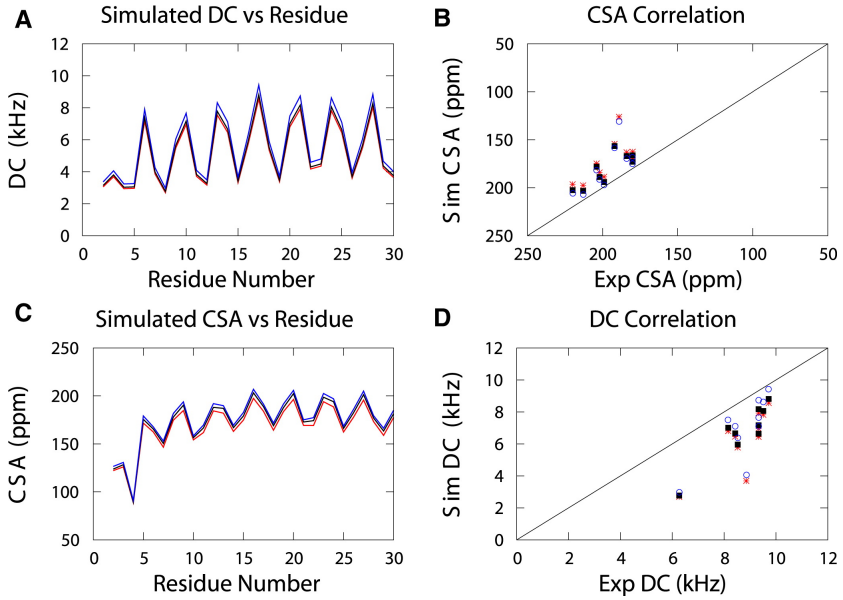


Figure 3.11: Effects of different values of the CSA tensors on the calculated PISEMA spectra. Values were obtained from Wu et al[23] for tensor 1 (red), from Straus et al[169] for tensor 2 (black), and from Page et al[137] for tensor 3 (blue). Tensor 1, $\sigma_{11} = 64, \sigma_{22} = 77, \sigma_{33} = 217, \nu_{||} = 9.75kHz$; tensor 2, $\sigma_{11} = 56.3, \sigma_{22} = 79, \sigma_{33} = 224, \nu_{||} = 10.05kHz$; tensor 3, $\sigma_{11} = 57.3, \sigma_{22} = 81.2, \sigma_{33} = 228, \nu_{||} = 10.74kHz$. Error bars are omitted for clarity.

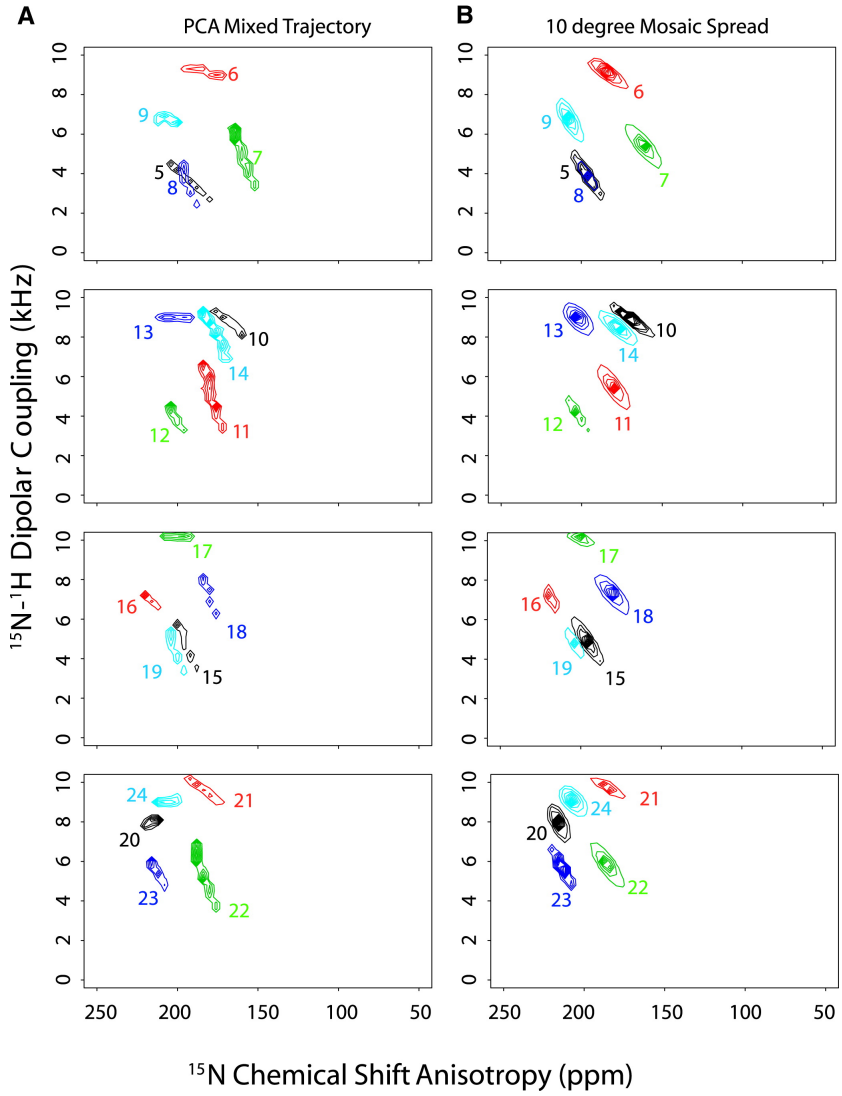


Figure 3.12: (A) Simulated contours of different residues using combined principal-component trajectories. (B) Simulated contour of 10° mosaic spread of protein orientation around the average position.

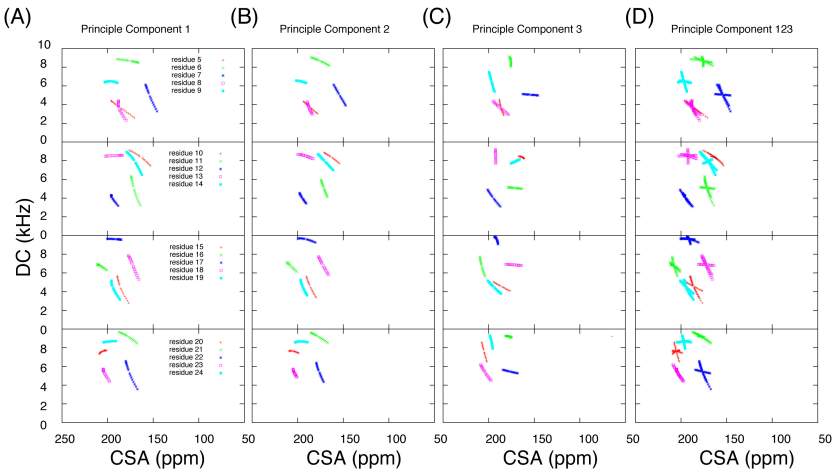


Figure 3.13: Effects of principal component on PISEMA. (A) Effects of PC1 on the appearance of PISEMA spectrum (B) Effects of PC2 on the PISEMA (C) Effects of PC3 on PISEMA spectrum (D) The effects of combined PC1, PC2 and PC3 trajectories on PISEMA spectrum

# Chapter 4

# Paramagnetic-based NMR restraints lift residual dipolar coupling degeneracy in multidomain detergent-solubilized membrane proteins



## 4.1 Introduction

Residual dipolar couplings (RDCs) constitute an excellent source of structural and dynamic information[236, 237, 238, 239]. Their use spans from structural refinement[236, 240, 241] to the characterization of unfolded proteins[242, 243, 244, 245, 246] and excited states of biomolecules[247, 248]. RDCs are also valuable NMR parameters to determine

the relative orientation of multidomain proteins and protein complexes, where interdomain or interprotein nuclear Overhauser effects (NOEs) are scarce[239, 167, 249].

However, the inherent degeneracy of RDCs (ghost orientations) complicates the extraction of orientational information[250]. It was originally thought that the RDC equation resulted in 8-fold degeneracy for peptide plane orientations[250] but recently it was shown that the analytical solution contains a 16-fold degeneracy[251]. The solutions are reduced to 4 if one considers the regular patterns of the dipolar couplings (i.e., dipolar waves) in secondary structures[243, 154, 252, 30, 253, 155]. In favorable cases, this inherent degeneracy is resolved using two or more alignment media[254]. Nonetheless, local geometry can still be ill-defined[251] and further confounded by the presence of conformational dynamics[255, 256, 163, 161]. Therefore, a major issue is to eliminate ghost orientations from the true orientations. To this extent, Bertini and co-workers have proposed the use of paramagnetic-based NMR restraints[257]. These authors have developed a new analysis of PREs to give a comprehensive view of the different conformations and dynamics of calmodulin as well as the calmodulin-$\alpha$-synuclein complex[257]. A similar approach has been utilized for resolving the solution conformation of the ternary complex of the E. coli Hsp70 chaperone[258].

Membrane proteins solubilized in detergent micelles are not immune to these challenges. Polytopic membrane proteins are often organized in multiple domains (independent or partially independent) that facilitate intra- and intercellular communication[210]. Membrane protein tertiary structures are often defined by domain interactions with the lipid membrane, that is, topology[259]. Several small and medium size membrane proteins involved in regulatory function of ion pumps fall into this category[259]. While membrane proteins are amenable to modern solution NMR techniques, it is still a challenging task to obtain long-range distance restraints from NOE data[260] especially for helical membrane proteins. Side-chain methyl labeling schemes can help determine the overall fold of membrane proteins, and new protocols have been developed to introduce a variety of different probes[261, 262, 263, 264]. RDCs constitute a viable alternative to obtaining long-range distance restraints for membrane proteins. To measure RDCs, membrane proteins need to be aligned in an anisotropic medium[265, 266, 267] or bound

to a lanthanide ion through adventitious sites[268] or engineered tags[269, 270]. In several instances, RDCs were crucial for improving the resolution of membrane protein structures solubilized in micelles[260, 271, 272].

In this Article, we show that the ghost orientations generated after RDC refinement of the detergent-solubilized protein phospholamban (PLN) can be eliminated by using paramagnetic-based restraints derived from site-directed spin-labeling using MTSSL (1-oxyl-2,2,5,5-tetramethyl-D3-pyrroline-3-methyl methanethiosulfonate).

## 4.2 Material and methods

### 4.2.1 PLN expression, purification, and mutagenesis

The overexpression and purification of AFA-PLN (i.e., cysteine-null monomer, C36A, C41F, C46A) was carried out as described in Buck et al[174]. For the A24C mutant, the plasmid encoding cysteine-null PLN was used as a template to introduce a single point mutation (A24C) by site directed mutagenesis using a Quick-change kit (Stratagene, San Diego, CA). The primers were designed as follows: forward 5'-GCCGCAGCAGTGCCGCCAGAACCTGC-3', reverse 5'-GCAGGTCTGGCGGCACTGCTGCGGC-3' (the mutated codon is underlined). The steps of the PCR protocol were reported previously[174].

The mutated plasmid was amplified in XL1-blue competent cells (Stratagene, San Diego, CA) and purified using QIAprep Spin kit (Qiagen, Carlsbad, CA). The sequence was confirmed by DNA sequencing (ABI PRISM 3130xl Genetic Analyzer, Biomedical Genomic Center, Minneapolis, MN). BL21(DE3) (Stratagene, San Diego, CA) E. coli cells were transformed with 100 ng of purified plasmid and selected on LB agar plates containing ampicillin (100 mg/mL). The L7C and R9C mutants were designed in a similar way. Protein expression and purification were carried out using a combination of affinity chromatography and HPLC. The steps of the PCR protocol were reported previously[174].

### 4.2.2 A24C-AFA-PLN and L7C-AFA-PLN spin-labeling with MTSSL

The spin labeling reaction protocol was optimized by Kirby et al[273]. One milligram of lyophilized A24C-AFA-PLN (or L7C-AFA-PLN) protein was solubilized in 1 mL of spin labeling buffer (60 mM Tris-HCl at pH 7.0 and 0.2% SDS). The final concentration of PLN was 0.2 mM. A 10-fold molar excess of MTSSL (1-oxyl-2,2,5,5-tetramethyl-D3-pyrroline-3-methyl methanethiosulfonate, Toronto Research Chemicals, North York, ON) was added to the reaction mix and incubated overnight at 4 °C in the dark. The unreacted spin label was removed by reversed-phase HPLC. Lyophilized MTSSL spin-labeled A24C-AFA-PLN (or L7C-AFA-PLN) was dissolved in 300 mM DPC, 20 mM phosphate buffer (pH = 6.0), 120 mM NaCl, 0.01% NaN3, and 10% D2O. The final concentrations of A24C-AFA-PLN and L7C-AFA-PLN were 0.1 mM. Labeling efficiency was measured by EPR[273].

### 4.2.3 Sample preparation for RDC measurements

AFA-PLN [U-$^{13}C$, $^{15}N$] was reconstituted into 100 mM deuterated DPC (Cambridge Isotope Laboratories), 25 mM phosphate buffer (pH 6.0), 120 mM NaCl, 10% D2O, and 0.1% NaN3 to give a final protein concentration of 0.8 mM. The stretched gels were polymerized from a 5.7 mm diameter cylinder under the following conditions: 100 mM Tris-HCl (pH 8.0), 5.1% acrylamide, 1.3% bis-acrylamide, 0.1% ammonium persulfate, and 0.33% TEMED. The gels were washed twice with 50 mM NaH2PO4/Na2HPO4 pH 6.5 (8 h/wash) and then twice with H2O (8 h/wash) as described previously[274]. The protein/detergent mixture was added to the dried polymerized gel and incubated at 37 °C for 24 h. After being stretched in a 5 mm Shigemi tube, the length of the gel increased by a factor of 1.7. The stretching apparatus was purchased from New Era Enterprises, Inc[274].

### 4.2.4 NMR spectroscopy

#### 4.2.4.1 RDC measurements

NMR experiments were carried out at 37 °C using a Varian spectrometer operating at a $^{1}H$ Larmor frequency of 599.54 MHz. 2D TROSY-based ($^{1}H$, $^{15}N$) pulse sequences

described by Permi and Annila[275] were used to measure $^{13}C'$-$^{15}N$ and $^{15}N$-$^{1}H$ one-bond J-couplings (and J-coupling + RDC). Each 2D experiment was acquired in the presence and absence of the stretched gel, where the difference in splitting allowed for the calculation of the RDCs. The total acquisition times in t1 and t2 were 62 and 77 ms with spectral widths of 1300 and 6600 Hz in the $^{15}N$ and $^{1}H$ dimensions, respectively. To measure the $^{13}C'$-$^{13}C_{\alpha}$ one-bond couplings, we performed a 3D uncoupled HNCO experiment in the presence and absence of stretched gel. Experiments were performed with 64 scans, spectral widths of 10000 Hz ($^{1}H$), 1000 Hz ($^{13}C$), and 1200 Hz ($^{15}N$), and total acquisition times of 83.2, 40, and 26.7 ms in the $^{1}H$, $^{13}C$, and $^{15}N$ dimensions, respectively. A recycle delay of 1.3 s was used for all of the experiments. All data were processed with NMRPipe[276] and analyzed using NMRVIEW[277]. The $^{15}N$-$^{1}H$, $^{13}C'$-$^{15}N$, and $^{13}C'$-$^{13}C_{\alpha}$ RDCs versus residue are shown in Figure 4.1A.

#### 4.2.4.2 PRE measurements

[$^{1}H$, $^{15}N$] heteronuclear single quantum coherence (HSQC) spectra were acquired in the presence of MTSSL (paramagnetic) and the diamagnetic analogue at positions 7 and 24 of PLN. The intensity reductions of the amide resonances were converted into $R_2^{SP}$ values and then to distance restraints using eq (4.1) and (4.2):[17]

$$\frac{I_{ox}}{I_{red}} = \frac{R_2 e^{-R_2^{sp} t}}{R_2 + R_2^{sp}} \tag{4.1}$$

$$r = [\frac{K}{R_2^{sp}}(4\tau_c + \frac{3\tau_c}{1+\omega_H^2\tau_c^2})]^{\frac{1}{6}} \tag{4.2}$$

where $I_{ox}$ and $I_{red}$ are the peak intensities corresponding to the oxidized and reduced MTSSL samples, respectively, $r$ is the distance between the nuclear spins and the unpaired electron, $\tau_c$ is the correlation time for the electron/nuclear spin interaction, $\omega_H$ is the proton Larmor frequency, $R_2$ and $R_2^{sp}$ are the transverse relaxation rates in the absence and presence of spin label, respectively, and K is a constant ($1.23 \times 10^{-32} cm^6 s^{-2}$)that depends on the gyromagnetic ratio, electronic g factor, and the Bohr magneton. For PLN, we used two different correlation times ($\tau_c$): 8.2 ns for the cytoplasmic domain and 15.4 ns for the transmembrane domain[217]. For peaks with intensity retention greater than 90% no upper limit distances were used. For the

other peaks, the intensity retentions were converted into distances according to eq (4.2) and implemented with upper and lower limit distances of ±4Å[17]. The PRE data and the calculated distance restraints for PLN are shown in Figure 4.2.

### 4.2.5 Calculation protocol

We used the standard energy target function implemented in XPLOR-NIH:[39]

$$E = E_{empirical} + E_{solution} \tag{4.3}$$

where $E_{empirical}$ is the sum of the energy terms for the covalent geometry (bond distance, bond angles, improper angles) and VDW interactions:

$$E_{empirical} = w_{bonds}E_{bonds} + w_{angles}E_{angles} + w_{impropers}E_{impropers} + w_{vdw}E_{vdw} \tag{4.4}$$

$E_{solution}$ includes the penalty functions for distance and orientational restraints (NOEs, dihedral angles, RDCs, and PREs):

$$E_{solution} = w_{NOE}E_{NOE} + w_{CDIH}E_{CDIH} + w_{DB}E_{DB} + w_{RDC}E_{RDC} + w_{PRE}E_{PRE} \tag{4.5}$$

All of the NOE-derived distance restraints used in the present study were taken from the previously published structure of PLN in DPC micelles[278]. A total of 373 distances (142 intra-residue and 231 inter-residue) were included. In addition, 58 hydrogen bonds derived from H/D exchange measurements were used. Finally, we obtained 38 dihedral angle restraints using the program TALOS[14] based on the chemical shifts of $H_\alpha$, $C_\alpha$, $C_\beta$, $C'$, $N$, and $H_N$[279]. All of the restraints used in the calculations are listed in Table 4.1.

To illustrate the impact of RDC and PRE restraints in the structural refinement of PLN, three simulated annealing protocols were used to generate three different structural ensembles[280]. In the first protocol, the restraints derived from typical solution NMR experiments, including NOE-derived distances, dihedral angle restraints and hydrogen bonds were used[278]. The second protocol included the restraints from the first protocol plus RDCs derived from the partial alignment of PLN in stretched gels. Finally, the third protocol was constructed by adding MTSSL distance restraints determined from PRE data to all constraints used in the first two protocols. Details for the calculation are presented in the Supporting Information.

## 4.3 Results

### 4.3.1 Structural refinement with NOE and dihedral angle restraints (Protocol One)

To test our approach, we used phospholamban (PLN), a 52-residue integral membrane protein in the sarcoplasmic reticulum (SR) that regulates the SR $Ca^{2+}$-ATPase (SERCA)[71]. PLN exists as a pentamer in the SR and depolymerizes to monomers upon encountering SERCA[281, 282]. To induce the monomeric state, we mutated the three transmembrane Cys residues (C36A, C41F, C46A) to obtain a stable monomer with activity identical to that of wild-type PLN[283]. The NMR structure of PLN monomer in dodecylphosphocholine micelles has been solved by our group[278]. For the structure determination, we used distance restraints derived from $^{15}N$-edited NOESY spectra, dihedral restraints from chemical shifts[14] and hydrogen bonds from H/D exchange factors[15]. On the basis of the calculations and subsequent backbone dynamics studies, we identified four dynamic domains: $\alpha$-helical domain Ia (residues 1-16), $\alpha$-helical region spanning domains Ib (residues 23-30) and II (residues 31-52), and a loop connecting the two helical regions (residues 17-22)[217].

Given the limited amount of distance and angular restraints found between the loop and its adjacent domains, the final structural ensemble resulted in many different conformers with poor convergence for the overall backbone conformation and topology. In our previous paper[278] we mapped solvent accessibility using $Mn^{2+}$ ions and the insertion of PLN in the micelles using 5- and 16-DSA (doxyl stearic acids), and we manually eliminated structures that contradicted the paramagnetic mapping. We concluded that PLN adopts a helical L-shaped conformation with the cytoplasmic amphipathic domain Ia adsorbed on the surface of the micelle. These results were also supported by solid-state NMR experiments carried out in lipid bilayers[284]. While these studies represent the initial characterization of PLN in membrane mimicking environments, they were limited by two factors: (a) the absence of topological restraints (the selection of conformers was carried out manually), and (b) the absence of the relative orientation of the two helical domains.

### 4.3.2 Structural refinement with RDCs and NOEs (Protocol Two)

To improve the solution NMR structural ensemble of PLN, we introduced restraints from three sets of RDCs obtained from the partial alignment of PLN in a stretched gel system[265, 285]. Figure 4.1 shows the three sets of RDCs used in the calculations. As expected from the helical secondary structure of PLN, the values of the RDCs oscillate periodically. There is a significant change in the pattern from residues 20 to 30 for the $^{15}N$-$^{1}H$ RDCs. These residues belong to domain Ib, a helical region that is more dynamic and solvent exposed than the transmembrane domain II[217]. When fitting these dynamically averaged RDCs to an average structure, a pronounced kink is observed between domain Ib and domain II. CPMG relaxation dispersion measurements obtained from PLN in DPC indicate the presence of chemical exchange (at least two conformations) for domain Ib[217, 84] and for this reason we excluded RDCs from domain Ib during our structural calculation. Dynamic model or ensemble simulations are needed to explain the discontinuity of the RDC pattern in this domain.

One of the challenges for structural refinement with RDCs is to accurately determine the axial ($D_a$) and rhombic (R) components of the alignment tensor. Several methods are available to estimate these values such as the maximum likelihood method by Moore and Warren[286] and the histogram method by Bax and co-workers[287]. To use these methods, large numbers of RDCs are needed, and their reliability is dependent on the accuracy of the experimental measurements. Because of the increased size of membrane proteins within detergent micelles, experimental RDCs have larger errors than their soluble counterparts. Furthermore, the errors for $^{13}C'$-$^{15}N$ and $^{13}C'$-$^{13}C_{\alpha}$ RDCs are much larger after scaling to those of the $^{15}N$-$^{1}H$ RDCs. To overcome this, $D_a$ and R were allowed to vary in the simulated annealing procedure[241]. This introduced two extra parameters into the structural fitting in addition to the tensor orientations. Four hundred structures were generated and clustered based on two groups of $D_a$ and R values (Figure 4.8). The solutions with $D_a$ = -8 Hz and $R$= 0.667 resulted in poor fitting of the RDCs. These structures were excluded from further analysis. The structures generated from the second group (shown red in Figure 4.3) with $D_a = 8.6$ ($\pm 0.3$) and $R = 0.52$ ($\pm 0.04$) were used in the following calculations. The high rhombicity $R$ is consistent with the RDC histogram shown in Figure 4.1B. Note that all of the RDCs

were implemented in the simulated annealing protocol using a flat-well potential[39].

The force constants for RDCs during the structure refinement were determined using the $R$ factor method ($R_{RDC}$) as described by Clore and co-workers[280]. Two variables were monitored as a function of the RDC force constant: (1) the $R$ factor for different sets of RDCs and (2) the penalties from energy terms other than RDCs. During the force constant ramping, the tensor values $D_a$ and $R$ were fixed (determined above to be $D_a = 8.6$ and $R = 0.52$). The best value for the force constant was 0.5 $kcal \cdot mol^{-1} Hz^{-2}$ (see Figure 4.8), which gives reasonable RDC agreement without large penalties from other structural and experimental restraints.

After optimizing $D_a$, R, and the force constants, we generated 300 structures and selected the 100 lowest energy structures for further analysis (Figure 4.4A,B). The resulting structures were clustered into four distinct families (Figure 4.5B). These structures differ in the relative orientation of the two helical domains, giving rise to a 4-fold degeneracy (Figure 4.5B). Each structural ensemble shows a backbone RMSD less than 1.6 Å, with good correlations between experimental and calculated RDC values (Figure 4.9). When viewed in the alignment tensor frame (Figure 4.3C), 16 different solutions for two helical domain structures are possible. A similar result has been obtained by Bax and co-workers[288] for the monomeric subunit of KcsA solubilized in detergent micelles.

The solutions can be distinguished by analyzing the orientation angles ($\theta$, $\varphi$) for each helical domain in the alignment tensor frame (Figure 4.3) as well as the interhelical angle $\chi$ (angle between domains Ia and II, red and blue, respectively). The orientations correspond to the following angles: ($\theta$, $\varphi$), ($\theta$, $\pi + \varphi$), ($\pi - \theta$, $\pi - \varphi$), and ($\pi - \theta$, $2\pi - \varphi$) (Figure 4.3B). The degenerate solutions agree with theoretical equations describing RDC periodicity[30, 253, 155]. The hydrophobic residues of the cytoplasmic domain Ia are oriented toward the transmembrane domain for families I and II and away from it for families III and IV (Figure 4.5A). Family I has a $\chi$ angle of 90°, while families III and IV display $\chi$ angles of 70°. Family II has a nearly antiparallel orientation between the two helical domains ($\chi \sim 140°$). This orientation is similar to that obtained by Oxenoid and Chou[289].

### 4.3.3 Structural refinement with NOEs, RDCs, and PREs (Protocol Three)

To reduce the degeneracy from the RDC solutions and remove the translational degree of freedom between the two domains, we implemented PRE distance restraints from two MTSSL-PLN constructs (L7C-AFA-PLN and A24C-AFA-PLN). These restraints were included in the calculations using the convention introduced by Battiste and Wagner[17](see Materials and Methods). To restrict the flexibility of the spin label, we used a dihedral angle potential (sinusoid potential) for the $\chi_1$, $\chi_2$, and $\chi_3$ angles of the spin label, which adopt defined values as determined from the crystal structures[290].

In protocol two (RDC+NOE refinement), the four families are energetically degenerate ($E_{RDC}$ and $E_{NOE}$) (Figure 4.4A,B), while in protocol three, these three families have different values of $E_{PRE}$ (Figure 4.4C,D). The structures within family I display the lowest PRE and NOE energies and thus were selected as the final ensemble (Figure 4.5C), with a backbone RMSD of 1.2 Å and good agreement with experimental RDCs (Figure 4.11) and PREs (Figure 4.12).

The energy differences between families I, III, and IV are very small and cannot be used as a criterion for selection. However, only family I has the hydrophobic residues facing toward the transmembrane helix, that is, facing the micelle interior (Figure 4.4C). As an attempt to further support our selection criteria to eliminate families III and IV, we labeled PLN with MTSSL at position 9 (R9C-PLN). When the MTSSL is engineered, the long chain of MTSSL cross-linked to Cys 9 inserts in the detergent micelle (see quenching of the resonances in the transmembrane domain in Figure 4.13C), resulting in an incorrect positioning of the side chain and faulty interpretation of the structural topology of the protein. Therefore, the flexibility and hydrophobic nature of the MTSSL spin label calls for special care when engineering site-specific mutants in membrane proteins when those sites are proximal to the membrane[291].

### 4.3.4 RDC analysis for the PLN pentamer

The structure of the PLN pentamer heavily relied on the use of RDCs in the structure refinement[289]. In fact, the structure of individual monomers were built prior to

assembly the symmetric pentamer. In the calculation of the monomeric units, the alignment tensors ($D_a = 9.00$, $R = 0.33$) were determined from singular value decomposition (SVD)[38] using $^{15}N$-$^{1}H$, $^{13}C'$-$^{15}N$, and $^{13}C'$-$^{13}C_\alpha$ RDCs. The alignment tensor was slightly different from the value we obtained for monomeric AFA-PLN. A comparison of RDC restraints between the pentamer (Oxenoid and Chou[289]) and monomer (this Article) showed a good correlation (Figure 4.7), especially with the $^{15}N$-$^{1}H$ RDCs. We then performed a similar structural calculation using Oxenoid and Chou's data[289]. We obtained four degenerate solutions (Figure 4.6A,B) as for the AFA-PLN monomer. On the basis of these results, we speculate that the authors[289] selected one of the solutions from the structural ensembles to build the pentamer.

Aside from the degeneracy issue, the choice of the tensor value for a symmetric pentamer is also debatable. It is well-known[249] that for oligomeric proteins the symmetry axis should coincide with the long axes of the alignment tensor, resulting in $R = 0$ (zero rhombicity). It was argued that asymmetric tensor values could result from protein dynamics;[289] however, the dynamic averaging of RDCs needs to be interpreted using a dynamic ensemble rather than fitting data to an average structure. We also refined the monomeric unit using RDCs measured by Oxenoid and Chou[289] with a symmetric tensor ($R = 0$). The $D_a$ value obtained was 6.5, using the methods described earlier (Figure 4.8). Figure 4.6C and D show that the angles do not have defined minima. This is also in agreement with a theoretical study, which showed that an infinite number of solutions exist for $R = 0$[251].

## 4.4 Discussion

To address the challenges of high-resolution structure determination of membrane proteins in micelles, the classical NOE-driven approach has been supplemented with RDCs. However, internal protein dynamics and intrinsic degeneracy of the RDC solutions complicate the data interpretation, resulting in ghost orientations. Here, we show that the degeneracy problem can be addressed by including RDC restraints with PREs from covalent attachment of the MTSSL spin label, which provides a fast and efficient method for determination of both membrane protein structure and topology.

We applied this method to a small multidomain membrane protein, PLN, which regulates the enzymatic activity of the SR $Ca^{2+}$-ATPase in cardiac muscle. While structural biologists agree on the secondary structure content of the transmembrane domain of PLN, there is an active debate about the topological arrangement of this protein in the lipid membrane[71]. The structure calculations carried out with NOE-based distance restraints do not provide a high-resolution picture of PLN due to the lack of restraints between the helical domains. In our original work, the correct orientations of the helical domains of PLN were selected manually based on the PRE data from Mn2+ and doxyl stearic acids[278]. The introduction of RDCs improves the resolution of the ensemble, but exemplifies four degenerate solutions, with ghost orientations of the cytoplasmic domain Ia. The combination of paramagnetic-based restraints with RDCs and NOEs enabled us to resolve the orientation that agrees best with all available data. This solution (family I) is similar to the recently determined structural ensemble of monomeric PLN in lipids[279, 292]. This is not surprising because several recent reports have shown that the structures of membrane proteins in micelles are similar to those determined in lipid bilayers[293]. The slight discrepancy in the average rotation angles between the hybrid ensemble and that determined by combining NOEs, RDCs, and PREs (Figure 4.7) has several possible origins. First, there are approximations in estimating the alignment tensor. Second, the cytoplasmic domain of PLN is rather dynamic[217, 84] and our interpretation of RDCs and PREs[294] does not take this into account. Third, the nonplanar surface of the micelle can cause bends and curvature to malleable domains of membrane proteins and peptides[274]. Fourth, RDCs provide information of bond orientations with respect to three alignment tensor axes in an asymmetric tensor. Orientational data from PISEMA experiments in oriented samples[279] only encode bond orientations with respect to one of the tensor axis (z axis), similar to RDCs with a symmetric tensor.

Because the studies of membrane proteins in detergent micelles are likely to continue to offer insightful information, we conclude that for multidomain membrane proteins the use of RDCs is not sufficient to define their conformational space and topology. Rather, orientational restraints need to be supplemented with long-range paramagnetic restraints from spin labels covalently linked to proteins to uniquely define their

topology[260, 295].

Finally, we would like to point out that this Article addresses only the geometric ambiguities derived from RDCs. In fact, intrinsic dynamics complicates the interpretation of both PREs and RDCs for structure determination. To overcome these problems, molecular dynamics methods are being developed for the variability of alignment tensors[158, 296] as well as for the modulations of the PRE effects on long-range interactions[294, 297]. A more comprehensive approach for flexible domains of membrane proteins will require the combined use of PREs and RDCs using ensemble molecular dynamics methods[294].

## 4.5 Supporting information

### 4.5.1 Structure calculation

In the first protocol, we started from a fully extended structure of PLN and carried out simulated annealing calculations starting at 6000 K and cooled to 0 K in Cartesian and torsion-angle space. The NOEs were modeled using a soft-square potential with a force constant ramping from 2 to 50 $kcal \cdot \text{Å}^{-2}$, while torsion angles were restrained by a square-well potential with a force constant of 200 $kcal \cdot rad^{-2}$. In addition, the torsion angle database potential 'RAMA' was used to bias the conformational search in the most allowed regions of the Ramachandran plot.

For the second protocol, residual dipolar coupling data were introduced during the lower temperature simulated annealing (3000 K) starting from the folded PLN structures obtained from protocol 1. The system was cooled down to 0 K with a cooling rate of 5 K/step using 4000 steps of internal dynamics and 4000 steps of Cartesian dynamics. Only RDCs derived from residues in the helical regions with order parameters were used ($S^2 > 0.6$). The force constants for RDC restraints were ramped from 0.05 to 0.5 $kcal \cdot Hz^{-2}$, which is determined by compromising RDC agreement and geometrical penalties. After obtaining the tensor magnitude, only tensor orientations were allowed to vary during the internal dynamics and Cartesian dynamics.

For the third protocol, additional MTSSL derived distance restraints are added to the restrains used in the second protocol. The MTSSL restraints were imposed by a force constant ramped from 2 to 50 $kcal \cdot \text{Å}^{-2}$. The parameter and topology files for MTSSL were taken from the library available in XPLOR-NIH. All the calculations were performed on a Linux cluster at the Minnesota Supercomputing Institute.

Table 4.1: NMR and Structural Refinement Statistics

| | **NOE+DIHE** | **RDC Ensemble** | | | | **PRE Ensemble** |
|---|---|---|---|---|---|---|
| | | **Family I** | **Family II** | **Family III** | **Family IV** | |
| **R.M.S. Deviation from experimental restraints** | | | | | | |
| NOE/H-bond (Å)(431) | 0.058 | 0.043 | 0.043 | 0.043 | 0.043 | 0.047 |
| Torsion angle (°)(38) | 0.072 | 0.739 | 0.724 | 0.746 | 0.741 | 1.160 |
| MTSSL PRE (Å)(43) | N/A | N/A | N/A | N/A | N/A | 0.039 |
| **RDC R-factor (%)** | | | | | | |
| $^1D_{NH}$ (Hz)(33) | N/A | 10.23 | 10.19 | 9.20 | 9.88 | 10.68 |
| $^1D_{NC'}$ (Hz)(36) | N/A | 22.04 | 22.16 | 21.97 | 22.13 | 23.55 |
| $^1D_{CCA}$ (Hz)(33) | N/A | 24.75 | 22.68 | 25.33 | 25.09 | 25.86 |
| **R.M.S. deviations from idealized covalent geometry** | | | | | | |
| Bond (Å) | 0.003 | 0.005 | 0.005 | 0.005 | 0.005 | 0.005 |
| Angle (°) | 0.397 | 0.655 | 0.647 | 0.648 | 0.650 | 0.736 |
| Impropers (°) | 0.278 | 0.629 | 0.612 | 0.619 | 0.621 | 0.684 |
| **Measure of structural quality** | | | | | | |
| % residues in most favored region | 92.5 | 90.7 | 91.7 | 87.1 | 90.1 | 85.3 |
| % residues in additional allowed region | 7.5 | 9.1 | 7.5 | 12.3 | 9.6 | 10.2 |
| % residues in generously allowed region | 0.0 | 0.2 | 0.8 | 0.6 | 0.3 | 4.5 |
| % residues in disallowed region | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| **Precision of atomic coordinates (Å)** | | | | | | |
| Backbone (residue 3-50) | 4.402 | 1.6 | 1.5 | 0.9 | 0.9 | 1.6 |
| Helix 1 (residue 3-18) | 0.621 | 0.2 | 0.1 | 0.1 | 0.1 | 0.2 |
| Helix 2 (residue 24-50) | 0.657 | 0.5 | 0.3 | 0.3 | 0.3 | 1.0 |

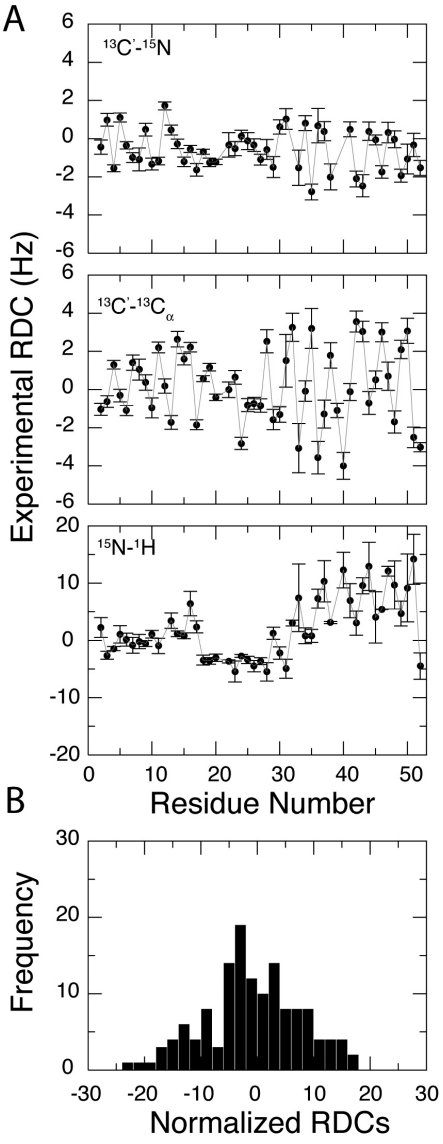


Figure 4.1: RDCs versus residue for AFA-PLN weakly oriented in stretched gels. (A) Top, $^{13}C'$-$^{15}N$ RDCs; middle, $^{13}C'$-$^{13}C_\alpha$ RDCs; bottom, $^{15}N$-$^{1}H$ RDCs. We reported only the RDC values for well-resolved peaks. (B) RDC histogram with all the data in (A) scaled to the $^{15}N$-$^{1}H$ RDCs.

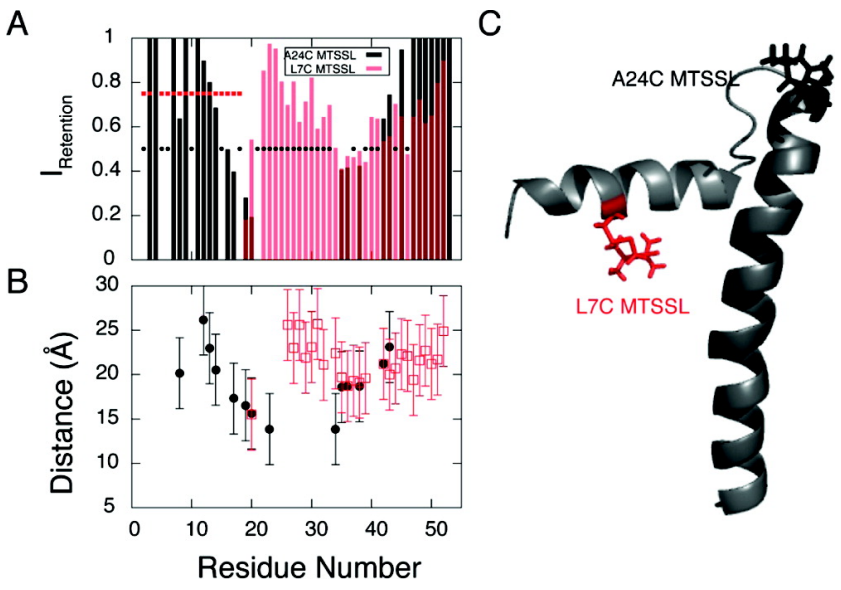


Figure 4.2: PRE data obtained from A24C-AFA-PLN (black) and L7C-AFA-PLN (red) cross-linked with MTSSL. (A) Intensity retention plot for A24C-AFA-PLN and L7C-AFA-PLN labeled with MTSSL. Unresolved peaks or completely quenched peaks are marked with asterisks. (B) Distances derived from PREs (see Materials and Methods). Only peaks with an intensity retention ratio less than 0.95 were used in the calculations. Also, some residues lacking $R_2$ values were omitted. (C) Cartoon representation of PLN with the MTSSL label at C24 and C7.

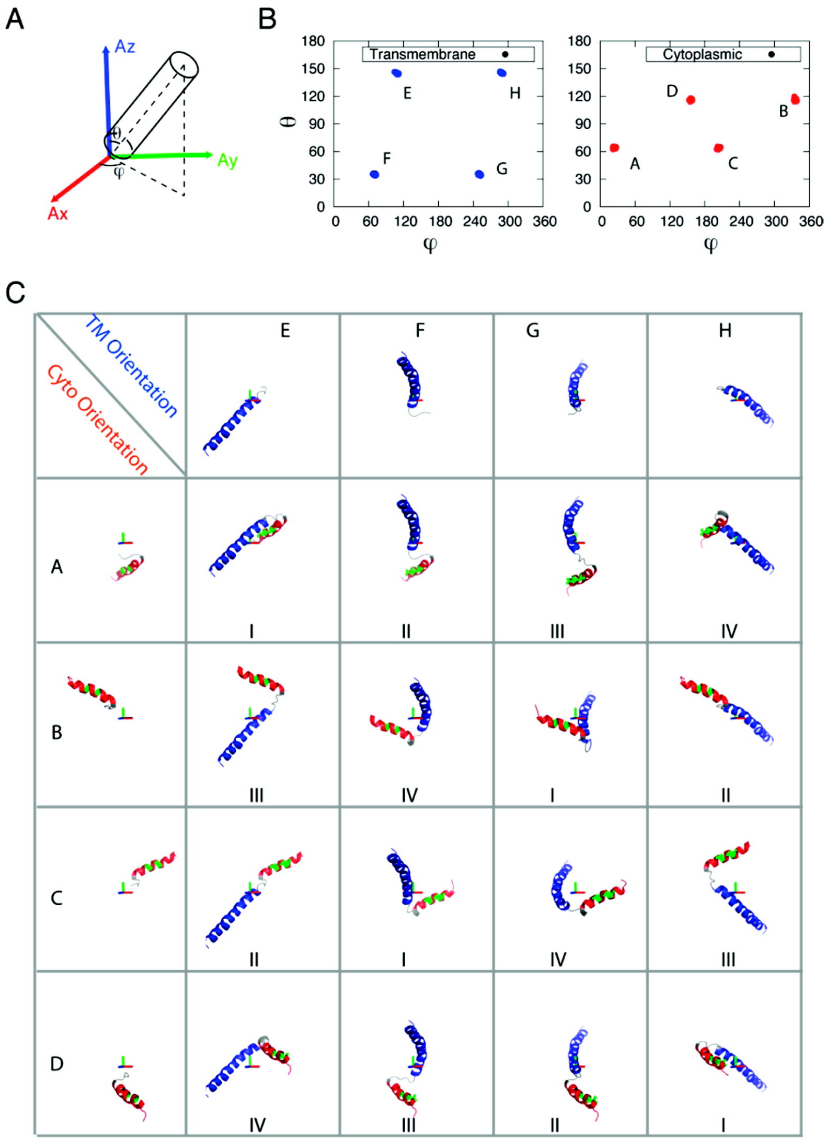


Figure 4.3: Orientation analysis for RDC structure ensemble. (A) Definition of orientation angles $\theta$ and $\varphi$ for a rigid body helix in the alignment frame. (B) $\theta$ and $\varphi$ plots for the cytoplasmic (Cyto, red) and transmembrane (TM, blue) domain orientations in the alignment frame. (C) Orientation of TM and Cyto domains with each letter corresponding to the $\theta$ and $\varphi$ angles defined in (B). Different combinations of orientations result in different families (labeled in the bottom as I, II, III, and IV). The green color on the cytoplasmic domain indicates hydrophobic residues.

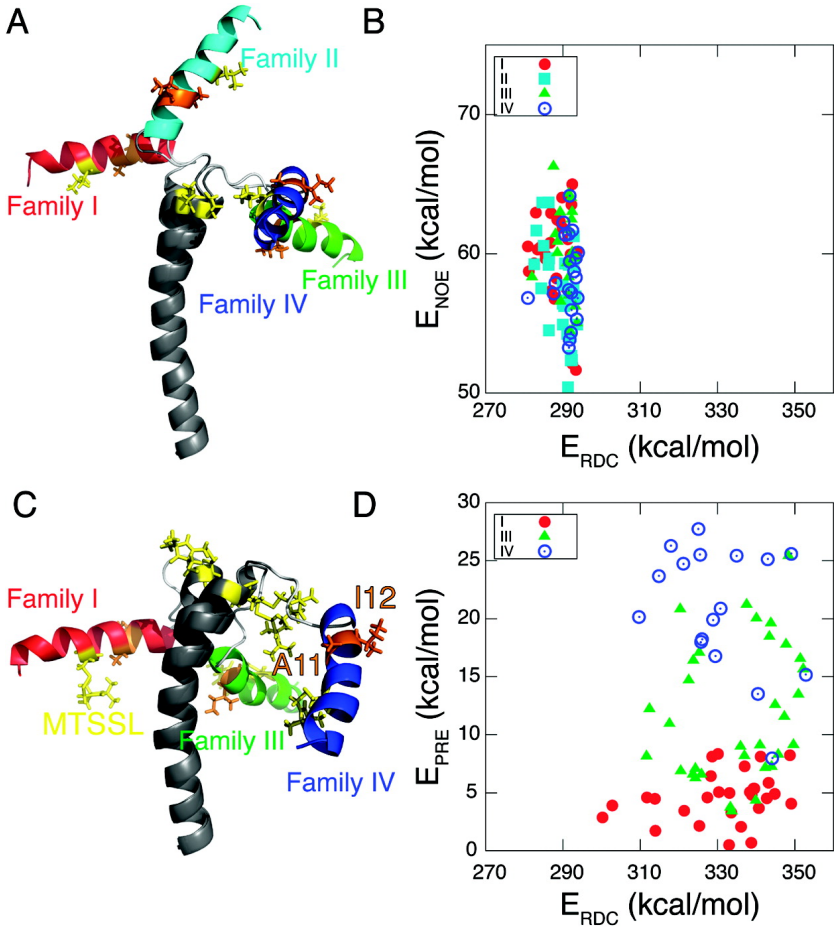


Figure 4.4: Using PRE to distinguish the degeneracy of RDC ensemble. (A) Comparison between representative structures from each family refined with protocol two. Side chains of the MTSSL spin label (yellow) and hydrophobic residues (A11 and I12, orange) are shown and labeled. (B) $E_{NOE}$ and $E_{RDC}$ of the four families of structures from protocol two. (C). Similar to (A) except for protocol three. (D) $E_{PRE}$ and $E_{RDC}$ of the three families of structures from protocol three.

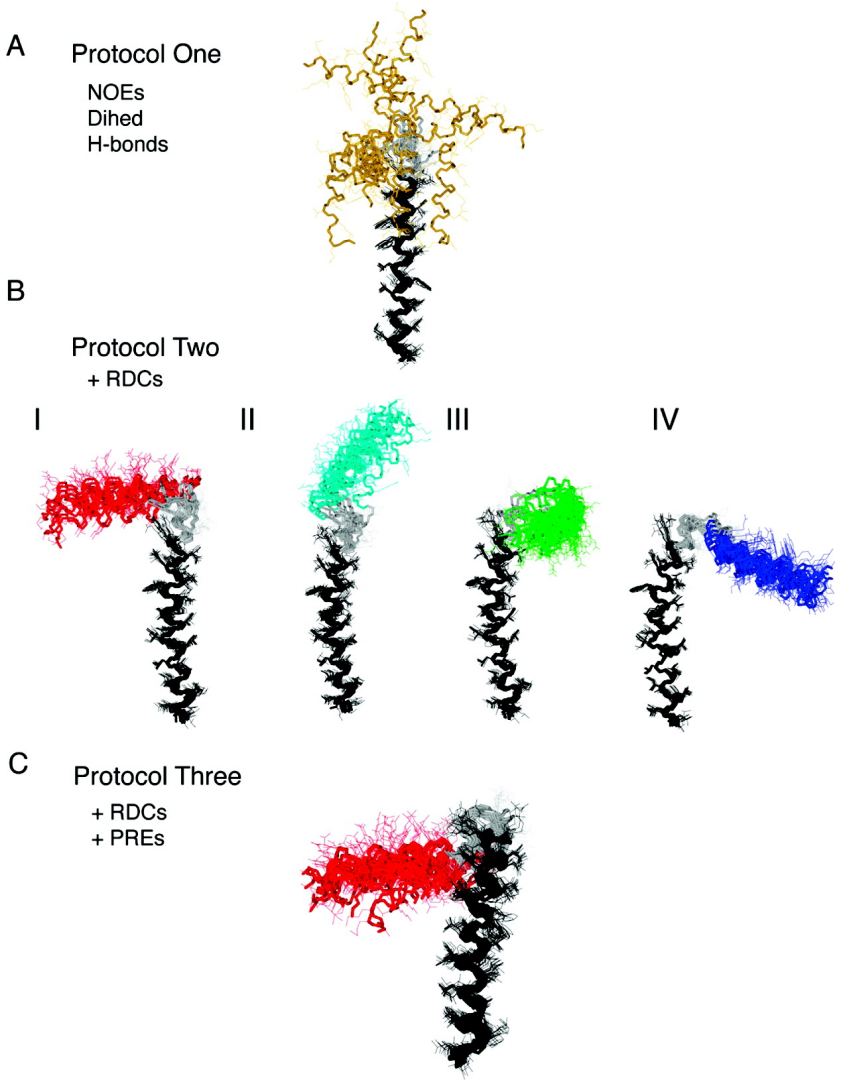


Figure 4.5: Overlay of different AFA-PLN conformational ensembles obtained by superimposing backbone atoms from residues 24-50 using MOLMOL. (A) Protocol one: ensemble obtained from NOEs, torsion angles, and hydrogen bonds (20 monomers). (B) Protocol two: RDC ensemble (15 monomers) with $^{15}N$-$^{1}H$, $^{13}C'$-$^{15}N$, and $^{(}13)C'$-$^{13}C_{\alpha}$ RDCs. (C) Protocol three: PRE ensemble (20 monomers) that utilizes PREs from MTSSL in addition to the restraints used in protocol two.

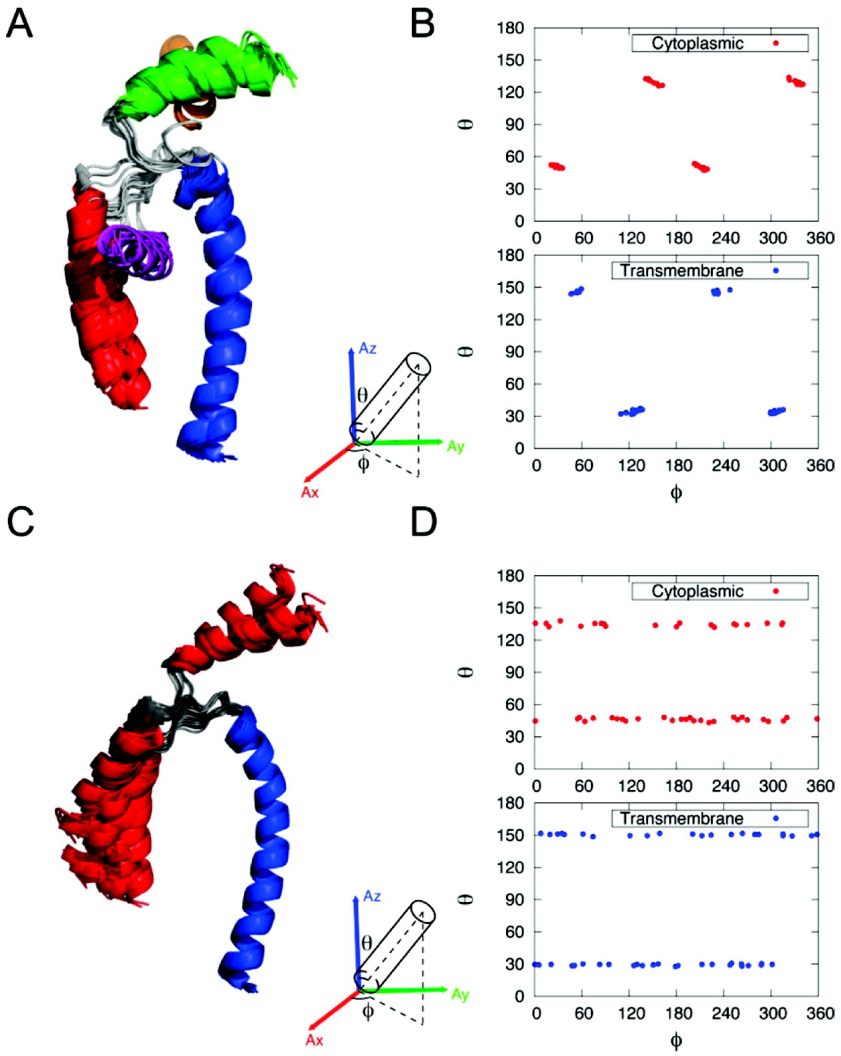


Figure 4.6: Orientation of rigid helix in different alignment tensor. (A) Overlay of monomer structures calculated using wt-PLN RDCs[289] with an asymmetric tensor ($R \neq 0$). (B) Orientation angles ($\theta$,$\phi$) of helical rigid bodies in the asymmetric alignment frame ($R \neq 0$). (C) Overlay of monomer structures calculated using wt-PLN RDCs(69) with a symmetric tensor ($R = 0$). (D) Orientation angles ($\theta$,$\phi$) of helical rigid bodies in the symmetric alignment frame ($R = 0$).

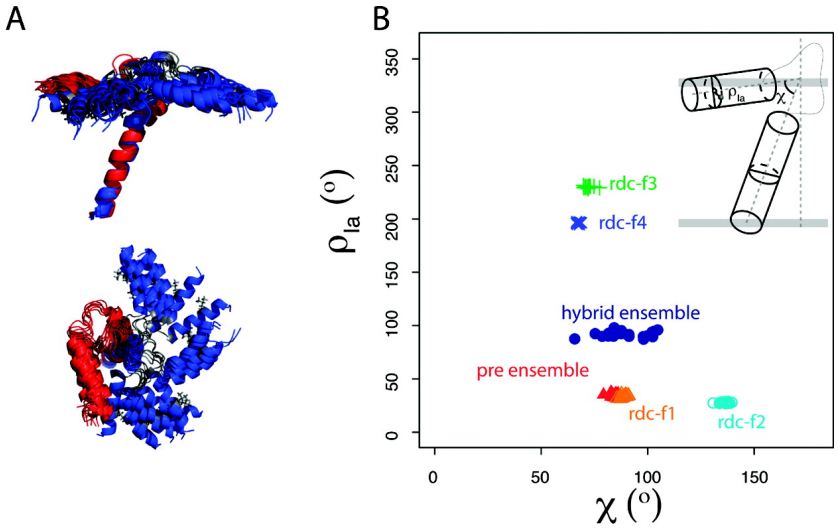


Figure 4.7: Compare micelle ensemble with hybrid ensemble. (A) Overlay of residues 31-52 from the PRE (red, protocol three) and hybrid[279] (blue) ensembles. (B) Plot of the interhelical angle ($\chi$) with the rotation angle of domain Ia ($\rho$) from the RDC-degenerate families (f1, orange; f2, cyan; f3, green; and f4, blue), PRE ensemble from protocol three (red), and hybrid ensemble (blue). The definitions of the $\chi$ and $\rho$ angles are shown within panel B.

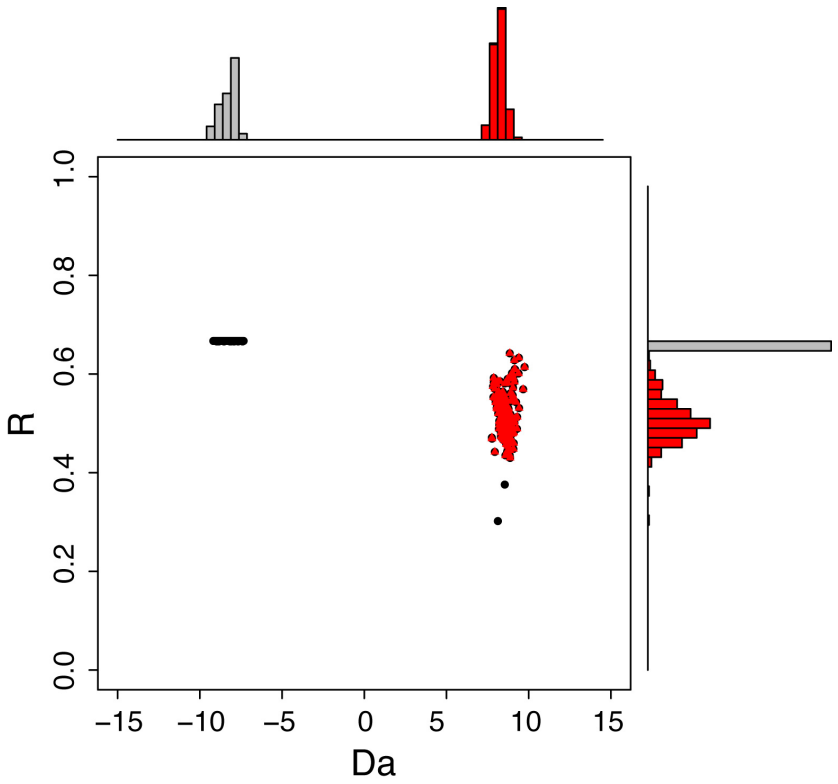


Figure 4.8: The ($D_a$, $R$) clustering (black) of structures refined with NOE and RDCs with variable tensor magnitude. Structures with $D_a > 0$ and $R > 0.4$ are selected for further structural calculation and analysis (red).

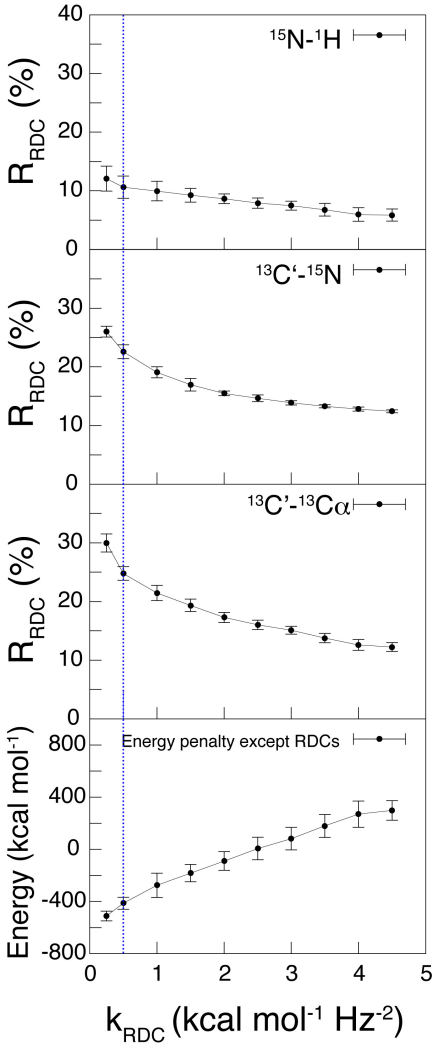


Figure 4.9: RDC R factors and energy value as functions of force constant used in the simulated annealing. The force constant for RDC was ramped from 0.25 $kcal \cdot mol^{-1}Hz^{-2}$ to 4.5 $kcal \cdot mol^{-1}Hz^{-2}$, while the force constants for other energy terms such as NOEs and torsion angles were fixed. RDCs are modeled using flat-well potentials and the relative weight between $^{1}H$-$^{15}N$, $^{13}C'$-$^{15}N$ and $^{13}C'$-$^{13}C_{\alpha}$ was set to be 1:1:1. The energy values are summations from different terms in Equation 1 other than RDCs and the magnitude indicates penalties from other solution NMR data (NOEs and torsion angles) and ideal geometry (bond, angle etc.).

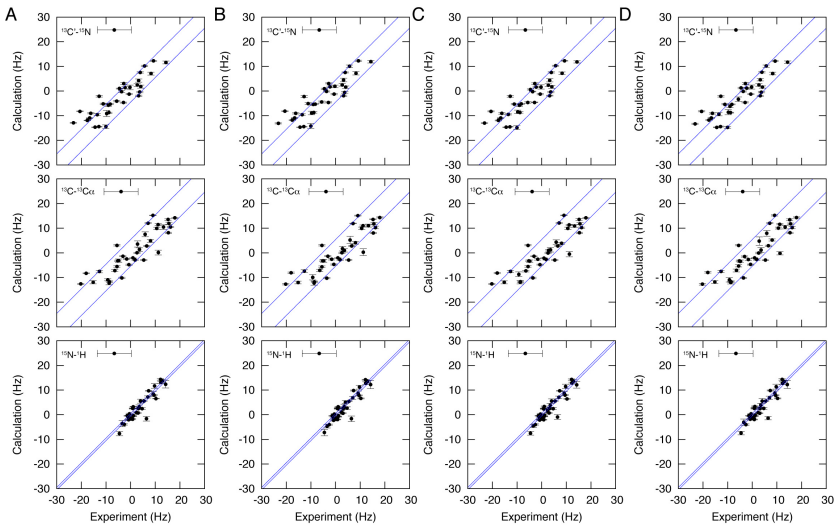


Figure 4.10: Correlation of experimental versus calculated RDCs of RDC ensemble family I,II, III and IV (15 monomers). Experimental errors of 0.5 Hz, 3.0 Hz and 4.5 Hz are used for $^{13}C'$-$^{15}N$ (A), $^{13}C'$-$^{13}C_\alpha$ (B) and $^{1}H$-$^{15}N$ (C) respectively.

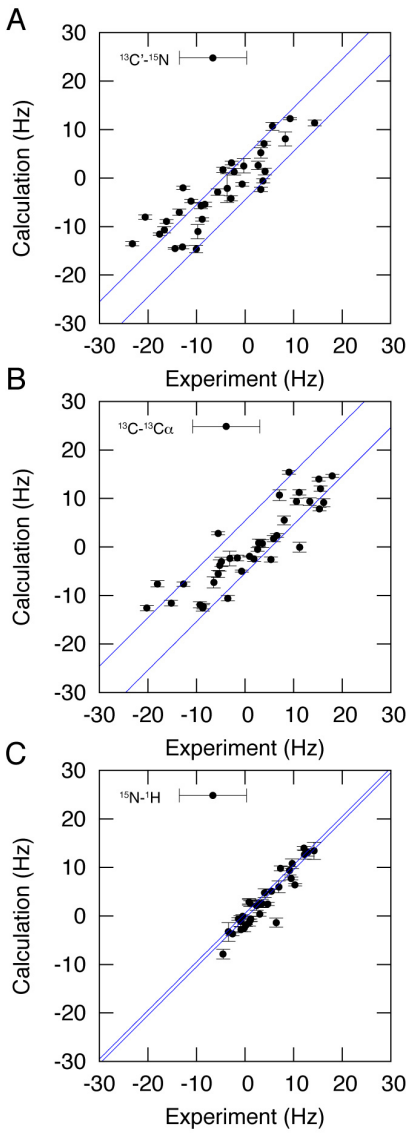


Figure 4.11: Correlation of experimental versus calculated RDCs from PRE ensemble (20 monomers). Experimental errors of 0.5 Hz, 3.0 Hz and 4.5 Hz are used for $^{13}C'$-$^{15}N$ (A), $^{13}C'$-$^{13}C_\alpha$ (B) and $^{1}H$-$^{15}N$ (C) respectively.

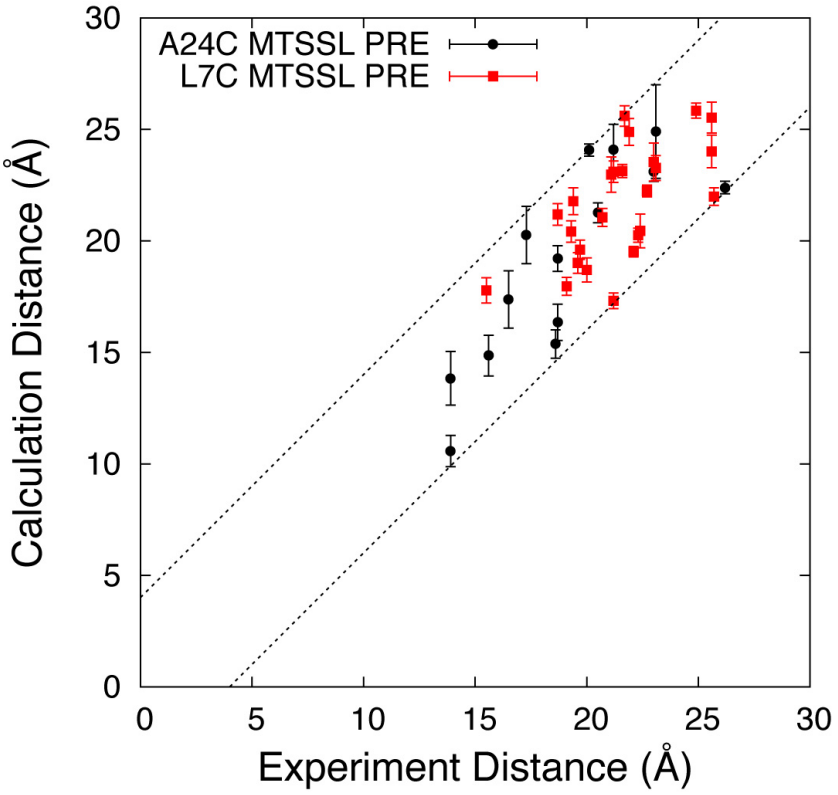


Figure 4.12: Correlation of experimental distances obtained from PREs (A24C MTSSL in black and L7C MTSSL in red) with calculated distances measured from the PRE ensemble (20 monomers). Experimental errors of 4 Å are plotted and calculated distances are shown with average and one standard deviation.

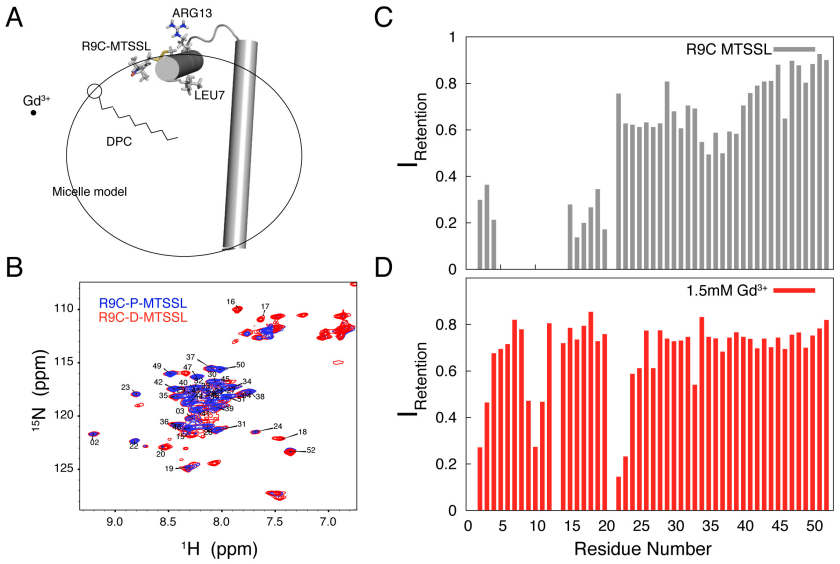


Figure 4.13: R9C-MTSSL spin label experiment. (A) Cartoon representation of R9C-MTSSL labeled AFAPLN in a simplified micelle model. The model is supported PRE data from NMR (B). The PRE quenching pattern in (C) indicates that the hydrophobic spin label prefers to insert to the micelle, quenching deeply inserted residues around 35. The $Gd^{2+}$ quenching, however, showed that the cytoplasmic domain lays on top of the micelle with preferential orientation for the hydrophobic residue (such as LEU7) and hydrophilic residues (such as ARG13).

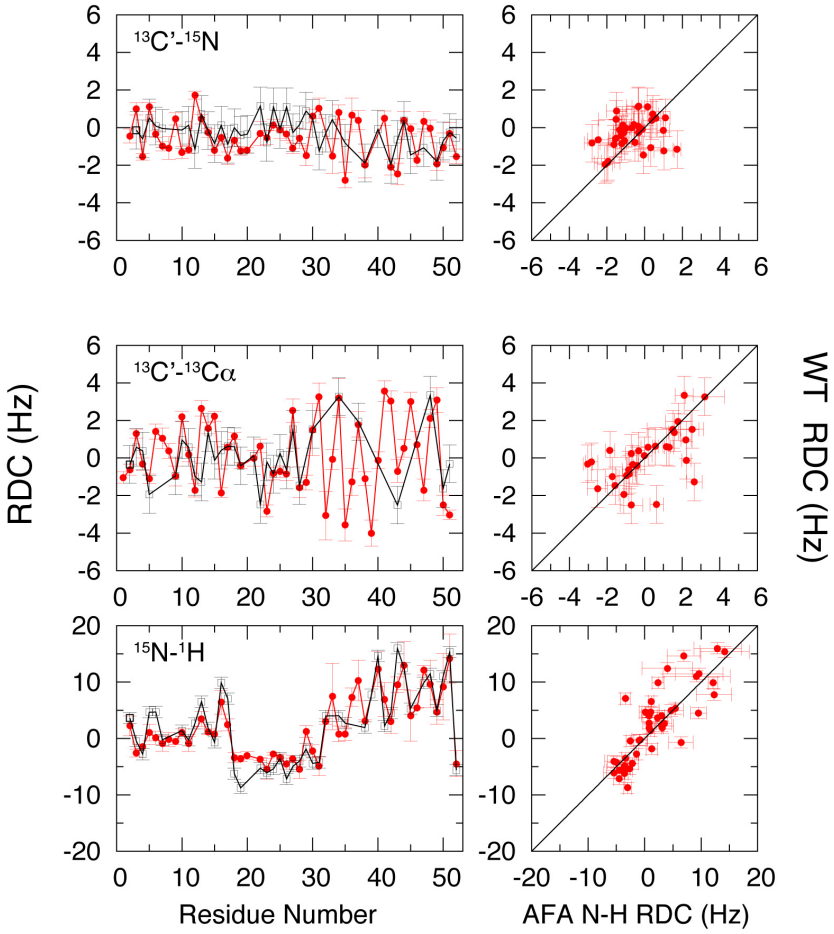


Figure 4.14: Comparison of the $^{1}H$-$^{15}N$, $^{13}C'$-$^{13}C_{\alpha}$ and $^{13}C'$-$^{15}N$ RDCs of AFA-PLN (monomer, red) with those of wt-PLN (pentamer, black) by residue number (left), with their correlations shown on the right.

# Chapter 5

# A refinement protocol to determine structure, topology, and depth of insertion of membrane proteins using hybrid solution and solid-state NMR restraints



## 5.1 Introduction

Nearly all high-resolution membrane protein structures deposited in the protein data bank (PDB) have been determined by X-ray crystallography, solution NMR, and cryoEM. While these techniques offer unparalleled atomic resolution that guides the interpretation of membrane protein biological function, the coordinates generated by the refinement procedures do not include a topological dimension. There are now several

examples in the literature, including mechanosensitive channels[298, 299], nucleobase-cation-symport-1 transporters[300], small multidrug resistant proteins[301], and potassium ion channels, that show protein function to be dictated by changes in the relative orientation of secondary structural elements with respect to the membrane (structural topology) rather than changes in the protein's secondary structure. Furthermore, other classes of membrane proteins do not even adopt a compact tertiary structure; rather their fold space is determined by the intrinsic interactions with the lipid membrane. Since membrane protein topology plays a fundamental role in protein function, the failure to describe such interactions results in incomplete structural characterization[259].

To mimic membrane protein environments, solution NMR spectroscopists utilize several different hydrophobic environments, including organic solvents, short chain lipid micelles, detergent micelles, or more recently, isotropic bicelles[19]. The deleterious effects of organic solvents on membrane protein function are well-documented and researchers have now almost completely abandoned this avenue. On the other hand, detergent micelles present a viable alternative. There have been several examples of micellar reconstitutions that preserve membrane protein functional integrity, giving rise to high-quality NMR spectra for structural and interaction studies[71, 261, 135]. However, the intrinsic curvature of micelles, which may impact membrane protein structure and topology, is a significant concern[274]. Small membrane proteins (30-200 residues), which account for most of the membrane proteins in several genomes, are especially vulnerable to changes induced by the membrane mimicking environment. While isotropic bicelles constitute an excellent membrane mimic and an alternative to micelles, the large size of the bicelle/protein complex results in substantially broader spectra than in detergent micelles[302, 303, 304].

Solid-state NMR of membrane proteins in oriented lipid bilayers is one of the most accurate methodologies to directly investigate protein topology. Improvements in sample preparation and NMR methodology have provided higher resolution and sensitivity leading to topological and dynamic studies of multispan membrane proteins[148, 305, 306, 19, 149]. However, these experiments rely on backbone measurements of chemical shift

anisotropy (CSA) and dipolar coupling (DC) using separated local field (SLF) experiments [reviewed in[178]] and heteronuclear correlation (HETCOR) spectroscopy[307]. The lack of side chain restraints has prevented complete structural characterization of membrane proteins using this technique.

A number of recent reports show that structure, topology, and oligomerization of membrane proteins are preserved both in lipid bilayers and in detergent micelles[293, 308, 309, 310]. In some instances X-ray structures of membrane proteins are almost superimposable to NMR structures determined in micelles[260]. Therefore, a logical solution to the structure determination problem is to combine information from different experimental approaches into one unique structural refinement protocol that provides structure and topology simultaneously. Based on these considerations, we propose a method that combines the high-resolution information obtained from solution NMR on membrane protein samples reconstituted in detergent micelles with the structural data obtained by solid-state NMR on proteins reconstituted into lipid bilayers. In this work, we illustrate how to implement both distance and orientational restraints into a single target function. A depth of insertion potential from DeGrado and co-workers is also employed to embed helical segments into a virtual bilayer[311]. The energy landscape is explored using the simulated annealing protocol implemented in XPLOR-NIH software[39] to determine the high-resolution structure and topology of membrane proteins simultaneously. We show the application of this method to monomeric phospholamban (PLN), a single pass membrane protein involved in cardiac muscle physiology[71, 261]. To obtain monomeric PLN, we mutated the three cysteines in the transmembrane domain (C36A, C41F, and C46A). These mutations maintain the functional integrity of the protein[278]. The solution NMR data was taken from the dodecylphosphocholine (DPC) studies from[278], while the solid-state NMR restraints were derived from the work in mechanically aligned lipid bilayers by[26, 279] and[284].

## 5.2 Methods and results

### 5.2.1 Energy terms in structural refinement

The hybrid solution and solid-state NMR target function ($E_{total}$) is formulated as a linear combination of geometrical ($E_{chem}$), solution NMR ($E_{sol-NMR}$), and solid-state NMR ($E_{ssNMR}$) terms:

$$E_{total} = E_{chem} + E_{sol-NMR} + EssNMR \tag{5.1}$$

The geometrical and solution NMR potentials were available in XPLOR-NIH force field version 2.18[39]. $E_{chem}$ is the sum of bonding ($E_{bonds}$, $E_{angles}$, $E_{improper}$) and non-bonding ($E_{vdw}$) interactions, with adjustable weighting factors ($w$),

$$E_{chem} = w_{bonds}E_{bonds} + w_{angles}E_{angles} + w_{improper}E_{improper} + w_{vdw}E_{vdw} \tag{5.2}$$

$E_{sol-NMR}$ is the sum of restraints from solution NMR experiments such as NOEs ($E_{NOE}$), hydrogen bonds ($E_{HBON}$), torsion angles ($E_{CDIH}$), and an empirical torsion angle potential from a database ($E_{DB}$;[312];[37]):

$$E_{sol-NMR} = w_{NOE}E_{NOE} + w_{CDIH}E_{CDIH} + w_{DB}E_{DB} + w_{HBON}E_{HBON} \tag{5.3}$$

To include both CSA and DC data obtained from SLF solid-state NMR experiments such as PISEMA (polarization inversion spin exchange at the magic angle)[313], we used the penalty function $E_{ssNMR}$:

$$E_{ssNMR} = w_{CSA}E_{CSA} + w_{DC}E_{DC} = w_{PISEMA}(E_{CSA} + w_rE_{DC}) \tag{5.4}$$

where $w_r$ is the relative weighting between CSA and DC and $w_{PISEMA}$ is the weighting of PISEMA potential both for CSA and DC. To implement both CSA ($E_{CSA}$) and DC ($E_{DC}$) potentials, we used flat-well penalty functions as reported by[314]:

$$E_{CSA} = \begin{cases} \sum \left(|\sigma_{c,i} - \sigma_{o,i}| - \sigma_{Error,i}\right)^2 & if\ |\sigma_{c,i} - \sigma_{o,i}| > \sigma_{Error,i} \\ 0 & if\ |\sigma_{c,i} - \sigma_{o,i}| < \sigma_{Error,i} \end{cases} \tag{5.5}$$

$$E_{DC} = \begin{cases} \sum \left(|\nu_{c,i} - \nu_{o,i}| - \nu_{Error,i}\right)^2 & if\ |\nu_{c,i} - \nu_{o,i}| > \nu_{Error,i} \\ 0 & if\ |\nu_{c,i} - \nu_{o,i}| < \nu_{Error,i} \end{cases} \tag{5.6}$$

$sigma_{o,i}(sigma_{Error,i})$ and $nu_{o,i}$ ($nu_{Error,i}$) are the experimental data (error) for CSA and DC and $sigma_{c,i}$, $nu_{c,i}$ are calculated values associated with residue $i$. In our implementation, this potential energy function allows for structural refinement using different experimental errors and CSA tensor values for individual residues. Similar energy functions for CSA and DC have been used to refine the backbone structures of membrane proteins[231, 222, 315, 316, 31].

### 5.2.2 Calculation protocol

**Step 1: implementation of solution NMR restraints** In step 1, we used solution NMR data with geometrical restraints to define the secondary structural elements of monomeric PLN (Figure 5.1).

**A.** The starting structure of monomeric PLN (the AFA-PLN mutant was used; C36A, C41F, C46A) was in an extended configuration, according to the simulated annealing protocol defined by[317].

**B.** Simulated annealing was carried out using the $E_{chem}$ and $E_{sol-NMR}$ potentials from Eq. 1. The system was cooled from 6,000 to 0 K in 5 K increments using 200 steps of torsion angle dynamics at each temperature. Both temperature and number of steps were optimized to achieve the lowest number of violations in the conformers generated. Note: the dihedral angle restraints for these calculations were generated from N, $H_N$, $H_\alpha$, C', $C_\alpha$, $C_\beta$ PLN chemical shifts using TALOS version 98.040.21.02[14].

**C.** 200 steps of molecular dynamics were performed in torsion angle space followed by minimization in both torsion angle and Cartesian spaces.

**D.** At the conclusion of step 1, we generated 200 structures of AFA-PLN, where the 100 lowest energy structures were used for conformational analysis.

Similar to the conclusions by[278], step 1 generates an ensemble of conformers with good convergence for the secondary structure elements, but that lacks precision in the structural overlay (see companion paper in[279]).

**Step 2: implementation of solid-state NMR restraints** In step 2, we refined the structures obtained in step 1 with the solid-state NMR data (DC and CSA) using

tensor 1 defined in Table 5.1.

**A.** Starting from the ensemble of 100 lowest energy structures, simulated annealing calculations were carried out using all potential energy terms in Eq. 1 ($E_{chem}$, $E_{sol-NMR}$, $E_{ssNMR}$). The initial annealing temperature was 3,000 K. Here, only torsion angle dynamics were applied to optimize the orientation of each individual peptide plane of AFA-PLN with respect to the Z-axis (corresponding to the lipid bilayer normal).

**B.** Cross-validation of CSA and DC restraints. To avoid overfitting the experimental data, we first optimized the solid-state NMR weighting factors ($w_{PISEMA}$ and $w_r$) in Eq. (4) based on the cross-validation factor R described by Cross and co-workers[318].

$$R = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{|calculated_i - observed_i|}{error_i}\right)^2 \tag{5.7}$$

N is the number of data points used in the refinement protocol, while $calculated_i$, $observed_i$ and experimental $error_i$ are the calculated value, experimental value, and experimental error in CSA and DC for residue $i$. We fixed the weighting factors for $E_{chem}$ and $E_{sol-NMR}$, and randomly partitioned the experimental data into two datasets: working (80% of data) and free (20% of data). The working dataset is included in the simulations, while the free dataset is not used and is back-calculated from the model. The work $R$ ($R_{work}$) indicates the fitting quality of the data included in the refinement, while the free $R$ ($R_{free}$) indicates the agreement between experimental and calculated data for residues with no CSA or DC restraints in the refinement protocol. In Figure 5.9, we varied two factors: (1) $w_{PISEMA}$ from 0.1 to 1,000 for a total of ten independent calculations and (2) $w_r$, the relative weight ratio of the DC/CSA potential, from 1/1 to 9/1. The $R_{work}$ and $R_{free}$ for CSA and DC were analyzed based on statistics for the 20 lowest energy structures from each of the calculations, as shown in Figure 5.9. As $w_{PISEMA}$ becomes too large (>10), the quality of the structures decreases ($E_{bonds}$, $E_{angles}$, $E_{improper}$, and $E_{vdw}$ become high) and the improvement in

the R values is minimal. Similarly, when $w_r$ is >3, the geometrical and solution NMR penalties increase dramatically without improving the PISEMA $R$ factor. When $w_r = 1$, there is a large DC $R_{free}$ factor, indicating poor correlation with the experiments. Based on Figure 5.9, the optimal weighting factor were found to be 5 for $w_{PISEMA}$ and 3 for $w_r$, which is in agreement with the 10 ($w_{PISEMA}$) determined by Cross and co-workers[318].

**C.** Using optimized weighting factor, a total of 400 structures were generated with the 100 lowest energy conformers selected for further analysis.

To evaluate the effects of the solid-state NMR restraints on refinement of the solution NMR structural ensemble, we defined three angles shown in Figure 5.2: $\theta$ (tilt angle with respect to the Z axis or membrane normal), $\rho$ (azimuthal rotation angle around helix axis), and $\chi$ [interhelical angle between domain Ia (residues 1-16) and the helix comprised of domains Ib and II (residues 23-52)]. The distributions of $\theta$ and $\rho$ for the two helical domains in the 100 lowest energy conformers are shown in Figure 5.3a. Noticeably, the orientation of the helix described by domains Ib and II is well-defined, with $\theta_{Ib,II} = 23 \pm 3°$ and $\rho_{Ib,II} = 205 \pm 3°$. The tilt angle of domain Ia is somewhat less defined with $\theta_{Ia} = 95 \pm 7°$. While the solution NMR ensemble (Figure 5.3b) fails to define the $\chi$ angle, the incorporation of CSA and DC restraints drastically reduces the conformational space allowed, confining $\chi$ between 70° and 125°. In contrast, the rotation angle ($\rho_{Ia}$) for domain Ia is not well-defined using PISEMA data alone. This originates from several different factors: (a) helices with tilt angles ∼90° have inherently clustered CSA (65-82 ppm) and DC (4-5 kHz) values, (b) helices with $\theta \sim 90°$ result in degenerate PISEMA spectra for $\rho = 0°$ or 180°, and (c) the dynamic nature of domain Ia results in broad resonances[319, 26]. While these spectroscopic difficulties were partially overcomed using several selectively labeled samples (see[279]), in order to determine the depth of insertion of the protein and fully resolve the orientation of the $\rho_{Ia}$ angle, we needed to incorporate a knowledge-based $E_z$ potential (step 3).

**Step 3: incorporation of the knowledge-based $E_z$ depth of insertion potential**

The knowledge-based $E_z$ potential by Degrado and coworkers[311] was used to

embed the helical domains of AFA-PLN into a virtual bilayer. The $E_z$ pseudo-energy was derived from the propensity of each amino acid in helical membrane proteins to insert into a lipid bilayer. This knowledge-based potential was derived from the statistical analysis of 24 crystal structures of helical membrane proteins. In our calculation, the helical domains were independently optimized under the knowledge-based $E_z$ potential using rigid-body minimization, where the only degree of freedom is the translation of an entire domain along the Z-axis. Note that the loop connecting domain Ia with domain Ib was allowed to move and adjust the depth of insertion for all regions of AFA-PLN. The incorporation of the $E_z$ term into the total energy function for simulated annealing calculations resulted in erroneous insertion of the protein in the virtual membrane irrespective of the weighting factor used (data not shown). This results from the step-function shape of the knowledge-based $E_z$ potential which acts on each residue only along the Z axis. While this potential applied to single residues is not sufficient to embed the protein in the virtual membrane, the sum of the $E_z$ potentials for all of the residues in a helix results in a unique minimum. Below we summarize the steps used to incorporate this knowledge-based potential into our calculations.

**A.** Definition of virtual bilayer. A virtual bilayer was built with the membrane normal parallel to the Z-axis and the origin in the center of the hydrocarbon core. The Z coordinate of the COM (center of mass) for each helical domain (using $C_\beta$ atoms since $E_z$ potential depends only on $C_\beta$ coordinates) describes its insertion into the virtual bilayer.

**B.** Starting with the 100 lowest energy structures obtained after step 2, we minimized the domain insertion along the Z coordinate by rigid-body minimization using the $E_z$ potential (i.e., the helical regions of domain Ia (residues 3-15) and domains Ib and II (residues 23-47) were held rigid). The use of the $E_z$ potential helped to resolve the rotation angle $\rho_{Ia}$ ambiguity. Only structures with $\rho_{Ia}$ within the range of 60-150° gave reasonable minimized structures (Figure 5.10). These angles are consistent with results from paramagnetic quenching experiments and NOEs connecting the hydrophobic face of domain Ia to the detergent micelle[71, 261, 278]. The introduction of

the $E_z$ potential confined the trans-membrane helix into a well-defined minimum, while the depth of insertion of the cytoplasmic domain was less defined with respect to the virtual membrane (Figure 5.4). Regardless, the translational degeneracy was removed, resulting in one population of conformers with the hydrophilic residues pointing toward the bulk water, in agreement with the amphipathic nature of the cytoplasmic helix. Note that domain Ib in the N-terminal portion of the transmembrane domain protrudes outside the virtual membrane. This is in agreement with the hydrophilic nature of this domain and explains the hydrogen/deuterium exchange data[278]. The flexible loop is also exposed to the solvent, which is in agreement with the dynamics probed by NMR spin relaxation measurements[217].

**C.** In order to relax local geometries and improve the structure quality, we subjected the remaining structures to low temperature (300 K) simulated annealing using torsion angle and Cartesian molecular dynamics with full van der Waals interactions and all of the other restraints, but excluding the $E_z$ potential. Although the knowledge-based $E_z$ potential was not used during this annealing step, we verified that the structures did not deviate from the depth-of-insertion minima shown in Figure 5.5b. From this ensemble, we selected 20 low energy conformers of PLN with no restraint violations, and deposited the structures into the protein data bank (PDB 2KB7,[279]). The biological relevance of structures is discussed in our companion paper[279]. Structural statistics are provided in Table 5.2.

### 5.2.3 Effects of different tensor components

One debate in the field surrounds the $^{15}N$ CSA and $^{1}H$-$^{15}N$ DC tensor components used to model solid-state NMR data obtained from PISEMA-type experiments. To address this issue, we carried out a systematic conformational search for AFA-PLN domain Ib and II utilizing different tensor components shown in Table 5.1. We performed Steps 1-2 with low temperature annealing as discussed in the "Methods and results" section for each of the three tensors given in Table 5.1. Following energy minimization, the 20 lowest energy structures are selected with good agreement between the experimental and calculated DC and CSA using the three different tensors (Figure 5.6).

While small changes in local structure may occur, differences in the tensor values only marginally affect the overall protein topology. As shown in Table 5.1, changing the tensor components (tensors 2 and 3) from those used for the majority of this work (tensor 1), caused the tilt ($\theta$) and azimuthal ($\rho$) angles to vary by 2.3° and 1.1°, respectively. Given the approximations necessary to calculate these angles for the transmembrane domain of PLN, these variations are negligible. In addition to the overall orientation, we also examined effects of different tensors on the local structures. Due to the nature of structural calculation, the bond angle, bond length and peptide plane planarity are all within acceptable errors of peptide geometry. We thus compared the Ramachandran angle ($\phi$, $\psi$) for the transmembrane domain shown in Figure 5.11. Although structures remained helical from all sets of simulations, there were considerable variations in these angles. These are likely due to the inclusion of different CSA and DC tensors. Note that for glycine residues, a different set of tensor components needs to be used[169].

### 5.2.4 Sensitivity to PISEMA misassignments

Solid-state NMR methods to obtain orientational information from assigned PISEMA resonances are well-established. However, the assignment procedures for real membrane protein samples are still in their infancy. Although PISA wheels are a common way to assign resonances in well-dispersed spectra, most experimental PI-SEMA spectra contain inhomogenous broadening of resonances (due to mosaic spread and protein dynamics) making the assignments challenging. In spite of several computational attempts, these effects have not been fully rationalized[168, 169]. Therefore, selective labeling techniques constitute an important resource for resonance assignment. Unfortunately, if residues are on the same face of a helix, degenerate frequencies will be expected, further hampering the assignment procedure. To assess the sensitivity of our method to this problem, we have carried out two different calculations with two equiprobable assignments for the [$^{15}N$-Ile] AFA-PLN PISEMA spectrum shown in Figure 5.7b (see[279]). The assignment was carried out using a combinatorial search routine as described by[27]. Using our hybrid method, we found that the two resonance assignments result in two conformational ensembles with similar tilt angles ($24 \pm 3°$ and $22 \pm 2°$), but different azimuthal rotation angles ($203 \pm 4°$ and $189 \pm 3°$). The conformational ensemble with the incorrect assignment (black assignment in Figure 5.7) displayed covalent geometry

violations at Ile45 for more than half of the structures, while also giving consistently higher conformational energy. This demonstrates that our hybrid method is sensitive to misassignment, and that it can be used to assist the assignment process.

### 5.2.5 Static helix approach

PISEMA data has been used to calculate the tilt and rotation angles of helices based on a static ideal helix[233]. In Figure 5.8, we show the agreement between the tilt and rotational angles with the PISEMA data using a static approach. The potential energy landscape represented in Figure 5.8 was obtained by calculating the least square fit between the experimental PISEMA data and an ideal poly-alanine $\alpha$-helix. The lowest energy conformers generated with the hybrid method are located at the bottom of the minima defined by tilt and rotation angles for domains Ia and the helix comprised of domains Ib and II.

This indicates that (1) the lowest energy conformers generated with the hybrid method are in remarkable agreement with the experimental data and (2) there is negligible discrepancy between the static fitting of the PISEMA data and a more refined conformational search procedure such as simulated annealing. A corollary to these considerations is the ideal nature of transmembrane helices embedded in lipid bilayers, demonstrated in several studies by Cross and co-workers using numerous membrane proteins[137]. Nevertheless, compared to the fitting with a rigid, ideal helix, the hybrid approach has the advantage of obtaining a high-resolution structure for both the backbone and side chains.

## 5.3 Discussion

The correlation between structure and function of complex biological systems requires the combination of several different techniques (biophysical and biochemical) to validate any mechanistic conclusions. The combination of methodologies, hybrid methods, is becoming quite common in structural biology[320]. For instance, Clore and co-workers have used X-ray crystallographic data synergistically with solution NMR restraints to refine the structure of soluble proteins[321]. More recently, solution NMR and small

angle X-ray scattering (SAXS) data have been combined to refine the structures of large complexes and to resolve the intrinsic ambiguities of residual dipolar couplings[322, 323]. Others have combined long-range distances from EPR with shorter range solution NMR distances for structure determination[324]. Finally, sparse solid-state NMR data have been used to complement X-ray structures to orient $\beta$-barrels within lipid bilayers[325].

In this work, we combined structural restraints from solid-state and solution NMR into a global target function, which includes a depth of insertion potential[311]. We applied this method to monomeric PLN, a small membrane protein that adopts an L-shaped conformation in lipids and whose tertiary fold is dictated by its strong interactions with the lipid membrane (see[279]). Solid-state and solution NMR have been previously used to study several membrane protein systems, demonstrating remarkable similarities of membrane protein structures in detergent micelles and lipid bilayers[293, 140, 147, 19, 142]. Consistent with these findings, our calculations show that the structural restraints for both solution and solid-state NMR are compatible and can be used synergistically to define the structure and topology of PLN in lipid membranes. In addition, the depth of insertion potential gives a clear picture of the membrane location with respect to the protein that can be used as a starting point to embed it into an explicit lipid bilayer for further refinement. In fact, this final refinement step of the structural ensemble has already been adopted for soluble proteins[326, 327, 328, 329], and will be especially important for membrane proteins to (1) better define the fold space and (2) image protein-lipid interactions at the atomic level. Table 5.2 reports the statistics on the conformers generated with the hybrid method and a comparison with the conformational ensemble derived from the solution NMR data alone. It is clear that the introduction of the solid-state NMR data does not significantly affect the covalent geometry. The changes in the ideal geometry are within the allowed errors and all of the conformers have passed the conformational filters of the PDB deposition site. Also, the resolution of each single helical domain does not change. However, the overall resolution of the backbone changes dramatically. While in the conformational ensemble generated from the solution NMR data alone the RMSD for all of the backbone atoms was $\sim 4.4$Å, the introduction of the solid-state NMR data drastically reduced the RMSD to $\sim 2.3$Å. It is apparent that the interactions of PLN with the lipid membrane exemplified by the

solid-state NMR data and the knowledge-base $E_z$ potential limit the conformational space defining the structural topology of PLN. Nonetheless, the resolution achieved underscores the dynamic nature of PLN, a property that allows this small membrane protein to interact with several different binding partners.

Several other restraints obtained from biophysical and biochemical techniques can be easily incorporated in this protocol along the lines of the HADDOCK approach[330]. In particular, it is important to measure depth of insertion experimentally using either paramagnetic quenching[331] or protein/lipid REDOR measurements[139, 332, 333] to complete the structural and topological characterization of membrane proteins within lipid bilayers. Although this approach is demonstrated for a small membrane protein, we are currently applying this method to the PLN pentamer (30 kDa), which will demonstrate the applicability to other multispan membrane proteins and complexes.

## 5.4 Conclusions

We developed a new computational protocol that is based on a total objective energy function ($E_{total}$) that incorporates solution NMR data (distances, torsion angles, and hydrogen bonds), solid-state NMR data (orientational information: CSA and DC) and a depth of insertion knowledge-based potential ($E_z$) to determine the high-resolution structure and topology of membrane proteins. This hybrid energy function has been implemented into an XPLOR-NIH protocol and made available for general use. While this method is demonstrated for a small single pass membrane protein, the rapid progress made in sample preparation and spectroscopy of medium size membrane proteins will allow application to larger systems. The XPLOR-NIH scripts for the hybrid method are available for download at the authors' website (www.chem.umn.edu/groups/veglia).

Table 5.1: Different tensor values (tensor 1[334] and 2[137]) and the resulting topological angles

| **Name** | **Tensor values** | | | **Topological angles Ensemble** |
|---|---|---|---|---|
| | $(\sigma_{11}, \sigma_{22}, \sigma_{33})(ppm)$ | $\nu_{NH}$**(kHz)** | $\theta_{Ib,II}(^\circ)$ | $\rho_{Ib,II}(^\circ)$ |
| Tensor 1 | (64,77,217) | 9.75 | 24±2.6 | 203±3.7 |
| Tensor 2 | (57.3,81.2,227.8) | 10.735 | 26±2.9 | 204±3.2 |
| Tensor 3 | (64,77,217) | 10.735 | 24±2.6 | 202±3.9 |

Table 5.2: NMR Refinement Statistics of AFA-PLN for solution NMR ensemble and hybrid ensemble

| **Structure quality Analysis** | **Solution NMR Ensemble** | **Hybrid Ensemble** |
|---|---|---|
| **R.M.S. Deviation from experimental restraints** | | |
| NOE/H-bond (Å)(422) | $0.06 \pm 0.004$ | $0.06 \pm 0.006$ |
| Torsion angle (°)(38) | $0.1 \pm 0.08$ | $0.4 \pm 0.22$ |
| **PISEMA R-factor (%)** | | |
| CSA(39/53) | N/A | $0.9 \pm 0.02$ |
| DC (39/53) | N/A | $1.0 \pm 0.03$ |
| **R.M.S. deviations from idealized covalent geometry** | | |
| Bond (Å) | 0.003±0.0004 | 0.006±0.0004 |
| Angle (°) | 0.4±0.01 | 0.7±0.03 |
| Impropers (°) | 0.3±0.02 | 0.2±0.002 |
| **Measure of structural quality** | | |
| % residues in most favored region | 93±2.7 | 90.0±2.1 |
| % residues in additional allowed region | 7.0±2.7 | 8.0±2.7 |
| % residues in generously allowed region | 0.0 | 2.0±2.5 |
| % residues in disallowed region | 0.0 | 0.0 |
| **Precision of atomic coordinates (Å)** | | |
| Backbone (residue 3-50) | 4.4 | 2.3 |
| Helix 1 (residue 3-18) | 0.6 | 0.6 |
| Helix 2 (residue 24-50) | 0.7 | 0.6 |
| **Helix Topology** | | |
| Helix 1 (residue 3-18) | | |
| Tilt ($\theta$) (°) | N/A | 102±2.2 |
| Rotation ($\rho$) (°) | N/A | 92±2.8 |
| Insertion ($Z$) (Å) | N/A | 16±0.8 |
| Helix 2 (residue 24-50) | | |
| Tilt ($\theta$) (°) | N/A | 24±2.5 |
| Rotation ($\rho$) (°) | N/A | 204±3.8 |
| Insertion ($Z$) (Å) | N/A | 5.5±0.5 |

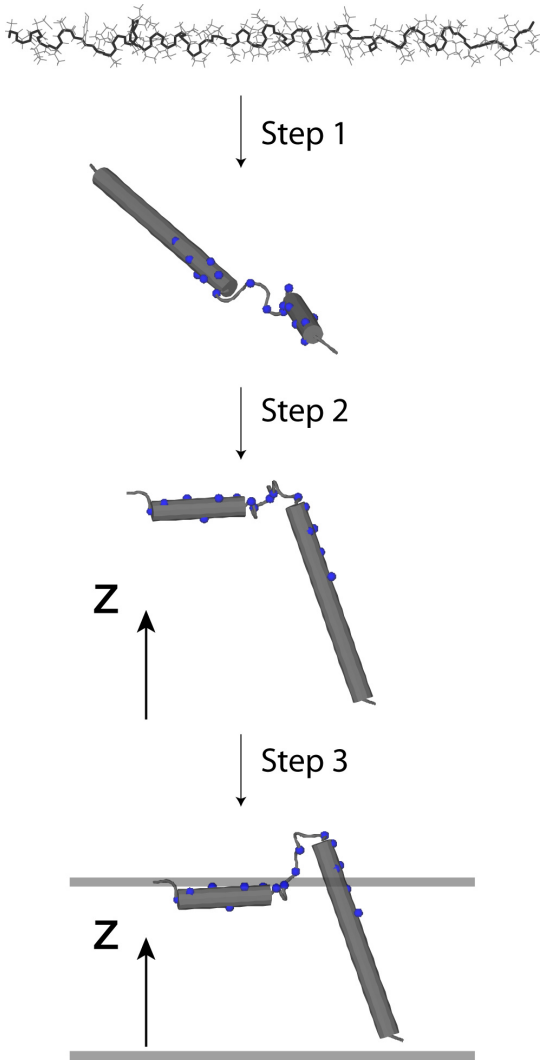


Figure 5.1: Overview of the hybrid refinement protocol for the simultaneous determination of structure and topology of membrane proteins. Step 1: Starting from an extended structure, the simulated annealing protocol minimizes a target function containing only solution NMR data (NOEs, torsion angles, and hydrogen bonds). Step 2: The orientational constraints derived from solid-state NMR are included together with solution NMR restraints to obtain the correct orientation along the Z direction. Step 3: Depth of insertion is determined using rigid body minimization in the presence of the depth of insertion potential, keeping the helical orientation with respect to Z fixed. The resulting structural ensemble is refined using low temperature simulated annealing

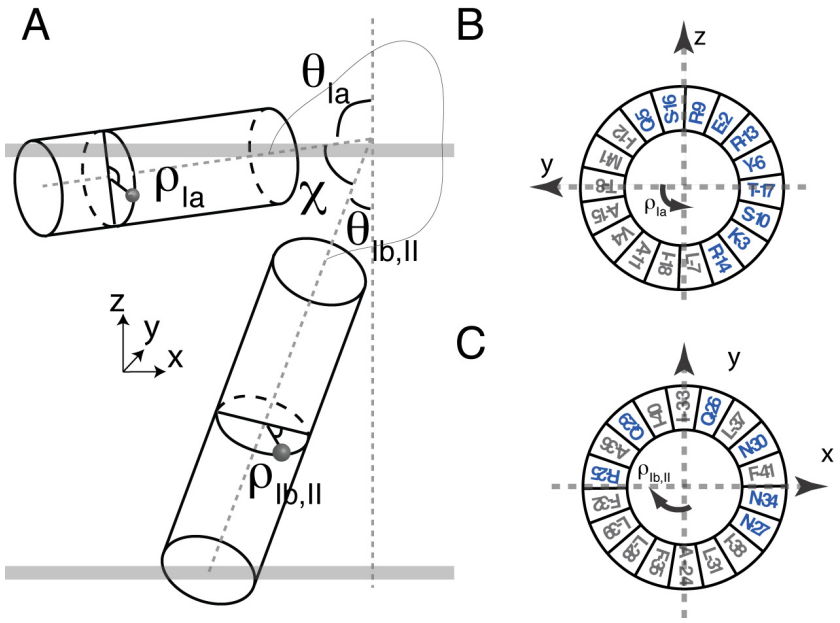


Figure 5.2: A Definition of ($\theta$, $\rho$) describing the orientation of helix with respect to membrane normal Z. ($\theta_{Ib,II}$, $\rho_{Ib,II}$) are the tilt and rotation angles for the domain Ib,II helix, while ($\theta$Ia, $\rho$Ia) are the tilt and rotation angles for the domain Ia helix. The interhelical angle between the two domains is described by $\chi$. B. Helical wheel representation of the reference orientation of domain Ia where $\rho_{Ia}$ is defined to be zero. The N atom of T8 is aligned to +y axis. $\rho$Ia rotates counterclockwise viewing from the top of y-z plane. At $\sim$ 90 degrees, the hydrophilic residues in blue point into bulk solution. C. Reference orientation of transmembrane domain where $\rho$Ib,II is defined to be zero

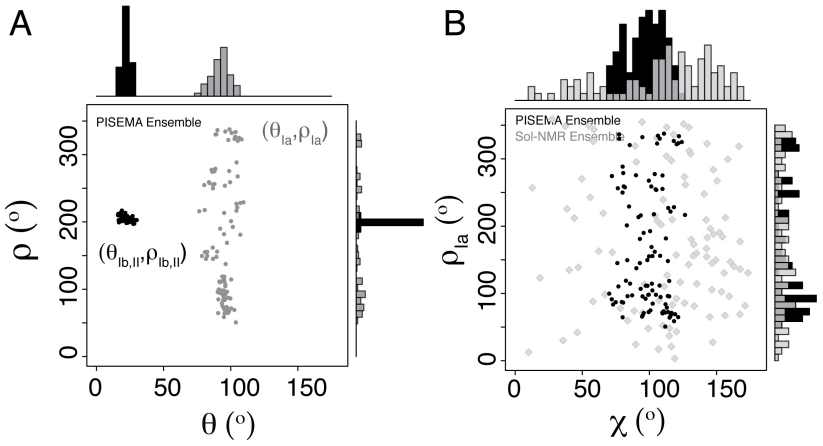


Figure 5.3: A. Distribution of $\theta$ and $\rho$ angles derived from domain Ia and domains Ib and II. B. Distribution of interhelical angle $\chi$ and $\rho_{Ia}$ and the comparison with the solution NMR ensemble

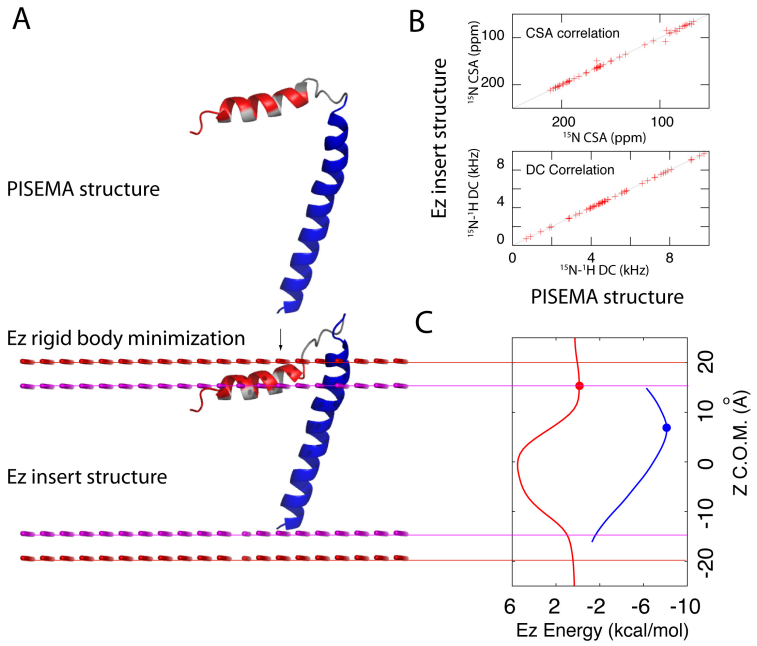


Figure 5.4: Rigid-body minimization using the knowledge-based $E_z$ potential. A. Backbone cartoon representation of a selected PLN conformer before (upper panel) and after rigid-body minimization. B. Comparison of the simulated CSA and DC before and after $E_z$ minimization. The only discrepancies are due to residues located in the dynamic loop, which do not have CSA and DC restraints. C. Representation of the $E_z$ potential energy function for the domain Ib, II (blue) and domain Ia (red) helices. After minimization, both domains reside in the minima

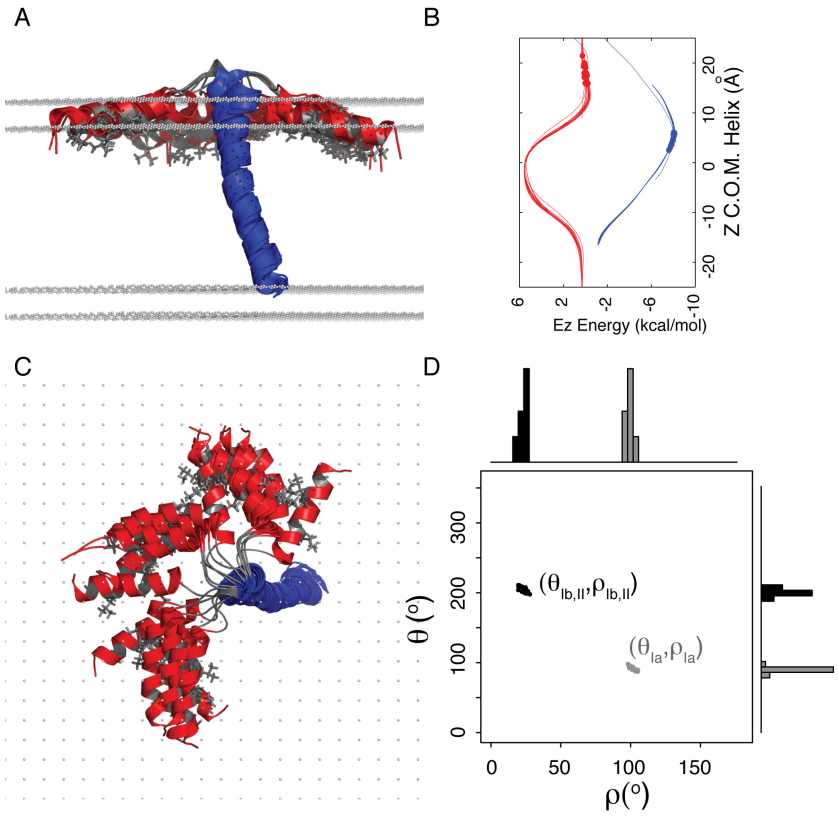


Figure 5.5: A. Conformational ensemble (structures and topologies) representing the 20 lowest energy structures in the virtual bilayer. Domain Ia is colored in red, and domains Ib and II are in blue. Hydrophobic side chains of domain Ia are shown in grey. Structure overlay is performed by rotating along Z and translating along X and Y, resulting in no changes in $E_z$ energy and PISEMA data. B. Position of the cytoplasmic and transmembrane domains with respect to the depth of insertion $E_z$ potential. C. Top view of a. D. Distribution of $\theta$ and $\rho$ angles in the final structural ensemble

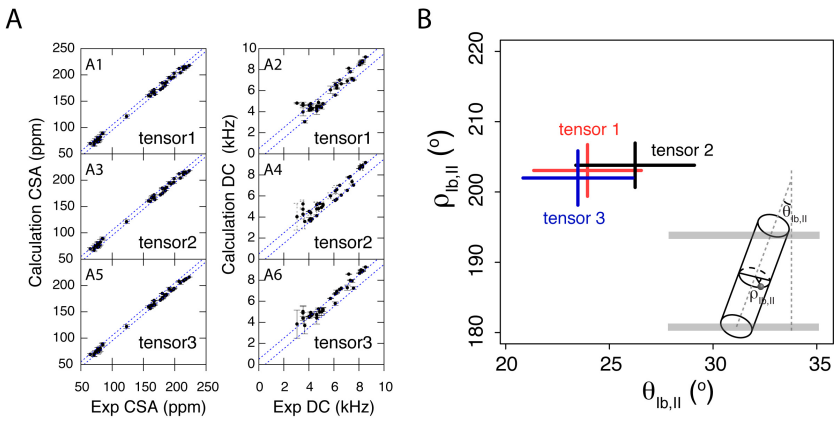


Figure 5.6: A. Comparison of the experimental and calculated CSA and DC for different values of the CSA tensor components shown in Table 5.1. The diagonal errors indicate ranges of±5 ppm and±0.5 kHz for CSA and DC, respectively. B. Effects of the different CSA tensor components on the tilt and rotation angles for the helix defined by domains Ib and II of PLN. Tensors 1[334], 2[137], and 3 are shown in red, black, and blue, respectively

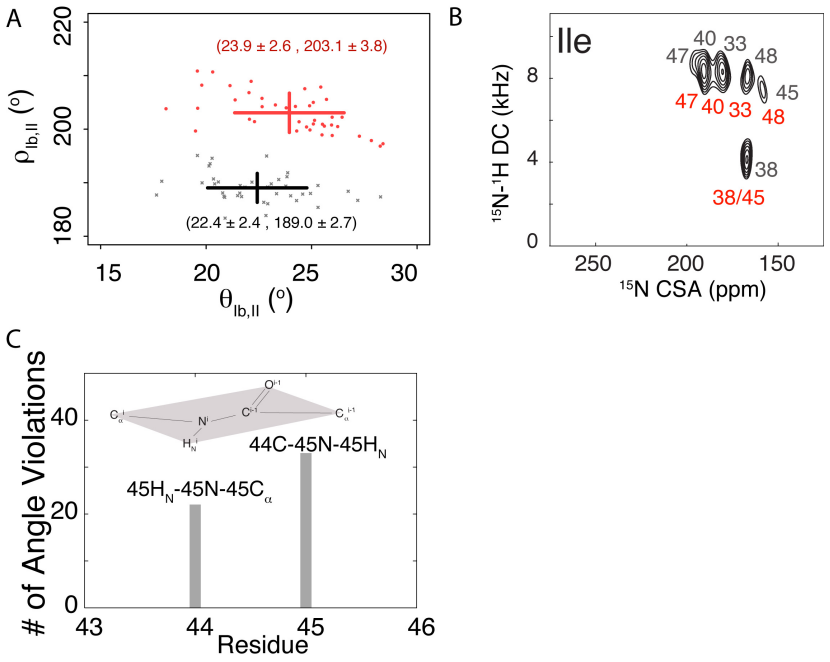


Figure 5.7: Effects of the resonance misassignment on the calculation of the structural topology of PLN. a Plots of the distribution of the rotation angles for the two ensembles obtained from simulated annealing calculations. The correct assignments give rise to an average rotation angle of 203°, while the incorrect assignment gives rise to an average rotation angle of 189°. b Experimental PISEMA spectra with the two equiprobable assignments for the isoleucine residues of PLN derived from the combinatorial assignment procedures (see text). c Histogram representing the number of violations obtained for residues 44 and 45 in the conformational ensemble generated with the incorrect assignments

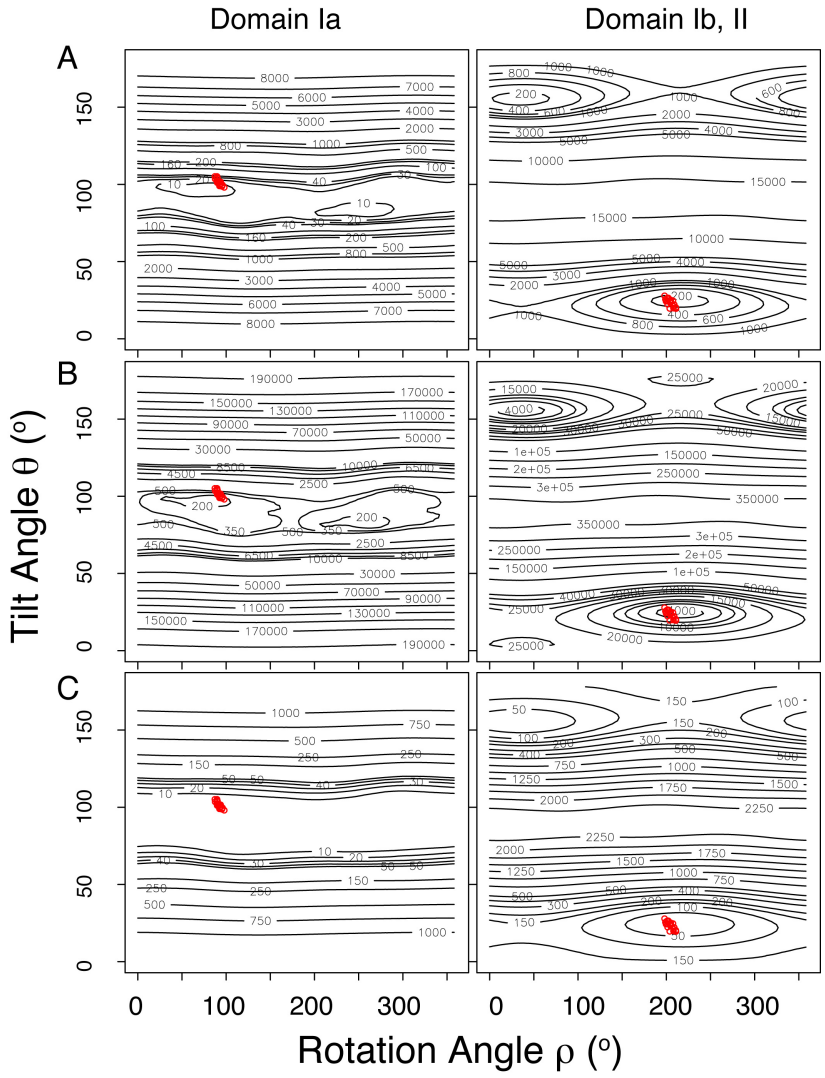


Figure 5.8: Projection of ($\theta$, $\rho$) of the hybrid ensemble (20 lowest) for the two helical domains of PLN onto the PISEMA potential surfaces obtained using rigid helix fitting. Three different scoring functions are used to generate the potential surfaces: These plots demonstrated that the topologies derived from the hybrid method correspond to the lowest energy minima identified in all of the PISEMA potential surfaces

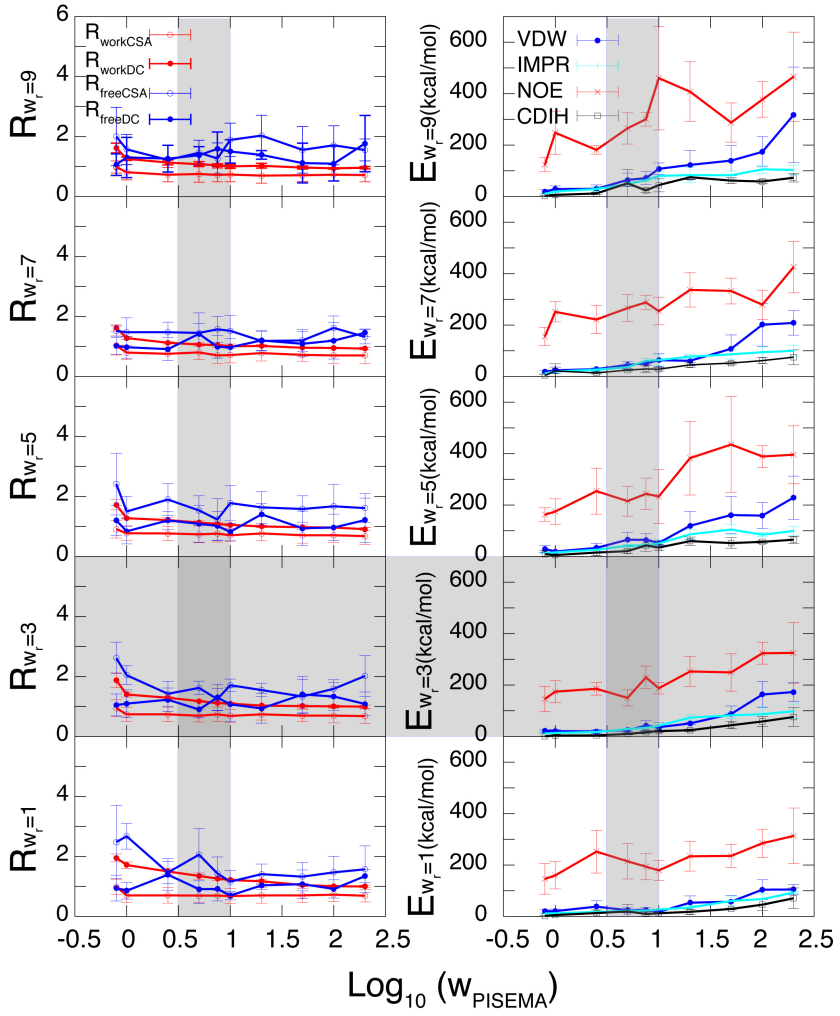


Figure 5.9: Optimization of the weighting factor $w_{PISEMA}$ for PISEMA restraints and the relative weighting between DC and CSA $w_r$ using cross-validation. $R_{workCSA}$, $R_{workDC}$, $R_{freeCSA}$, $R_{freeDC}$ indicate the work R factors for CSA, DC and free R factors for CSA and DC. VDW, IMPR, NOE and CDIH are the output energies for non-bond interactions, improper angle, NOE restraints and dihedral angle restraints. The subscript $w_r$ denotes different weighting factors used in the simulations. The shaded regions show the optimal regions for $w_{PISEMA}$ and $w_r$.

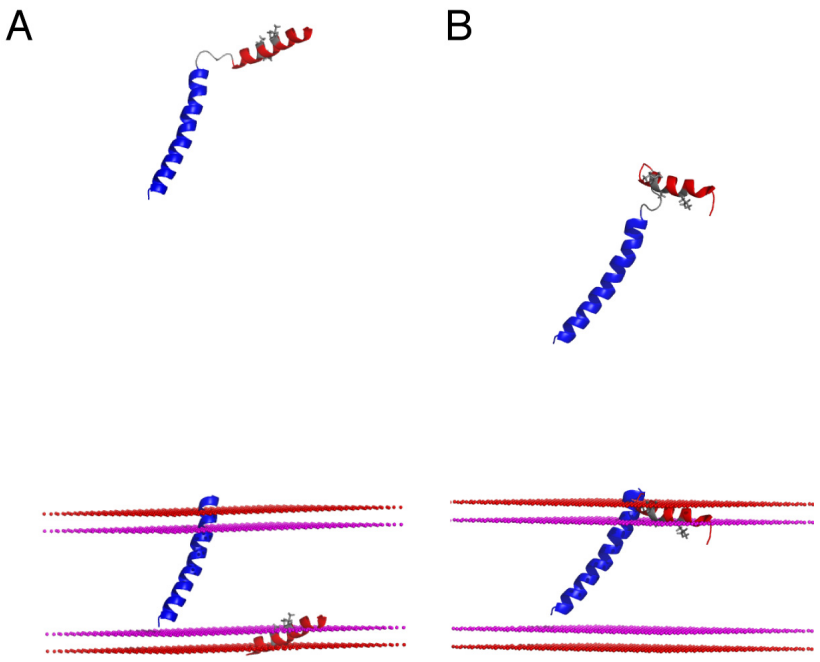


Figure 5.10: Selection criteria based on the $E_z$ potential for the PLN conformers reported in Figure 5.11. (A) Representative conformer of the PLN structures with $\rho_{Ia}$ angle greater than 150°. These conformershave the hydrophobic residues facing towards the bulk solution. The only possibility to minimize their insertion is to break the loop connecting the two helical domains. This is energetically prohibited. (B) Representative conformer of the PLN structures with $\rho_{Ia}$ angles less than 150°. The insertion of these conformers in the virtual membrane with energetically favorable $E_z$ energy does not require the disruption of the connecting loop. These conformers were selected for further analysis.

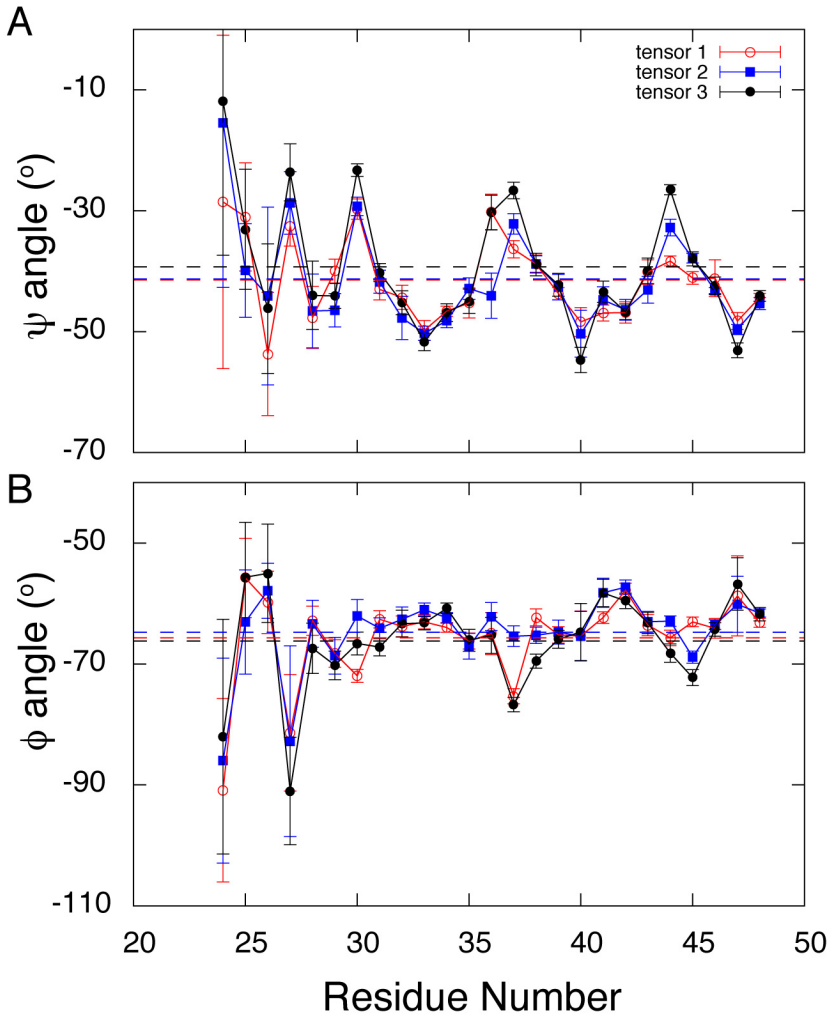


Figure 5.11: Effects of different tensor values on the backbone torsion angles of domain Ib and II. Their averages are shown in dotted lines.

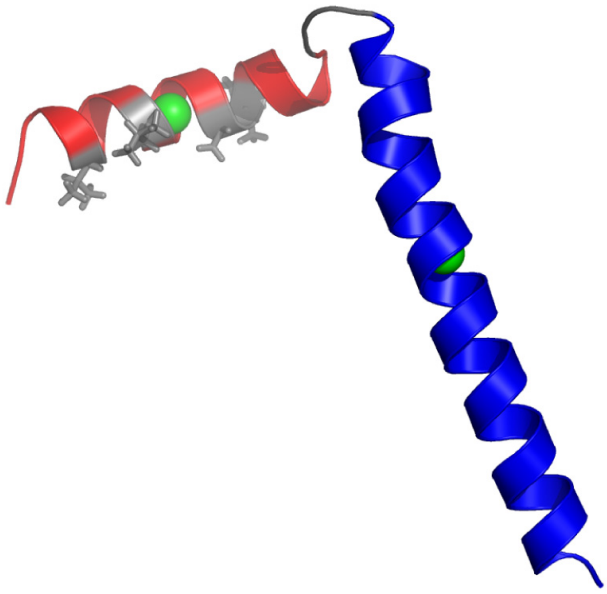

Figure 5.12: Cartoon representation of one AFA-PLN structure. The centers of masses ($C_\beta$ atoms of domain Ia (residue 4:16) and domain Ib and II (residue 24:48)) are shown by green spheres.

# Chapter 6

# Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach



## 6.1 Introduction

Structure and topology are central to membrane protein function[259]. Recently determined high-resolution structures reveal the compact folds for several membrane proteins, such as electron and proton-conducting proteins involved in photosynthesis and respiration ($http://blanco.biomol.uci.edu/Membrane\_Proteins\_xtal.html$). However, a significant population of membrane proteins does not possess a compact tertiary fold, but has its fold space defined through interactions of secondary structure elements (helices, turns, and loops) with the lipid membrane, i.e., the topology[259]. This

is the case for phospholamban (PLN), a mammalian protein that is essential in the regulation of cardiac muscle contractility[335], and that has recently become a major target for gene therapy to ameliorate cardiomyopathies[336, 337]. PLN is located in the sarco(endo)plasmic reticulum (SR) of cardiac myocytes, inhibiting the SR $Ca^{2+}$-ATPase (SERCA) by shifting its relative $Ca^{2+}$ affinity[338]. In vitro and in vivo experiments have shown PLN to exist as a homopentamer that deoligomerizes into active monomers that bind SERCA in a 1:1 molar ratio[281]. The monomeric form of PLN exists in equilibrium between a dynamically disordered R state and a more restricted T state[135, 339] and has 3 structural domains (helix-loop-helix) and 4 dynamic domains: Ia (residues 1-16), loop (residues 17-22), Ib (residues 23-30), and II (residues 31-52)[71]. As measured by NMR[217, 319] and EPR spectroscopies[135, 339, 340, 146], the T state is predominant ($\sim$84%) in both micelles and lipid bilayers[341].

The lack of PLN tertiary structure (noncompact fold) and the semiflexible loop that connects domain Ia to Ib represent major obstacles for structural studies of PLN by solution NMR. In particular, the paucity of NOEs and the degeneracy of residual dipolar coupling solutions prevented the complete characterization of the structural topology[278, 155]. Here, we show how solution NMR angular and distance restraints derived from detergent micellar studies, together with solid-state NMR orientational restraints obtained in mechanically aligned lipid bilayers, are used to simultaneously define the high-resolution structure and topology of monomeric PLN. The various restraints are combined into a single energy objective function and used jointly with a membrane immersion depth potential[311]. The average structure from the calculated conformational ensemble is then embedded into an explicit 1,2-dioleoyl-glycero-3-sn-phosphocholine (DOPC) lipid bilayer and equilibrated, revealing the high-resolution T state structure of monomeric PLN (L-shaped) and the protein-lipid interactions at the atomic level. The resulting fold of monomeric PLN is defined by its strong interactions with lipid membranes. Unlike the previously determined structural ensemble[278], this ensemble more precisely specifies the position of each structural domain with respect to the lipid membrane, with domain Ib protruding toward the membrane/water interface and ready to make hydrogen bonds and salt bridges with the juxtamembrane region

of SERCA. Domain II matches the hydrophobicity of the DOPC bilayer, ready to interact with transmembrane helices M2, M5, and M9 of SERCA's binding groove. The amphipathic domain Ia is helical and adsorbed on the surface of the membrane bilayer. Taken with the recent findings from several other groups[342, 343], we propose that the reversible binding of this domain with the lipid bilayer represents a further regulatory element of PLN function.

## 6.2 Materials and methods

### 6.2.1 Sample preparation

[U-$^{15}N$] AFA-PLN was expressed in Escherichia coli bacteria and purified as described[174]. Most of the selectively labeled samples were also expressed in E. coli (Table 6.4). PLN samples that showed isotopic scrambling, [$^{15}N$-Ser] and [$^{15}N$-Gln], were synthesized by using Fmoc solid-phase peptide synthesis as reported[344, 283]. For solution NMR experiments, AFA-PLN was solubilized in 300 mM DPC, 120 mM NaCl, and 20 mM NaH2PO4 (pH 6.0). Solid-state NMR samples were prepared in either 4/1 DOPC/1,2-dioleoyl-glycero-3-sn-phosphoethanolamine (DOPE) or DOPC lipid bilayers as described (no difference in spectra was discernable with or without DOPE)[26]. Samples were hydrated for 4-5 days at 40 °C in a saturated solution of Na2HPO4 (the [U-$^{15}N$] AFA-PLN in D2O).

### 6.2.2 NMR spectroscopy

Solution NMR experiments were collected in 300 mM DPC (pH 6) and at 37 °C. Solid-state NMR experiments PISEMA[307] and SAMPI4[345] were acquired at 30 °C. Spectra of [U-$^{15}N$] samples were acquired with 1 k scans and 30 $t_1$ increments, whereas selectively labeled samples required 4 to 12 k scans and ∼8-16 increments. Experiments were performed at $^{1}H$ frequencies of 600 and 700 MHz using Bruker (DMX) and Varian (VNMRS) spectrometers, respectively, equipped with low electric field flat-coil probes[346]. Additional details are given in the SI Text.

### 6.2.3 Structural calculations

The structure calculations were carried out by using XPLOR-NIH software[39]. We defined a hybrid solution and solid-state NMR target function ($E_T$), which is formulated as a combination of geometrical ($E_{chem}$), solution NMR ($E_{solNMR}$), and solid-state NMR ($E_{ssNMR}$) terms:

$$E_T = E_{chem} + E_{solNMR} + E_{ssNMR}$$

The potential energy functions (E) are expressed in the form $w_E$, where w represents the relative weight for each energy term, which has been optimized as reported by Shi et al[292]. All of the restraints are approximated by harmonic functions. Specifically, both CSA and DC potentials were implemented as flat-well potential functions according to Bertram and coworkers[347]. The experimental errors for CSA and DC were set to 5 ppm and 0.5 kHz, respectively. A total of 500 monomers were generated. One hundred conformers displayed no violations and were selected for statistical analysis (Table 6.1). Further details regarding the calculations and molecular dynamics simulations in DOPC are described in the SI Text.

## 6.3 Results

### 6.3.1 Solution and solid-state NMR restraints

A previous solution NMR structure of AFA-PLN (C36A, C36F, C41A; fully functional monomeric mutant) was determined using protein solubilized in 600 mM DPC, pH 4.2, 50 °C[278]. Here, we repeated these experiments on AFA-PLN solubilized in 300 mM DPC, pH 6.0, 37 °C, which are conditions that preserve SERCA's enzymatic activity. The NOE patterns at 37 °C and 50 °C were essentially the same. We also performed H(CCO)NH and C(CO)NH TOCSY experiments to fully assign the side chains, which enabled interresidue NOEs to Pro-21 to be identified and incorporated into the structural calculations. In addition, we measured $^3J$-coupling constants to restrain the $\chi_1$ rotamers (Table 6.3).

For solid-state NMR spectroscopy, the AFA-PLN samples were reconstituted into DOPC lipids at a lipid/protein ratio of ∼160/1. The solid-state NMR restraints and

residue assignments for the $^{15}N$ chemical-shift anisotropy (CSA) and $^{1}H$-$^{15}N$ dipolar coupling (DC) of domain Ib and II were obtained from PISEMA experiments conducted at 30 °C[307]. For smaller $^{1}H$-$^{15}N$ DC values (i.e., for residues located in the PLN cytoplasmic domain), we carried out SAMPI4 experiments[345] with selectively labeled PLN protein. A higher level of hydration and temperature optimization led to a substantial increase in resolution and sensitivity with respect to our previous sample preparations, without jeopardizing the macroscopic alignment of the lipid membranes (Figure 6.5). A uniformly $^{15}N$-labeled [U-$^{15}N$] AFA-PLN sample now gives a complete PISA-wheel spectrum for the helix comprised of domains Ib and II (Figure 6.1 C). To complete the resonance assignments, we also collected PISEMA spectra on several selectively labeled PLN samples. Figure 6.1 shows PISEMA and SAMPI4 spectra for [$^{15}N$-Leu, Ile, Ala, Met, Phe, Gln-26/Gln-29, Ser, Arg ($^{15}N_{\alpha}$), and Asn] AFA-PLN. The orientational restraint dataset consists of 44 assigned residues (82 total restraints), which correspond to ∼85% assignment of the backbone.

Evidence for the dynamic N terminus and loop region is seen in the [$^{15}N$-Met] AFA-PLN PISEMA spectrum that shows 2 resonances with $^{15}N$ CSA values of 123 and 101 ppm, corresponding to Met-1 and Met-20 (Figure 6.6). These $^{15}N$ chemical shifts appear at essentially isotropic values, indicating that the observed anisotropic chemical shifts are dynamically averaged for both the loop and N terminus. Based on these results and the dynamic characterization of PLN by solution NMR[217], the DC and CSA values for residues located in the most dynamic regions of PLN (Glu-2, Lys-3, Glu-19, and Pro-21) were not determined.

The solid-state NMR assignments (Table 6.2) were carried out by using extensive selective labeling and a combinatorial algorithm for minimizing the resonance positions based on the periodic nature of the helices[27]. Figure 6.2 shows the CSA and DC plotted vs. the residue number in AFA-PLN. As expected for helical proteins, both CSA and DC oscillate with a period of ∼3.6 (Figure 6.8). The amplitude of the oscillation is larger for domains Ib and II than for domain Ia, reflecting the different orientations with respect to the external magnetic field. These oscillations can be interpreted in terms of tilt ($\theta$) and rotation ($\rho$) angles with respect to the membrane bilayer normal

(parallel to B0) for the different protein domains. An estimate of alignment using a static ideal helix to fit the PISA-wheel patterns[155, 29, 28] gives $\theta$ and $\rho$ angles of $\sim$24° and $\sim$198°, respectively, for the helix comprised of domains Ib and II and a $\theta$ angle of $\sim$97° for domain Ia (Figure 6.9). These angles are in close agreement with those published using CSA and DC for selectively labeled [$^{15}N$-Leu, Ala, and Ile] AFA-PLN alone[26], and reinforce the previously proposed L-shaped topology for monomeric PLN[278, 348].

### 6.3.2 Structure calculations

The technical details of the structure determination protocols and the validation methods for the hybrid approach are reported in ref. 27. In the first stage of the structure calculation, we started with PLN in an extended configuration, and carried out simulated annealing using only the solution NMR restraints [NOEs, $^{3}J$-coupling, dihedral angles from TALOS[14] and hydrogen bonds]. Because there are still very few restraints in the loop region (residues 17-22), this step generated 200 PLN conformers that converged into well defined secondary structure elements but lacked precision in the 3D structure (Figure 6.10B). In the second stage of the refinement, we introduced the solid-state NMR orientational restraints (total of 82) using torsion angle dynamics and simulated annealing calculations. This protocol allowed for the secondary structural elements to be maintained during the overall orientation of the protein according to the alignment tensor (Bo, the membrane normal is fixed along the z axis), avoiding possible integration problems of the simulated annealing algorithm. The third stage consisted of rigid-body minimization of the conformational ensemble into a virtual membrane potential (i.e., the depth of insertion potential, $E_z$) according to the energy function described by DeGrado and coworkers[311]. A final step of Cartesian molecular dynamics refinement was used to relax the PLN conformers.

### 6.3.3 Convergence and validation of the hybrid conformational ensemble

The conformers were selected according to the following criteria: NOE violations $< 0.5$Å, covalent bond violations $< 0.01$Å, and bond angle violations $< 5°$[39]. Figure

6.2 shows the agreement between calculated and experimental CSA and DC, which is excellent given the large number of structural restraints (solution and solid-state NMR restraints). Some small deviations (∼5 ppm) that were observed could be due to imperfect CSA tensors for each amino acid residue as shown by Wang et al[29]; however, these deviations do not substantially affect the structure or domain orientations (see ref. 27). A summary of the parameters analyzed to assess the quality of the structures is reported in Table 6.1.

### 6.3.4 Comparison of the hybrid conformational ensemble with the conformational ensemble obtained from solution NMR data alone

Because our hybrid ensemble has an angular dependence, it is not appropriate to perform a simple overlay of the backbone atoms as the criterion to calculate the precision of the ensemble. Rather, the 2 ensembles were compared on the basis of 3 angles (Figure 6.3A): $\chi$ (the angle between the helix comprising domains Ib and II and the helix comprising domain Ia), $\rho_{Ia}$ (the rotation angle for domain Ia), and $\rho_{Ib,II}$ (rotation angle for domains Ib and II). Figure 6.3C shows a plot of $\rho_{Ia}$ vs. the $\chi$ angle for the solution NMR ensemble alone and the final ensemble generated with the introduction of the orientational restraints of solid-state NMR. It is apparent that the solution NMR ensemble has a broad distribution of angles. Among the different conformers obtained in detergent micelles, we originally chose the L-shaped conformations on the basis of paramagnetic quenching experiments[278]. Using the hybrid method, we do not have to select the structures manually; rather the orientational restraints derived from CSA and DC reduce the conformational space allowed, resulting in a structural ensemble with precise structures (Figure 6.10A) and topologies (Figure 6.3) with respect to the lipid bilayer. It is important to note that the deviations of the angles (the errors reported for $\rho$ and $\theta$ in Table 6.1) reflect the quality of the fit and not the mosaic spread of the sample.

### 6.3.5 Equilibration of DOPC lipids around the average hybrid structure

To fully describe the atomic-level interactions of PLN with lipids, we placed the average hybrid structure into an equilibrated DOPC lipid bilayer using CHARMM version 33a2[34]. We then performed 30 ns of molecular dynamics simulation using NAMD version 2.6[349] to allow the lipids to equilibrate around the PLN structure, which was harmonically restrained with a force constant of 20 $kcal \cdot mol^{-1}$Å$^{-2}$ throughout the simulation. This ensured that the experimental restraints would be satisfied, and provided a detailed picture of the specific interactions the AFA-PLN monomer makes with DOPC lipids (Figure 6.4B). The setup of the molecular dynamics calculations is reported in SI Text.

## 6.4 Discussion

The complexity and challenges presented by membrane protein structure determination call for interdisciplinary approaches[320]. X-ray crystallography and solution NMR have been widely used to determine high-resolution structures for several membrane proteins[7, 320, 7, 350, 351, 288]. These techniques are able to give atomic resolution information about the backbone and side chains but fall short in the elucidation of membrane protein topology. In contrast, solid-state NMR using oriented membrane protein samples can give molecular details regarding both backbone structure and topology. Because a number of membrane protein structures recently determined show similar folds in micelles, lipid bilayers, and crystals[350, 352, 353, 293, 354, 260], we propose to combine the restraints from these techniques into a unique protocol with the goal of obtaining the high-resolution structure (backbone and side chains) and topology within the lipid bilayer. In the literature, there are several examples of backbone structure determination of membrane protein using solid-state NMR data alone[147, 316] and a few examples reported for the combined use of solution and solid-state NMR information in a qualitative fashion[293, 355, 356]. However, there is no precedent for the combination of these restraints into a total energy function, which includes distance, torsion angle, and orientational restraints with a semiempirical depth of insertion potential.

We applied this hybrid approach to the monomeric active form of PLN, which, despite its size, is a challenging system. In fact, PLN does not possess a compact fold, but instead, has its 3D architecture dictated by interactions with the lipid bilayer. Several in vivo and in vitro studies have shown that PLN adopts a remarkably similar structure both in micelles and lipid bilayers[339, 340, 278, 342, 343, 26, 357, 358], justifying the combination of restraints. In the event that restraints obtained in lipid bilayers and detergent micelles were incompatible, it is recommended that those from lipids be weighted more heavily, due to potential limitations of micelles.

Our structural ensemble [Protein Data Bank (PDB) ID code 2KB7], generated from simulated annealing procedures, shows PLN to exist in an L-shaped conformation (T state), while giving a qualitative view of its immersion depth within the lipid bilayer (Figure 6.4). In close agreement with the static fit carried out on the initial solid-state NMR data[26, 284], the structural ensemble of PLN shows that the membrane embedded helix (domains Ib and II) crosses the membrane bilayer with a tilt angle of $24 \pm 2^\circ$ and an azimuthal (rotation) angle of $204 \pm 4^\circ$. This indicates that a specific face of the helix, i.e., residues Arg-25, Gln-29, Phe-32, Ala-36, Leu-43, Ile-47, and Met-50, is on average oriented toward the SR lumen. Domain Ia lies on the surface of the lipid bilayer (Figure 6.4), adopting an angle of $102 \pm 2^\circ$ with respect to the membrane plane. Although the PISA wheel pattern for a helix at 102° is exactly the same as that for 78°, the CSA and DC plotted vs. residue are able to distinguish these angles within the data. The rotation angle of this domain is challenging to resolve using PISEMA or SAMPI4 experiments, because of the small dispersion of CSA and DC values for helices positioned at ∼90° with respect to the bilayer normal. However, the inclusion of the $E_z$ potential[311] helped overcome these ambiguities by identifying 2 energetically different ensembles. One ensemble had higher energy, with the hydrophobic face of the helix pointed toward the bulk solvent, and a second more energetically favorable ensemble, which had the hydrophobic face toward the hydrocarbon region. The latter agrees with several topological studies that we have carried out by using various hydrophilic and hydrophobic paramagnetic probes[278], and the direct observation of NOEs between the detergent and the methyl groups of Val-4 and Leu-7 in domain Ia[261].

For the final stage of refinement, we chose one of the structures from the conformational ensemble (model 6 from PDB ID code 2KB7) and embedded it into an explicit DOPC lipid bilayer for molecular dynamics calculations. Figure 6.4 shows a snapshot from the simulation, which allows for appreciation of the atomic interactions of PLN within the lipid bilayer. Specifically, the hydrophobic side chains of residues Val-4, Leu-7, Ala-11, Ile-12, and Ala-15 are in contact with the hydrophobic interior (methylene region) of the membrane. Glu-2, Lys-3, and Tyr-6 are positioned in the interfacial region of the bilayer, and make hydrogen bonds with the DOPC lipid headgroup. Of the 3 arginine residues within domain Ia, Arg-9 and Arg-13 are solvent-exposed and accessible for interaction with PLN binding partners, whereas Arg-14 snorkels into the lipid bilayer headgroup region. Preliminary studies also indicate that the $N_\epsilon$ of Arg-14 has the highest-order parameter of all arginines within AFA-PLN, which is consistent with increased solvent protection. This is noteworthy, because mutations at Arg-9 and Arg-14 have been linked to the abnormal regulation of SERCA, resulting in dilated cardiomyopathy[359, 360]. Ser-16 and Thr-17 (the 2 phosphorylatable sites) are exposed to the bulk solvent, prone to interact with their respective kinases (i.e., protein kinase A and CamKII kinase, respectively). The amphipathic interactions of domain Ia with the lipid bilayer are also present in the oligomeric state of PLN, as measured by solution NMR in detergent micelles[341, 359], solid-state NMR in lipid bilayers[146, 361], EPR saturation transfer experiments in lipid bilayers and detergent micelles[339, 340, 146], and fluorescence resonance energy transfer in intact cells[357] and detergent reconstituted systems[362].

The topology of PLN determined with this approach also reveals that part of domain Ib extends outside the lipid bilayer. This reconciles the structural model of PLN with several datasets derived from solution NMR[319, 278, 132] and saturation transfer EPR measurements[339]. Specifically, Gln-23 and Ala-24 are solvent exposed[278], whereas Arg-25, Gln-26 and Asn-27 are located within the choline/glycerol interface of the lipid headgroup region, in agreement with paramagnetic quenching and hydrogen/deuterium exchange experiments[278]. Leu-28, Gln-29, and Asn-30 reside in the glycerol/carbonyl region, making hydrogen bonds with the carbonyl groups. Because this domain forms key inhibitory interactions with SERCA via hydrogen bonds and salt bridges[363], the

depth of insertion may be an important mechanism regulating SERCA and PLN function.

Domain II is the most hydrophobic and rigid part of PLN. Hydrogen/deuterium exchange experiments show that the amide protons of this region exchange with the solvent on the timescale of days[278]. This domain crosses the membrane bilayer with an angle of 24°, with the proposed SERCA binding face pointing away from the cytoplasmic domain Ia and prone to the interaction with the enzyme (Figure 6.4D). Remarkably, domain II does not cross the 2 membrane leaflets completely; rather the C-terminal leucine (Leu-52) stops short in the luminal leaflet with its side chain in the methylene region of the bilayer and the carboxyl terminus snorkeling into the carbonyl region of the DOPC headgroups in contact with transient water molecules. Unlike the N-terminal domain Ia, the interactions of the C-terminal carboxyl group constitute the only anchor of domain II in the luminal leaflet of the membrane, an element that may confer the necessary mobility (piston-like motion) for PLN to be lifted up from the membrane to interact with SERCA[340].

The presence of amphipathic cytoplasmic domains is emerging as a common motif for single-pass regulatory and viral membrane proteins. In fact, the L-shaped membrane architecture of PLN is reminiscent of the structures and topologies of the FXYD proteins (a family of membrane associated proteins that serve as subunits to Na,K-ATPases)[354], the fd coat protein (responsible for viral assembly)[364], the VpU protein from HIV-1[365], and the M2 channel from the influenza A virus[366]. Amphipathic helix motifs are known to be functionally important in a number of capacities, ranging from stabilizing protein structures to sensing changes in the physical properties of the membrane[367, 368].

In our allosteric model of regulation[135], the amphipathic helix of PLN drives the conformational equilibrium toward the T state (the resting state), whereas the transition to the more pliable R-state is necessary for PLN to mold into SERCA's binding groove. The dynamic nature of the R-state was underscored by EPR analysis[135, 339, 340] and probably detected by magic angle spinning solid-state NMR experiments[358]. Based on these data, it is likely that the interaction between lipids and the amphipathic domain

Ia is modulated by the composition of the lipid membrane, and that these interactions may constitute an additional mechanism of SERCA regulation. Because PLN has at least 6 different binding partners (SERCA, protein kinase A, CamKII kinase, protein phosphatase I, A kinase anchoring protein, and the antiapoptotic HS-1 associated protein X-1), the reversible binding to the membrane surface and the interplay between the T and R states may give the necessary flexibility in binding these proteins.

## 6.5 Supporting information

### 6.5.1 Materials and methods

#### 6.5.1.1 Sample preparation

[U-$^{15}N$] AFA-PLN was expressed in Escherichia coli bacteria and purified as described[174]. Most of the selectively labeled samples were also expressed in E. coli. PLN samples that showed isotopic scrambling, [$^{15}N$-Ser] and [$^{15}N$-Gln], were synthesized using Fmoc solid-phase peptide synthesis as reported previously[344, 283]. The composition of each peptide was verified by amino acid analysis and its molecular mass was verified by MALDI-TOF. For solution NMR experiment, AFA-PLN was solubilized in 300 mM DPC, 120 mM NaCl, 20 mM NaH2PO4 (pH 6.0). Solid-state NMR samples were prepared in either 4/1 DOPC/DOPE or DOPC lipid bilayers as previously described (no difference in spectra was discernable with or without DOPE)[26]. Lipids (80 mg) were dissolved in 4 ml chloroform (Avanti Polar Lipids) and added to PLN (∼4 mg) that was first solubilized in 100 $\mu$L trifluoroethanol. Solvents were first evaporated under a stream of N2 (g) and then placed in a vacuum desiccator overnight to ensure complete removal of organic solvents. The lipid/protein film was resuspended in 40 ml of H2O and vortexed. Small unilamellar vesicles were prepared by extruding the lipid mixture through polycarbonate filters of decreasing pore size (200, 100, 50$\mu$) using a bench-top extruder (Northern Lipids Inc. Burnaby, BC Canada). This suspension was concentrated to 2 ml in volume and placed evenly onto ∼25 glass plates (Marienfeld, $5.7mm \times 12mm \times 80\mu$m). Samples were hydrated for 4-5 days at 40 °C in a saturated solution of Na2HPO4. The [U-$^{15}N$] AFA-PLN sample was hydrated in a saturated solution of Na2HPO4 in D2O. Solid-state NMR data were acquired at 30 °C.

#### 6.5.1.2 NMR spectroscopy

Solution NMR resonance assignments, NOEs, and $^3J$-coupling constants were collected under conditions of 300 mM DPC (pH 6) and at 37 °C. For solid-state NMR, 2D PISEMA[313] and SAMPI4[345] experiments were performed with SPINAL64 decoupling during acquisition[369]. A phase modulated Lee-Goldberg (PMLG) was used in the PISEMA indirect dimension to decouple proton-proton interactions while allowing for the evolution of $^1H$-$^{15}N$ DC[370, 371]. The initial 90° $^1H$ pulse, cross-polarization, PMLG ($^1H$ effective field), and SPINAL decoupling during acquisition were applied at ∼60 kHz RF field strength. For SAMPI4 experiments, an RF field strength of 50 kHz was used. Spectra of [U-$^{15}N$] samples were acquired with 1 k scans and 30 $t_1$ increments, while selectively labeled samples required 4-12 k scans and ∼8-16 increments. Experiments were performed using both Bruker and Varian spectrometers. The DMX Bruker spectrometer (National High Magnetic Field Laboratory) operating at a $^1H$ frequency of 600 MHz was equipped with a low-E probe built by the RF program[346]. The Varian VNMRS spectrometers operating at $^1H$ frequencies of 600 and 700 MHz were equipped with a BioStatic probe from Varian, Inc.

$\chi^1$ Rotamers from $^3J$ Couplings. $\chi^1$ rotamers were determined by measuring $^3J_{C'C_\gamma}$ and $^3J_{NC_\gamma}$ with quantitative 2D spin echo difference spectroscopy by Bax and coworkers[372]. For Thr, Val, and Ile residues $^3J_{C'C_\gamma}$ and $^3J_{NC_\gamma}$ were measured using 2D experiments described previously[311, 342, 343]. Aromatic $^3J_{C'C_\gamma}$ couplings for Phe and Tyr were measured using the pulse sequence described by Hu et al[373]. All of the spectra were acquired using a Varian DirectDrive VNMRS spectrometer operating at $^1H$ frequency of 600 MHz.

#### 6.5.1.3 Structural calculations

The structure calculations were carried out using X-PLOR-NIH software[39] as more completely described in the companion paper[292]. We defined a hybrid solution and solid-state NMR target function ($E_T$), which is formulated as a combination of geometrical ($E_{chem}$), solution NMR ($E_{solNMR}$), and solid-state NMR ($E_{ssNMR}$) terms:

$$E_T = E_{chem} + E_{solNMR} + E_{ssNMR}$$

The potential energy functions (E) are expressed in the form $w_E$, where w represents the relative weight for each energy term, which has been optimized as reported by Shi et al[292]. All of the restraints are approximated by harmonic functions. Specifically, both CSA and DC potentials were implemented as flat-well potential functions according to Bertram and coworkers[347]. The experimental errors for CSA and DC were set to 5 ppm and 0.5 kHz, respectively.

The hybrid conformational ensemble was generated starting from AFA-PLN in an extended configuration. In the first stage, we implemented only solution NMR restraints (NOE, torsion angles, and hydrogen bonds). The temperature was decreased from 6000 K to 0 K in steps of 5 K. At each temperature, 20 ps of torsion angle dynamics and several steps of minimization in both torsion angle and Cartesian space were performed. This step folds PLN secondary structure elements. The second structural ensemble was generated by including orientational restraints obtained from PISEMA data. Starting from this folded configuration, solid-state NMR restraints (both CSA and DC) are combined with the solution NMR data into a simulated annealing protocol. To satisfy CSA and DC data, we carried out simulated annealing starting from 3000 K using only torsion angle dynamics. The force constants for the solid-state NMR restraints were optimized using cross-validation[318]. To orient PLN with respect to the z direction (corresponding to a fictitious lipid bilayer normal), we used rigid body minimization under orientational restraints with the $E_z$ potential. The depth-dependent potential term ($E_z$) has been described by DeGrado and coworkers, and is based on the propensity of amino acid residues in a helix to have depth-dependence insertions into the bilayer[374]. In the final XPLOR-NIH refinement protocol, the calculation was performed starting from 300 K using both torsion angle and Cartesian molecular dynamics to relax local geometries and improve the structure quality without significantly changing the orientation of each conformer generated. A total of 500 monomers were generated. 100 conformers displayed no violations and were selected for statistical analysis (Table 6.1). The full descriptions of the target function and calculation protocols are available in ref. 12.

#### 6.5.1.4 Molecular dynamics

The final stage of the structural refinement was carried out using MD simulations in DOPC lipids with the intent of relaxing the PLN structure in an explicit environment that mimics the interactions of PLN within a cell membrane. The information derived from the $E_z$ potential was used to embed conformer #6 from PDB 2KB7 into a preequilibrated DOPC bilayer. To insert the protein in the bilayer, the overlapping lipids around the protein were deleted. The system was then slowly heated from 10 K to 310 K over 300 ps with the protein harmonically restrained with a force constant of 20 $kcal \cdot mol^{-1}$Å$^2$. The restraints were decreased to 5 $kcal \cdot mol^{-1}$Å$^2$ during heating and maintained for 10 ns of equilibration. Following equilibration, the restraints were removed and the system was further simulated for 100 ns.

The set-up and heating were done using CHARMM Version 33a2[34], whereas the equilibration and sampling were performed with NAMD version 2.6. The PARAM27 force field with the CMAP correction[27, 29, 28] was used throughout the simulation. The system consisted of AFA-PLN (+3 charge), 122 DOPC molecules equally distributed between the top and bottom layer and TIP3P waters. Counter ions were added to achieve electroneutrality and an ionic strength near physiology conditions. Hexagonal periodic boundary conditions were used with a unit cell of $68 \times 68 \times 88$Å. The particle-mesh-Ewald (PME) method with a mesh of $\sim$1 Å was used to describe the electrostatic interactions[375]. Van der Waals interactions were evaluated using a cutoff scheme feathered to zero between 9 and 11 Å with a switching function. Constant pressure (1 atm) and temperature (310 K) were maintained during MD simulations through Nosé-Hoover Langevin piston[376, 377] and Langevin dynamics. The RATTLE algorithm was applied to covalent hydrogen bonds and an integration time step of 2 fs was used.

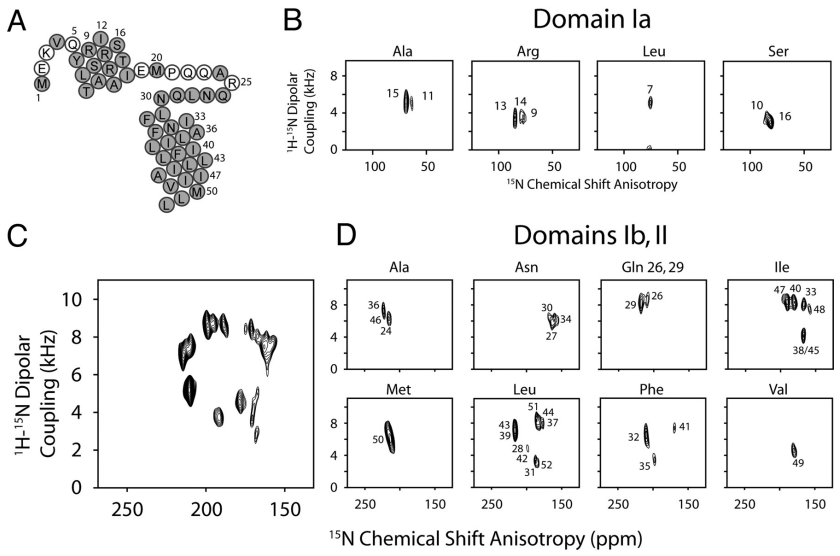


Figure 6.1: Solid-state NMR spectra of AFA-PLN in lipid bilayers. (A) Primary sequence of AFA-PLN showing assigned residues in gray (CSA and DC assignments in Table 6.2). (B) Assigned PISEMA (Leu, Ala) and SAMPI4 spectra (Ser, Arg) of domain Ia. (C) [U-$^{15}N$] PISEMA spectrum of PLN domain Ib and II displaying a uniform intensity across the PISA wheel. (D) Selectively labeled PLN PISEMA spectra of domain Ib and II. Overlay shown in Figure 6.7.

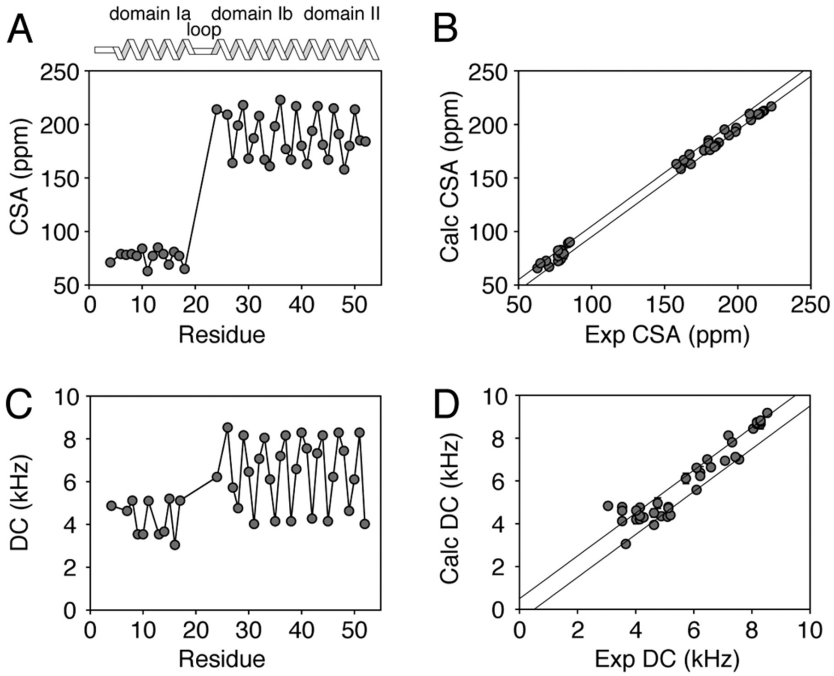


Figure 6.2: Experimental CSA (A) and DC (C) values from Figure 6.1 plotted vs. residue number. (B and D) show correlation plots of calculated vs. experimental CSA and DC for the hybrid ensemble, respectively. The parallel lines in B and D indicate the experimental errors used in the structural calculations.

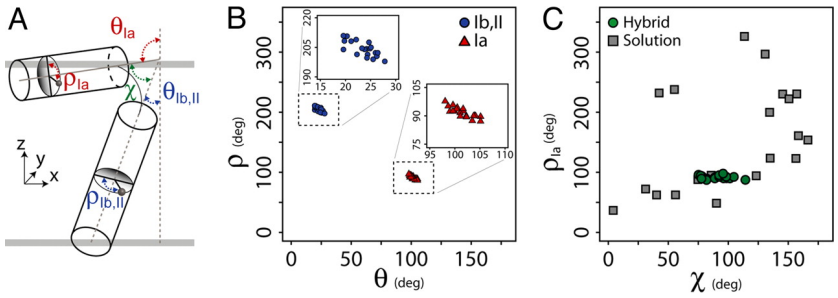


Figure 6.3: Comparison between the solution NMR ensemble and the conformational ensemble generated with the hybrid protocol. (A) Angles used to describe the topology of monomeric AFA-PLN. (B) Rotation ($\rho$) versus tilt angle ($\theta$) for the 20 structures in the deposited ensemble. The Insets better allow for variations to be seen in the angles describing the topology with respect to the membrane normal for domains Ib and II (Left, circles) and domain Ia (Right, triangles). (C) Comparison of the ensemble of structures generated from solution NMR restraints alone (squares) with the hybrid solution and solid-state NMR method (circles).

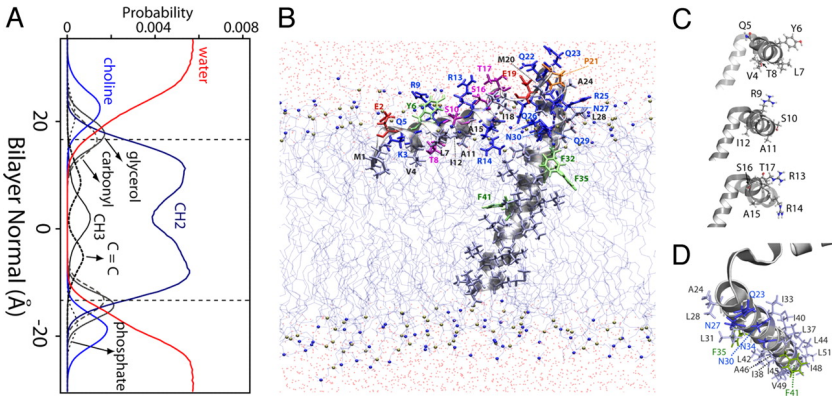


Figure 6.4: Structural topology of monomeric PLN in a DOPC lipid bilayer. (A) Probability distribution profile for the chemical groups of DOPC in the molecular dynamics simulation of PLN within the bilayer. (B) Structure of monomeric AFA-PLN in a DOPC lipid bilayer. (C) Detailed images of the residues within the amphipathic domain Ia helix show that the hydrophobic residues (Val-4, Leu-7, Ala-11, Ile-12, Ala-15) face the interior of the bilayer. (D) Structure highlighting the face of PLN within domains Ib and II that have been shown by mutagenesis and cross-linking data to interact with SERCA (56). Residues in D shown with side chains in sticks and labeled on the structure reside on the same helix of PLN, i.e., they are positioned for binding SERCA. In both B and D, the colors reflect the hydrophobicity: light gray, hydrophobic; blue/red/purple, hydrophilic; green, aromatic (Phe); specifically Ser and Thr residues are shown in purple, Glu is shown in red, and Asn and Gln are shown in blue.

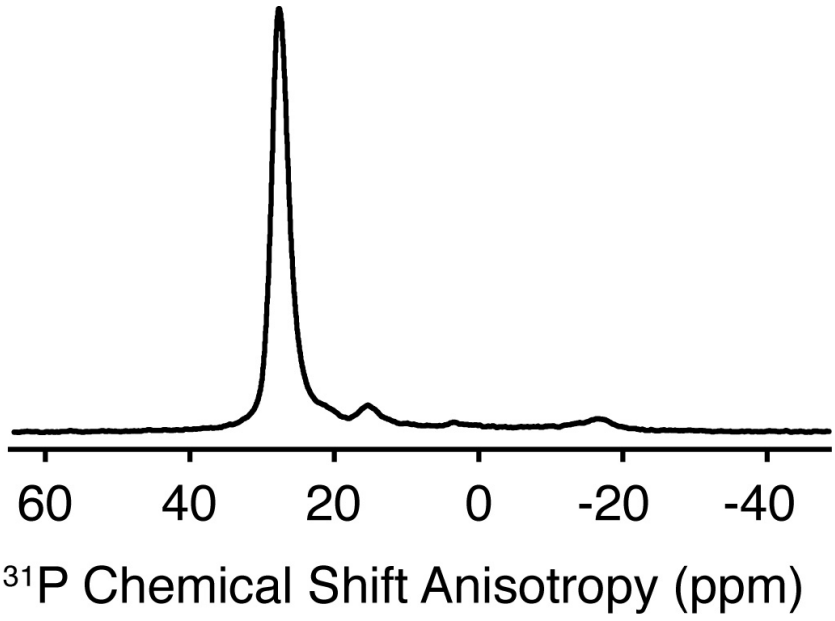


Figure 6.5: $^{31}P$ spectrum of the [U-$^{15}N$] AFA-PLN sample in Figure 6.1c. A single-pulse on $^{31}P$ was acquired with proton decoupling to verify the alignment of the lipid bilayer. This spectrum is representative of the selectively labeled samples shown in Figure 6.1.

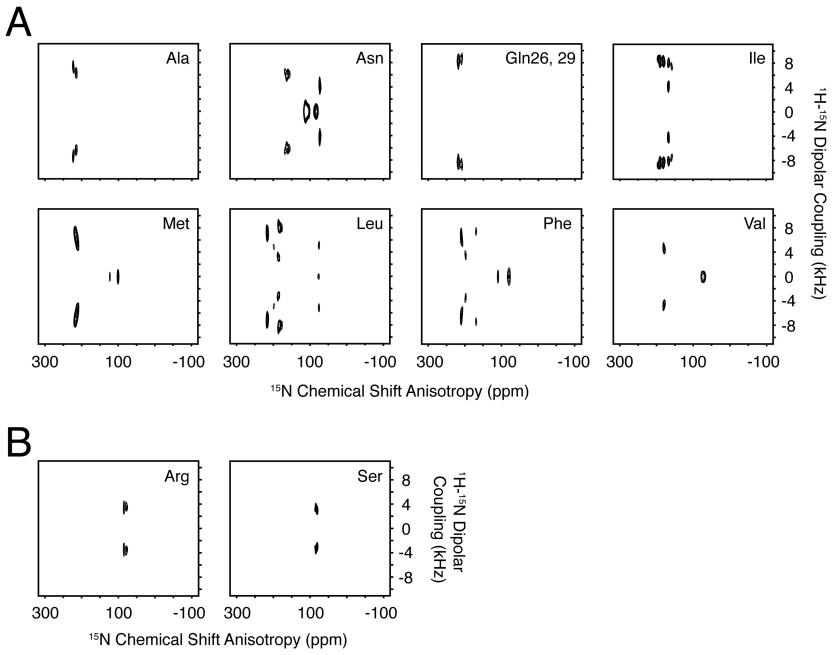


Figure 6.6: Supporting information for Figure 6.1 in the manuscript showing full PISEMA spectra. Note that the Asn and Phe spectra shows resonances upfield of those plotted in Figure 6.1 due to the side chain $^{15}N$ resonances of Asn (the HSQC of the Phe sample showed scrambling to $^{15}N$ Asn side chains). Due to the RF offsets ($^{15}N$ - 200 ppm, $^{1}H$ - 5-6 ppm), some of the cytoplasmic domain peaks in the 2D PISEMA spectra (A) show up with 1) lower intensity than the peaks in domain Ib and II or 2) at zero-frequency indicating that those residues have an amide proton off-resonance from that used in the experiment (5-6 ppm). For several cytoplasmic domain resonances, SAMPI4 experiments (B) were run to obtain these frequencies (CSA and DC) more accurately.

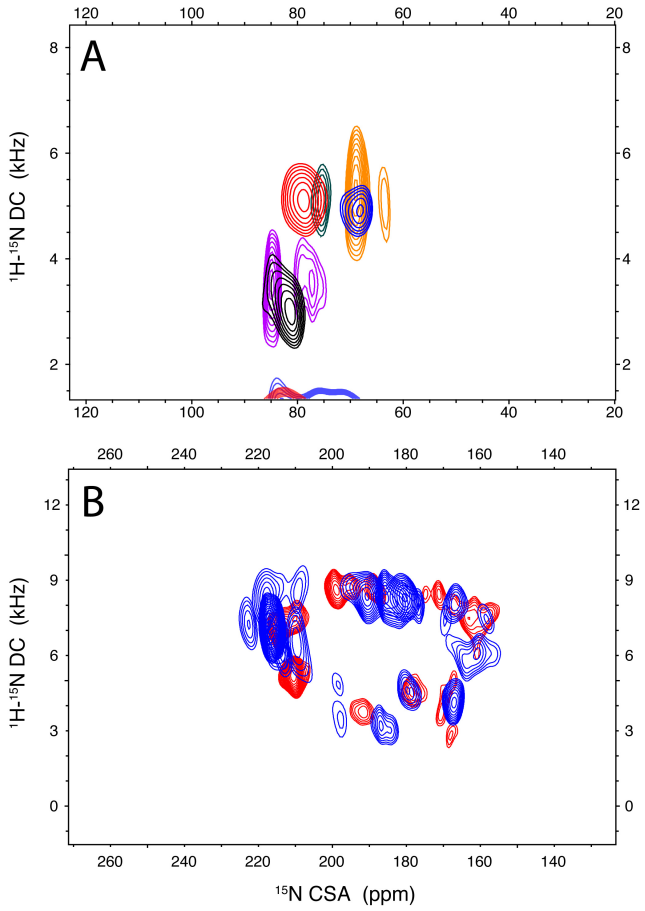


Figure 6.7: Overlay of spectra for domain Ia (A) and domain Ib and II (B). Panel A is color-coded with residue-type: orange-[$^{15}N$-Ala], red-[$^{15}N$-Thr], blue-[$^{15}N$-Val], purple-[$^{15}N_{\alpha}$-Arg], light blue-[$^{15}N$-Leu], and black-[$^{15}N$-Ser]. The Arg and Ser spectra were acquired using the SAMPI4 pulse sequence, while all others were acquired using the PISEMA experiment. The appearance of zero-frequency peaks in the spectra of [$^{15}N$-Val] and [$^{15}N$-Thr] are due to a mismatched $^{1}H$-offset. Panel B is an overlay of all selectively labeled PISEMA spectra shown in Figure 6.1D (blue) with the [U-$^{15}N$] spectrum equilibrated in D2O (blue). The small changes between these spectra arise from peaks missing in the [U-$^{15}N$] spectrum due to D2O exchange, as well as slight differences in CSA and DC in sample-to-sample preparations. Error bars in the CSA and DC restraints reflect the effects.

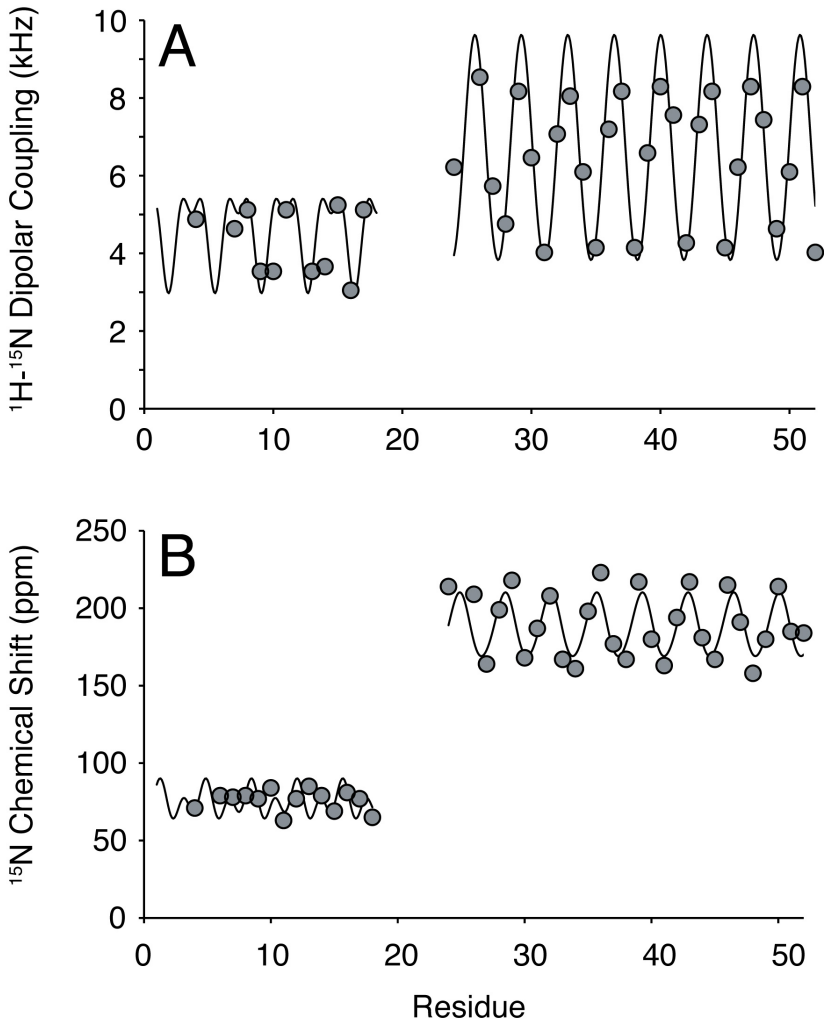


Figure 6.8: Supplement to Figure 6.2 within the text. Data shown is the same as in Figure 6.2A and C, except with ideal waves plotted with the experimental CSA and DC data.

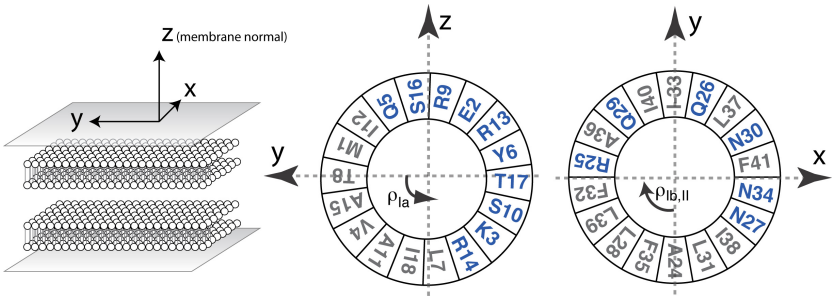


Figure 6.9: Description of the angles used in Figure 6.3. For the domain Ia rotation angle, ?Ia is defined to be zero when the vector described by the amide nitrogen and the center of the helical axis is co-linear with the y axis. Likewise the ?Ib,II is similar, with the exception that the amide nitrogen of Ala24 is used.

Table 6.1: Summary of structural ensemble statistics

| | |
|---|---|
| **R.M.S. Deviation from experimental restraints** | |
| NOE/H-bond (Å)(422) | $0.06 \pm 0.006$ |
| Torsion angle (°)(99) | $0.4 \pm 0.22$ |
| **PISEMA R-factor (%)** | |
| CSA(43/53) | $0.87 \pm 0.02$ |
| DC (39/53) | $0.99 \pm 0.03$ |
| **R.M.S. deviations from idealized covalent geometry** | |
| Bond (Å) | 0.006±0.0004 |
| Angle (°) | 0.70±0.03 |
| Impropers (°) | 0.21±0.002 |
| **Measure of structural quality** | |
| % residues in most favored region | 90±2 |
| % residues in additional allowed region | 8±3 |
| % residues in generously allowed region | 2±2 |
| % residues in disallowed region | 0 |
| **Precision of atomic coordinates (Å)** | |
| Backbone (residue 4-50) | 2.3 |
| Helix 1 (residue 4-18) | 0.6 |
| Helix 2 (residue 24-50) | 0.6 |
| **Protein Topology (domain in subscript)** | |
| Tilt ($\theta_{Ia}$) (°) | 102±2 |
| Rotation ($\rho_{Ia}$) (°) | 92±3 |
| Insertion ($Z_{Ia}$) (Å) | 16.2±0.8 |
| Tilt ($\theta_{Ib,II}$) (°) | 24±2 |
| Rotation ($\rho_{Ib,II}$) (°) | 204±3.8 |
| Insertion ($Z_{Ib,II}$) (Å) | 5.5±0.5 |

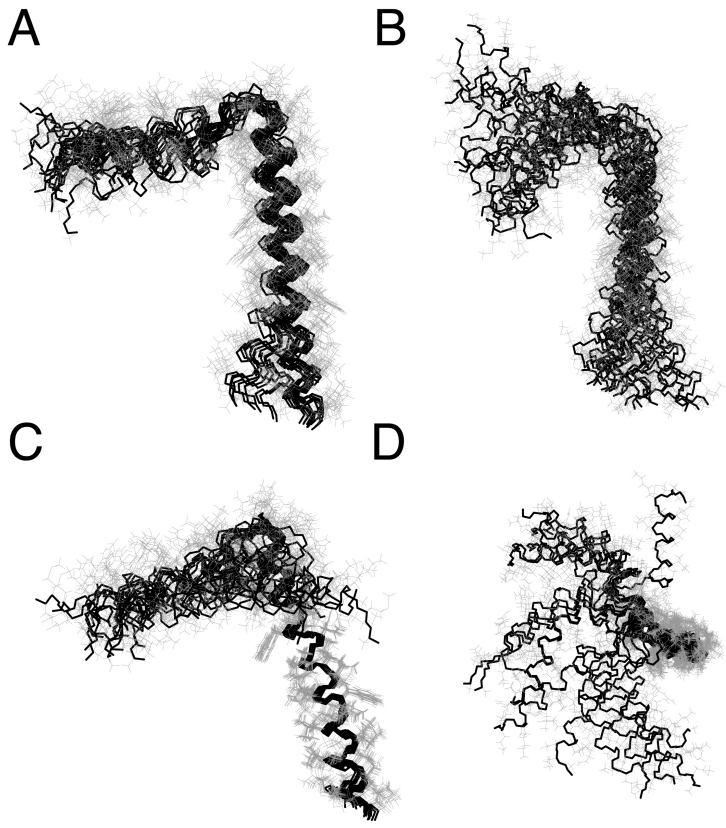


Figure 6.10: Ensemble overlay of residues 4-50 for the hybrid (A) and solution NMR ensembles (B). (C) and (D) show side and top views of the hybrid ensemble where residues 24-50 in domain Ib and II are superimposed.

Table 6.2: Summary of the NMR restraints from solid-state NMR.

| **Residue** | $^{15}N$ **CSA (ppm)** | $^{1}H-^{15}N$ **DC (kHz)** |
|---|---|---|
| 1 | 101/123 | 0 |
| 4 | 71 | 4.9 |
| 6 | 79 | |
| 7 | 78 | 4.6 |
| 8 | 79 | 5.1 |
| 9 | 77 | 5.1 |
| 10 | 84 | 5.1 |
| 11 | 63 | 5.1 |
| 12 | 77 | 5.1 |
| 13 | 85 | 5.1 |
| 14 | 79 | 5.1 |
| 15 | 69 | 5.1 |
| 16 | 81 | 5.1 |
| 17 | 77 | 5.1 |
| 18 | 65 | 5.1 |
| 20 | 101/123 | |
| 24 | 214 | 6.2 |
| 26 | 209 | 8.5 |
| 27 | 164 | 5.7 |
| 28 | 199 | 4.8 |
| 29 | 218 | 8.2 |
| 30 | 168 | 6.5 |
| 31 | 187 | 4.0 |
| 32 | 208 | 7.1 |
| 33 | 167 | 8.0 |
| 34 | 161 | 6.1 |
| 35 | 198 | 4.1 |
| 36 | 223 | 7.2 |
| 37 | 177 | 8.2 |
| 38 | 167 | 4.1 |
| 39 | 217 | 6.6 |
| 40 | 180 | 8.3 |
| 41 | 163 | 7.6 |
| 42 | 194 | 4.3 |
| 43 | 217 | 7.3 |
| 44 | 181 | 8.2 |
| 45 | 167 | 4.1 |
| 46 | 215 | 6.2 |
| 47 | 191 | 8.3 |
| 48 | 158 | 7.4 |
| 49 | 180 | 4.6 |
| 50 | 214 | 6.1 |
| 51 | 185 | 8.3 |
| 52 | 184 | 4.0 |

Note Met1 and Met20 cannot be distinguished in the combinatorial assignment algorithm

Table 6.3: Summary of $\chi^1$ Restraints Used in the Hybrid Calculation

| **Residue** | $\chi^{1\circ}$ |
|---|---|
| 4 | 180 |
| 6 | 180 |
| 7 | -60 |
| 17 | 60 |
| 28 | -60 |
| 31 | -60 |
| 32 | 180 |
| 33 | -60 |
| 35 | 180 |
| 41 | 180 |
| 43 | -60 |
| 44 | -60 |
| 45 | -60 |
| 47 | -60 |
| 48 | -60 |
| 49 | 180 |

Table 6.4: Summary of synthetic or bio-expressed AFA-PLN samples used for solid-state NMR experiments.

| **AFA-PLN Label** | **Synthetic or Bio-expressed PLN** |
|---|---|
| $[^{15}N - Ala]$ | Bio |
| $[^{15}N - Arg]$ | Bio |
| $[^{15}N - Asn]$ | Bio |
| $[^{15}N - Ile]$ | Bio |
| $[^{15}N - Met]$ | Bio |
| $[^{15}N - Leu]$ | Bio |
| $[^{15}N - Phe]$ | Bio |
| $[^{15}N - Val]$ | Bio |
| $[^{15}N - Ser]$ | Synthetic |
| $[^{15}N_\alpha - Gln26, Gln29]$ | Synthetic |

# Chapter 7

# Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method



## 7.1 Introduction

The membrane protein complex formed by $Ca^{2+}$-ATPase (SERCA) and PLN regulates $Ca^{2+}$ concentration within the sarcoplasmic reticulum (SR), thereby controlling muscle excitation-contraction coupling[176, 338]. PLN is a 52-residue transmembrane (TM) protein highly conserved across mammals[338]. Its helix-loop-helix secondary structure is further subdivided into four dynamic domains: domain Ia (residue 1-16), loop (residue 17-22), domain Ib (residue 23-30), and domain II (residue 31-52)[278, 84]. The hydrophobic TM domain II is the most conserved and responsible for SERCA inhibition, while the cytoplasmic domain harbors two phosphorylation sites that reverse PLN inhibitory function[338]. PLN has a direct role in the pathophysiology of the

heart muscle, with three lethal mutations linked to dilated cardiomyopathy in humans (R9C-PLN, R14del, and L39-truncated-PLN)[360]. In both synthetic and cell membranes, PLN forms pentamers that dissociate into monomers upon interacting with SERCA[176, 379]. While the stoichiometry of the SERCA/PLN complex has been assessed[176, 379], both the role and the structure of the PLN pentamer remain a matter of active debate. Since PLN expression in both atria and ventricles is higher than SERCA, it is likely that oligomerization may participate in SERCA regulation[380]. Insights into PLN organization in the membrane have come from biochemical and biophysical data[338, 379, 381]. Initial electrophysiological measurement indicated that PLN formed $Ca^{2+}$channels[382]. However, more recent electrochemical studies concluded that PLN does not conduct $Cl^{-}$ or $Ca^{2+}$ ions[383].

Divergent structural models for the PLN pentamer have been proposed in the literature[381]. While very similar in the secondary structure content, these models differ in the membrane topology. The two most recent models are a solution NMR structure of PLN in detergent micelles (dodecylphosphocholine, DPC) with an unusual bellflower topology[289], and a low-resolution pinwheel model derived from fluorescence anisotropy data[362]. The bellflower model, with its funnel-shaped TM pore, would support an active role of the pentamer in transporting $Ca^{2+}$ or $Cl^{-}$ ions, contributing to the ion homeostasis in the SR. In contrast, the absence of a pore in the pinwheel topology would support the role of the pentamer assembly as a reservoir for active monomers and would constitute a supplementary regulatory mechanism for SERCA via monomer/pentamer equilibrium[384].

In this Article, we present the high-resolution structure and topology of the PLN pentamer in its T-state[135] as determined by a hybrid solution and solid-state NMR method[292, 84]. In its T-state, PLN adopts a pinwheel topology with the cytoplasmic helices making extensive contacts with the lipid bilayer surface. The TM helices traverse the membrane at an angle of $\sim 11^{\circ}$ with respect to the bilayer normal and have a narrow hydrophobic pore (average radius $\sim$2Å), which tapers to $\sim$1Å at the juxtamembrane region (residues 23-30).

## 7.2 Materials and methods

PLN was expressed in E. coli bacteria and purified as previously described[84]. For solution NMR experiments, PLN was solubilized in 300 mM deuterated DPC, 120 mM NaCl, 20 mM NaH2PO4 (pH ∼6), and 50 mM 2-mercaptoethanol. For solid-state NMR spectroscopy, PLN was reconstituted in 4/1 DOPC/DOPE lipid bilayers and spread onto glass plates as previously described[84] or packed into a 3.2 mm MAS rotor. NMR experiments were performed at $^{1}H$ frequencies of 600 and 700 MHz using Bruker (DMX) and Varian (VNMRS) spectrometers, respectively, equipped with low electric field flat-coil probes[346]. Structure calculations were carried out with XPLOR-NIH (v.2.23) software, using the hybrid solution and solid-state NMR method[292, 84]. For details on MD simulations see Supporting Information.

## 7.3 Results

### 7.3.1 Structure and topology of PLN in detergent micelles and lipid bilayers

The overall strategy for the hybrid NMR method consists of combining secondary structure information from solution NMR with orientational restraints obtained from solid-state NMR in lipid membranes[292, 84]. Here, we selected micelle and lipid conditions, where PLN predominantly adopts a T-state conformation as probed by EPR and NMR[26, 284, 340]. For solution NMR studies, we refolded bacterially expressed PLN (rabbit primary sequence) into DPC micelles. In DPC, PLN forms stable pentamers that have a melting temperature of ∼70 °C[385], giving a homogenous and well-dispersed 2D [$^{1}H$,$^{15}N$]-TROSY (transverse relaxation optimized spectroscopy)[10] spectrum (Figure 7.5A). To assign the PLN backbone and side chain resonances, we utilized standard triple-resonance experiments[386, 387]. The chemical shift assignments are in complete agreement with those reported for the human PLN pentamer[289], with the exception of Asn27, which is Lys27 in the rabbit sequence. Spatial correlations and other long-range distance restraints between nuclear spins were obtained from the quantification of a series of 3D [$^{1}H$, $^{1}H$, $^{15}N$]-NOESY-HSQC experiments.

We detected a total of 1945 intra-protomer NOEs (389 NOEs per protomer) (Figure 7.5B). The chemical shift indices and the NOE patterns confirmed the existence of the helix-loop-helix structural motifs (Figure 7.5B-C) also present in monomeric PLN[261]. Although micelle studies give useful insight into the topology of PLN, they may not reproduce the architecture of this protein in cell membranes[388]. A more rigorous approach is to supplement micelle studies with solid-state NMR techniques in lipid bilayers[137, 19, 139, 138, 389]. Therefore, we carried out solid-state NMR experiments in mechanically aligned lipid bilayers[84]. PLN was reconstituted into 1,2-dioleoyl-sn-glycero3-phosphocholine (DOPC)/1,2-dioleoyl-sn-glycero3-phosphoethanolamine (DOPE) lipid bilayers at 4/1 (mol/mol) ratio, which maximizes the activity of SERCA[390] and closely mimics the SR membrane composition[391]. To avoid non-specific PLN aggregation, we used a lipid-to-protein molar ratio of 200/1. Orientational restraints were derived from PISEMA[343] and SAMPI4[345] experiments, measuring dipolar couplings (DC) and anisotropic chemical shifts (CSA) on a series of selectively $^{15}N$ labeled samples (see Supporting Information). Figure 7.1E shows the overlay of several oriented solid-state NMR spectra located in two distinct regions: 1) resonances at 200 ppm that correspond to the TM domains and 2) resonances at 65-100 ppm, assigned to the cytoplasmic domain Ia. We measured a total of 29 DC and 31 CSA values (Table 7.2) and converted them into restraints for the hybrid structure calculations, where both solution and solid-state NMR restraints are implemented into the target energy function[292, 84].

### 7.3.2 Interprotomer distance restraints

Interprotomer distance restraints were derived from NOESY experiments using an asymmetric isotopic labeling scheme that allowed for the unambiguous assignment of interprotomer methyl-methyl NOEs, defining the quaternary structure[261]. We prepared two different mixed pentamer samples. The first was a 1/1 molar mixture of Ile-$^{13}C^{\delta_1}H_3$/Leu-$^{13}C^{\delta 1/2}H_3$, Val-$^{13}C^{\gamma 1/2}H_3$ in a U-[$^2H$, $^{12}C$] background, where interprotomer NOEs were measured between methyl protons[261]. The second sample consisted of a 1/1 mixture of U-$^{15}N$/Ile-$^{13}C^{\delta 1}H_3$ in a U-[$^2H$, $^{12}C$] background. This sample gave quaternary restraints between Ile methyl protons and backbone amide protons. We used a total of 75 interprotomer NOEs (15 NOEs per protomer) in the structure calculation

(Figure 7.1 and Table 7.3). Although several other NOEs are present in the spectral region corresponding to the Leu residues in the Ile/Leu zipper (e.g. Leu37, Leu44, Leu51), these peaks are highly overlapped and we did not use them in the calculations. To validate the interprotomer distance restraints obtained in micelles, we carried out dipolar assisted rotational resonance recoupling (DARR) magic-angle-spinning (MAS) experiments in 4/1 DOPC/DOPE lipid bilayers (see SI methods). These experiments (Figure 7.1B and Figure 7.9) utilized a Leu/Ile asymmetric labeling method and confirmed the dipolar connectivities measured by NOESY experiments (Figure 7.9C). The DARR cross-peaks are also in good agreement with the Ile/Leu zipper present in the pentamer. To define the quaternary arrangement in the juxtamembrane domain Ib, we carried out two sets of calculations with and without the following two restraints: 1) an MAS REDOR contact between the nitrogen side chain of Gln29 of PLN found by Smith and co-workers[392] and 2) an interprotomer NOE between Ile33$H^{\delta 1}$ and Gln29$H^{\epsilon}$ we detected in micelles (Figure 7.1B). The two structural ensembles calculated without the REDOR and NOE restraints are almost identical with an RMSD between the TM domains $< 0.8$Å (Figure 7.11), demonstrating that these contacts were not crucial to define the overall fold of the pentamer. In addition, the agreement between the distances obtained by solution and solid-state NMR is further evidence that the T-state structures in micelles and bilayers are very similar, and supports using the restraints for structure calculations. Finally, an EPR derived restraint between spin labels at Lys3 in each monomer[84] was used in order to better define the arrangement of the cytoplasmic helices.

### 7.3.3 Positioning of PLN in micelle and lipid bilayers

To position the two helical segments (1-16 and 23-52), we used planar restraints (see SI methods) derived by measuring NOEs between the amide backbone of PLN and the detergent hydrocarbon tails in a 3D [$^1H$, $^1H$, $^{15}N$]-NOESY spectrum (Figure 7.1A and Figure 7.8). Since PLN was fully deuterated (except the back-exchanged amide protons), the presence of cross-peaks at 1-2 ppm in the $^1H$ dimension indicated a close proximity between the amide-hydrogen and the methylene and methyl protons of the DPC micelle. These contacts between PLN and the micelle were observed throughout

the pentamer with the exception of residues 21 to 23 within the dynamic and solvent-exposed loop (Figure 7.1A and Figure 7.8). Residues in the cytoplasmic domains Ia and Ib displayed NOE contacts with the micelle hydrocarbon tails, indicating a strong association with the DPC micelle. These results confirmed unambiguously that the amphipathic domain Ia is absorbed on the micelle surface, with the most polar residues facing the bulk water and the hydrophobic residues pointing toward the inner core of the micelle, mirroring previous paramagnetic mapping studies[84].

### 7.3.4 Hybrid PLN structural ensemble

The calculation protocol is illustrated in Figure 7.6. Symmetry distance restraints[393] were used to maintain the pentamer geometry during the simulated annealing calculations. The relative position of the two domains (depth-of-insertion) was constrained with planar restraints derived from protein/micelle NOEs[145] (See SI methods). Figure 7.2A-B displays the final conformational ensemble of PLN, with a well-defined pinwheel topology. The low energy conformers show the convergence of the TM domains with an RMSD of ∼0.5 Å (Table 7.1). We assessed the quality of our structural ensemble with the MolProbity software[394]. Our 2KYV ensemble obtained a total score of 1.8 (94 percentile), showing an excellent van der Waals packing and local covalent geometry (Table 7.1). The PLN protomers are held together by a tight Leu/Ile zipper, which is the driving force for PLN oligomerization that confers high thermostability[338]. The interlocking of alternating Leu and Ile residues gives rise to a left-handed coiled-coil assembly (Figure 7.2C)[381]. The interior of the TM domain is lined with residues that are sensitive to mutations (Figure 7.2E). Single mutations of these residues lead to PLN de-oligomerization and enhanced inhibitory function (Figure 7.2E)[176]. Other crucial residues for PLN oligomerization are the three cysteines (Figure 7.2D and Figure 7.7), which participate in quaternary contacts[381, 395]. Indeed, we detected long-range interprotomer NOEs between the amide hydrogen of Cys41 and the methyl protons of Ile40 (Figure 7.1). We also observed short-range intraprotomer NOEs between the sulfhydryl hydrogen of Cys36 and the amide hydrogens of Phe35 and Leu37, as well as the sulfhydryl hydrogen of Cys41 and the amide hydrogen of Ile40 (Figure 7.7B). These experimental observables confirmed that Cys36 and Cys41 are positioned at the interface between the protomers, which is consistent with their chemical reactivity[395].

The intraprotomer NOEs resulting in hydrogen bonding distances between the Cys backbone amides and the (i - 4) carbonyl oxygens (Figure 7.7B) are in agreement with mutagenesis experiments and computational modeling[381, 395]. Overall, the hybrid ensemble agrees with a wealth of structural data from different techniques and various membrane mimicking systems. The only inconsistency is relative to the dynamic loop region (Table 7.4).

### 7.3.5 Comparison between structural models

The hybrid PLN ensemble differs significantly from the model proposed by Chou and Oxenoid[289]. In the bellflower model, the five domain-Ia helices are oriented almost perpendicularly to the bilayer normal, which probably resulted from the selection of one of the ghost orientations derived from the analytical solutions of the orientational dependent residual dipolar couplings (RDCs)[235]. In the hybrid ensemble, domains Ia are embedded within the top leaflet of the lipid bilayer (Figure 7.3A and Figure 7.4C), with the juxtamembrane domain Ib more tightly packed, resulting in straight $\alpha$-helices crossing the membrane at an angle of $\sim 11^{\circ}$. Conversely, the bellflower model (Molprobity score 3.5, 9 percentile) presents a peculiar bending in the TM helices, diverging from one another at the juxtamembrane region (Figure 7.3B). The distorted $\alpha$-helical domain Ib obtained in the bellflower arrangement contrasts with the hydrogen bonds found in this region[392] and is probably due to dynamics[84] in this domain that averages the RDC values[235]. Also, the torsion angle predictions from TALOS[14] agree remarkably well with our conformational ensemble, while deviating substantially in the bellflower ensemble (Figure 7.10). The most striking difference for the two models can be appreciated in the back-calculation of the PISEMA spectra (Figure 7.12). While the structural ensemble for the bellflower model differs considerably from the spectra determined in lipid membranes, the pinwheel structure agrees well with the experimental data. The CSA values and the scaling of the DC (Figure 7.1E-F) in domain Ia are strong evidence that this domain has a perpendicular orientation with respect to the bilayer normal.

Our hybrid structural ensemble has an average width of the pentameric pore that is too narrow to fit a fully hydrated ion ( 3Å, Figure 7.3C). Unlike other pentameric ligand-gated ion channels, whose inner pores are interspersed with hydrophilic residues[396, 397], the PLN inner pore, which encompasses domain II ( 20 Å), is coated with hydrophobic residues, making it energetically improbable for a hydrated ion to diffuse through it (Figure 7.3D). Theoretical calculations support our findings, showing that the free energy for crossing the PLN pore for chloride ions is $\sim$ 20 kcal/mol, while for calcium ions it is more than 40 kcal/mol[383]. In addition, the three cysteine residues of the TM $\alpha$-helices (Figure 7.2C) do not face the interior of the pore[381], precluding their participation in the ion transport.

### 7.3.6 Pentamer to monomer topological changes

Overall, the structure of each protomer within the pentameric assembly of PLN is similar to that of the monomeric PLN (AFA-PLN, 2KB7)[84]. However, the two forms differ noticeably in the membrane topology (Figure 7.4A). In monomeric PLN, the tilt angle is $\sim$24°, whereas in the pentamer it is $\sim$11°. This difference is probably due to the driving forces for folding and stability of the two species. For the monomer, the driving force is the hydrophobic mismatch between protein and bilayer to optimize lipid-protein interactions, causing the TM domain to adopt a more pronounced tilt. In contrast, the orientation of each monomer within the pentamer is driven by interprotomer interactions (Leu/Ile zipper in domain II and hydrogen bonds in the domain Ib), which dominate over the hydrophobic mismatch, giving rise to a smaller tilt angle.

### 7.3.7 Molecular dynamic (MD) simulations of PLN in DOPC bilayers

The lowest energy conformer from the final conformational ensemble was embedded in a DOPC lipid bilayer (See SI for details) and the pentamer dynamics were analyzed for 150 ns. Unlike previous MD simulations that used the bellflower model as starting conformations[398, 399], our data show that the pinwheel topology of PLN does not significantly deviate from the initial configuration (Figure 7.4B). During the simulations, the tilt angles of the TM domains fluctuate around 13° in agreement with the solid-state NMR data (Figure 7.4B). The cytoplasmic domains, however, undergo larger

fluctuations without detaching from the bilayer surface (Figure 7.4B), in agreement with the dynamics determined for these domains[217]. The simulations reveal several interesting protein/protein and protein/lipid interactions that contribute to the stability of the pinwheel topology. Specifically, the guanidinium group of Arg14, snorkeling on the surface of the membrane, makes several hydrogen bonds with the oxygens of the lipid's glycerol moiety (Figure 7.4E). These interactions are functionally important since the deletion of this residue generates a loss-of-function species that is unable to properly regulate SERCA[360]. Additionally, the side chain of Arg25 in domain Ib forms a salt bridge with the side chain of Glu19, which is persistent throughout most of the simulations (Figure 7.4F). Several other interprotomer interactions take place in the juxtamembrane portion of PLN, with Gln29 forming a hydrogen bond with the backbone carbonyl oxygen of Asn27 of the neighboring monomer (Figure 7.4G). In contrast, MD studies showed that the bellflower topology[289] is highly unstable, with domains Ia (constituting the putative vestibule of the bellflower channel) bending after a few nanoseconds and interacting readily with the lipid bilayers[383, 398, 399]. Finally, MD trajectories show that the bellflower pore closes after a few nanoseconds, excluding water and ions and preventing transport across the bilayer.

## 7.4 Discussion

The T-state of PLN is the most populated in living cells[384]. It has been hypothesized that its interactions with lipid membranes are important for both oligomerization[400] and SERCA regulation[401]. The T-state structure of PLN reported here reveals that both interprotomer interactions and lipid/protein interactions contribute to the overall folding and stability. The cytoplasmic domains contain well-defined helices, with the hydrophobic faces absorbed onto the membrane surface (Figure 7.4C). This topology agrees with rotational dynamics measurements carried out by EPR[341] and solid-state NMR[84]. MD calculations show that the pinwheel structure does not deviate significantly from the initial topology, with the domains Ia maintaining a tight association with the lipid bilayer (Figure 7.4C-D). The pinwheel topology dismisses the previous models and agrees with a plethora of structural, biophysical, biochemical, and mutagenesis data[84, 392, 401, 402, 403]. The narrow pore within the TM domain is not

consistent with the proposed ion-channel function of PLN[382].

What is the function of the pentameric assembly? Are the interactions with lipids significant for PLN biological function? Lipid-mediated clustering and oligomerization is one of the emerging aspects defining the biological function of membrane proteins. These mechanisms are responsible for allosteric signaling, ion channel formation, and enzyme regulation[404]. Here, we propose that the ground state of PLN is in a quiescent T-state, with the amphipathic helices absorbed into the membrane. Upon membrane detachment, the cytoplasmic domains unfold[84] into a higher energy R-state, potentially acting as a mechanism by which the PLN pentamer dissociates one or more subunits to interact with SERCA[135].

Although in vivo and in vitro data show that promoting PLN de-oligomerization increases SERCA inhibition[176, 379, 405], we propose that the PLN pentamer actively participates in the regulatory mechanism, tuning the extent of monomeric species available for SERCA inhibition. The latter is supported by cardiac-specific simultaneous over-expression studies of pentameric and monomeric PLN[406]. It was found that the PLN monomer (C41F) is able to reduce $Ca^{2+}$ affinity of SERCA to the same extent as the PLN pentamer, but it was not able to reduce the rate of myocardial relaxation[406]. Moreover, PLN phosphorylation enhances the population of the pentamer[379], while an increase in SERCA expression shifts the equilibrium toward the monomeric species[282]. Importantly, this equilibrium plays a crucial role in muscle pathophysiology. Our recent studies show that an increased stability of the pentamer by the R9C mutation results in a decreased inhibition of SERCA and hampers PLN phosphorylation at Ser16 by PKA, which may lead to dilated cardiomyopathy[385].

In addition to the pentamer/monomer equilibrium, the conformational interconversion from the more structured T-state to a more dynamic, unfolded, and membrane detached R-state[135, 340, 136] is a salient aspect of PLN regulatory function. The T-state is the most populated in both micelles and lipid membranes[340, 84]. However, this equilibrium can be shifted toward the R-state by phosphorylation[340, 132], binding partners[77], lipid composition[401, 407], or protein aggregation. The tendency

of monomeric PLN domain Ia to unfold into the R-state was also detected by solid-state NMR[358]. These authors showed that in DMPC (20:1 lipid to protein ratio), the TM domain of PLN is helical, while the domain Ia is mostly unfolded. Remarkably, we found a direct correlation between the extent of R-state and SERCA regulatory function[136, 408]. These studies indicate a potential avenue to design PLN analogs for gene therapy in heart failure[336].

From a biophysical standpoint, our study reveals a novel aspect of TM oligomerization. The packing of the PLN pentamer prevails over the hydrophobic mismatch of the lipid membrane, causing a change in the topology of the TM domain relative to the monomer. Given the dynamic nature of lipids, changes in TM topology within the greasy bilayer may be energetically favorable over folding and unfolding of secondary structure domains, and may represent the way protein-protein interactions take place within lipid membranes.

## 7.5 Supporting information

### 7.5.1 Pentameric PLN expression and purification

PLN expression and purification were carried out as previously reported[174, 84]. Briefly, PLN was expressed in E. coli BL21(DE3) (Novagen) as a fusion protein with maltose binding protein (MBP, New England Biolabs). After amylose affinity chromatography, PLN was cleaved from MPB fusion by incubation with tobacco etch virus (TEV) protease. The final purification step was achieved by reverse phase HPLC. After lyophilization, protein masses were assessed by SDS-PAGE electrophoresis and MALDI-TOF mass spectrometry.

For U- $^{15}N$ and $^{13}C$ labeled PLN, the cells were grown in M9 minimal media containing $^{15}N$H4Cl (Cambridge Isotope Laboratory) and $^{13}C$-glucose (Sigma-Aldrich) as the sole nitrogen and carbon sources, respectively. To label specific residues in PLN, the cells were grown in media containing all of the amino acids at natural abundance ($\sim$99.6% 14N), with the exception of the amino acids of interest, which were $^{15}N$ labeled at the backbone amide nitrogen. For MAS experiments (DARR), the amino acids were

U-$^{13}C$ labeled. For the asymmetrically labeled sample preparations, we followed the protocol reported in Ref[84].

For the asymmetric NOESY experiments, we utilized two different samples: one for the detection of methyl-methyl NOEs, the second for the detection amide-methyl NOEs. The first sample was obtained from mixing two different protein preparations in $^{2}H$-glucose incorporating the precursors a) 2-keto3-(methyl-$^{2}H_3$)-butyric acid4-$^{13}C$, 3-$^{2}H_1$ (Leu and Val) and b) 2-ketobutyric acid4-$^{13}C$, 3, 3-$^{2}H_2$ (Ile)[261]. The first fusion protein was U-$^{2}H$ with $^{13}C$ methyl labeling at Leu$^{\delta 1/2}$ and Val$^{\gamma 1/2}$, while the second fusion protein was U-$^{2}H$, with $^{13}C$ methyl labeling at Ile$^{\delta 1}$ only. The second sample was prepared by mixing 50% of a fusion protein U-$^{2}H$, $^{15}N$ labeled and 50% U-$^{2}H$ with $^{13}C$ methyl labeled at Ile$^{\delta 1}$, Leu$^{\delta 1/2}$, and Val$^{\gamma 1/2}$. Note that the two fusion proteins were purified separately and then mixed (1:1 molar ratio) before cleavage with TEV protease to ensure uniform mixing. The mixed PLN pentamers were purified as described above. Asymmetric labeled samples for MAS-NMR experiments were prepared using the same approach. In this case, the two fusion proteins were labeled with U-$^{13}C$ Leucine and $^{13}C$ Isoleucine, respectively, were mixed (1:1 molar ratio) prior to TEV cleavage and HPLC purification.

### 7.5.2 NMR sample preparation

#### 7.5.2.1 Solution NMR

The samples were prepared according to the procedure reported in Ref[84]. PLN was reconstituted in a phosphate buffer (120 mM Na2HPO4 pH 6.0, 120 mM NaCl, 3 mM NaN3, 50mM 2-mercaptoethanol, and 5% D2O) containing 300 mM dodecylphosphocoline (DPC) (Sigma-Aldrich) (or protonated DPC for the protein-DPC NOESY experiment). To avoid aggregation, all samples were denatured with 6 M guanidine-HCl and extensively dialyzed against NMR buffer prior to acquisition of the NMR spectra. For all samples, concentration of pentameric PLN was ∼0.3 mM.

#### 7.5.2.2

Mechanically oriented samples were prepared as previously described[84]. First, small unilamellar vesicles in a 4:1 molar ratio of 1,2-dioleoyl-sn-glycero3-phosphocholine (DOPC) and 1,2-dioleoyl-sn-glycero3-phosphoethanolamine (DOPE) (Avanti Polar Lipids, Alabaster, AL). Then, we dissolved 2-4 mg of PLN in 10% SDS and added to the SUVs, followed by one freeze/thaw cycle. SDS was removed by extensive dialysis against water in the presence of 250 $\mu$M DTT. Lipid vesicles containing PLN were concentrated to $\sim$2 mg/mL and spread onto 40 glass slides (5.7 mm $\times$ 12 mm $\times$ 0.030 mm) (Matsunami Glass Ltd, Osaka, Japan). The glass slides containing the lipid/protein preparations were incubated at 45 °C for 3 hours, and subsequently stacked on top of each other and re-hydrated for 3-4 days in a humidity chamber (45 °C and 99% humidity). The samples were sealed in glass cells and used for NMR measurements. For all samples, the final lipid/protein molar ratio was $\sim$200/1. MAS samples for DARR experiments were prepared by co-dissolving 20 mg of DOPC/DOPE and 1mg of $^{13}C$-Leu/$^{13}C$-Ile PLN in organic solvent (TFE/Chloroform). The solvent was removed by evaporation under N2 stream. The preparations were re-hydrated with 10 $\mu$L of aqueous buffer (50$\mu$M 2-mercaptoethanol, 120 mM Na2HPO4) and transferred into 3.2 mm zirconia thin-walled MAS rotors.

### 7.5.3 NMR experiments

#### 7.5.3.1 Solution NMR

. For the assignment of intramonomer NOEs, two $^{15}N$-edited NOESY spectra were acquired on a Varian spectrometer operating at 800 MHz ($^{1}H$ Larmor frequency, NMRFAM; Madison, WI) and equipped with a cold probe. Each experiment was recorded with 512, 80, and 60 complex points in the three dimensions, with spectral widths of 9000 Hz, 9000 Hz, and 1560 Hz, respectively. The acquisition times were 57 ms, 9 ms, and 38 ms in $t_3$, $t_2$, and $t_1$ dimension, respectively. NOE mixing times of 90 and 220 ms were acquired with a recycle delay of 1.5 s and 8 transients for each increment. The temperature was set to 30 °C.

For the assignment of inter-monomer NOEs, two NOESY spectra were used. The first is described in Ref[261], whereas the second consisted of a simple $^{15}N$-edited NOESY in which the $\omega_1$-spectral window (indirect $^1H$ dimension) was reduced to 1300 Hz, resulting in aliasing of the resonances and a substantial improvement in resolution without increasing the experimental acquisition time. This was possible because of the absence $^1H$ resonances due to the deuterated background of the protein. Unambiguous assignment of the inter-monomer NOEs was achieved using chemical shift assignments of $^{13}C$H3-Ile$^{\delta 1}$ and $^{15}N$H groups. The experiment consisted of a matrix of 787, 70, and 28 complex points, and spectral widths of 10000 Hz, 4000 Hz, and 1300 Hz. The acquisition times were 79 ms, 18 ms, and 21 ms for $t_3$, $t_2$, and $t_1$, respectively. A mixing time of 400 ms was used with a recycle delay of 1.3 s utilizing 24 transients for each increment. The topology of the pentamer in micelles was probed by measuring NOEs between PLN and DPC[409]. In this experiment PLN was labeled U-$^2H$ and $^{15}N$, with the detergent fully protonated. NOEs were detected using a $^{15}N$-edited NOESY with 852, 60, and 30 complex points, spectral widths of 10000 Hz, 6600 Hz, and 1300 Hz and acquisition times of 85 ms, 9 ms, and 23 ms in $t_3$, $t_2$, and $t_1$, respectively. As with all the NOESY experiments, the mixing times were optimized using build-up curves to avoid spin diffusion. $^1H$ chemical shifts were referenced directly to 3-(trimethylsilyl)-1-propane-sulfonic acid sodium salt (Aldrich, Milwaukee, WI) at 0 ppm, whereas $^{15}N$ and $^{13}C$ were indirectly referenced[410]. Complete Carbon shifts ($C_\alpha$, $C_\beta$, $C_o$) were determined using triple resonance experiments HNCO, CBCA(CO)NH[386, 387]. Secondary structure prediction (SSP) scores were obtained using the software SSP by Lewis Kay (University of Toronto: http://abragam.med.utoronto.ca/software.html.)

#### 7.5.3.2 Solid-State NMR

Oriented solid-state NMR spectra were acquired at a 14.1 T magnetic field strength ($^1H$ frequency of 600.1 MHz) using Bruker DMX (National High Magnetic Field Laboratory, NHMFL; Tallahassee, FL) and Varian VNMRS spectrometers. The 2D polarization inversion spin exchange at the magic angle (PISEMA) experiment[343] was performed at $\gamma B_1/(2\pi)$ strength of 60 kHz for the $^1H$ 90° pulse, cross-polarization, and SPINAL64 decoupling[369]. Low-E probes with a doubly tuned, low-inductance resonator built by the RF program at the NHMFL were used[346]. PISEMA experiments

were acquired with 4 k scans and 12 $t_1$ increments for $^{15}N$-Leu, $^{15}N$-Cys, $^{15}N$-Ile, and $^{15}N$-Val PLN, 12 k scans and 8 $t_1$ increments for $^{15}N$-Ala, $^{15}N$-Asn and $^{15}N$-Arg PLN, and 20 k scans and 8 $t_1$ increments for $^{15}N$-Thr. All experiments used a recycle delay of 4 s. The spectra were processed using NMRPipe[411] and viewed with NMRVIEW software[277]. An exponential window function was applied by using 100-300 Hz line broadening along the $^{15}N$ chemical-shift dimension ($t_2$). No window function was used to process the data in the indirect dipolar coupling dimension. After Fourier transformation and zero-filling, the data consisted of a total matrix size of 2 k × 1 k points in the direct and indirect dimensions, respectively.

Dipolar assisted rotational resonance (DARR)[412] experiments were run at a 14.1 T magnetic field strength ($^1H$ frequency of 600.1 MHz) using a Varian VNMRS spectrometer equipped with a 3.2 mm bioMAS probe. The DARR mixing time was set to 200 ms and utilized a radiofrequency application to $^1H$ that corresponded to the n=1 condition (15 kHz). Experiments were acquired with 1024 scans and 50 $t_1$ increments. Temperature was set at -25° C and the sample was spun at 8 kHz or 15 kHz (note that no differences in spectra quality were detected). The DARR experiment was repeated at 5° C. The pattern for the dipolar connectivities is identical to -25° C. The signal to noise ratio was lower for the DARR acquired at 5° C.

### 7.5.4 Structure calculation of hybrid pentamer

The structure calculations were carried out with XPLOR-NIH (v.2.23) software, using the hybrid solution and solid-state NMR methods reported in Refs[84, 292] A target function ($E_T$) was formulated as a linear combination of geometrical ($E_{chem}$), solution NMR ($E_{solNMR}$), and solid-state NMR ($E_{ssNMR}$) terms:

$$E_T = E_{chem} + E_{solNMR} + E_{ssNMR}$$

The hybrid energy $E_T$ function was minimized with simulated annealing protocols using combined minimization and molecular dynamics. To generate the pentamer ensemble, a five-step structure calculation protocol was used as in Figure 7.6.

**Step 1:** A hybrid monomer ensemble was generated starting from a solution-NMR ensemble. After folding the secondary structure elements, solid-state NMR restraints

(CSA and DC) were included in the simulated annealing protocol, starting at a temperature of 3000 K with torsion angle dynamics. The force constants for the solid-state NMR restraints were optimized using cross-validation as described previously[318]. A total of 200 monomer structures were generated.

**Step 2:** An initial symmetric pentamer was assembled with five monomer replicates, each of which was rotated by 72° (to avoid the clash of cytoplasmic helices) and placed at a distance of 50 Å. The pentamer was subjected to a rigid-body minimization protocol using the inter-monomer NOEs.

**Step 3:** Solid-state and solution NMR restraints were then combined using a high-temperature annealing protocol similar to step 1. An initial temperature of 5,000 K was used. Non-crystallographic symmetry (NCS) was used to ensure that each monomer was identical in the pentamer, imposing a force constant of 50 kcal/mol. To maintain five-fold symmetry, 832 symmetry restraints were added as NOE-like distance constraints with a force constant of 10 kcal/mol[393]. Planar distance restraints[145] were added to determine an initial depth of insertion with respect to the micelle and the relative positioning between the two helical domains of PLN. The restraints were derived from a normalization of the protein-detergent NOEs intensities (Figure 7.9) using the following equation:

$$I_{H_2O}/(I_{DPC} + I_{H_2O})$$

where IH2O is the intensity of the water cross-peak, IDPC is the intensity of the protein-DPC cross peak. Since Leu7, Arg14, and Asn30 residues have the same normalized values, they were chosen to lie on the same plane.

**Step 4:** Starting from a temperature of 500 K and using torsion angle dynamics and Cartesian molecular dynamics, the conformer local geometries were relaxed to improve the structure quality. A full VDW potential was introduced, which resulted in better side chain packing. Twenty lowest energy conformers were selected for further structural analysis.

**Step 5:** To better define the depth of insertion in a realistic model membrane, one of the twenty conformers from step 4 was selected for structural refinement using MD

simulations in explicit DOPC lipid bilayers. To place PLN within the membrane bilayer, we used rigid body minimization with the Ez potential, maintaining a fixed orientation with respect to the z-axis. The potential term (Ez) is database potential for insertion of helical segments into a fictitious bilayer. The final system was then set up using CHARMM-GUI[36, 413], utilizing the PARAM27 force field with the CMAP correction[185, 414]. The calculations were performed using NAMD version 2.6[349].

The optimal number of 1,2-dioleoyl-sn-glycero3-phosphocholine (DOPC) lipids per leaflet was determined using P21 periodic boundary conditions (PBC) with a tetragonal unit cell[415], where the lipids can freely equilibrate between the two layers. Starting with 176 on the bottom leaflet and 151 on the top leaflet, the number of lipids converged to 167 at the bottom leaflet (lumen) and 160 on top leaflet (cytoplasm) after 3.5 ns of equilibration. Once the number of lipids converged, the system was rebuilt following the method described by Woolf and Roux[186], with P1 tetragonal PBC (112 $\times$ 112 $\times$ 112 Å) and 327 lipids. We also included 75 $K^+$ and 90 $Cl^-$ ions to model 150 mM ionic strength and to maintain the electroneutrality of the system. The overall system consisted of 140,667 atoms. The dynamics of the system were simulated as an isothermal-isobaric ensemble using 310 K and at a pressure of 1 atm. The temperature was kept constant using the Langevin dynamics algorithm and the pressure of the system was controlled anisotropically, allowing the z-axis to fluctuate independently from lipid plane X-Y, which was restrained to a constant ratio of 1:1. The electrostatic interactions were calculated using particle-mesh Ewald (PME)[416] method, while a Lennard-Jones potential used for the non-bonding interactions (feathered to zero in the range of 9 to 11 Å). Hydrogen atom covalent bonds were restrained using the RATTLE algorithm using a multiple integration time-step with 6 fs for full electrostatic evaluation and 2 fs for all other energy terms. In total, we performed 10 ns of equilibration with PLN backbone atoms harmonically restrained (using a force constant of 0.5 $kcal \cdot mol^{-1}$Å$^{-2}$), followed by 140 ns of unconstrained MD simulations.

Table 7.1: Summary of structural ensemble statistics

| | |
|---|---|
| **R.M.S. Deviation from experimental restraints** | |
| NOE (Å)( total: 1945 intramonomer, 75 intermonomer) | $0.05 \pm 0.003$ |
| Torsion angle (°)(85 per monomer) | $0.46 \pm 0.16$ |
| **PISEMA RMS (%)** | |
| CSA(31 per monomer, ppm) | 2.8 |
| DC (29 per monomer kHz) | 0.3 |
| **R.M.S. deviations from idealized covalent geometry** | |
| Bond (Å) | 0.005±0.0005 |
| Angle (°) | 0.60±0.01 |
| Impropers (°) | 0.16±0.006 |
| **Measure of structural quality** | |
| % residues in most favored region | 93.9±2.5 |
| % residues in additional allowed region | 5.8±2.2 |
| % residues in generously allowed region | 0.3±0.7 |
| % residues in disallowed region | 0 |
| **Precision of atomic coordinates (Å)** | |
| Backbone (residue 1-52) | 4.2 |
| Helix 1 (residue 2-18) | 3.6 |
| Helix 2 (residue 24-50) | 0.5 |
| **Average MolProbity score**[1] | 1.8 (94 percentile) |
| **Protein Topology (domain in subscript)** | |
| Tilt ($\theta_{Ia}$) (°) | 96±3 |
| Rotation ($\rho_{Ia}$) (°) | 78±16 |
| Tilt ($\theta_{Ib,II}$) (°) | 11±1 |
| Rotation ($\rho_{Ib,II}$) (°) | 223±6 |

[1]MolProbity scores are determined using the validation tool MolProbity[394]

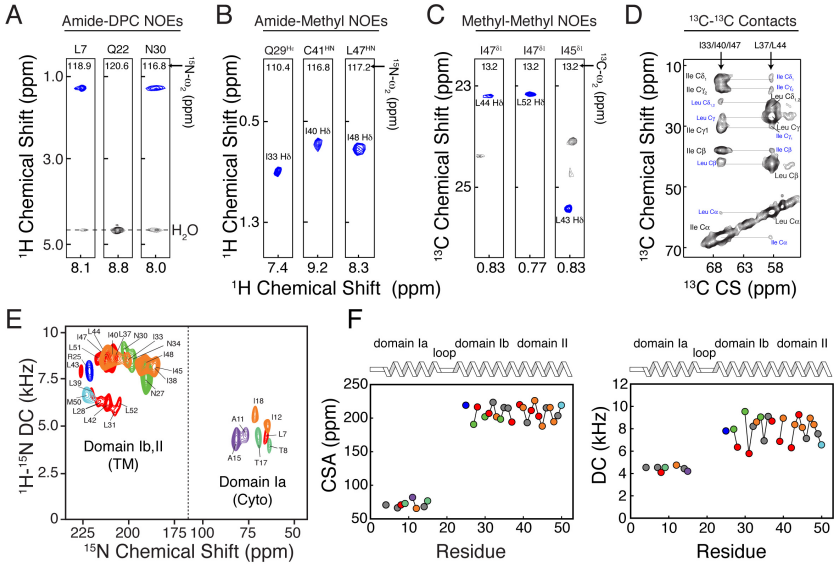


Figure 7.1: Multidimensional solution and solid-state NMR spectra defining the structural topology of PLN in DPC micelles, planar lipid bilayers and lipid vesicles. (A) 2D planes from a 3D [$^1H$, $^1H$, $^{15}N$]-NOESY-HSQC spectrum (300 ms mixing time) showing NOEs between amide protons of PLN and methyl and methylene groups of DPC. Note the peaks at ∼4.5 ppm are exchange cross peaks between the protein amide resonances and the water signal. (B) 2D planes from 3D [$^1H$, $^1H$, $^{15}N$]-NOESY-HSQC (400 ms mixing time) on a mixed PLN sample with 1/1 of U-[$^2H$-$^{15}N$]/U-[$^2H$-14N]-$^{13}C$H3-Ile$^{\delta 1}$ PLN. (C) 2D planes from 3D [$^1H$, $^{13}C$, $^{13}C$]-HSQC-NOESY-HSQC experiment performed on a sample containing 1/1 of U-[$^2H$-14N]-$^{13}C$H3-Ile$^{\delta 1}$ and U-[$^2H$-14N]-$^{13}C$H3-Leu$\delta$1/Val$\gamma$1 PLN. (D) 2D-DARR experiments (200 ms mixing time) on a 1/1 sample of U-$^{13}C$-Leu/U-$^{13}C$-Ile PLN sample. Intra-residue and interprotomer cross-peaks are labeled in black and blue, respectively. (E) Overlay of 2D [$^1H$, $^{15}N$]-PISEMA spectra of selectively labeled PLN in oriented DOPC/DOPE (4/1) lipid bilayer. (F) Plot of the CSA (left) and DC (right) values versus residue.

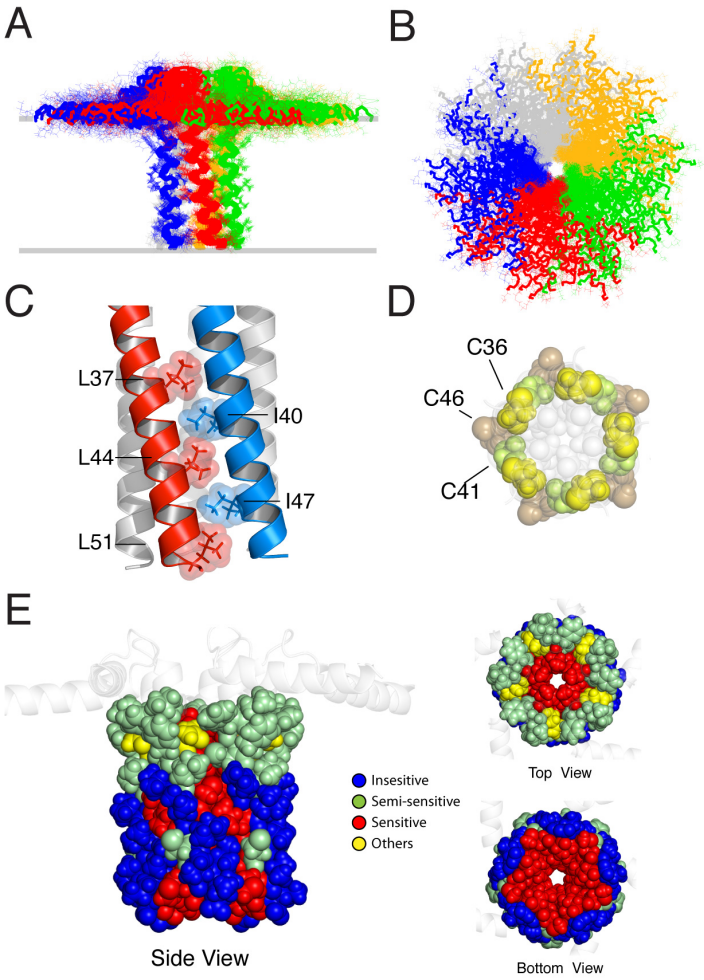


Figure 7.2: Pinwheel architecture of pentameric PLN in lipid bilayers. (A) and (B) PLN hybrid conformational ensemble. Overlay of the 20 lowest energy structures. The conformers were aligned using heavy atoms from residues 24 to 48. See Table 7.1 for structural statistics. (C) Domain II Leu/Ile zipper motif. Residues are shown using space filling model with atomic radii. (D) Spatial arrangement of the three cysteines in the PLN TM domain. The hydrophobic residues lining the inner pore are shown in light gray. Cys41 is at the interface between protomers. (E) Mapping of the mutagenesis studies on the PLN pentamer. Residue sensitive to mutation and oligomerization[176] are indicated in red. Left: side view of the PLN pentamer with residues from 24 to 52 shown in space filling model.

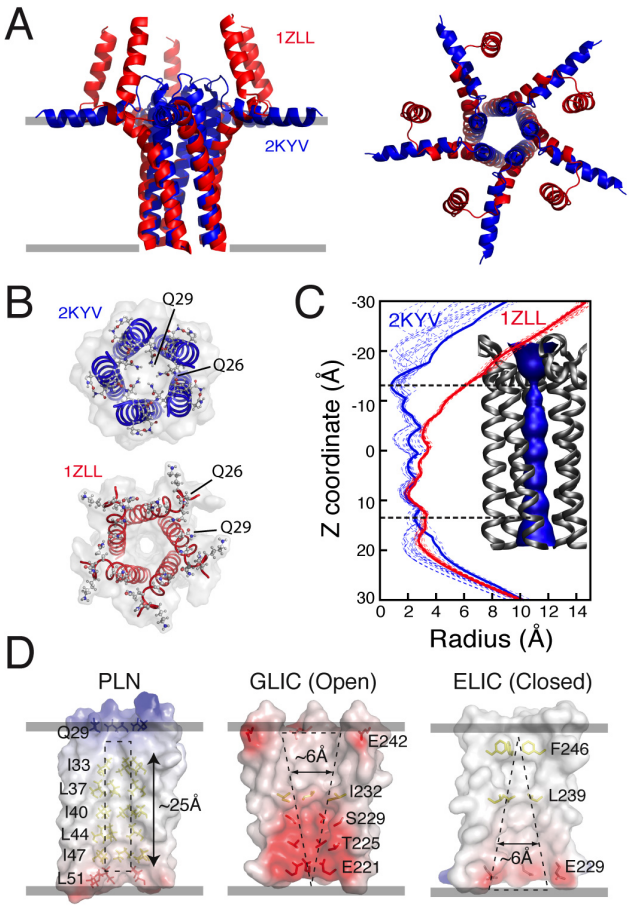


Figure 7.3: Comparison between 2KYV and 1ZLL structural models. (A) Overlay of the TM regions of the 1ZLL (red) and 2KYV (blue) models. Protein backbones were aligned using heavy atoms for residues 24 to 48. (B) Top view of the 1ZLL and 2KYV models. (C) Width of the central pore calculated with the program HOLE2[417] for the 2KYV (blue dashed traces) and 1ZLL (red dashed traces) ensembles. Thick solid lines refer to the structures shown in (A). Inset: surface rendering of the pore for the 2KYV model. (D) Electrostatic surface potentials of two pentameric ligand-gated ion channels (ELIC and GLIC)[396, 397] and PLN (2KYV).

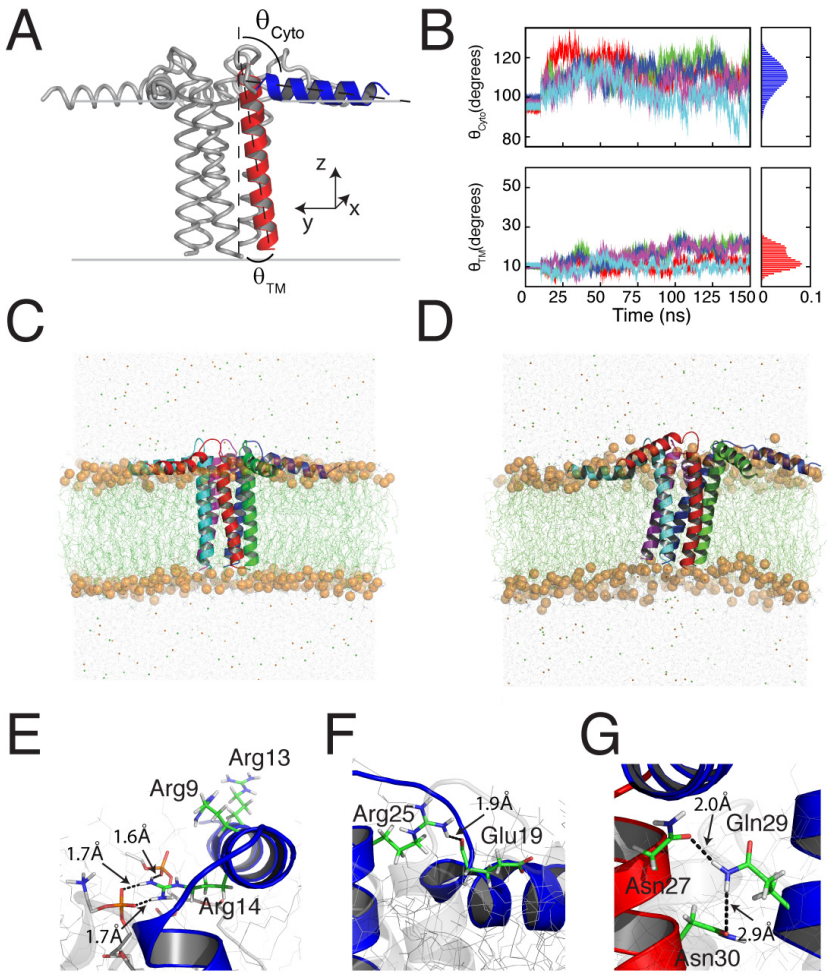


Figure 7.4: MD simulations of PLN in explicit DOPC lipids. (A) Average structure of PLN from the hybrid NMR conformational ensemble with the TM domain (red) tilted by 11o. (B) Time course of the tilt angles of both the TM and cytoplasmic helices of PLN pentamer. (C-D) Snapshots of PLN from the MD trajectories taken at 10 ns and 150 ns. (E) Orientation of the three Arg in domain Ia. The guanidinium group of Arg14 forms transient hydrogen bond with several DOPC lipid molecules. (F) Arg25 and Glu19 salt bridge. (G) Interprotomer contacts between the side chain of Gln29 and Asn27 and Asn30.

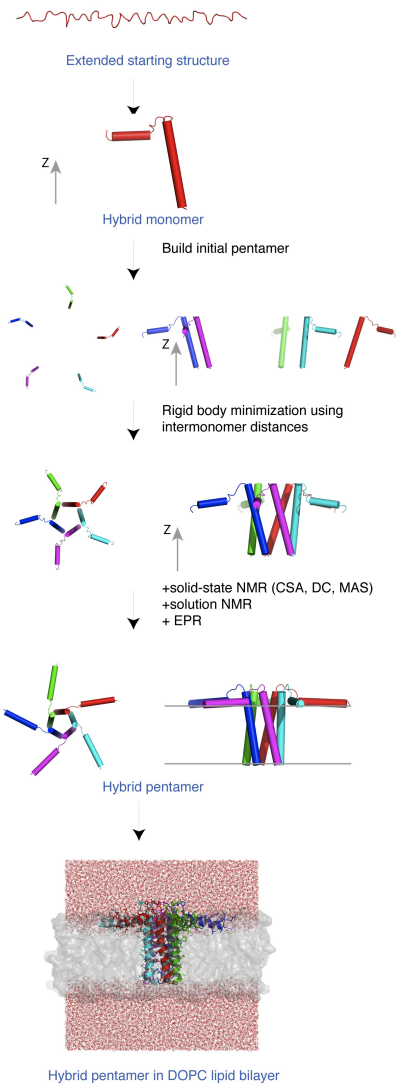


Figure 7.5: (A) $[^1H, ^{15}N]$-TROSY HSQC spectrum of U-$^{15}N$ labeled PLN reconstituted in 300 mM deuterated DPC. (B) Plot of the short range backbone NOEs obtained from $[^1H,^{15}N]$ NOESY-HSQC experiments. (C) Chemical shift index histograms of PLN obtained from triple resonance experiments. Overlapping or missing residues are not shown.

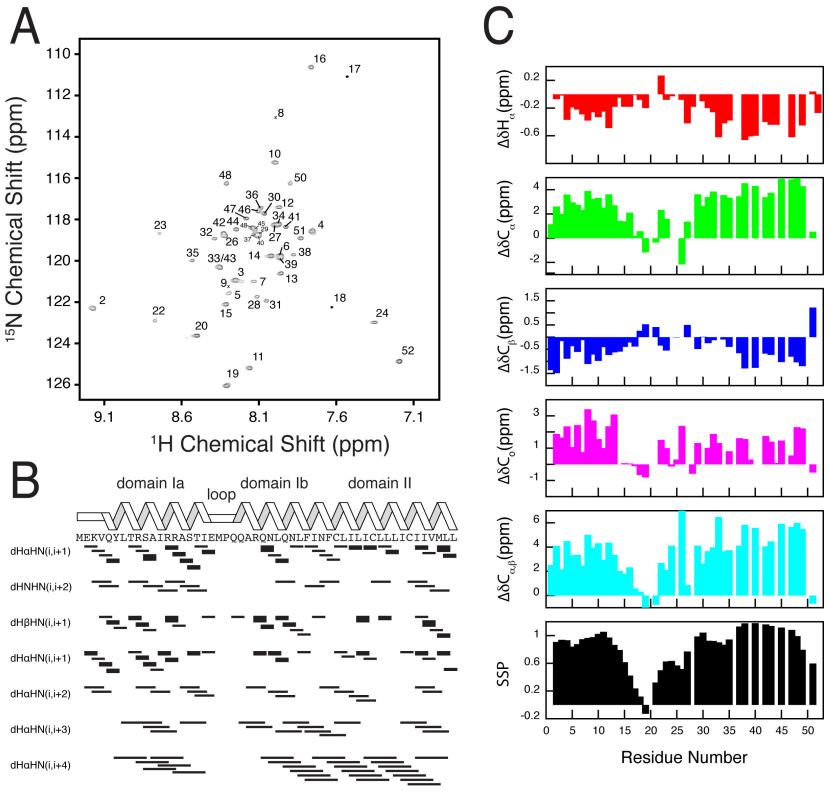


Figure 7.6: Overview of the hybrid refinement protocol for the simultaneous determination of structure and topology of pentameric PLN. (See SI-Methods for details).

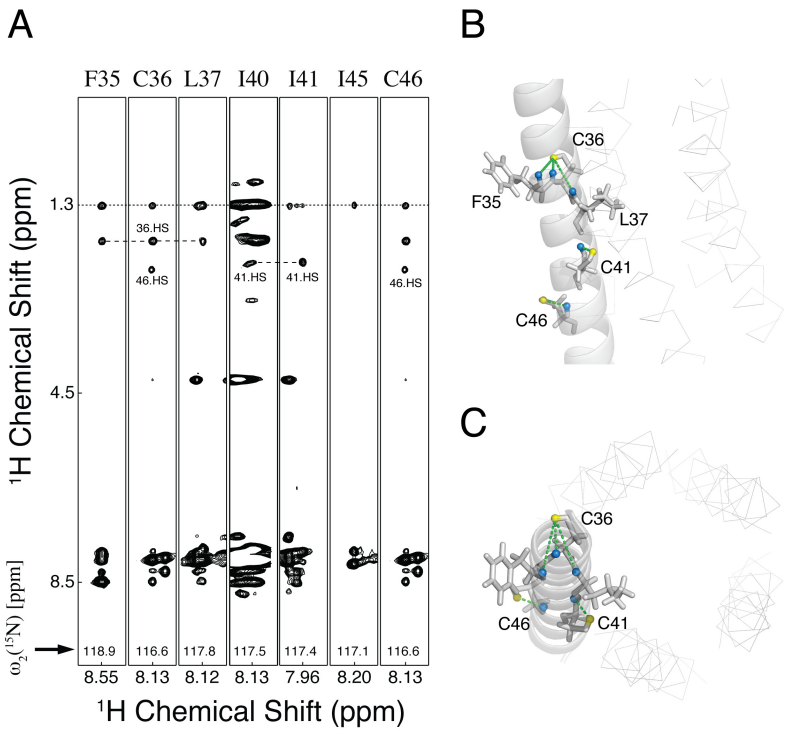


Figure 7.7: NOE patterns for Cys residues in PLN. (A) Strip plots from a 3D $^{15}N$-edited NOESY experiment on a sample of U-$^{15}N$ PLN reconstituted in 300 mM DPC taken at the $^{15}N$ frequencies for specific residues. Although Cys36 and Cys46 have similar $^{15}N$ chemical shift, the residues were identified using sequential assignments. This shows that the $H_\gamma$ chemical shifts are different for the three Cys residues, supporting the differential involvement in pentamer packing shown in panels B and C. Note that L37 and I40 are partially overlapped.

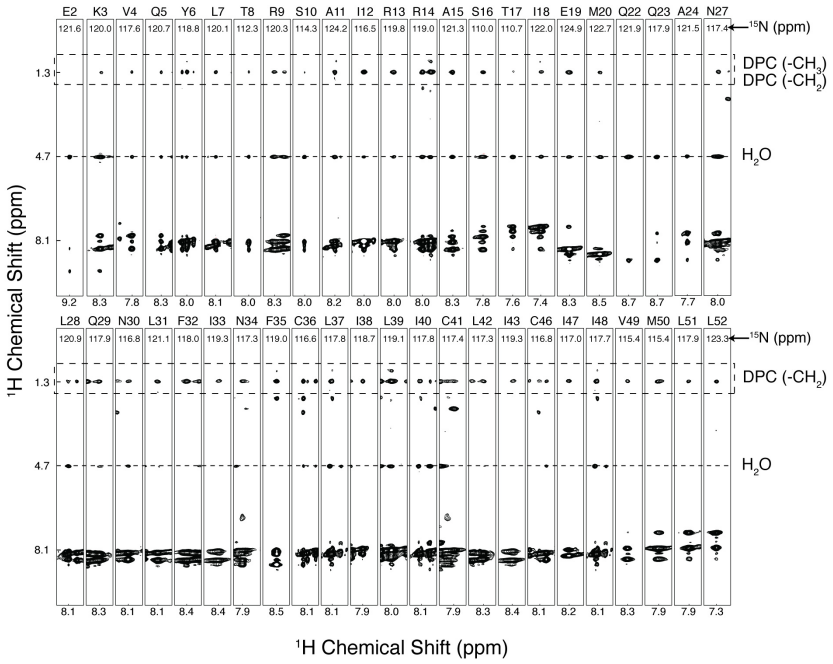


Figure 7.8: NOE strip plots from the $^{15}N$-edited 3D NOESY of U-$^{2}H$, U-$^{15}N$ PLN in protonated 300 mM DPC. Overlapping residues are not shown. The peak at 4.7 ppm is assigned to the water signal and the ratios between the cross peaks and the diagonal peaks from the amide resonances was used as an indication of the water exposure of each residue (See SI methods).

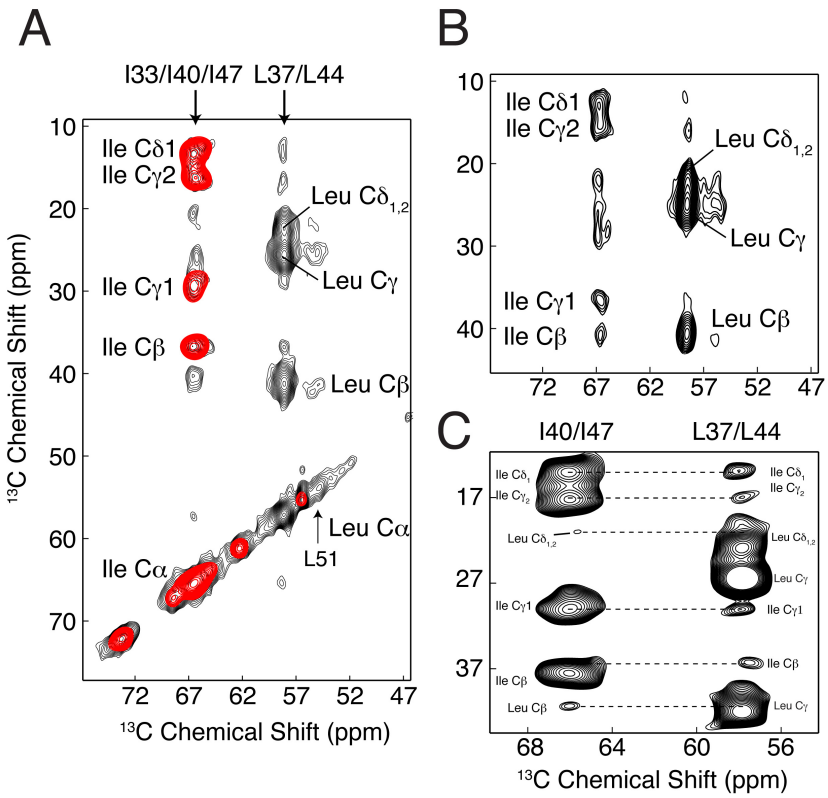


Figure 7.9: 2D-DARR experiments on $^{13}C$-PLN samples in DOPC/DOPE lipid vesicles. (A) Spectra overlay of the mixed samples with 50% U-$^{13}C$-Leu/ 50% U-$^{13}C$-Ile (black) and control sample with 100% U$^{13}C$-Ile PLN (red) acquired at acquired at -25 °C. Note Leu51 has a distinct chemical shift and can be unambiguously assigned based on the chemical shift found by solution NMR. (B) The spectrum of the mixed sample with 50% U$^{13}C$-Leu/ 50% U-$^{13}C$-Ile PLN was acquired at 5 °C. (C) The spectrum of the mixed sample of 50% U-$^{13}C$-Leu37/44 and 50% U-$^{13}C$-Ile40/47 synthesized by solid phase FMOC synthesis. The residue-specific assignments for all of the Leu and Ile residues were accomplished by acquiring 2D DARR experiments on single site $^{13}C$ labeled PLN samples.

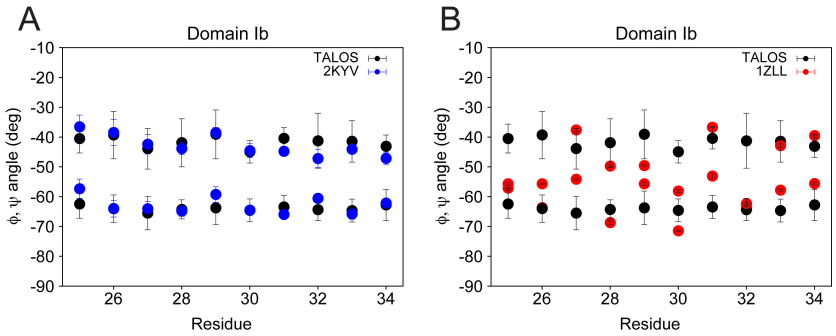


Figure 7.10: Agreements between the predicted dihedral angles using TALOS and the dihedral angles obtained in the calculated ensemble of the bellflower and pinwheel structural models for domain Ib (residues 25 to 34). (A) Dihedral angle calculated using TALOS[14] (black) and dihedral angles back calculated from the hybrid (2KYV) ensemble (blue). (B) Dihedral angle calculated using TALOS (black) and dihedral angles back calculated from the bellflower (1ZLL) ensemble (red).

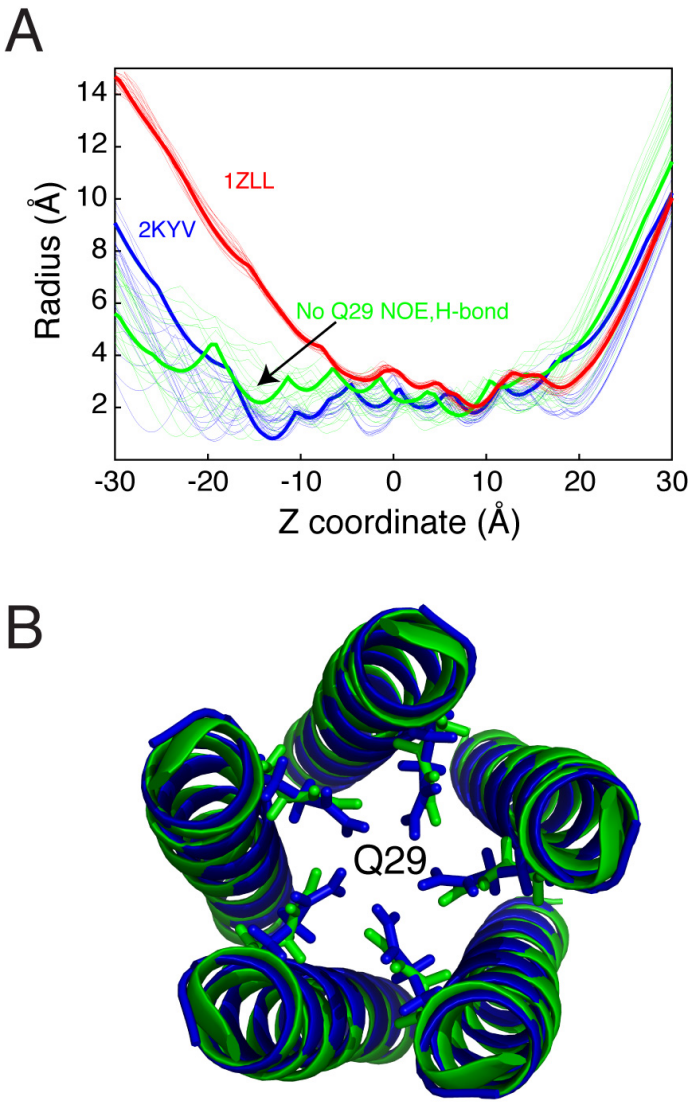


Figure 7.11: Effect of removing Q29 inter-monomer restraints on the packing of PLN. (A) Comparison of the width of the PLN pore in the hybrid ensemble with (blue) and without (green) the Q29-I33 and Q29 H-bond restraints. Red traces refer to the width of the PLN pore in the bellflower ensemble. Thick lines correspond to the lowest energy structure in each ensemble. (B) Top view of the juxtamembrane region of PLN and orientation of the Q29 side chain in the presence (blue) and absence (green) of the two intermonomer restraints.

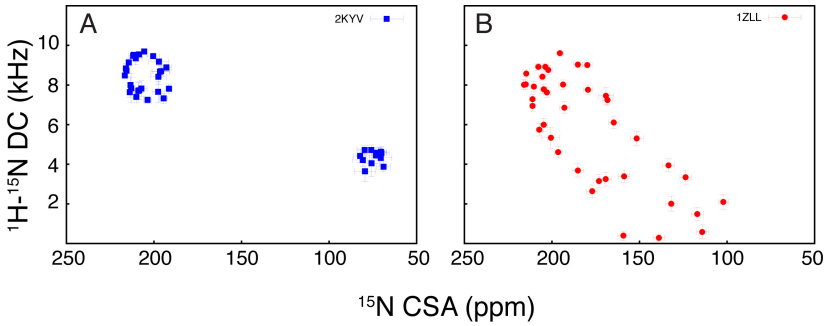


Figure 7.12: Comparison between PISEMA patterns back calculated from the pinwheel and bellflower ensemble. (A) Back calculation of PISEMA pattern from 2KYV ensemble for residues 4-16 and 25-50. Notice the two well defined PISA wheels in the 2KYV ensemble, corresponding to domain Ia (∼70 ppm) and domain II (∼200 ppm). (B) Back calculation of PISEMA pattern from 1ZLL ensemble (residues 4-16 and 25-50). The 1ZLL ensemble was aligned to one of the 2KYV PDBs using backbone atoms of residue 32 to 50.

Table 7.2: Summary of the NMR restraints from solid-state NMR. Note Met1 and Met20 cannot be distinguished in the combinatorial assignment algorithm.

| **Residue** | $^{15}N$ **CSA (ppm)** | $^{1}H-^{15}N$ **DC (kHz)** |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | 70.5 | 4.5 |
| 5 | | |
| 6 | 80.2 | 4.5 |
| 7 | 66.8 | 4.5 |
| 8 | 71.0 | 4.1 |
| 9 | 72.8 | 4.5 |
| 10 | | |
| 11 | 82.0 | |
| 12 | 65.6 | 4.7 |
| 13 | | |
| 14 | 68.1 | 4.4 |
| 15 | 77.0 | 4.2 |
| 25 | 219.1 | 7.8 |
| 26 | | |
| 27 | 190.6 | 7.9 |
| 28 | 216.6 | 6.3 |
| 29 | | |
| 30 | 202.1 | 9.5 |
| 31 | 207.4 | 5.8 |
| 32 | 223.5 | 8.2 |
| 33 | 201.7 | 8.6 |
| 34 | 198.3 | 9.0 |
| 35 | 215.8 | 6.9 |
| 36 | 214.9 | 9.0 |
| 37 | 194.5 | 8.7 |
| 38 | | |
| 39 | 220.3 | 6.9 |
| 40 | 215.8 | 8.9 |
| 41 | 193.0 | |
| 42 | 211.4 | 6.3 |
| 43 | 226.3 | 8.3 |
| 44 | 202.7 | 9.2 |
| 45 | 187.8 | 8.1 |
| 46 | 215.3 | 7.1 |
| 47 | 215.8 | 8.9 |
| 48 | 194.6 | 8.3 |
| 49 | 203.7 | 7.5 |
| 50 | 218.9 | 6.5 |
| 51 | | |
| 52 | | |

Table 7.3: Inter-monomer restraints (NOE)

| **Monomer** $i$ | **Monomer** $i+1$ | **Distance (Å)** |
|---|---|---|
| Gln29 $H_{\epsilon 2\#}$ | Ile33 $H_{\delta 1\#}$ | 1.8-5.0 |
| Ile40 $H_{\delta 1\#}$ | Ile40 $H_{\delta 1\#}$ | 1.8-5.0 |
| Ile40 $H_{\delta 1\#}$ | Cys41 $H_N$ | 1.8-5.0 |
| Leu43 $H_{\delta\#}$ | Ile48 $H_{\delta 1\#}$ | 1.8-5.0 |
| Leu43 $H_{\delta\#}$ | Ile45 $H_{\delta 1\#}$ | 1.8-5.0 |
| Leu43 $H_{\delta\#}$ | Ile45 $H_N$ | 1.8-5.0 |
| Leu43 $H_{\delta\#}$ | Ile45 $H_\alpha$ | 1.8-5.0 |
| Leu44 $H_N$ | Leu44 $H_{\delta\#}$ | 1.8-5.0 |
| Ile47 $H_{\delta 1\#}$ | Leu44 $H_{\delta\#}$ | 1.8-5.0 |
| Ile47 $H_{\delta 1\#}$ | Leu44 $H_{\beta\#}$ | 1.8-5.0 |
| Ile47 $H_{\delta 1\#}$ | Ile48 $H_N$ | 1.8-5.0 |
| Ile47 $H_N$ | Ile48 $H_{\delta 1\#}$ | 1.8-5.0 |
| Ile47 $H_{\delta 1\#}$ | Leu51 $H_{\delta\#}$ | 1.8-5.0 |
| Ile47 $H_{\delta 1\#}$ | Leu52 $H_{\delta\#}$ | 1.8-5.0 |
| Gln29 $N_{\epsilon 2}$ | Gln29 $C_\delta$ | 1.8-5.0* |

*Distance from REDOR-NMR experiment[392]

Table 7.4: Experimental measured distance in PLN

| **Ref** | **Techniques** | **Atome pair** | **Distance (Å)** | **Lipid System** | **Temp ( °C)** | **Hybrid** |
|---|---|---|---|---|---|---|
| Smith et al[418] | Rotational Resonance | 1-$^{13}C$-L7/3-$^{13}C$-A11 | 4.6±0.2 | DMPC | -10 | 4.6±0.1 |
| | | 1-$^{13}C$-P21/3-$^{13}C$-A24 | 4.3±0.2 | | | 6.7±1.3 |
| | | 1-$^{13}C$-L42/2-$^{13}C$-C36 | 4.7±0.2 | | | 5.0±0.1 |
| | REDOR | 1-$^{13}C$-I18/$^{15}N_H$-Q22 | 4.7±0.2 | | | 8.3±0.9 |
| Liu et al[392] | REDOR | 5-$^{13}C$-Q22/$^{15}N_\epsilon$-Q22 | >5.5 | POPC/POPS | -10 | 16.2±2.8 |
| | | 5-$^{13}C$-Q26/$^{15}N_\epsilon$-Q26 | ∼5.5 | | | 5.5±1.3 |
| | | 5-$^{13}C$-Q29/$^{15}N_\epsilon$-Q29 | 4.1±0.2*** | | | 3.6±0.5 |
| Ahmed et al[419] | Rotational Resonance | 1-$^{13}C$-F32/2-$^{13}C$-A36* | 4.3±0.2** | DMPC | -50 | 4.8±0.1 |
| Middleton W et al[420] | REDOR | 1-$^{15}N$-L39/1-$^{13}C$-L37 | 3.2±0.2 | DMPC | -50 | 3.2±0.02 |
| Hughes and Middleton W et al[421] | REDOR | 2-$^{13}C$-A24/$^{15}N_H$-Q26 | >5.0 | DMPC | -40 | 4.2±0.1 |
| | Rotational Resonance | 1-$^{31}C$-P21/2-$^{13}C$-A24 | 4.65±0.25 | DMPC | -40 | 6.3±1.1 |
| Robia et al[362] | Fluorescence Energy Transfer | N-terminus/N-terminus | 46.1±2.1 | Insect Cells microsomes | RT | 45.4±3.8 |
| | | IA/IA | 37.9±3.8 | | | 33.3±3.4 |
| | | IB/IB | <24 | | | 11.5±0.9 |
| Traaseth et al[338] | DEER | Tempo-K3/Tempo-K3 | 47±10*** | DOPC/DOPE | -208 | 45.0±4.9 |

*In this sample, Cys36 was substituted with Alanine
**Rotational resonance condition n=1
***Included as restraints in the structure calculation

# Chapter 8

# Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy



## 8.1 Introduction

Posttranslational phosphorylation is among the most common mechanisms of cell signaling both in eukaryotes and prokaryotes[72]. Phosphorylation is orchestrated by kinases, which are involved in many vital cellular functions including metabolism, growth, cell differentiation, and target substrates localized in several compartments, including cytoplasm, mitochondria, plasma membrane, sarcoplasmic reticulum membrane, nucleus, microtubules, and actin filaments[73]. The human kinome (the collection of all human protein kinases) accounts for ∼2% of the entire genome and, therefore, encompasses the largest family of enzymes[74].

In addition to differential expression in various cellular sites, kinases are activated or deactivated and localized by cofactors and ancillary regulatory proteins to achieve precise control over time and space. Protein kinase A (PKA) is considered the prototype for the protein kinase family[75]. PKA exists as an inactive heterotetrameric assembly with two catalytic subunits (PKA-C) bound to a dimer of regulatory (R) subunits. PKA-C is unleashed upon $\beta$-adrenergic stimulation, which disassembles the heterotetramer. The R subunits are responsible for the localization of PKA-C via interactions with A Kinase Anchoring Proteins (AKAPs) to achieve spatial control[72].

The heat stable protein kinase inhibitor (PKI) also controls PKA activity and localization[72]. This small protein comprises a high affinity pseudosubstrate region that binds competitively to the substrate binding groove, as well as a nuclear export signal, directing PKA to locations outside the nucleus[76]. Binding of PKI to PKA-C forms a complex in which the enzyme is poised for catalysis (correct positioning of atoms), but remains locked due to the lack of a hydroxyl group acceptor[74].

The bean-shaped fold of PKA-C is highly conserved[75], with two lobes (small and large) undergoing structural rearrangements during the substrate recognition and product release steps of catalysis[74]. Three conformational states of PKA-C have been identified by x-ray crystallography along various stages of the catalytic cycle[74]: open (apo form), intermediate (binary form), and closed (ternary complex). NMR dynamic measurements[77] linked the major conformational states along the reaction coordinates, showing the critical role of intrinsic dynamics (the equilibrium fluctuations that allow the exploration of the free energy landscape) for protein kinase function. Specifically, the rates of conformational fluctuations are correlated with the transition of closed to open conformations. These fluctuations are synchronous with the enzyme rate-limiting step (product release), underscoring the prominent role of conformational dynamics in substrate recognition and catalysis[77].

Understanding how PKA-C interacts with both substrates and inhibitors from both a structural and dynamics perspective will define general criteria for activation and deactivation of protein kinases, with obvious repercussion in the design of new drugs.

Here, we examined the effects of inhibitors on the conformational dynamics of the enzyme: a competitive peptide inhibitor corresponding to the inhibitory region of PKI ($PKI_{5-24}$)[78] and magnesium, which under high concentrations behaves kinetically as a noncompetitive inhibitor of PKA-C[79, 80]. The combination of high $Mg^{2+}$ concentrations with PKI has implications for the cellular control of PKA-C activity to arrest transcription during mitosis[81, 82]. By comparing the results with those obtained in the presence of a peptide substrate[77] which competes with $PKI_{5-24}$ at the binding groove of PKA-C, we found that both $PKI_{5-24}$ and excess $Mg^{2+}$ restrict the enzyme dynamics on a fast (ps-ns) and slow ($\mu$s-ms) NMR timescale without drastically changing the conformation of the ternary complex. Inhibitor binding modifies the energy landscape by restricting the motions of the enzyme backbone[83]. These findings unveil a relatively unexplored role of magnesium in protein kinase regulation and establish a new paradigm for the design of protein kinase inhibitors.

## 8.2 Materials and methods

### 8.2.1 Protein expression and purification

PKA-C was expressed and purified from E. coli according to procedures previously published[87, 423]. The enzyme concentration was determined spectrophotmetrically by A280 ($A_{0.1\%} = 1.2$) and its activity was tested using the standard substrate Kemptide. Peptide synthesis of $PKI_{5-24}$ was performed using standard Fmoc chemistry on a CEM Liberty/Discover microwave synthesizer and purified using reverse-phase high pressure liquid chromatography.

### 8.2.2 ITC measurements

ITC data were acquired on a MCS-ITC microcalorimeter (MicroCal Inc., Northampton, MA). Stock solutions of PKA-C, $PLN_{1-20}$, and $PKI_{5-24}$ were dissolved in 20 mM phosphate buffer solution (pH 6.5) containing 180 mM KCl and 4 mM MgCl2, and degassed. Titrations were conducted at 27 °C using 0.1 mM PKA-C in the absence or presence of 6 mM AMP-PNP and with a stock of synthetic peptide (1.8 mM). The samples were stirred at 410 rpm. 20 injections were separated by 300 s of equilibration

(5 $\mu$L for the first, followed by 10 $\mu$L for each of the remaining). A one-site binding model was assumed and the data were fit using MicroCal Origin software (version 5.0).

### 8.2.3 Circular Dichroism

Concentrated stocks of PKA-C were diluted to 5 $\mu$M in CD buffer (10 mM PIPES, pH 7.0, 150 mM NaCl), in the presence of 0.48 (low $Mg^{2+}$, 1:1.2 $Mg^{2+}$:nucleotide) or 3.0 mM MgCl2 (high $Mg^{2+}$, 5:1 $Mg^{2+}$:nucleotide), 0.60 mM AMP-PNP (selected samples) and 15 $\mu$M peptide (selected samples). Samples were incubated over a range of 25 to 75 °C at 1 °C/min in a rectangular quartz cuvette in a Jasco J-815 spectropolarimeter. Spectra were acquired at 222 nm following an equilibration time of 10 s. A blank consisted of all reaction components except PKA-C and was subtracted from each spectrum. The data were fitted to a two-state sigmoidal unfolding model using Origin 8.0 (Microcal) using $T_m$ as the midpoint. Errors were derived from non-linear least squares fitting of the data.

### 8.2.4 Acquisition of NMR data

NMR samples consisted of ~500 $\mu$M PKA-C, 90 mM KCl, 20 mM KH2PO4, 10 mM MEGA-8, 20 mM DTT, 12 mM AMP-PNP, and 5% 2H2O (uncorrected pH=6.5). Low (1:1.2 $Mg^{2+}$:nucleotide) or high (5:1 $Mg^{2+}$:nucleotide) $Mg^{2+}$ conditions were done under 10 mM or 60 mM $Mg^{2+}$, respectively. Experiments were carried out on a Varian instrument operating at 800.29 MHz $^1H$ Larmor frequency with an inverse triple axis gradient cryoprobe at 33 °C. The data were processed with NMRPIPE[411] and analyzed with SPARKY[424]. Relaxation experiments were performed using the pulse sequences by Kay et al[425], with TROSY-detection[10] and a spectral width of 10,500 Hz (2,200 Hz) for the $^1H$ ($^{15}N$) dimension. $R_{1\rho}$ measurements used a 1,500 Hz spin-lock field strength centered at the $^{15}N$ carrier frequency. $R_{1\rho}$ values were converted into R2 as described[426].

$R_{ex}$ was measured using the method introduced by Palmer[427], with a Hahn echo period of $2/J_{NH}$[77]. The detection of $\alpha$, $\beta$, or longitudinal two-spin order magnetization ($I_{zz}$) during this Hahn echo period ($2/J_{NH}$=10.8 ms) was used to measure $R_{ex}$ with

the following relationship[428]:

$$R_{ex} = C_{zz} ln(\rho_{zz}) + C_{\beta} ln(\rho_{\beta})$$

where $Czz = (2\tau)^{-1}$, $C_{\beta} = (<\kappa> -1)(4\tau) - 1$, $\kappa = 1 - 2 * ln\rho_{zz}/ln\rho_{\beta}$, $\rho_{zz} = I_{zz}/I_{\alpha}$, and $\rho_{\beta} = I_{\beta}/I_{\alpha}$. Three 2-D experiments were recorded in triplicate, and in an interleaved manner to obtain $I_{zz}$, $I_{\alpha}$, and $I_{\beta}$. The value $<\kappa>$ was obtained from the trimmed mean of all $^{15}N$ resonances which did not exhibit chemical exchange.

Verification of $R_{ex}$ was done by measuring inverse peak heights ($I$) at decreasing temperatures[429] as described in the support information. Resonances which experienced conformational exchange diminished in peak intensity more significantly with decreasing temperatures[429] (Figure 8.7C-E).

### 8.2.5 MD simulations

MD simulations were setup using CHARMM c36a1 and run with NAMD using the crystal structures described in the Supplementary Methods. All structures were solvated in a TIP3 water box with $K^+$ and $Cl^-$ added as counter ions to reach an ionic strength of ∼150 mM. Following an initial equilibration, 75 ns MD simulations for each system in constant temperature and pressure (NPT) ensemble were performed. RMSD, RMSF and PCA for all simulations were performed as described in the Supplementary Methods.

## 8.3 Results

### 8.3.1 Thermodynamics of binding and enzyme stability

To compare the effects of binding substrates and inhibitors to PKA-C, we synthesized two peptides: the first peptide corresponding to the cytoplasmic domain of phospholamban ($PLN_{1-20}$), an endogenous inhibitor for the sarcoplasmic reticulum Ca-ATPase and native substrate of PKA-C in cardiac muscle[84]; and the second peptide corresponding to $PKI_{5-24}$ (Figure 8.5). Crystallographic data and enzyme kinetic assays indicate that these peptides compete for the same binding site in PKA-C[75, 77]. Enzyme thermostability and thermodynamics of binding experiments were carried out at low and high concentrations of MgCl2. At low concentrations of MgCl2, PKA-C binds

one cation, the primary $Mg^{2+}$ ion, essential for enzyme activity[79]. At high concentrations of MgCl2, PKA-C binds a secondary $Mg^{2+}$ ion, which inhibits catalytic activity in a noncompetitive manner[79].

To measure the dissociation constants for PKA-C bound to substrate or inhibitor peptides, we employed isothermal titration calorimetry (ITC). The ITC measurements reveal that $PKI_{5-24}$ binds to PKA-C 25 times more tightly than $PLN_{1-20}$ (Figure 8.1A and Table 8.1). This difference in affinity is amplified by ADP to 40 times, reflecting the strong cooperative effect between PKI and nucleotide[85, 86]. Remarkably, $PKI_{5-24}$ and $PLN_{1-20}$ have different thermodynamics of binding (Figure 8.1A). $PLN_{1-20}$ binding to PKA-C is dominated by a favorable overall entropy contribution to the binding free energy ($\Delta$G). The presence of ADP increases the favorable entropy of binding, resulting in an enhanced binding affinity, and indicating a positive cooperativity similar to that of Kemptide[87]. In contrast, the binding of $PKI_{5-24}$ to PKA-C is enthalpically driven, overcoming an unfavorable entropic contribution to $\Delta$G. Although the crystal structures of PKA-C complexes containing either $PLN_{1-20}$ or $PKI_{5-24}$ indicate more favorable enthalpy would be expected upon binding the inhibitor (more favorable intermolecular interactions), the structures alone are not adequate enough to predict these opposite thermodynamic driving forces to binding. The presence of ADP reduces the unfavorable entropy of binding $PKI_{5-24}$, leading to a higher binding affinity with a more pronounced positive cooperativity ($\sim$800 times).

Thermostability measurements revealed that $PKI_{5-24}$ conferred significantly greater stability to PKA-C than the substrate. Thermal melting of PKA-C bound to nucleotide, substrate, and inhibitor was monitored with circular dichroism (CD) spectroscopy (Figure 8.1B) at low and high $Mg^{2+}$ concentrations. Apo PKA-C melted at 47 °C, while under low $Mg^{2+}$ the non-hydrolyzable nucleotide, AMP-PNP, shifted the melting temperature ($\Delta T_m$) by $\sim 1$°C and the ternary complex with $PKI_{5-24}$ shifted $\Delta T_m$ by $\sim 2$°C, consistent with previous reports using ATP$\gamma$S[88, 89]. However, unlike the increased stability for $PKI_{5-24}$, addition of $PLN_{1-20}$ to the binary complex had a negligible effect on $\Delta T_m$. High $Mg^{2+}$ shifted $\Delta T_m$ by an additional $\sim 1$°C for both the nucleotide bound form and ternary complex with $PLN_{1-20}$. In contrast, $\Delta T_m$ was shifted the furthest in

the presence of $PKI_{5-24}$ and high $Mg^{2+}$ ($\Delta T_m \sim 4°C$), indicating a synergistic effect of stabilizing the enzyme when both inhibitors are present.

Taken together, the calorimetry and thermostability measurements indicate that $PKI_{5-24}$ binds to PKA-C with an opposite enthalpic/entropic balance than $PLN_{1-20}$. While substrate binding is entropically driven and confers little thermostability, inhibitor binding is enthalpically driven and significantly enhances the thermostability of the enzyme.

### 8.3.2 Chemical shift perturbations map the transitions from open to closed state

The residue-specific changes in the backbone of PKA-C upon binding AMP-PNP, and followed by substrate or inhibitor peptides at low magnesium were analyzed with amide chemical shifts measurements from $^1H/^{15}N$-TROSY-HSQC NMR spectroscopy. As demonstrated with the smaller seven-residue peptide Kemptide[87], the overall trend in chemical shift changes from these titrations correlate with the displacement of $C_\alpha$ atoms observed by x-ray crystallography (Figure 8.6 A-C). However, the chemical shift perturbations were generally small ($< \Delta\delta > \sim 0.04$ ppm), with AMP-PNP binding accounting for the majority of the differences (Figure 8.1C and Figure 8.6 D-G). Both $PLN_{1-20}$ and $PKI_{5-24}$ binding to the binary complex gave similar $\Delta\delta$ relative to the apo-state ($< \Delta\delta > \sim 0.04$ ppm, Figure 8.1 C,D). Binding of a secondary $Mg^{2+}$ ion to the PKA-C/AMP-PNP/$PKI_{5-24}$ complex resulted in no appreciable effects in the enzyme fingerprint as demonstrated by the correlation plot in Figure 8.1E. These data indicate that (a) once the nucleotide is bound to the enzyme, only minimal structural changes occur for the transitions from intermediate to closed state, and (b) the conformation of the enzyme bound to the peptide inhibitor closely resembles the substrate bound conformation. However, small but significant chemical shift differences ($\Delta\delta \sim 0.01$-$0.04$ ppm) are present in catalytically important regions of the enzyme (Figure 8.1F,G), such as the glycine-rich, activation, DFG, and peptide positioning loops, which are conserved throughout all protein kinases[74]. Strikingly, chemical shifts changes were linear from apo PKA-C to binary and ternary complexes with substrate or inhibitor peptide, and finally to the ternary complex with inhibitor peptide at high magnesium (Figure 8.1F).

This linearity between the different complexes indicates that the enzyme undergoes fast exchange between the major conformations and that the populations are shifted by ligand binding. Nucleotide binding shifts the conformational ensemble of open toward the intermediate state, and substrate or inhibitor peptide binding skews the population toward the closed state. While there is a clear distinction between the apo and nucleotide bound forms, the differences in chemical shifts between various ternary complexes are minimal (Figure 8.1F). The latter indicates that both substrate and inhibitors shift the population toward closed states, which are structurally quite similar.

### 8.3.3 Inhibitor binding quenches the dynamics in the enzyme backbone

The combination of thermodynamics and chemical shift perturbations show that the ternary complexes formed by the enzyme with substrate or inhibitor differ in thermostability, but not significantly in structure. To probe the internal dynamics of the enzyme, we used nuclear spin relaxation measurements for fast, ps-ns ($T_1$, $T_2$, and H-X NOE), and slow, $\mu$s-ms ($R_{ex}$), dynamics on the NMR time scale for each complex[90]. Our data show that the PKA-C/AMP-PNP/PKI$_{5-24}$ complex with two $Mg^{2+}$ ions bound is the most compact with the fastest overall tumbling rate ($T_1/T_2$ ratios), and least flexible (increased in H-X NOEs and longer $T_2$ values, Figure 8.2 and Figure 8.7A-B, and Table 8.2). This finding is in agreement with the previous fluorescence anisotropy studies[91, 92]. More importantly, a comparison of the ternary complexes with one $Mg^{2+}$ ion reveals that the PKA-C/AMP-PNP/PKI$_{5-24}$ complex has substantially decreased ps-ns dynamics with respect to the PKA-C/AMP-PNP/PLN$_{1-20}$ complex (Figure 8.2 and Table 8.2). This decrease becomes more apparent under high $Mg^{2+}$ concentrations. The quenched fast dynamics for the PKA-C/AMP-PNP/PKI$_{5-24}$ complex indicates a decrease in conformational entropy of the enzyme. Although this cannot be directly compared to the macroscopic methods of thermodynamics, a decrease in conformational entropy agrees qualitatively with the trend of unfavorable overall entropy of binding and the enhanced enzyme thermostability.

The analysis of inverse peak heights at temperatures ranging between 22-33 oC suggests a marked decrease in conformational exchange (i.e., slow dynamics on the NMR

time scale) for PKA-C upon inhibitor binding (Figure 8.7C-E). We quantified the $\mu$s-ms conformational exchange rates ($R_{ex}$) across the enzyme backbone and found diminished values for the ternary complex with $PKI_{5-24}$ compared to the ternary complex with substrate[77] (see Materials and Methods, Table 8.2 and Figure 8.2). Addition of the inhibitory $Mg^{2+}$ ion to the ternary complex with $PKI_{5-24}$ resulted in nearly absent $\mu$s-ms dynamics (Figure 8.2). This is particularly evident for regions surrounding the conserved loops at the active site of PKA-C.

### 8.3.4 PKA-C energy landscape

To define the energy landscape and interpret the dynamics for each form of PKA-C along stages of the catalytic cycle, we carried out full atomistic molecular dynamics (MD) simulations of PKA-C in water using the apo, binary (nucleotide bound), and ternary complexes (containing $PLN_{1-20}$ or $PKI_{5-24}$, see Supplementary Methods for experimental setup and data analysis). In each case, the simulations were carried out in the presence of either one or two bound $Mg^{2+}$ ions to model the low and high $Mg^{2+}$ conditions used experimentally[74]. Root mean square deviations (RMSD) showed that the simulations reached equilibration after ~20 ns (Figure 8.8A), while root mean square fluctuations (RMSF) confirmed that the most dynamic regions of the enzyme reside in the proximity of the catalytically important loops[74, 77] (glycine-rich, activation, magnesium, and peptide positioning loops), as well as some of the structural elements such as the B-helix, H-helix, F-helix, with RMSFs detected for the N- and C-termini (Figure 8.8B-C).

The conformational interconversion of the enzyme through different states identified by x-ray crystallography (Table 8.3) were monitored using principal component analysis (PCA)[93]. The PCA results show the directionality and amplitude of protein motions, in which the first several principal components are correlated with large conformational changes and have been used to interpret conformational variations observed experimentally[94, 95, 96]. This provides a reasonable method to extrapolate a relatively short MD trajectory to provide a qualitative description of motions occurring over a longer timescale[97], such as the opening and closing events probed by NMR spectroscopy[77].

We calculated the PCA for the PKA-C complex with $PLN_{1-20}$, and found the first two components account for ∼60% of variance in coordinates during the MD simulation (Figure 8.8D). The first principal component ($PC_1$) accounts for the opening and closing of the two lobes of the enzyme, while the second component ($PC_2$) describes shearing between the lobes (Figure 8.3A and Figure 8.8E-F). To probe the opening and closing of the active site cleft[75, 77], we monitored the interatomic distance between Ser53 and Gly186 (dS53-G186)[74] during MD trajectories (Figure 8.3A) with respect to $PC_1$. This 2D surface (Figure 8.3B) describes relative motion between the two lobes and the opening and closing of the active site. As points of reference, the crystallographic conformational states were also plotted on this surface to determine the degree of opening, while the RMSD from these structures for each simulation is provided in Table 8.4.

The apo enzyme (orange trace) samples a broad distribution of conformations identified by x-ray crystallography (Figure 8.3B), while those sampled by the binary form with one $Mg^{2+}$ have distributions similar to open and intermediate conformations. Addition of the substrate (ternary complex, blue trace) does not change $PC_1$ observed for the binary form, but changes the distribution of dS53-G186 towards the closed state. In contrast, the distribution of conformations for the ternary complex with $PKI_{5-24}$ (red trace) is shifted, where $PC_1$ encompasses conformations similar to intermediate and open states. However, the dS53-G186 distance distribution indicates that the enzymes active site is not completely open. This is in agreement with the NMR data showing residual dynamics for the PKA-C/AMP-PNP/$PKI_{5-24}$ complex at low $Mg^{2+}$ concentrations.

Addition of the second $Mg^{2+}$ ion resolves the degeneracy of the conformational states and defines localized minima (Figure 8.3B): the distribution describing the binary form occupies the intermediate conformational states, the ternary complex with substrate has dS53-G186 values clustered near the intermediate and closed states, and, strikingly, the ternary complex with inhibitor shows a quantitative agreement with the crystallographic structures (Figure 8.3B, bottom right panel).

The results from MD simulations corroborate the NMR and thermodynamics data, providing a framework to interpret the experimental results in terms of the energy landscape. Although the linchpin role of $Mg^{2+}$ ions in both stabilization of the pyrophosphate group of ATP and local rigidification of the glycine rich loop has been previously reported[98], our simulations reveal that the effect of $Mg^{2+}$ is a global phenomenon, affecting the overall dynamics of the enzyme and trapping it in inert states.

## 8.4 Discussion

The transitions between open and closed states in PKA-C are driven by internal dynamics, which plays a major role for substrate recognition and turnover[77]. The nucleotide acts as a dynamic and allosteric effector, causing a population shift of PKA-C from open to closed states and selecting for dynamically committed states- which are defined as nearly isoenergetic conformations compatible with catalysis that interconvert rapidly. The rates of the structural fluctuations between these states are sensitive to nucleotide binding, which increases the substrate binding cooperatively via a conformational selection mechanism[77]. More importantly, the rates of the enzyme conformational fluctuations are synchronous with the enzyme rate-limiting step (i.e. ADP release), underscoring the prominent role of conformational dynamics in substrate recognition and catalysis[77]. The novel aspect about the current study is the link between restricted conformational dynamics and enzyme inhibition, establishing a new paradigm for controlling protein kinase activity.

Based on our previous data[77, 87] and this work, we propose that the energy landscape of PKA-C comprises dynamically uncommitted, committed, and quenched states. Ligand binding (via nucleotide, substrates, or inhibitors) drives the enzyme to select these different states (Figure 8.4). The apo enzyme is characterized by a dynamically uncommitted state- it can explore the energy landscape to access open and closed conformations, but does not transition between these conformations on a timescale relevant for catalysis[77]. The dynamically committed states of the protein kinase occupy conformational space in which the active site cleft opens and closes at a timescale optimal for turnover. These fluctuations are induced by nucleotide binding, which acts as

a dynamic and allosteric activator[77], coupling the small and large lobes, completing the catalytic spine of the enzyme[99, 100], and preparing the active site for substrate binding. This differentiates protein kinases from small molecule kinases, such as adenylate kinase, since the assembly of the spine architecture is a prerequisite for protein kinase activation[99, 100]. Substrate bound PKA-C retains sufficient conformational motions, maintaining the enzyme in a dynamically committed state for product release and turnover[77]. Although the dynamics are on a slower time scale than the chemical step (phosphoryl transfer), they position and prepare the substrate for catalysis. In contrast, dynamically quenched states of inhibited PKA-C are characterized by narrow, deep wells with high energy barriers between conformations and hindered opening/closing of the active cleft. The two inhibitors analyzed here ($PKI_{5-24}$, and excess $Mg^{2+}$) raise the energy barriers of the conformational landscape, generating discrete minima with quenched dynamics. This phenomenon may explain why, in general, crystal structures of protein kinases have been more accessible in their inhibited forms while substrate bound forms have been elusive to such analyses.

More importantly, these results support a possible role of magnesium for regulation and localization of PKA-C in the cell. Reports indicate that $Mg^{2+}$ concentrations can change depending on the cell compartment[101] and phase of the cell cycle[81, 102]. Specifically, nuclear concentrations of $Mg^{2+}$ increase substantially during mitosis[81]. This triggers binding of PKA-C to PKI with high affinity[80], causing PKA-C to be exported from the nucleus, and arresting transcription[82]. Such a mechanism of inhibition and localization exploits the enhanced affinity of PKI in the presence of high $Mg^{2+}$[80]. Interestingly, enhanced binding affinity in the presence of excess $Mg^{2+}$ was also reported for the type-I (but not the type-II) regulatory subunit of PKA[80]. Since the regulatory subunits are localized differently within the cell where $Mg^{2+}$ concentrations may vary[103, 104], this would offer another form of localization and control. Therefore, $Mg^{2+}$ can potentially act as a rheostat for the strength of PKA regulation, dictate its compartmental localization, and influence the overall effect on biological function.

Allosteric enzymes often exist in dynamic ensembles of conformations, where catalytic efficiency is achieved by excited states able to cross inherently low energy barriers

between major conformations[105, 106, 107, 108, 109, 110, 111, 112, 113, 430]. Although the dynamics measured by NMR do not influence the chemical step of catalysis directly[114], it has been shown that intrinsic protein dynamics is the driving force for crossing potential energy barriers with ligands, such as co-factors or substrates, activating or deactivating enzyme dynamics to modulate biological function[105, 111, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124].

Although our findings are limited to the system studied here, the mechanism of inhibition via quenched dynamics may play a role in other systems. In fact, recent NMR studies on dihydrofolate reductase[125], triosephosphate isomerase[126], ribonuclease A[127], HIV-1 reverse transcriptase[128], and imidazole glycerol phosphate synthase[129] show that slow $\mu$s-ms motions are severely dampened when these enzymes are inhibited[125, 128, 129] or when catalytically hindering mutations are introduced[126, 127].

Our results demonstrate that the inhibition of PKA-C is marked by changes in intrinsic dynamics. These findings are timely, as they suggest a mode by which inhibitors could be designed to modulate PKA activity. Protein kinase inhibition is an important therapeutic avenue for treating diseases[431, 130, 432]. While many therapeutic agents targeting protein kinases are competitors for the nucleotide binding site, there is a growing interest in developing inhibitors that bind at remote sites (i.e., Abl kinase), working through an allosteric mechanism of inhibition[432, 433]. While we have previously described an allosteric network of dynamic communication in PKA-C[77, 87], our new results reveal a role of this network in its inhibition. Gaining control of PKA-C through competitive or allosteric modulation of dynamics is an exciting possibility for further research.

## 8.5 Supporting information

### 8.5.1 Methods

#### 8.5.1.1 Production of synthetic peptides

Peptides were synthesized using standard Fmoc chemistry on an CEM microwave synthesizer, starting with Fmoc-Glu(OtBu)-PEG-PS resin (0.4 g, 0.5 mmol/g). Side

chain protecting groups were 2,2,5,7,8-pentamethylchroman-6-sulfonyl (Pmc) for arginine, $N^{\omega}$-triphenylmethyl (Trt) for asparagine and glutamine, tert-butyl ester (OtBu) for glutamic acid, and tert-butyl ethers (tBu) for serine, threonine, and tyrosine. Deprotection of the resin-bound peptide was done using Reagent K (82.5% TFA, 5% phenol, 5% thioanisole, 2.5% 1,2-ethandiol, and 5% water) for 3 hours at 298 K. The resin mixture was washed three times (2 ml each) using the same cocktail and filtrate was collected. The peptide was precipitated overnight at 273 K in 80 ml of diethyl ether, then collected by centrifugation and washed 3 times with 30 ml of diethyl ether. The crude peptide was then purified by preparative HPLC using a Waters C18 reversed-phase cartridge (2.5 $\times$ 10 cm, 15 $\mu$m, 300 Å) with 0.1% TFA and CH3CN as eluents with detection at 220 nm. A linear gradient of 100:0 to 70:30 over 30 minutes at 5 ml/min was used. Purities of pooled fractions were >97% as determined by analytical HPLC using a Vydac C18 column (0.46 $\times$ 25 cm) and confirmed by ESI-TOF; for PKI4-25 a calculated 2222.4, m/z 2222.1 found. Stock aliquots were prepared in water and lyophilized. Concentrations of these solutions were verified by amino acid analysis.

#### 8.5.1.2 NMR spectroscopy

Mapping of slow conformational interconversion rates ($R_{ex}$) along the amide backbone was performed using methods established by Wang et al[427]. Briefly, a modified TROSY sequence allowed the detection of the $\alpha$, $\beta$, or the longitudinal two-spin order magnetization ($I_{zz}$) during a Hahn echo period set to $2/J_{NH}$ (10.8 ms). This allowed the measurement of the cross-correlated relaxation rates between dipole-dipole relaxation and chemical shift anisotropy ($\eta_{xy}$), the in-phase nitrogen amide transverse relaxation rate ($R_{2\alpha}$), and the longitudinal two spin order magnetization which can be used to measure $R_{ex}$ using the following relationship[428]:

$$R_{ex} = C_{zz} ln(\rho_{zz}) + C_{\beta} ln(\rho_{\beta})$$

where $Czz = (2\tau)^{-1}$, $C_{\beta} = (<\kappa> -1)(4\tau) - 1$, $\kappa = 1 - 2 * ln\rho_{zz}/ln\rho_{\beta}$, $\rho_{zz} = I_{zz}/I_{\alpha}$, and $\rho_{\beta} = I_{\beta}/I_{\alpha}$. Three 2-D experiments were recorded in triplicate, and in an interleaved manner to obtain $I_{zz}$, $I_{\alpha}$, and $I_{\beta}$. The value $<\kappa>$ was obtained from the mean of all amide resonances within one standard deviation (trimmed mean) and did not exhibit chemical exchange. Due to site to site variations in magnitude and orientation of the

chemical shift anisotropy tensor, fluctuations of $R_{ex}$ distributed around zero is typical for nonexchanging residues[427]. However, this method does not require the high pulsing rates inherent to CPMG-based sequences which have been shown to cause instabilities when used on cryogenic probes for samples containing high salt concentrations[7].

#### 8.5.1.3 Molecular dynamics simulations

Several different simulations were performed, including PKA apo form, binary form with ATP, ternary form with the inhibitor $PKI_{5-25}$, and ternary form with the substrate $PLN_{1-20}$. To study the effects of $Mg^{2+}$, PKA with one $Mg^{2+}$ bound was also simulated, including binary and ternary complexes. The starting structure for the apo form came from the crystal structure PDB:1CMK, from which a PKI peptide was removed. The binary complex and PKI-bound complex came from PDB:1ATP. The ternary complex with PLN was obtained from an NMR docking study, which was obtained from chemical shift perturbations. During the docking, PKA structure was taken from PDB:1ATP and limited flexibility was allowed for the backbone atoms. The backbone RMSD between 1CMK and 1ATP is around 0.3 Å while the PLN-bound PKA structure has an RMSD of 0.6 Å with respect to 1ATP. In total, nine different simulations were performed. Each system was solvated in a cubic box of $80 \times 80 \times 80$ $Å^3$. Counter ions ($K^+$ and $Cl^-$) were added to reach neutrality and an ionic strength of ∼150 mM. $Mg^{2+}$ ions were not added to the solvent since negligible free $Mg^{2+}$ would be present under low $Mg^{2+}$ conditions used for NMR analysis ($Mg^{2+}$:AMP-PNP was 1:1.2)

All simulations were set up using CHARMM c36a1[36] and production dynamics were performed with parallel and scalable program NAMD 2.7b1[349]. CHARMM27 force field (with CMAP correction) was used to describe all the molecular systems with the TIP3P water model. After initial minimization, the system was gradually heated from 10 K to 310 K every 30 K using 10 ps NPT simulations at each temperature. During minimization and heating, harmonic restraints were kept on non-water and non-counter ions and gradually decreased from 25 to 3 kcal/mol*$Å^2$. The harmonic restraints force constant (3 kcal/mol*$Å^2$) was applied to all heavy atoms of proteins during the initial 5 ns simulations in NAMD and removed afterwards. All simulations in NAMD were performed at 310 K employing Langevin dynamics with a damping constant of 1.0

$ps^{-1}$ and at 1 atm with Nosé-Hoover Langevin piston pressure control. Long-range electrostatic interactions were accounted by particle-mesh Ewald (PME) summation. A cutoff of 9 Å was used when evaluating non-bond interactions. Shake algorithms were applied to all bonds involving hydrogen atoms. The equations of motions were integrated using r-RESPA multiple step scheme with a time step of 2 fs.

Each system was simulated for 75 ns and harmonic restraints were applied to backbone atoms in the first 5 ns. We analyzed the root mean square deviation (RMSD) of individual components of the systems as a function of time (Figure 8.8A). The initial structure after heating up from each simulation was used as a reference. To study the fluctuations of the PKA along the sequence, we used root mean square fluctuations (RMSF), which is the time average of RMSD for each residue (with reference to the average structure) during the simulation (Equation 2). The last 40 ns trajectory (saved every 1 ps) was used for RMSF analysis.

$$RMSF(i) = \frac{1}{T}\sum_{t_j=1}^{T}\sqrt{(x_i(t_j) - x_{i,ave})^2}$$

where $T$ is the total number of structures used and $x_{i,ave}$ is the average coordinates for residue $i$ after superimposing all backbone atoms.

The RMSD reports the overall fluctuations compared to the starting structure. The RMSDs of each simulation are shown in Figure 8.8A. The backbone RMSD of PKA oscillated ∼2.2 Å for APO form and slightly decreased to below 2 Å in the binary and ternary forms. The inhibitor PKI bound ternary complex had a lower RMSD (∼ 1Å) compared with the substrate PLN bound (∼ 2Å) during the simulation. The RMSD for the ligand ATP is less than 1 Å in both the binary and ternary PKI forms and increased slightly in the ternary PLN case. The RMSD for ATP increases in all cases in the presence of one $Mg^{2+}$. The PLN backbone RMSD (∼ 3Å) is much smaller after $Mg^{2+}$ removal, indicating reduced fluctuations. The other RMSD values are similar.

The backbone fluctuations of each residue around its average position are characterized by its RMSF value, shown in Figure 8.8B,C. The RMSF values are independent of the reference structure used in the calculation. Using starting structure or the last

snapshot as a reference gives almost identical RMSF values. In Figure 8.8B,C, we compared the fluctuations of the apo form, the binary and ternary forms (both PKI and PLN bound) in presence of two $Mg^{2+}$ in the simulations. Within each simulation, the RMSF values were generally large for unstructured loops compared with $\alpha$-helices and $\beta$-sheets. Overall, the apo form of PKA has the largest fluctuations and the inhibitor (PKI) bound form of PKA is most stable (lowest RMSF values). ATP bound and PLN bound forms of PKA are in the middle with similar fluctuations. After removing the inhibitory $Mg^{2+}$, the flexibility of PKA backbone increased almost in all cases (Figure 8.8A). Apo form remains the most flexible, while binary, PKI-bound and PLN-bound have similar fluctuations due to increase in PKI-bound PKA. Although difference can be found between different binding partners, the relative order is dependent on the sequence positions.

In order to extract the large scale motions, principal component analysis (PCA) was applied to PKA backbone atoms (N, $C_\alpha$, C and O) from MD trajectory. It has been widely used to study conformational space sampling of proteins using MD simulations trajectories, X-ray structure or NMR structures[76, 77]. We chose MD simulations of PKA bound to ATP/$PLN_{1-20}$ in the presence of two $Mg^{2+}$ ions to perform PCA. Snapshots saved every 10 ps (7500 snapshots) were used to construct a covariance matrix C, as in Equation 3.

$$C_{ij} =< (x_i - < x_i >)(x_j - < x_j >) >$$

Before this, translational and rotational motions were removed by overlaying backbone atoms of two helical regions in the big lobe, including helix E (residue 140 to 160) and helix F (residue 217 to 233). In addition, the N-terminal and C-terminal residues were excluded in the PCA analysis and only backbone atoms from residue 50 to 300 were used.

By diagonalizing the covariance matrix $C$, a set of eigenvalues $P$ and the corresponding eigenvectors $V$ are obtained. $P_i$ corresponds to mean displacement along directions $v_i$. The first several principal components (PCs), ranked according to their eigenvalues, account for a large percentage of the fluctuations. The first two PCs are shown in Figure 8.8D, accounting for $\sim$60% of the overall fluctuations. $PC_1$ corresponds to the open

and close motion between the small lobe and big lobe. $PC_2$, instead, depicts a shearing motion of the small lobe orthogonal to the open-close motion.

The projection $Proj(M, PC_i)$ of any structure (trajectory) M onto the $i$th PC was calculated using equation 4.

$$Proj[M, Vi] = M_{bb} \cdot V_i$$

where Mbb is the backbone coordinates after overlaying M with the reference structure using the same two helix. We used the R and its Bio3d library for the PCA analysis[204].

Since we used the large lobe helices as a reference, the large lobe did not show large motions in each PCs. When all backbone atoms are used for PCA and the covariance prepared by superimposing all backbone atoms, the PCA results are much noisier. The first 5 PCs accounts for nearly equal amount of variance ($\sim$10%), with the first 10 PCs accounts for about 55% (Figure 8.8D). Although visual inspection clearly showed the existence of open-close between the small and big lobe, the N-terminus showed large fluctuations in the first several PCs. A similar choice of references for PCA analysis has also been reported in the literature[96].

Table 8.1: Thermodynamic parameters of PKA-C binding to the substrate, $PLN_{1-19}$, or the inhibitor, $PKI_{5-24}$

| **Complex** | **Δ H (kcal/mol)** | **Δ S (kcal/mol*T)** | **Δ -TΔ S (kcal/mol)** | $K_d(\mu M)$ |
|---|---|---|---|---|
| $PLN_{1-20}$ to apo PKA | -0.6±0.1 | 0.017± 0.001 | -5.1±0.2 | 49±8 |
| $PLN_{1-20}$ to ADP:PKA | -1.2±0.1 | 0.019± 0.001 | -5.7±0.1 | 10±4 |
| $PKI_{5-24}$ to apo PKA | -12.2±0.3 | -0.0145± 0.0003 | 4.3±0.1 | 1.8±0.4 |
| PKI5-20 to apo PKA | -12.3±0.5 | -0.0089± 0.0003 | 2.4±0.1 | 0.06±0.03 |

Table 8.2: Nuclear spin relaxation parameters on PKA-C in the free-state and bound to ligands

| **Parameter** | **Apo** | **AMP-PNP Bound (Low $Mg^{2+}$)** | **AMP-PNP/$PLN_{1-20}$ Bound (Low $Mg^{2+}$)** | **AMP-PNP/$PKI_{5-24}$ Bound (Low $Mg^{2+}$)** | **AMP-PNP/$PKI_{5-24}$ Bound (High $Mg^{2+}$)** |
|---|---|---|---|---|---|
| $T_1$ (s) | 1.55±0.18 | 1.47±0.19 | 1.35±0.38 | 1.35±0.48 | 1.19±0.21 |
| Trimmed $T_1$ (s) | 1.58±0.14 | 1.49±0.15 | 1.41±0.32 | 1.33±0.35 | 1.21±0.10 |
| $T_2$ (s) | 0.036±0.002 | 0.029±0.005 | 0.033±0.003 | 0.031±0.001 | 0.038±0.006 |
| Trimmed $T_2$ (s) | 0.019±0.001 | 0.020±0.003 | 0.025±0.002 | 0.024±0.009 | 0.029±0.007 |
| Trimmed $T_1/T_2$ (s) | 82.05±0.10 | 74.50±0.18 | 57.09±0.24 | 55.42±0.42 | 41.75±0.31 |
| $\tau_c$ (s) | 19.8±0.7 | 15.6±0.5 | 16.1±0.6 | 14.2±0.8 | - |
| H-X NOE | 0.71±0.32 | 0.74±0.28 | 0.70±0.32 | 0.75±0.28 | 0.79±0.23 |
| Trimmed H-X NOE | 0.76±0.13 | 0.78±0.13 | 0.74±0.16 | 0.79±0.13 | 0.81±0.10 |
| % of Residues with $R_{ex}$ >8 Hz | 5% | 16% | 13% | 11% | 5% |

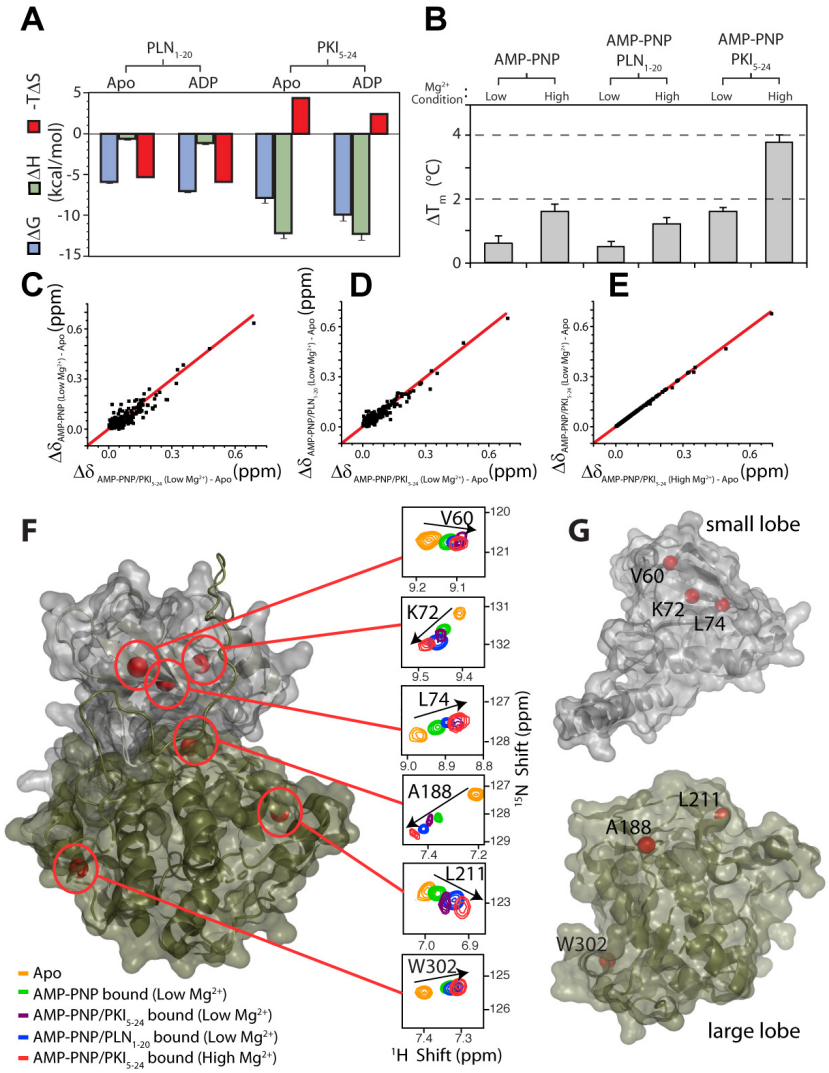


Figure 8.1: Thermodynamic and NMR analysis of PKA-C. (A) Thermodynamics of PKA-C binding to substrate and inhibitor, with or without ADP. The binding of $PLN_{1-20}$ is dominated by favorable entropy, whereas $PKI_{5-24}$ is enthalpy driven, overcoming an entropic penalty. (B) Melting measurements showed that $PKI_{5-24}$ confers the greatest thermostability to PKA-C as represented by the shift in $T_m$ (relative to apo-PKA-C, $\Delta T_m$). High $Mg^{2+}$ confers slightly higher stability to each complex. (C-E) Correlation of chemical shift perturbations ($\Delta\delta$) in PKA-C between the different forms. The majority of perturbation occurs upon nucleotide binding (C), whereas formation of the ternary complexes were quite similar to one another (D and E). (F) Linearity of chemical shifts between the apo form and the ternary form (AMP-PNP/$PKI_{5-24}$ with high $Mg^{2+}$) is observed, indicating that the enzyme opens and closes on a fast timescale. (G) Enzymes views from the active site surface formed by the small and large lobes.

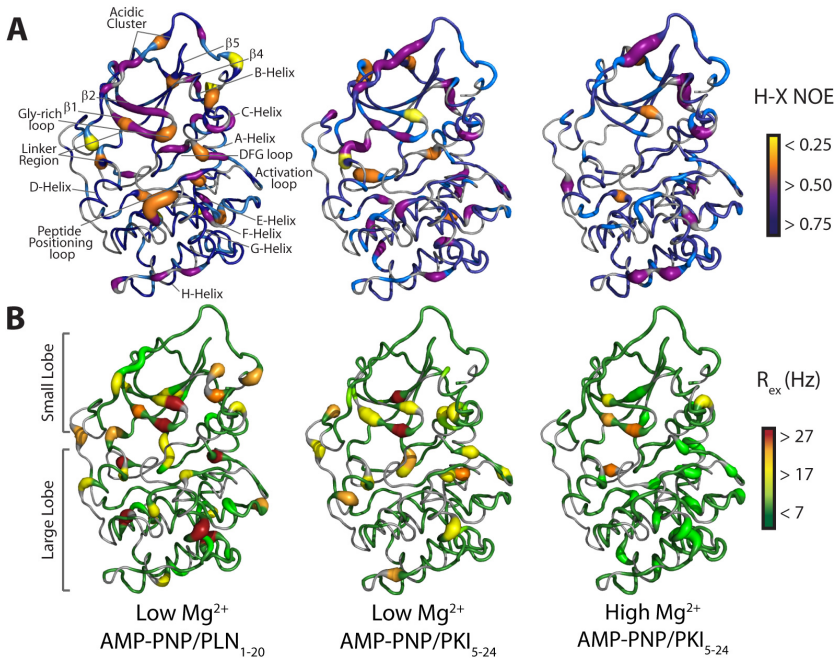


Figure 8.2: Backbone dynamics of PKA-C in different ternary complexes. Mapping of (A) fast and (B) slow backbone dynamics show that, upon inhibition with $PKI_{5-24}$ and with high $Mg^{2+}$, a decrease of picosecond to millisecond dynamics occurs throughout the backbone. For the comparison, the previously published dynamics of PKA-C with the substrate $PLN_{1-20}$ is shown[434].

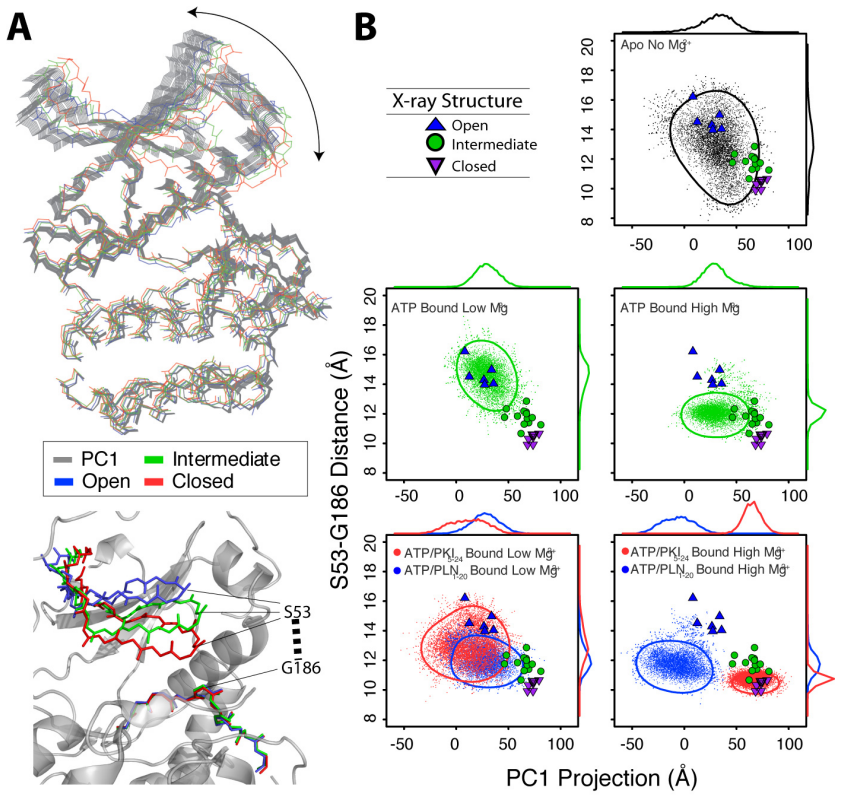


Figure 8.3: Comparison of MD simulations for PKA-C. (A) Global motions suggested by PCA analysis of MD trajectories correspond to opening and closing of the active site ($PC_1$), which compared well with the distances between residues S53 and G186 in crystal structures of open (1CMK), intermediate (1BX6), and closed (1ATP) conformations. (B) A map of the interatomic distances vs. the $PC_1$ from MD simulations indicate that PKA-C accessed the major crystallographic conformations frequently, except in the presence of inhibitors.

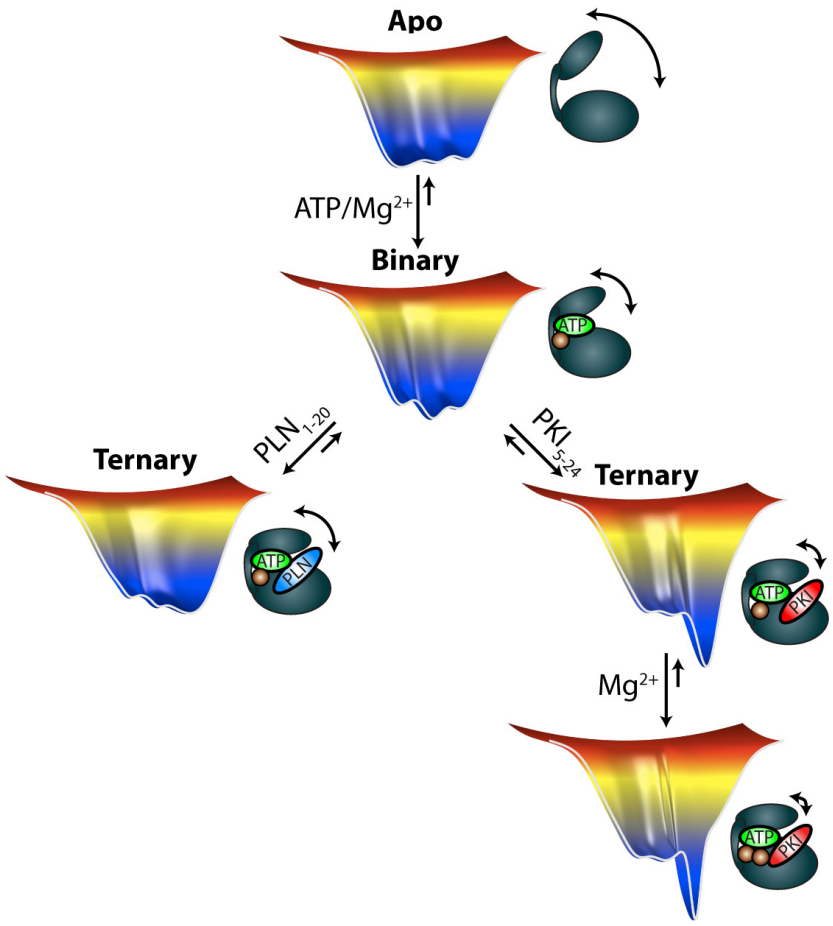


Figure 8.4: The energy landscape of PKA-C is modulated by ligand binding. The apo state is dynamically uncommitted, having dynamics which are not tuned to turnover. Nucleotide binding induces motions which are synchronized to turnover (dynamically committed) and are persistent in the ternary complex with substrate. However, $PKI_{5-24}$, or excess $Mg^{2+}$ and $PKI_{5-24}$, induces favorable enthalpy which lowers the energy of one or more conformational states and raises the energy barriers in the conformational landscape. This phenomenon hinders conformational fluctuations, inhibits turnover, and results in a dynamically quenched enzyme.

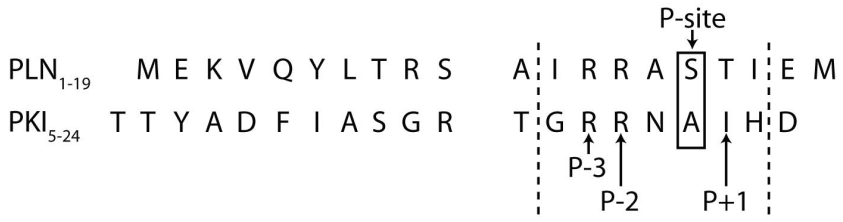


Figure 8.5: Sequences of $PLN_{1-20}$ and $PKI_{5-24}$. The seven residue canonical recognition sequence (X-X-R-R-X-S/T-$\Phi$, where X is any amino acid and $\Phi$ has hydrophobic character) is highlighted.

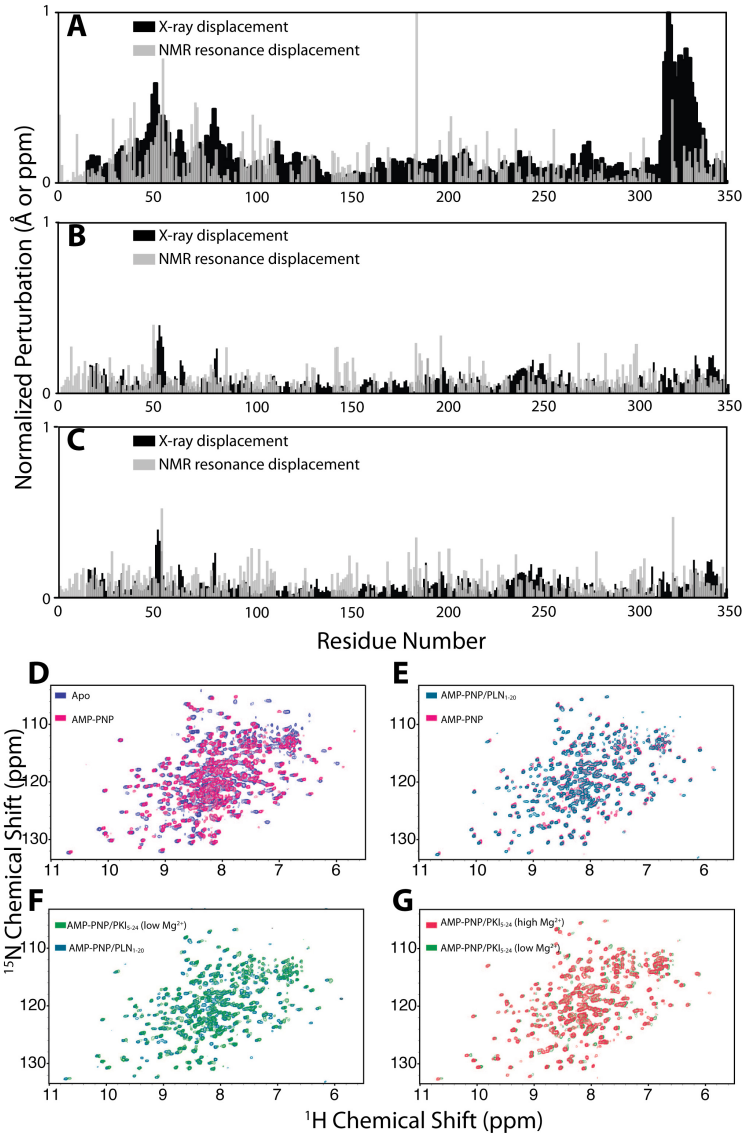


Figure 8.6: (A-C) Overlay of normalized displacements of $C_\alpha$ atoms (derived from X-ray crystallography) and NMR chemical shift perturbations. The overall agreement in the trends from both indicate that the NMR data, in part, probes the structural transitions observed by X-ray crystallography. The perturbations probed are (A) apo → nucleotide bound, (B) nucleotide bound→nucleotide/$PLN_{1-20}$ , and (C) nucleotide bound→nucleotide/$PKI_{5-24}$ bound. (D-G) Overlay of NMR spectra of the PKA-C in different forms. The overlays are done such that they can be used to compare the transition of apo→nucleotide bound→nucleotide/substrate (or inhibitor) bound.

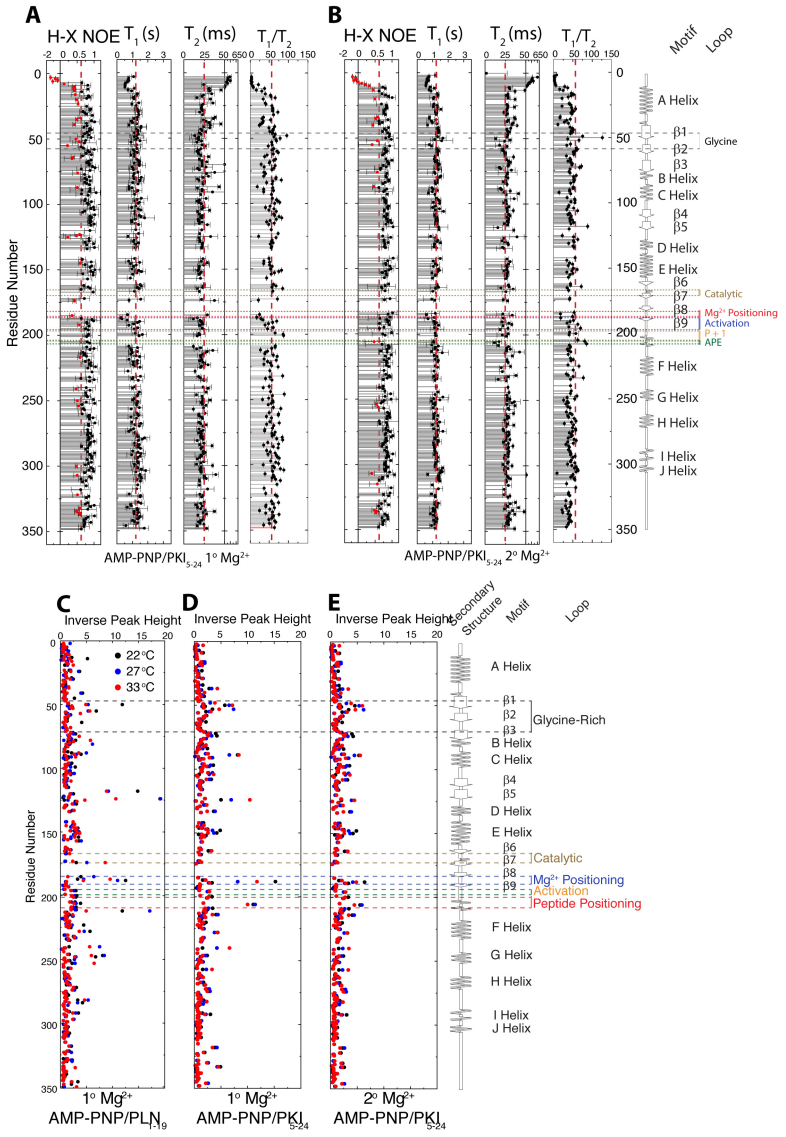


Figure 8.7: (A and B) NMR analysis of picosecond to nanosecond dynamics in the PKA-C as function of $Mg^{2+}$ ions. Heteronuclear NOE (H-X NOE) values show a systematic increase in the presence of high $Mg^{2+}$ (black data points), which is accompanied by longer $T_2$ times. The enzyme also becomes more compact and tumbles faster in solution, as reflected by the decrease $T_1/T_2$ values. (C-E) Inverse peak heights of PKA-C when bound to (C) AMP-PNP/PLN$_{1-19}$, (D) AMP-PNP/PKI$_{5-24}$ with low $Mg^{2+}$, and (E) AMP-PNP/PKI$_{5-24}$ with high $Mg^{2+}$. The trend in slow dynamics decreases from C to E.

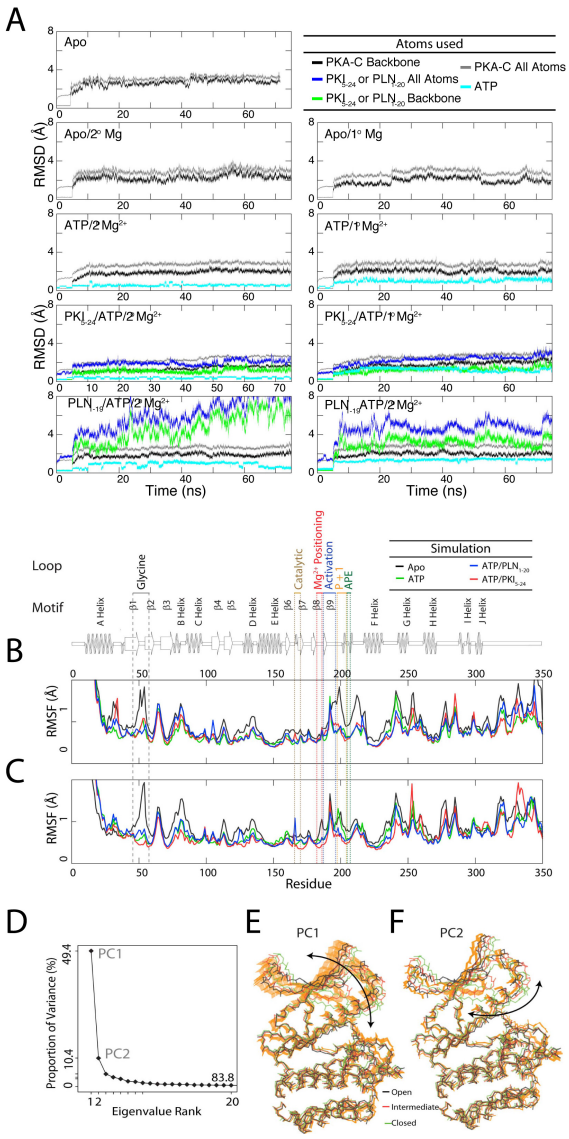


Figure 8.8: (A) Backbone and all-atom RMSD of PKA-C, ATP, PKI, and PLN during nine different simulations. In each simulation, the number of $Mg^{2+}$ ions during the simulation is specified. The backbone atoms include heavy atoms N, Ca, C, and O. For ATP, all atoms are used to calculate the RMSD values. (B and C) Rmsf of backbone atoms for each residue in simulations bound to different ligands with (B) 1 $Mg^{2+}$ or (C) 2 $Mg^{2+}$. The RMSF values for each residue are averaged over the last 40 ns of the trajectory. (D-F) PCA of the MD trajectories of PKA-C bound to ATP/$PLN_{1-20}$ in the presence of two $Mg^{2+}$ ions. (D) Plot describing the contribution to the overall motions based on each eigenvalue. (E and F) Domain motions described by the first two principle components correspond to opening and closing of the active site ($PC_1$) and shearing between the two lobes ($PC_2$) of the enzyme. The crystal structures corresponding to open (1CMK), intermediate (1BX6), and closed (1ATP) states are shown in black, red, and green, respectively.

Table 8.3: PDB accession numbers for the structures used to define the dS53-G186 conformational space.

| **State** | **PDB#** |
|---|---|
| Open | 1CMK |
| | 1CTP |
| | 1J3H |
| | 1SYK |
| | 2QVS |
| Closed | 1APM |
| | 1ATP |
| | 1CDK |
| | 1FMO |
| | 1JBP |
| | 1L3R |
| | 1RDQ |
| | 1U7E |
| | 2CPK |
| Intermediate | 1BKX |
| | 1BX6 |
| | 1JLU |
| | 1RE8 |
| | 1REJ |
| | 1REK |
| | 1STC |
| | 1YDR |
| | 1YDS |
| | 1YDT |
| | 3IDB |
| | 3IDC |
| | 3DND |
| | 3DNE |

Table 8.4: Average and standard deviation of backbone heavy atom RMSDs during the last 20 ns of MD simulations for each form of PKA-C. Representative crystal structures from each form of PKA-C, are used as reference state to measure these values. Values in bold indicate the forms with the best fit.

| **Complex form** | **open** | **Intermediate** | **Close** |
|---|---|---|---|
| Apo No $Mg^{2+}$ | **2.18±0.25** | 2.72±0.23 | 2.83±0.26 |
| ATP Bound Low $Mg^{2+}$ | 2.52±0.13 | **1.71±0.18** | 1.98±0.20 |
| ATP Bound High $Mg^{2+}$ | 2.61±0.11 | **1.87±0.11** | 2.03±0.14 |
| ATP/PLN Bound Low $Mg^{2+}$ | 2.26±0.12 | **1.81±0.14** | 1.97±0.17 |
| ATP/PLN Bound High $Mg^{2+}$ | 2.41±0.09 | **1.79±0.13** | 1.90±0.16 |
| ATP/PKI Bound Low $Mg^{2+}$ | 2.39±0.09 | **1.93±0.15** | 2.14±0.17 |
| ATP/PKI Bound High $Mg^{2+}$ | 3.13±0.12 | 1.74±0.10 | **1.61±0.08** |

# References


[1] E. Wallin and G. von Heijne. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*, 7(4):1029–1038, 1998.

[2] W. C. Wimley. The versatile beta-barrel membrane protein. *Current Opinion in Structural Biology*, 13(4):404 – 411, 2003.

[3] Y. Arinaminpathy, E. Khurana, D. M. Engelman, and M. B. Gerstein. Computational analysis of membrane proteins: the largest class of drug targets. *Drug Discovery Today*, 14(23-24):1130 – 1135, 2009.

[4] H. Belrhali, P. Nollert, A. Royant, C. Menzel, J. P. Rosenbusch, E. M. Landau, and E. Pebay-Peyroula. Protein, lipid and water organization in bacteriorhodopsin crystals: a molecular view of the purple membrane at 1.9 Å resolution. *Structure*, 7(8):99 – 917, 1999.

[5] C. Toyoshima, M. Nakasako, H. Nomura, and H. Ogawa. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature*, 405(6787):647–55, 2000.

[6] F. Fanelli, P. G. De Benedetti, F. Raimondi, and M. Seeber. Computational modeling of intramolecular and intermolecular communication in GPCRs. *Current Protein & Peptide Science*, 10(2):173–185, 2009.

[7] L. K. Tamm and B. Liang. NMR of membrane proteins in solution. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 48(4):201–201–210, 2006.

[8] S. N. Jonathan and D. M. Engelman. Introduction to the membrane protein reviews: The interplay of structure, dynamics, and environment in membrane protein function. *Annual Review of Biochemistry*, 75(1):707–712, 2006.

[9] J. Cavanagh, J. W. Fairbrother, A. G. Palmer III, and N. J. Skelton. *Protein NMR Spectroscopy, principle and practice*. Accademic Press, 1996.

[10] K. Pervushin, R. Riek, G. Wider, and K. Wuthrich. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences U. S. A.*, 94(23):12366–71, 1997.

[11] T. D. W. Claridge. *High Resolution NMR Techniques in Organic Chemistry*. Pergamon, New York, 1999.

[12] M. Karplus. Vicinal proton coupling in nuclear magnetic resonance. *Journal of the American Chemical Society*, 85(18):2870–2871, 1969.

[13] D. S. Wishart, B. D. Sykes, and F. M. Richards. Relationship between nuclear magnetic chemical shift and protein secondary structure. *Journal of Molecular Biology*, 222:311–333, 1991.

[14] G. Cornilescu, F. Delaglio, and A. Bax. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *Journal of Biomolecular NMR*, 13(3):289–302, 1999.

[15] G. Veglia, A. C. Zeri, C. Ma, and S. J. Opella. Deuterium/hydrogen exchange factors measured by solution nuclear magnetic resonance spectroscopy as indicators of the structure and topology of membrane proteins. *Biophysical Journal*, 82(4):2176–2183, 2002.

[16] M. Zweckstetter and A. Bax. Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. *Journal of the American Chemical Society*, 122(15):3791–3792, 2000.

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

[18] R. A. Laskowski, J. A. C. Rullman, M. W. MacArther, R. Kaptein, and J. M. Thornton. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, 8:477–486, 1998.

[19] S. J. Opella and F. M. Marassi. Structure determination of membrane proteins by NMR spectroscopy. *Chemical Reviews*, 104(8):3587–606, 2004.

[20] J. Torres, T. J. Stevens, and M. Samso. Membrane proteins: the 'wild west' of structural biology. *Trends in Biochemical Sciences*, 28(3):137 – 144, 2003.

[21] F. Marassi and S. J. Opella. NMR structural studies of membrane proteins. *Current Opinion in Structural Biology*, 8:640–648, 1998.

[22] J. K. Denny, J. Wang, T. A. Cross, and J. R. Quine. PISEMA powder patterns and PISA wheels. *Journal of Magnetic Resonance*, 152(2):217–226, 2001.

[23] C. H. Wu, A. Ramamoorthy, L. M. Geirasch, and S. J. Opella. Simultaneous characterization of the amide $^{1}$H chemical shift, $^{1}$H-$^{15}$N dipolar, and $^{15}$N chemical shift interaction tensor in a peptide bond by three-dimensional solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 117:6148–6149, 1995.

[24] J. Wang, S. Kim, F. Kovacs, and T. A. Cross. Structure of the transmembrane region of the M2 protein $H^+$ channel. *Protein Science*, 10(11):2241–2250, 2001.

[25] D. S. Thiriot, A. A. Nevzorov, L. Zagyanskiy, C. H. Wu, and S. J. Opella. Structure of the coat protein in Pf1 bacteriophage determined by solid-state NMR spectroscopy. *Journal of Molecular Biology*, 341(3):869–79, 2004.

[26] N. J. Traaseth, J. J. Buffy, J. Zamoon, and G. Veglia. Structural dynamics and topology of phospholamban in oriented lipid bilayers using multidimensional solid-state NMR. *Biochemistry*, 45(46):13827–13834, 2006.

[27] J. J. Buffy, N. J. Traaseth, A. Mascioni, P. L. Gor'kov, E. Y. Chekmenev, W. W. Brey, and G. Veglia. Two-dimensional solid-state NMR reveals two topologies of sarcolipin in oriented lipid bilayers. *Biochemistry*, 45(36):10939–10946, 2006.

[28] F. M. Marassi and S. J. Opella. A solid-state NMR index of helical membrane protein structure and topology. *Journal of Magnetic Resonance*, 144(1):150–5, 2000.

[29] J. Wang, J. Denny, C. Tian, S. Kim, Y. Mo, F. Kovacs, Z. Song, K. Nishimura, Z. Gan, R. Fu, J. R. Quine, and T. A. Cross. Imaging membrane protein helical wheels. *Journal of Magnetic Resonance*, 144(1):162–167, 2000.

[30] A. Mascioni and G. Veglia. Theoretical analysis of residual dipolar coupling patterns in regular secondary structures of proteins. *Journal of the American Chemical Society*, 125(41):12520–12526, 2003.

[31] A. A. Nevzorov and S. J. Opella. Structural fitting of PISEMA spectra of aligned proteins. *Journal of Magnetic Resonance*, 160(1):33–39, 2003.

[32] P. Guntert. Structure calculation of biological macromolecules from NMR data. *Quarterly Review of Biophysics*, 31(2):145–237, 1998.

[33] A. T. Brunger, G. M. Clore, A. M. Gronenborn, R. Saffrich, and M. Nilges. Assessing the quality of solution nuclear magnetic resonance structures by complete cross-validation. *Science*, 261(5119):328–331, 1993.

[34] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4(2):187–217, 1983.

[35] A. T. Brunger. *XPLOR Manual Version 3.1.* Yale University Press, New Haven, 1993.

[36] B. R. Brooks, C. L. Brooks 3rd, A. D. Mackerell Jr, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z.

Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry*, 30(10):1545–1614, 2009.

[37] J. Kuszewski, A. M. Gronenborn, and G. M. Clore. Improvements and extensions in the conformational database potential for the refinement of NMR and X-ray structures of proteins and nucleic acids. *Journal of Magnetic Resonance*, 125(1):171–177, 1997.

[38] J. A. Losonczi, M. Andrec, M. W. Fischer, and J. H. Prestegard. Order matrix analysis of residual dipolar coupling using singular value decomposition. *Journal of Magnetic Resonance*, 138:334–342, 1999.

[39] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, and G. M. Clore. The Xplor-NIH NMR molecular structure determination package. *Journal of Magnetic Resonance*, 160(1):65–73, 2003.

[40] J. Iwahara, C. D. Schwieters, and G. M. Clore. Ensemble approach for NMR structure refinement against $^{1}$H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *Journal of the American Chemical Society*, 126(18):5879–5896, 2004.

[41] G. M. Crippen and T. F. Havel. *Distance geometry and molecular conformation.* Research Studies Press, Taunton, UK, 1988.

[42] W. Braun and N. Go. Calculation of protein conformations by proton-proton distance constraints : A new efficient algorithm. *Journal of Molecular Biology*, 186(3):611 – 626, 1985.

[43] A. Jain, N. Vaidehi, and G. Rodriguez. A fast recursive algorithm for molecular dynamics simulation. *Journal of Computational Physics*, 106(2):258 – 268, 1993.

[44] C. D. Schwieters and G. M. Clore. Internal coordinates for molecular dynamics and minimization in structure determination and refinement. *Journal of Magnetic Resonance*, 152(2):288 – 302, 2001.

[45] H. Katz, T. Roderich Walter, and R. L. Somorjai. Rotational dynamics of large molecules. *Computers & Chemistry*, 3(1):25–32, 1979.

[46] A. K. Mazur, V. E. Dorofeev, and R. A. Abagyan. Derivation and testing of explicit equations of motion for polymers described by internal coordinates. *Journal of Computational Physics*, 92(2):261–272, 1991.

[47] R. A. Abagyan and A. K. Mazur. New methodology for computer-aided modelling of biomolecular structure and dynamics. 2. local deformations and cycles. *Journal of Biomolecular Structure Dynamics*, 6(4):833–845, 1989.

[48] P. Guntert, C. Mumenthaler, and K. Wuthrich. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *Journal of Molecular Biology*, 273(1):283–298, 1997.

[49] H. Abe, W. Braun, T. Noguti, and N. Go. Rapid calculation of first and second derivatives of conformational energy with respect to dihedral angles for proteins general recurrent equations. *Computers & Chemistry*, 8(4):239–247, 1984.

[50] H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690, 1984.

[51] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16):1668–1688, 2005.

[52] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen. Gromacs: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16):1701–1718, 2005.

[53] L. Verlet. Computer "experiments" on classical fluids. ii. equilibrium correlation functions. *Physical Review*, 165(1):201–214, Jan 1968.

[54] C. L. Brooks III, B. M. Pettitt, and M. Karplus. Structural and energetic effects of truncating long ranged interactions in ionic and polar fluids. *The Journal of Chemical Physics*, 83(11):5897–5908, 1985.

[55] P. J. Steinbach and B. R. Brooks. New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of Computational Chemistry*, 15:667–683, 1994.

[56] H. E. Alper, D. Bassolino, and T. R. Stouch. Computer simulation of a phospholipid monolayer-water system: The influence of long range forces on water structure and dynamics. *The Journal of Chemical Physics*, 98(12):9798–9807, 1993.

[57] T. Darden, L. Perera, L. Li, and L. Pedersen. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure*, 7(3):R55 – R60, 1999.

[58] R. Hockney and J. Eastwood. *Computer simulation using particles*. McGraw-Hill, New York, 1981.

[59] M. P. Allen and T. J. Tildesley. *Computer simulation of liquids*. Oxford University Press, 1999.

[60] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384–2393, 1980.

[61] S. Nose. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–519, 1984.

[62] S. Nose. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 100(1):191–198, 1984.

[63] W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3):1695–1697, Mar 1985.

[64] J. Gumbart, Y. Wang, A. Aksimentiev, E. Tajkhorshid, and K. Schulten. Molecular dynamics simulations of proteins in lipid bilayers. *Current Opinion in Structural Biology*, 15(4):423 – 431, 2005. Membranes/Engineering and design.

[65] M. A. Lomize, A. L. Lomize, I. D. Pogozheva, and H. I. Mosberg. OPM: Orientations of proteins in membranes database. *Bioinformatics*, 22(5):623–625, 2006.

[66] C. Kandt, W. L. Ash, and D. P. Tieleman. Setting up and running molecular dynamics simulations of membrane proteins. *Methods*, 41(4):475 – 488, 2007. Structural Biology of Membrane Proteins.

[67] F. Jahnig. What is the surface tension of a lipid bilayer membrane? *Biophysical Journal*, 71(3):1348–1349, 1996.

[68] S. E. Feller and R. W. Pastor. On simulating lipid bilayers with an applied surface tension: periodic boundary conditions and undulations. *Biophysical Journal*, 71(3):1350–1355, 1996.

[69] R. W. Benz, F. Castro-Roman, D. J. Tobias, and S. H. White. Experimental validation of molecular dynamics simulations of lipid bilayers: A new approach. *Biophysical Journal*, 88(2):805–817, 2005.

[70] J. B. Klauda nd R. M. Venable, J. A. Freites, J. W. O'Connor andD. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, and R. W. Pastor. Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *The Journal of Physical Chemistry B*, 114(23):7830–7843, 2010.

[71] N. J. Traaseth, K. N. Ha, R. Verardi, L. Shi, J. J. Buffy, L. R. Masterson, and G. Veglia. Structural and dynamic basis of phospholamban and sarcolipin inhibition of $Ca^{2+}$-ATPase. *Biochemistry*, 47(1):3–13, 2008.

[72] D. A. Walsh and S. M. Van Patten. Multiple pathway signal transduction by the cAMP-dependent protein kinase. *The FASEB Journal*, 8(15):1227–1236, 1994.

[73] J. B. Shabb. Physiological substrates of cAMP-dependent protein kinase. *Chemical Reviews*, 101(8):2381–2411, 2001.

[74] D. A. Johnson, P. Akamine, E. Radzio-Andzelm, M. Madhusudan, and S. S. Taylor. Dynamics of cAMP-dependent protein kinase. *Chemical Reviews*, 101(8):2243–2270, 2001.

[75] S. S. Taylor, J. Yang, J. Wu, N. M. Haste, E. Radzio-Andzelm, and G. Anand. PKA: a portrait of protein kinase dynamics. *Biochimica et Biophysica Acta*, 1697(1-2):259–269, 2004.

[76] G. D. Dalton and W. L. Dewey. Protein kinase inhibitor peptide (PKI): a family of endogenous neuropeptides that modulate neuronal cAMP-dependent protein kinase function. *Neuropeptides*, 40(1):23–34, 2006.

[77] L. R. Masterson, C. Cheng, T. Yu, M. Tonelli, A. P. Kornev, S. S. Taylor, and G. Veglia. Dynamics connect substrate recognition to catalysis in protein kinase A. *Nature Chemical Biology*, 6(11):821–828, 2010.

[78] D. B. Glass, H. C. Cheng, B. E. Kemp, and D. A. Walsh. Differential and common recognition of the catalytic sites of the cgmp-dependent and cAMP-dependent protein kinases by inhibitory peptides derived from the heat-stable inhibitor protein. *The Journal of Biological Chemistry*, 261(26):12166–12171, 1986.

[79] P. F. Cook, M. E. Neville, K. E. Vrana, F. T. Hartl, and R. Roskoshi. Adenosine cyclic 3',5'-monophosphate dependent protein kinase: Kinetic mechanism for the bovine skeletal muscle catalytic subunit. *Biochemistry*, 21:5794–5799, 1982.

[80] B. Zimmermann, S. Schweinsberg, S. Drewianka, and F. W. Herberg. Effect of metal ions on high-affinity binding of pseudosubstrate inhibitors to PKA. *The Biochemical Journal*, 413(1):93–101, 2008.

[81] R. Strick, P. L. Strissel, K. Gavrilov, and R. Levi-Setti. Cation-chromatin binding as shown by ion microscopy is essential for the structural integrity of chromosomes. *The Journal of Cell Biology*, 155(6):899–910, 2001.

[82] A. S. Stoykova, M. D. Dabeva, R. N. Dimova, and A. A. Hadjiolov. Ribosome biogenesis and nucleolar ultrastructure in neuronal and oligodendroglial rat brain cells. *Journal of Neurochemistry*, 45(6):1667–1676, 1985.

[83] R. V. Mauldin, M. J. Carroll, and A. L. Lee. Dynamic dysfunction in dihydrofolate reductase results from antifolate drug binding: modulation of dynamics within a structural state. *Structure*, 17(3):386–394, 2009.

[84] N. J. Traaseth and G. Veglia. Probing excited states and activation energy for the integral membrane protein phospholamban by NMR CPMG relaxation dispersion experiments. *Biochimica et Biophysica Acta*, 1798(2):77–81, 2010.

[85] S. Whitehouse and D. A. Walsh. Mg X ATP2-dependent interaction of the inhibitor protein of the cAMP-dependent protein kinase with the catalytic subunit. *The Journal of Biological Chemistry*, 258(6):3682–3692, 1983.

[86] J. Lew, N. Coruh, I. Tsigelny, S. Garrod, and S. S. Taylor. Synergistic binding of nucleotides and inhibitors to cAMP-dependent protein kinase examined by acrylodan fluorescence spectroscopy. *The Journal of Biological Chemistry*, 272(3):1507–1513, 1997.

[87] L. R. Masterson, A. Mascioni, N. J. Traaseth, S. S. Taylor, and G. Veglia. Allosteric cooperativity in protein kinase A. *Proceedings of the National Academy of Sciences U. S. A.*, 105(2):506–511, 2008.

[88] F. W. Herberg, M. L. Doyle, S. Cox, and S. S. Taylor. Dissection of the nucleotide and metal-phosphate binding sites in cAMP-dependent protein kinase. *Biochemistry*, 38(19):6352–6360, 1999.

[89] J. Yang, S. M. Garrod, M. S. Deal, G. S. Anand, V. L. Woods Jr, and S. Taylor. Allosteric network of cAMP-dependent protein kinase revealed by mutation of Tyr204 in the p+1 loop. *Journal of Molecular Biology*, 346(1):191–201, 2005.

[90] A. G. Palmer 3rd. NMR probes of molecular dynamics: overview and comparison with other techniques. *Annual Review of Biophysics and Biomolecular Structure*, 30:129–155, 2001.

[91] M. Gangal, S. Cox, J. Lew, T. Clifford, S. M. Garrod, M. Aschbaher, S. S. Taylor, and D. A. Johnson. Backbone flexibility of five sites on the catalytic subunit of cAMP-dependent protein kinase in the open and closed conformations. *Biochemistry*, 37(39):13728–13735, 1998.

[92] F. Li, C. Juliano, E. Gorfain, S. S. Taylor, and D. A. Johnson. Evidence for an internal entropy contributin to phosphoryl transfer: a study of domain clossure,

backbone flexibility, and the catalytic cycle of cAMP-dependent protein kinase. *Journal of Molecular Biology*, 315:459–469, 2002.

[93] M. Karplus and J. N. Kushick. Method for estimating the configurational entropy of macromolecules. *Macromolecules*, 14(2):325–332, 1981.

[94] O. Miyashita, J. N. Onuchic, and P. G. Wolynes. Nonlinear elasticity, protein-quakes, and the energy landscapes of functional transitions in proteins. *Proceedings of the National Academy of Sciences U. S. A.*, 100(22):12570–12575, 2003.

[95] P. Maragakis and M. Karplus. Large amplitude conformational change in proteins explored with a plastic network model: adenylate kinase. *Journal of Molecular Biology*, 352(4):807–822, 2005.

[96] K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus, and D. Kern. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature*, 450(7171):913–916, 2007.

[97] X. Cheng, I. Ivanov, H. Wang, S. M. Sine, and J. A. McCammon. Nanosecond-timescale conformational dynamics of the human alpha7 nicotinic acetylcholine receptor. *Biophysical Journal*, 93(8):2622–2634, 2007.

[98] I. V. Khavrutskii, B. Grant, S. S. Taylor, and J. A. McCammon. A transition path ensemble study reveals a linchpin role for $Mg^{2+}$ during rate-limiting ADP release from protein kinase A. *Biochemistry*, 48(48):11532–11545, 2009.

[99] A. P. Kornev, S. S. Taylor, and L. F. Ten Eyck. A helix scaffold for the assembly of active protein kinases. *Proceedings of the National Academy of Sciences U. S. A.*, 105(38):14377–14382, 2008.

[100] A. P. Kornev and S. S. Taylor. Defining the conserved internal architecture of a protein kinase. *Biochimica et Biophysica Acta*, 1804(3):440–444, 2009.

[101] M. Fatholahi, K. LaNoue, A. Romani, and A. Scarpa. Relationship between total and free cellular $Mg^{2+}$ during metabolic stimulation of rat cardiac myocytes and perfused hearts. *Archives of Biochemistry and Biophysics*, 374(2):395–401, 2000.

[102] E. Murphy. Mysteries of magnesium homeostasis. *Circulation Research*, 86(3):245–248, 2000.

[103] J. R. Mauban, M. O'Donnell, S. Warrier, S. Manni, and M. Bond. AKAP-scaffolding proteins and regulation of cardiac physiology. *Physiology*, 24:78–87, 2009.

[104] N. Alto, J. J. Carlisle Michel, K. L. Dodge, L. K. Langeberg, and J. D. Scott. Intracellular targeting of protein kinases and phosphatases. *Diabetes*, 51 Suppl 3:S385–8, 2002.

[105] S. Hammes-Schiffer and S. J. Benkovic. Relating protein motion to catalysis. *Annual Review of Biochemistry*, 75:519–541, 2006.

[106] D. D. Boehr, D. McElheny, H. J. Dyson, and P. E. Wright. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science*, 313(5793):1638–1642, 2006.

[107] K. K. Frederick, M. S. Marlow, K. G. Valentine, and A. J. Wand. Conformational entropy in molecular recognition by proteins. *Nature*, 448(7151):325–329, 2007.

[108] X. Yao, M. K. Rosen, and K. H. Gardner. Estimation of the available free energy in a LOV2-J alpha photoswitch. *Nature Chemical Biology*, 4(8):491–497, 2008.

[109] J. Gsponer, J. Christodoulou, A. Cavalli, J. M. Bui, B. Richter, C. M. Dobson, and M. Vendruscolo. A coupled equilibrium shift mechanism in calmodulin-mediated signal transduction. *Structure*, 16(5):736–746, 2008.

[110] T. Mittag, L. E. Kay, and J. D. Forman-Kay. Protein dynamics and conformational disorder in molecular recognition. *Journal of Molecular Recognition*, 23(2):105–116, 2010.

[111] P. E. Wright and H. J. Dyson. Linking folding and binding. *Current Opinion in Structural Biology*, 19(1):31–38, 2009.

[112] R. G. Smock and L. M. Gierasch. Sending signals dynamically. *Science*, 324(5924):198–203, 2009.

[113] S. R. Tzeng and C. G. Kalodimos. Dynamic activation of an allosteric regulatory protein. *Nature*, 462(7271):368–372, 2009.

[114] A. V. Pisliakov, J. Cao, S. C. Kamerlin, and A. Warshel. Enzyme millisecond conformational dynamics do not catalyze the chemical step. *Proceedings of the National Academy of Sciences U. S. A.*, 106(41):17359–17364, 2009.

[115] E. Freire. The propagation of binding interactions to remote sites in proteins: analysis of the binding of the monoclonal antibody D1.3 to lysozyme. *Proceedings of the National Academy of Sciences U. S. A.*, 96(18):10118–10122, 1999.

[116] S. Kumar, B. Ma, C. J. Tsai, N. Sinha, and R. Nussinov. Folding and binding cascades: dynamic landscapes and population shifts. *Protein Science*, 9(1):10–19, 2000.

[117] G. G. Hammes. Multiple conformational changes in enzyme catalysis. *Biochemistry*, 41(26):8221–8228, 2002.

[118] D. Zuiderweg Kern E. R. The role of dynamics in allosteric regulation. *Current Opinion in Structural Biology*, 13(6):748–57, 2003.

[119] M. A. Pufall, G. M. Lee, M. L. Nelson, H. S. Kang, A. Velyvis, L. E. Kay, L. P. McIntosh, and B. J. Graves. Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science*, 309(5731):142–5, 2005.

[120] J. F. Swain and L. M. Gierasch. The changing landscape of protein allostery. *Current Opinion in Structural Biology*, 16(1):102–108, 2006.

[121] N. Popovych, S. Sun, R. H. Ebright, and C. G. Kalodimos. Dynamically driven protein allostery. *Nature Structural and Molecular Biology*, 13(9):831–838, 2006.

[122] P. Li, I. R. Martins, G. K. Amarasinghe, and M. K. Rosen. Internal dynamics control activation and activity of the autoinhibited vav dh domain. *Nature Structural and Molecular Biology*, 15(6):613–618, 2008.

[123] D. D. Boehr, R. Nussinov, and P. E. Wright. The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chemical Biology*, 5(11):789–796, 2009.

[124] R. Das, S. Chowdhury, M. T. Mazhab-Jafari, S. Sildas, R. Selvaratnam, and G. Melacini. Dynamically driven ligand selectivity in cyclic nucleotide binding domains. *The Journal of Biological Chemistry*, 284(35):23682–23696, 2009.

[125] R. V. Mauldin and A. L. Lee. Nuclear magnetic resonance study of the role of M42 in the solution dynamics of Escherichia coli dihydrofolate reductase. *Biochemistry*, 49(8):1606–1615, 2010.

[126] R. B. Berlow, T. I. Igumenova, and J. P. Loria. Value of a hydrogen bond in triosephosphate isomerase loop motion. *Biochemistry*, 46(20):6001–6010, 2007.

[127] N. Doucet, E. D. Watt, and J. P. Loria. The flexibility of a distant loop modulates active site motion and product release in ribonuclease A. *Biochemistry*, 48(30):7160–7168, 2009.

[128] J. M. Seckler, M. D. Barkley, and P. L. Wintrode. Allosteric suppression of HIV-1 reverse transcriptase structural dynamics upon inhibitor binding. *Biophysical Journal*, 100(1):144–153, 2011.

[129] J. M. Lipchock and J. P. Loria. Nanometer propagation of millisecond motions in v-type allostery. *Structure*, 18(12):1596–1607, 2010.

[130] P. Cohen. Protein kinases-the major drug targets of the twenty-first century? *Nature Reviews.Drug Discovery*, 1(4):309–315, 2002.

[131] J. J. Buffy, B. A. Buck-Koehntop, F. Porcelli, N. J. Traaseth, D. D. Thomas, and G. Veglia. Defining the intramembrane binding mechanism of sarcolipin to calcium ATPase using solution NMR spectroscopy. *Journal of Molecular Biology*, 358(2):420–429, 2006.

[132] N. J. Traaseth, D. D. Thomas, and G. Veglia. Effects of Ser16 phosphorylation on the allosteric transitions of phospholamban/$Ca^{2+}$-ATPase complex. *Journal of Molecular Biology*, 358(4):1041–1050, 2006.

[133] L. Shi, A. Cembran, J. Gao, and G. Veglia. Tilt and azimuthal angles of a transmembrane peptide: a comparison between molecular dynamics calculations

and solid-state NMR data of sarcolipin in lipid membranes. *Biophysical Journal*, 96(9):3648–3662, 2009.

[134] S. H. White and G. von Heijne. The machinery of membrane protein assembly. *Current Opinion in Structural Biology*, 14(4):397–404, 2004.

[135] J. Zamoon, F. Nitu, C. Karim, D. D. Thomas, and G. Veglia. Mapping the interaction surface of a membrane protein: unveiling the conformational switch of phospholamban in calcium pump regulation. *Proceedings of the National Academy of Sciences U. S. A.*, 102(13):4747–4752, 2005.

[136] K. N. Ha, N. J. Traaseth, R. Verardi, J. Zamoon, A. Cembran, C. B. Karim, D. D. Thomas, and G. Veglia. Controlling the inhibition of the sarcoplasmic $Ca^{2+}$-ATPase by tuning phospholamban structural dynamics. *The Journal of Biological Chemistry*, 282(51):37205–37214, 2007.

[137] R. C. Page, S. Kim, and T. A. Cross. Transmembrane helix uniformity examined by spectral mapping of torsion angles. *Structure*, 16(5):787–797, 2008.

[138] M. Baldus. Magnetic resonance in the solid state: applications to protein folding, amyloid fibrils and membrane proteins. *European Biophysics Journal*, 36 Suppl 1:S37–48, 2007.

[139] M. Hong. Oligomeric structure, dynamics, and orientation of membrane proteins from solid-state NMR. *Structure*, 14(12):1731–1740, 2006.

[140] X. M. Gong, C. M. Franzin, K. Thai, J. Yu, and F. M. Marassi. Nuclear magnetic resonance structural studies of membrane proteins in micelles and bilayers. *Methods in Molecular Biology*, 400:515–529, 2007.

[141] U. H. Durr, L. Waskell, and A. Ramamoorthy. The cytochromes P450 and b5 and their reductases-promising targets for structural studies by advanced solid-state NMR spectroscopy. *Biochimica et Biophysica Acta*, 1768(12):3235–3259, 2007.

[142] R. C. Page, C. Li, J. Hu, F. P. Gao, and T. A. Cross. Lipid bilayers: an essential environment for the understanding of membrane proteins. *Magnetic Resonance in Chemistry*, 45(S1):S2–S11, 2007.

[143] A. Ramamoorthy, S. Thennarasu, D. K. Lee, A. Tan, and L. Maloy. Solid-state NMR investigation of the membrane-disrupting mechanism of antimicrobial peptides MSI-78 and MSI-594 derived from magainin 2 and melittin. *Biophysical Journal*, 91(1):206–216, 2006.

[144] S. H. Park, A. A. Mrse, A. A. Nevzorov, M. F. Mesleh, M. Oblatt-Montal, M. Montal, and S. J. Opella. Three-dimensional structure of the channel-forming transmembrane domain of virus protein u (Vpu) from HIV-1. *Journal of Molecular Biology*, 333(2):409–424, 2003.

[145] P. Teriete, C. M. Franzin, J. Choi, and F. M. Marassi. Structure of the Na,K-ATPase regulatory protein FXYD1 in micelles. *Biochemistry*, 46(23):6774–6783, 2007.

[146] N. J. Traaseth, R. Verardi, K. D. Torgersen, C. B. Karim, D. D. Thomas, and G. Veglia. Spectroscopic validation of the pentameric structure of phospholamban. *Proceedings of the National Academy of Sciences U. S. A.*, 104(37):14676–14681, 2007.

[147] J. Hu, T. Asbury, S. Achuthan, C. Li, R. Bertram, J. R. Quine, R. Fu, and T. A. Cross. Backbone structure of the amantadine-blocked trans-membrane domain M2 proton channel from Influenza A virus. *Biophysical Journal*, 92(12):4335–4343, 2007.

[148] A. A. De Angelis, S. C. Howell, A. A. Nevzorov, and S. J. Opella. Structure determination of a membrane protein with two trans-membrane helices in aligned phospholipid bicelles by solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 128(37):12256–12267, 2006.

[149] S. H. Park, S. Prytulla, A. A. De Angelis, J. M. Brown, H. Kiefer, and S. J. Opella. High-resolution NMR spectroscopy of a GPCR in aligned bicelles. *Journal of the American Chemical Society*, 128(23):7402–7403, 2006.

[150] P. C. van der Wel, E. Strandberg, J. A. Killian, and R. E. Koeppe 2nd. Geometry and intrinsic tilt of a tryptophan-anchored transmembrane alpha-helix determined by $^2$H NMR. *Biophysical Journal*, 83(3):1479–88, 2002.

[151] E. Strandberg, P. Wadhwani, P. Tremouilhac, U. H. Durr, and A. S. Ulrich. Solid-state NMR analysis of the pgla peptide orientation in DMPC bilayers: structural fidelity of $^{2}$H-labels versus high sensitivity of $^{19}$F-NMR. *Biophysical Journal*, 90(5):1676–1686, 2006.

[152] R. W. Glaser, C. Sachse, U. H. Durr, P. Wadhwani, S. Afonin, E. Strandberg, and A. S. Ulrich. Concentration-dependent realignment of the antimicrobial peptide pgla in lipid membranes observed by solid-state $^{19}$F-NMR. *Biophysical Journal*, 88(5):3392–3397, 2005.

[153] M. F. Mesleh, S. Lee, G. Veglia, D. S. Thiriot, F. M. Marassi, and S. J. Opella. Dipolar waves map the structure and topology of helices in membrane proteins. *Journal of the American Chemical Society*, 125(29):8928–8935, 2003.

[154] M. F. Mesleh, G. Veglia, T. M. DeSilva, F. M. Marassi, and S. J. Opella. Dipolar waves as NMR maps of protein structure. *Journal of the American Chemical Society*, 124(16):4206–4207, 2002.

[155] A. Mascioni, B. L. Eggimann, and G. Veglia. Determination of helical membrane protein topology using residual dipolar couplings and exhaustive search algorithm: application to phospholamban. *Chemistry and Physics of Lipids*, 132(1):133–144, 2004.

[156] J. Wang, J. D. Walsh, J. Kuszewski, and Y. X. Wang. Periodicity, planarity, and pixel (3P): a program using the intrinsic residual dipolar coupling periodicity-to-peptide plane correlation and phi/psi angles to derive protein backbone structures. *Journal of Magnetic Resonance*, 189(1):90–103, 2007.

[157] G. Bouvignies, P. R. Markwick, and M. Blackledge. Characterization of protein dynamics from residual dipolar couplings using the three dimensional gaussian axial fluctuation model. *Proteins*, 71(1):353–363, 2008.

[158] X. Salvatella, B. Richter, and M. Vendruscolo. Influence of the fluctuations of the alignment tensor on the analysis of the structure and dynamics of proteins using residual dipolar couplings. *Journal of Biomolecular NMR*, 40(1):71–81, 2008.

[159] P. R. Markwick, G. Bouvignies, and M. Blackledge. Exploring multiple timescale motions in protein GB3 using accelerated molecular dynamics and NMR spectroscopy. *Journal of the American Chemical Society*, 129(15):4724–4730, 2007.

[160] G. Bouvignies, P. Markwick, R. Bruschweiler, and M. Blackledge. Simultaneous determination of protein backbone structure and dynamics from residual dipolar couplings. *Journal of the American Chemical Society*, 128(47):15100–15101, 2006.

[161] N. A. Lakomek, T. Carlomagno, S. Becker, C. Griesinger, and J. Meiler. A thorough dynamic interpretation of residual dipolar couplings in ubiquitin. *Journal of Biomolecular NMR*, 34(2):101–115, 2006.

[162] Y. Ryabov and D. Fushman. Analysis of interdomain dynamics in a two-domain protein using residual dipolar couplings together with $^{15}$N relaxation data. *Magnetic Resonance in Chemistry*, 44 Spec No:S143–51, 2006.

[163] J. R. Tolman and K. Ruan. NMR residual dipolar couplings as probes of biomolecular dynamics. *Chemical Reviews*, 106(5):1720–1736, 2006.

[164] M. Blackledge. Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Progress in NMR Spectroscopy*, 46:23–61, 2005.

[165] A. Bax. Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Science*, 12(1):1–16, 2003.

[166] J. R. Tolman, H. M. Al-Hashimi, L. E. Kay, and J. H. Prestegard. Structural and dynamic analysis of residual dipolar coupling data for proteins. *Journal of the American Chemical Society*, 123(7):1416–1424, 2001.

[167] M. W. F. Fischer, J. A. Losonczi, J. L. Weaver, and J. H. Prestegard. Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. *Biochemistry*, 38(28):9013–9022, 1999.

[168] J. R. Quine, S. Achuthan, T. Asbury, R. Bertram, M. S. Chapman, J. Hu, and T. A. Cross. Intensity and mosaic spread analysis from PISEMA tensors in solid-state NMR. *Journal of Magnetic Resonance*, 179(2):190–198, 2006.

[169] S. K. Straus, W. R. Scott, and A. Watts. Assessing the effects of time and spatial averaging in $^{15}$N chemical shift/$^{15}$N-$^{1}$H dipolar correlation solid state NMR experiments. *Journal of Biomolecular NMR*, 26(4):283–295, 2003.

[170] D. W. Bleile, W. R. Scott, and S. K. Straus. Can PISEMA experiments be used to extract structural parameters for mobile beta-barrels? *Journal of Biomolecular NMR*, 32(2):101–111, 2005.

[171] S. Ozdirekcan, C. Etchebest, J. A. Killian, and P. F. Fuchs. On the orientation of a designed transmembrane peptide: toward the right tilt angle? *Journal of the American Chemical Society*, 129(49):15174–15181, 2007.

[172] S. Esteban-Martin and J. Salgado. The dynamic orientation of membrane-bound peptides: bridging simulations and experiments. *Biophysical Journal*, 93(12):4278–4288, 2007.

[173] V. V. Vostrikov, C. V. Grant, A. E. Daily, S. J. Opella, and R. E. Koeppe 2nd. Comparison of polarization inversion with spin exchange at magic angle and geometric analysis of labeled alanines methods for transmembrane helix alignment. *Journal of the American Chemical Society*, 130(38):12584–12585, 2008.

[174] B. Buck, J. Zamoon, T. L. Kirby, T. M. DeSilva, C. Karim, D. Thomas, and G. Veglia. Overexpression, purification, and characterization of recombinant Ca-ATPase regulators for high-resolution solution and solid-state NMR studies. *Protein Expression and Purification*, 30(2):253–61, 2003.

[175] A. Mascioni, C. Karim, G. Barany, D. D. Thomas, and G. Veglia. Structure and orientation of sarcolipin in lipid environments. *Biochemistry*, 41(2):475–82, 2002.

[176] D. H. Mac Lennan and E. G. Kranias. Phospholamban: a crucial regulator of cardiac contractility. *Nature Reviews*, 4:566–577, 2003.

[177] M. Periasamy and A. Kalyanasundaram. SERCA pump isoforms: Their role in calcium transport and disease. *Muscle & Nerve*, 2007.

[178] A. Ramamoorthy, Y. Wei, and L. Dong-Kuk. PISEMA solid-state NMR spectroscopy. *Annual Reports on NMR Spectroscopy*, 52:1–52, 2004.

[179] M. A. Lomize, A. L. Lomize, I. D. Pogozheva, and H. I. Mosberg. OPM: orientations of proteins in membranes database. *Bioinformatics*, 22(5):623–625, 2006.

[180] W. Im, M. Feig, and C. L. Brooks 3rd. An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophysical Journal*, 85(5):2900–2918, 2003.

[181] S. Tristram-Nagle, H. I. Petrache, and J. F. Nagle. Structure and interactions of fully hydrated dioleoylphosphatidylcholine bilayers. *Biophysical Journal*, 75(2):917–925, 1998.

[182] S. C. Costigan, P. J. Booth, and R. H. Templer. Estimations of lipid bilayer geometry in fluid lamellar phases. *Biochimica et Biophysica Acta - Biomembranes*, 1468(1-2):41 – 54, 2000.

[183] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry B*, 102(18):3586–3616, 1998.

[184] A. D. Mackerell Jr. Empirical force fields for biological macromolecules: overview and issues. *Journal of Computational Chemistry*, 25(13):1584–1604, 2004.

[185] A. D. Mackerell Jr, M. Feig, and C. L. Brooks 3rd. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry*, 25(11):1400–1415, 2004.

[186] T. B. Woolf and B. Roux. Structure, energetics, and dynamics of lipid-protein interactions: A molecular dynamics study of the gramicidin A channel in a DMPC bilayer. *Proteins*, 24(1):92–114, 1996.

[187] T. B. Woolf and B. Roux. Molecular dynamics simulation of the gramicidin channel in a phospholipid bilayer. *Proceedings of the National Academy of Sciences U. S. A.*, 91(24):11631–11635, 1994.

[188] B. A. Lewis and D. M. Engelman. Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. *Journal of Molecular Biology*, 166:211–217, 1983.

[189] B. A. Lewis and D. M. Engelman. Bacteriorhodopsin remains dispersed in fluid phospholipid bilayers over a wide range of bilayer thicknesses. *Journal of Molecular Biology*, 166(2):203–210, 1983.

[190] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, 1983.

[191] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23:327–341, 1977.

[192] P. Renner, C. G. Caratsch, P. G. Waser, and P. Lazarovici andN. Primor. Presynaptic effects of the pardaxins, polypeptides isolated from the gland secretion of the flatfish pardachirus marmoratus. *Neuroscience*, 23:319–325, 1987.

[193] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh Ewald method. *The Journal of Chemical Physics*, 103(19):8577–8593, 1995.

[194] C. Sagui and T. A. Darden. Molecular dynamics simulations of biomolecules: long-range electrostatic effects. *Annual Review of Biophysics and Biomolecular Structure*, 28:155–179, 1999.

[195] J. Dufourcq, J. F. Faucon, G. Fourche, J. L. Dasseux, M. Le Marie, and T. Gulik-Krzywicki. Morphological changes of phosphatidylcholine bilayers induced by melittin: Vesicularization, fusion, discoidal particles. *Biochimica et Biophysica Acta*, 859(1):33–48, 1986.

[196] A. Buckling and M. Brockhurst. Microbiology: Ramp resistance. *Nature*, 438(7065):170–171, 2005.

[197] K. Matsuzaki. Why and how are peptide-lipid interactions utilized for self defence? *Biochemical Society Transactions*, 29(Pt 4):598–601, 2001.

[198] T. Hara, H. Kodama, M. Kondo, K. Wakamatsu, A. Takeda, T. Tachi, and K. Matsuzaki. Effects of peptide dimerization on pore formation: Antiparallel disulfide-dimerized magainin 2 analogue. *Biopolymers*, 58(4):437–446, 2001.

[199] G. Lipari and A. Szabo. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. theory and range of validity. *Journal of the American Chemical Society*, 104:4546–4559, 1982.

[200] R. Cole and J. P. Loria. Fast-modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data. *Journal of Biomolecular NMR*, 26(3):203–213, 2003.

[201] A. M. Mandel, M. Akke, and A. G. Palmer. Backbone dynamics of Escherichia coli ribonuclease HI: correlations with structure and function in an active enzyme. *Journal of Molecular Biology*, 246:144–163, 1995.

[202] A. E. Garcia. Large-amplitude nonlinera motions in proteins. *Physical Review Letters*, 68:2696–2699, 1992.

[203] J. B. Clarage, T. Romo, B. K. Andrews, B. M. Pettitt, and G. N. Phillips Jr. A sampling problem in molecular dynamics simulations of macromolecules. *Proceedings of the National Academy of Sciences U. S. A.*, 92(8):3288–3292, 1995.

[204] B. J. Grant, A. P. Rodrigues, K. M. ElSawy, J. A. McCammon, and L. S. Caves. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*, 22(21):2695–2696, 2006.

[205] M. C. Wiener and S. H. White. Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of x-ray and neutron diffraction data. ii. distribution and packing of terminal methyl groups. *Biophysical Journal*, 61(2):428–433, 1992.

[206] W. C. Wimley and S. H. White. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Structural Biology*, 3(10):842–848, 1996.

[207] A. O. Gramolini, M. G. Trivieri, G. Y. Oudit, T. Kislinger, W. Li, M. M. Patel, A. Emili, E. G. Kranias, P. H. Backx, and D. H. Maclennan. Cardiac-specific overexpression of sarcolipin in phospholamban null mice impairs myocyte function that is restored by phosphorylation. *Proceedings of the National Academy of Sciences U. S. A.*, 103(7):2446–2451, 2006.

[208] A. C. Johansson and E. Lindahl. Amino-acid solvation structure in transmembrane helices from molecular dynamics simulations. *Biophysical Journal*, 91(12):4450–4463, 2006.

[209] S. H. White, W. C. Wimley, A. S. Ladokhin, and K. Hristova. Protein folding in membranes: determining energetics of peptide-bilayer interactions. *Methods in Enzymology*, 295:62–87, 1998.

[210] S. H. White and W. C. Wimley. Membrane protein folding and stability : Physical principles. *Annual Review of Biophysics Biomolecular Structures*, 28:319–365, 1999.

[211] F. N. Petersen, M. O. Jensen, and C. H. Nielsen. Interfacial tryptophan residues: a role for the cation-$\pi$ effect? *Biophysical Journal*, 89(6):3985–3996, 2005.

[212] A. O. Gramolini, T. Kislinger, M. Asahi, W. Li, A. Emili, and D. H. MacLennan. Sarcolipin retention in the endoplasmic reticulum depends on its C-terminal RSYQY sequence and its interaction with sarco(endo)plasmic $Ca^{2+}$-ATPases. *Proceedings of the National Academy of Sciences U. S. A.*, 101(48):16807–16812, 2004.

[213] J. Aqvist. A simple way to calculate the axis of an alpha-helix. *Computers & Chemistry*, 10(2):97–99, 1986.

[214] T. K. Rostovtseva, H. I. Petrache, N. Kazemi, E. Hassanzadeh, and S. M. Bezrukov. Interfacial polar interactions affect gramicidin channel kinetics. *Biophysical Journal*, 94(4):L23–5, 2008.

[215] S. H. Park and S. J. Opella. Tilt angle of a trans-membrane helix is determined by hydrophobic mismatch. *Journal of Molecular Biology*, 350(2):310–318, 2005.

[216] U. Harzer and B. Bechinger. Alignment of lysine-anchored membrane peptides under conditions of hydrophobic mismatch: A CD, $^{15}$N and $^{31}$P solid state NMR spectroscopy investigation. *Biochemistry*, 39:13106–13144, 2000.

[217] E. E. Metcalfe, J. Zamoon, D. D. Thomas, and G. Veglia. $^{1}$H/$^{15}$N heteronuclear NMR spectroscopy shows four dynamic domains for phospholamban reconstituted in dodecylphosphocholine micelles. *Biophysical Journal*, 87(2):1205–1214, 2004.

[218] P. J. Bond and M. S. Sansom. Membrane protein dynamics versus environment: simulations of ompa in a micelle and in a bilayer. *Journal of Molecular Biology*, 329(5):1035–1053, 2003.

[219] T. J. Tobias, K. C. Tu, and M. Klein. Atomic-scale molecular dynamics simulations of lipid membranes. *Current Opinion in Colloid & Interface Science*, 2:15–26, 1997.

[220] S. Berneche, M. Nina, and B. Roux. Molecular dynamics simulation of melittin in a dimyristoylphosphatidylcholine bilayer membrane. *Biophysical Journal*, 75(4):1603–1618, 1998.

[221] T. W. Allen, O. S. Andersen, and B. Roux. Structure of gramicidin a in a lipid bilayer environment determined using molecular dynamics simulations and solid-state NMR data. *Journal of the American Chemical Society*, 125(32):9868–9877, 2003.

[222] R. Ketchem, B. Roux, and T. Cross. High-resolution polypeptide structure in a lamellar phase lipid environment from solid state NMR derived orientational constraints. *Structure*, 5(12):1655–1669, 1997.

[223] T. B. Woolf and B. Roux. The binding site of sodium in the gramicidin A channel: comparison of molecular dynamics with solid-state NMR data. *Biophysical Journal*, 72(5):1930–1945, 1997.

[224] R. R. Ketchem, W. Hu, and T. A. Cross. High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR. *Science*, 261(5127):1457–1460, 1993.

[225] A. Ramamoorthy, S. K. Kandasamy, D. K. Lee, S. Kidambi, and R. G. Larson. Structure, topology, and tilt of cell-signaling peptides containing nuclear localization sequences in membrane bilayers determined by solid-state NMR and molecular dynamics simulation studies. *Biochemistry*, 46(4):965–975, 2007.

[226] U. Sternberg, R. Witter, and A. S. Ulrich. All-atom molecular dynamics simulations using orientational constraints from anisotropic NMR samples. *Journal of Biomolecular NMR*, 38(1):23–39, 2007.

[227] S. Hellstern, S. Pegoraro, C. B. Karim, A. Lustig, D. D. Thomas, L. Moroder, and J. Engel. Sarcolipin, the shorter homologue of phospholamban, forms oligomeric structures in detergent micelles and in liposomes. *Journal of Biological Chemistry*, 276(33):30845–30852, 2001.

[228] L. Becucci, R. Guidelli, C. B. Karim, D. D. Thomas, and G. Veglia. An electrochemical investigation of sarcolipin reconstituted into a mercury-supported lipid bilayer. *Biophysical Journal*, 93(8):2678–2687, 2007.

[229] M. B. Ulmschneider, J. P. Ulmschneider, M. S. Sansom, and A. Di Nola. A generalized born implicit-membrane representation compared to experimental insertion free energies. *Biophysical Journal*, 92(7):2338–2349, 2007.

[230] J. P. Ulmschneider, M. B. Ulmschneider, and A. Di Nola. Monte Carlo folding of trans-membrane helical peptides in an implicit generalized Born membrane. *Proteins*, 69(2):297–308, 2007.

[231] W. Im and C. L. Brooks 3rd. De novo folding of membrane proteins: an exploration of the structure and NMR properties of the fd coat protein. *Journal of Molecular Biology*, 337(3):513–519, 2004.

[232] F. M. Marassi. A simple approach to membrane protein secondary structure and topology based on NMR spectroscopy. *Biophysical Journal*, 80(2):994–1003, 2001.

[233] F. M. Marassi and S. J. Opella. Using pisa pies to resolve ambiguities in angular constraints from PISEMA spectra of aligned proteins. *Journal of Biomolecular NMR*, 23(3):239–242, 2002.

[234] S. Kim and T. A. Cross. Uniformity, ideality, and hydrogen bonds in transmembrane alpha-helices. *Biophysical Journal*, 83(4):2084–2095, 2002.

[235] L. Shi, N. J. Traaseth, R. Verardi, M. Gustavsson, J. Gao, and G. Veglia. Paramagnetic-based NMR restraints lift residual dipolar coupling degeneracy in multidomain detergent-solubilized membrane proteins. *Journal of the American Chemical Society*, 2011.

[236] N. Tjandra and A. Bax. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science*, 278(5340):1111–1114, 1997.

[237] A. Bax, G. Kontaxis, and N. Tjandra. Dipolar coupling in macromolecular structure determination. *Methods in Enzymology*, 339:127–174, 2001.

[238] J. H. Prestegard. New techniques in structural NMR-anisotropic interactions. *Nature Structural Biology*, 5:517–522, 1998.

[239] J. H. Prestegard, H. M. Al-Hashimi, and J. R. Tolman. NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Quarterly Review of Biophysics*, 33(4):371–424, 2000.

[240] N. Tjandra, J. Marquardt, and M. G. Clore. Direct refinement against proton-proton dipolar coupling in NMR structure of macromolecules. *Journal of Magnetic Resonance*, 142:393–396, 2000.

[241] M. G. Clore, M. A. Gronenborn, and N. Tjandra. Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *Journal of Magnetic Resonance*, 131:159–162, 1998.

[242] G. Nodet, L. Salmon, V. Ozenne, S. Meier, M. R. Jensen, and M. Blackledge. Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings. *Journal of the American Chemical Society*, 131(49):17908–17918, 2009.

[243] M. R. Jensen and M. Blackledge. On the origin of NMR dipolar waves in transient helical elements of partially folded proteins. *Journal of the American Chemical Society*, 130(34):11266–11267, 2008.

[244] S. Meier, M. Blackledge, and S. Grzesiek. Conformational distributions of unfolded polypeptides from novel NMR techniques. *The Journal of Chemical Physics*, 128(5):052204, 2008.

[245] A. K. Jha, A. Colubri, K. F. Freed, and T. R. Sosnick. Statistical coil model of the unfolded state: resolving the reconciliation problem. *Proceedings of the National Academy of Sciences U. S. A.*, 102(37):13099–13104, 2005.

[246] J. A. Marsh and J. D. Forman-Kay. Structure and disorder in an unfolded state under nondenaturing conditions from ensemble models consistent with a large number of experimental restraints. *Journal of Molecular Biology*, 391(2):359–374, 2009.

[247] P. Vallurupalli, D. F. Hansen, E. Stollar, E. Meirovitch, and L. E. Kay. Measurement of bond vector orientations in invisible excited states of proteins. *Proceedings of the National Academy of Sciences U. S. A.*, 104(47):18473–18477, 2007.

[248] A. J. Baldwin, D. F. Hansen, P. Vallurupalli, and L. E. Kay. Measurement of methyl axis orientations in invisible, excited states of proteins by relaxation dispersion NMR spectroscopy. *Journal of the American Chemical Society*, 131(33):11939–11948, 2009.

[249] H. M. Al-Hashimi, P. J. Bolon, and J. H. Prestegard. Molecular symmetry as an aid to geometry determination in ligand protein complexes. *Journal of Magnetic Resonance*, 142(1):153–158, 2000.

[250] G. A. Mueller, W. Y. Choy, D. Yang, J. D. Forman-Kay, R. A. Venters, and L. E. Kay. Global folds of proteins with low densities of NOEs using residual dipolar

couplings: application to the 370-residue maltodextrin-binding protein. *Journal of Molecular Biology*, 300(1):197–212, 2000.

[251] J. C. Hus, L. Salmon, G. Bouvignies, J. Lotze, M. Blackledge, and R. Bruschweiler. 16-fold degeneracy of peptide plane orientations from residual dipolar couplings: Analytical treatment and implications for protein structure determination. *Journal of the American Chemical Society*, 130(47):15927–15937, 2008.

[252] M. F. Mesleh and S. J. Opella. Dipolar waves as NMR maps of helices in proteins. *Journal of Magnetic Resonance*, 163(2):288–299, 2003.

[253] J. D. Walsh and Y. X. Wang. Periodicity, planarity, residual dipolar coupling, and structures. *Journal of Magnetic Resonance*, 174(1):152–162, 2005.

[254] H. M. Al-Hashimi, H. Valafar, M. Terrell, E. R. Zartler, M. K. Eidsness, and J. H. Prestegard. Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. *Journal of Magnetic Resonance*, 143(2):402–406, 2000.

[255] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard. Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proceedings of the National Academy of Sciences U. S. A.*, 92(20):9279–9283, 1995.

[256] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard. NMR evidence for slow collective motions in cyanometmyoglobin. *Nature Structural Biology*, 4(4):292–7, 1997.

[257] I. Bertini, Y. K. Gupta, C. Luchinat, G. Parigi, M. Peana, L. Sgheri, and J. Yuan. Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. *Journal of the American Chemical Society*, 129(42):12786–12794, 2007.

[258] E. B. Bertelsen, L. Chang, J. E. Gestwicki, and E. R. Zuiderweg. Solution conformation of wild-type e. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proceedings of the National Academy of Sciences U. S. A.*, 106(21):8471–8476, 2009.

[259] G. von Heijne. Membrane-protein topology. *Nature Reviews.Molecular Cell Biology*, 7(12):909–918, 2006.

[260] Y. Zhou, T. Cierpicki, R. H. Jimenez, S. M. Lukasik, J. F. Ellena, D. S. Cafiso, H. Kadokura, J. Beckwith, and J. H. Bushweller. NMR solution structure of the integral membrane enzyme DsbB: functional insights into DsbB-catalyzed disulfide bond formation. *Molecular Cell*, 31(6):896–908, 2008.

[261] N. J. Traaseth, R. Verardi, and G. Veglia. Asymmetric methyl group labeling as a probe of membrane protein homo-oligomers by NMR spectroscopy. *Journal of the American Chemical Society*, 130(8):2400–2401, 2008.

[262] V. Tugarinov and L. E. Kay. An isotope labeling strategy for methyl TROSY spectroscopy. *Journal of Biomolecular NMR*, 28(2):165–72, 2004.

[263] V. Tugarinov and L. E. Kay. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase g using a new labeling strategy and novel NMR methods. *Journal of the American Chemical Society*, 125(45):13868–78, 2003.

[264] J. D. Gross, V. M. Gelev, and G. Wagner. A sensitive and robust method for obtaining intermolecular NOEs between side chains in large protein complexes. *Journal of Biomolecular NMR*, 25:235–242, 2003.

[265] R. Tycko, F. J. Blanco, and Y. Ishii. Alignment of biopolymers in strained gels: A new way to create detectable dipole-dipole couplings in high-resolution biomolecular nmr. *Journal of the American Chemical Society*, 122(38):9340–9341, 2000.

[266] S. M. Douglas, J. J. Chou, and W. M. Shih. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proceedings of the National Academy of Sciences U. S. A.*, 104(16):6644–6648, 2007.

[267] J. Lorieau, L. Yao, and A. Bax. Liquid crystalline phase of G-tetrad DNA for NMR study of detergent-solubilized proteins. *Journal of the American Chemical Society*, 130(24):7536–7537, 2008.

[268] G. Veglia and S. J. Opella. Lanthanide ion binding to adventitious sites aligns membrane proteins in micelles for solution NMR spectroscopy. *Journal of the American Chemical Society*, 122(47):11733–11734, 2000.

[269] C. Ma and S. J. Opella. Lanthanide ions bind specifically to an added ef-hand and orient a membrane protein in micelles for solution NMR spectroscopy. *Journal of Magnetic Resonance*, 146(2):381–384, 2000.

[270] J. Wohnert, K. J. Franz, M. Nitz, B. Imperiali, and H. Schwalbe. Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. *Journal of the American Chemical Society*, 125(44):13338–13339, 2003.

[271] W. D. Van Horn, H. J. Kim, C. D. Ellis, A. Hadziselimovic, E. S. Sulistijo, M. D. Karra, C. Tian, F. D. Sonnichsen, and C. R. Sanders. Solution nuclear magnetic resonance structure of membrane-integral diacylglycerol kinase. *Science*, 324(5935):1726–1729, 2009.

[272] T. Cierpicki, B. Liang, L. K. Tamm, and J. H. Bushweller. Increasing the accuracy of solution NMR structures of membrane proteins by application of residual dipolar couplings. high-resolution structure of outer membrane protein a. *Journal of the American Chemical Society*, 128(21):6947–6951, 2006.

[273] T. L. Kirby, C. B. Karim, and D. D. Thomas. Electron paramagnetic resonance reveals a large-scale conformational change in the cytoplasmic domain of phospholamban upon binding to the sarcoplasmic reticulum Ca-ATPase. *Biochemistry*, 43(19):5842–52, 2004.

[274] J. J. Chou, J. D. Kaufman, S. J. Stahl, P. T. Wingfield, and A. Bax. Micelle-induced curvature in a water-insoluble HIV-1 Env peptide revealed by NMR dipolar coupling measurement in stretched polyacrylamide gel. *Journal of the American Chemical Society*, 124(11):2450–2451, 2002.

[275] P. Permi and A. Annila. Transverse relaxation optimised spin-state selective NMR experiments for measurement of residual dipolar couplings. *Journal of Biomolecular NMR*, 16(3):221–227, 2000.

[276] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR*, 6(3):277–293, 1995.

[277] B. A. Johnson. Using NMRview to visualize and analyze the NMR spectra of macromolecules. *Methods in Molecular Biology*, 278:313–352, 2004.

[278] J. Zamoon, A. Mascioni, D. D. Thomas, and G. Veglia. NMR solution structure and topological orientation of monomeric phospholamban in dodecylphosphocholine micelles. *Biophysical Journal*, 85(4):2589–2598, 2003.

[279] N. J. Traaseth, L. Shi, R. Verardi, D. G. Mullen, G. Barany, and G. Veglia. Structure and topology of monomeric phospholamban in lipid membranes determined by a hybrid solution and solid-state NMR approach. *Proceedings of the National Academy of Sciences U. S. A.*, 106(25):10165–10170, 2009.

[280] G. M. Clore and J. Kuszewski. Improving the accuracy of NMR structures of rna by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. *Journal of the American Chemical Society*, 125(6):1518–1525, 2003.

[281] Y. Kimura, K. Kurzydlowski, M. Tada, and D. H. MacLennan. Phospholamban inhibitory function is activated by depolymerization. *Journal of Biological Chemistry*, 272(24):15061–4, 1997.

[282] L. G. Reddy, L. R. Jones, and D. D. Thomas. Depolymerization of phospholamban in the presence of calcium pump: a fluorescence energy transfer study. *Biochemistry*, 38(13):3954–62, 1999.

[283] C. B. Karim, C. G. Marquardt, J. D. Stamm, G. Barany, and D. D. Thomas. Synthetic null-cysteine phospholamban analogue and the corresponding transmembrane domain inhibit the Ca-ATPase. *Biochemistry*, 39(35):10892–7, 2000.

[284] A. Mascioni, C. Karim, J. Zamoon, D. D. Thomas, and G. Veglia. Solid-state NMR and rigid body molecular dynamics to determine domain orientations of monomeric phospholamban. *Journal of the American Chemical Society*, 124(32):9392–9393, 2002.

[285] J. J. Chou, S. Gaemers, B. Howder, J. M. Louis, and A. Bax. A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles. *Journal of Biomolecular NMR*, 21(4):377–382, 2001.

[286] J. J. Warren and P. B. Moore. A maximum likelihood method ofr determining Da and R for sets of dipolar coupling data. *Journal of Magnetic Resonance*, 149:271–275, 2001.

[287] M. G. Clore, M. A. Gronenborn, and A. Bax. A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *Journal of Magnetic Resonance*, 133:216–221, 1998.

[288] J. H. Chill, J. M. Louis, F. Delaglio, and A. Bax. Local and global structure of the monomeric subunit of the potassium channel KcsA probed by NMR. *Biochimica et Biophysica Acta*, 1768(12):3260–3270, 2007.

[289] K. Oxenoid and J. J. Chou. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proceedings of the National Academy of Sciences U. S. A.*, 102(31):10870–10875, 2005.

[290] R. Langen, K. J. Oh, D. Cascio, and W. L. Hubbell. Crystal structures of spin labeled T4 lysozyme mutants: implications for the interpretation of EPR spectra in terms of structure. *Biochemistry*, 39(29):8396–8405, 2000.

[291] M. Sammalkorpi and T. Lazaridis. Modeling a spin-labeled fusion peptide in a membrane: implications for the interpretation of EPR experiments. *Biophysical Journal*, 92(1):10–22, 2007.

[292] L. Shi, N. J. Traaseth, R. Verardi, A. Cembran, J. Gao, and G. Veglia. A refinement protocol to determine structure, topology, and depth of insertion of membrane proteins using hybrid solution and solid-state NMR restraints. *Journal of Biomolecular NMR*, 44(4):195–205, 2009.

[293] C. M. Franzin, P. Teriete, and F. M. Marassi. Structural similarity of a membrane protein in micelles and membranes. *Journal of the American Chemical Society*, 129(26):8078–8079, 2007.

[294] M. M. Dedmon, K. Lindorff-Larsen, J. Christodoulou, M. Vendruscolo, and C. M. Dobson. Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *Journal of the American Chemical Society*, 127(2):476–477, 2005.

[295] B. Liang, J. H. Bushweller, and L. K. Tamm. Site-directed parallel spin-labeling and paramagnetic relaxation enhancement in structure determination of membrane proteins by solution NMR spectroscopy. *Journal of the American Chemical Society*, 128(13):4389–4397, 2006.

[296] A. De Simone, B. Richter, X. Salvatella, and M. Vendruscolo. Toward an accurate determination of free energy landscapes in solution states of proteins. *Journal of the American Chemical Society*, 131(11):3810–3811, 2009.

[297] L. Salmon, G. Nodet, V. Ozenne, G. Yin, M. R. Jensen, M. Zweckstetter, and M. Blackledge. NMR characterization of long-range order in intrinsically disordered proteins. *Journal of the American Chemical Society*, 132(24):8407–8418, 2010.

[298] V. Vasquez, M. Sotomayor, J. Cordero-Morales, K. Schulten, and E. Perozo. A structural mechanism for mscs gating in lipid bilayers. *Science*, 321(5893):1210–1214, 2008.

[299] W. Wang, S. S. Black, M. D. Edwards, S. Miller, E. L. Morrison, W. Bartlett, C. Dong, J. H. Naismith, and I. R. Booth. The structure of an open form of an e. coli mechanosensitive channel at 3.45 a resolution. *Science*, 321(5893):1179–1183, 2008.

[300] S. Weyand, T. Shimamura, S. Yajima, S. Suzuki, O. Mirza, K. Krusong, E. P. Carpenter, N. G. Rutherford, J. M. Hadden, J. O'Reilly, P. Ma, M. Saidijam, S. G. Patching, R. J. Hope, H. T. Norbertczak, P. C. Roach, S. Iwata, P. J. Henderson,

and A. D. Cameron. Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science*, 322(5902):709–713, 2008.

[301] D. C. Bay, K. L. Rommens, and R. J. Turner. Small multidrug resistance proteins: a multidrug transporter family that continues to grow. *Biochimica et Biophysica Acta*, 1778(9):1814–1838, 2008.

[302] S. F. Poget and M. E. Girvin. Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better. *Biochimica et Biophysica Acta*, 1768(12):3098–3106, 2007.

[303] S. F. Poget, S. M. Cahill, and M. E. Girvin. Isotropic bicelles stabilize the functional form of a small multidrug-resistance pump for NMR structural studies. *Journal of the American Chemical Society*, 129(9):2432–2433, 2007.

[304] C. R. Sanders, B. J. Hare, K. P. Howard, and J. H. Prestegard. Magnetically-oriented phospholipid micelles as a tool for the study of membrane-associated molecules. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 26(5):421–444, 1994.

[305] U. H. Durr, K. Yamamoto, S. C. Im, L. Waskell, and A. Ramamoorthy. Solid-state NMR reveals structural and dynamical properties of a membrane-anchored electron-carrier protein, cytochrome b5. *Journal of the American Chemical Society*, 129(21):6670–6671, 2007.

[306] K. J. Hallock, K. Henzler Wildman, D. K. Lee, and A. Ramamoorthy. An innovative procedure using a sublimable solid to align lipid bilayers for solid-state NMR studies. *Biophysical Journal*, 82(5):2499–2503, 2002.

[307] A. Ramamoorthy, C. H. Wu, and S. J. Opella. Experimental aspects of multidimensional solid-state NMR correlation spectroscopy. *Journal of Magnetic Resonance*, 140(1):131–140, 1999.

[308] K. R. Mackenzie. Folding and stability of alpha-helical integral membrane proteins. *Chemical Reviews*, 106(5):1931–1977, 2006.

[309] D. T. Moore, B. W. Berger, and W. F. DeGrado. Protein-protein interactions in the membrane: sequence, structural, and biological motifs. *Structure*, 16(7):991–1001, 2008.

[310] A. L. Stouffer, R. Acharya, D. Salom, A. S. Levine, L. Di Costanzo, C. S. Soto, V. Tereshko, V. Nanda, S. Stayrook, and W. F. DeGrado. Structural basis for the function and inhibition of an influenza virus proton channel. *Nature*, 451(7178):596–599, 2008.

[311] A. Senes, D. C. Chadi, P. B. Law, R. F. Walters, V. Nanda, and W. F. Degrado. E(z), a depth-dependent potential for assessing the energies of insertion of amino acid side-chains into membranes: derivation and applications to determining the orientation of transmembrane and interfacial helices. *Journal of Molecular Biology*, 366(2):436–448, 2007.

[312] J. Kuszewski, A. M. Gronenborn, and G. M. Clore. Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases. *Protein Science*, 5(6):1067–1080, 1996.

[313] C. H. Wu, A. Ramamoorthy, and S. J. Opella. Three-dimensional solid-state NMR experiment that correlates the chemical shift and dipolar coupling frequencies of two heteronuclei. *Journal Magnetic Resonance*, 109:270–272, 1994.

[314] R. Bertram, J. R. Quine, M. S. Chapman, and T. A. Cross. Atomic refinement using orientational restraints from solid-state NMR. *Journal of Magnetic Resonance*, 147(1):9–16, 2000.

[315] R. R. Ketchem, K. C. Lee, S. Huo, and T. A. Cross. Macromolecular structural elucidation with solid-state NMR-derived orientational constraints. *Journal of Biomolecular NMR*, 8(1):1–14, 1996.

[316] J. Lee, J. Chen, C. L. Brooks 3rd, and W. Im. Application of solid-state NMR restraint potentials in membrane protein modeling. *Journal of Magnetic Resonance*, 193(1):68–76, 2008.

[317] M. Nilges, A. M. Gronenborn, A. T. Brunger, and G. M. Clore. Determination of three-dimensional structures of proteins by simulated annealing with interproton

distance restraints. application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Engineering*, 2(1):27–38, 1988.

[318] S. Kim, J. R. Quine, and T. A. Cross. Complete cross-validation and R-factor calculation of a solid-state NMR derived structure. *Journal of the American Chemical Society*, 123(30):7292–7298, 2001.

[319] E. E. Metcalfe, N. J. Traaseth, and G. Veglia. Serine 16 phosphorylation induces an order-to-disorder transition in monomeric phospholamban. *Biochemistry*, 44(11):4386–4396, 2005.

[320] N. P. Cowieson, B. Kobe, and J. L. Martin. United we stand: combining structural methods. *Current Opinion in Structural Biology*, 18(5):617–622, 2008.

[321] B. Shaanan, A. M. Gronenborn, G. H. Cohen, G. L. Gilliland, B. Veerapandian, D. R. Davies, and G. M. Clore. Combining experimental information from crystal and solution studies: joint X-ray and NMR refinement. *Science*, 257(5072):961–964, 1992.

[322] F. Gabel, B. Simon, and M. Sattler. A target function for quaternary structural refinement from small angle scattering and NMR orientational restraints. *European Biophysics Journal*, 35(4):313–327, 2006.

[323] F. Gabel, B. Simon, M. Nilges, M. Petoukhov, D. Svergun, and M. Sattler. A structure refinement protocol combining NMR residual dipolar couplings and small angle scattering restraints. *Journal of Biomolecular NMR*, 41(4):199–208, 2008.

[324] L. K. Tamm, A. L. Lai, and Y. Li. Combined NMR and EPR spectroscopy to determine structures of viral fusion domains in membranes. *Biochimica et Biophysica Acta*, 1768(12):3052–3060, 2007.

[325] R. Mahalakshmi and F. M. Marassi. Orientation of the Escherichia coli outer membrane protein OmpX in phospholipid bilayer membranes determined by solid-state NMR. *Biochemistry*, 2008.

[326] J. R. Calhoun, W. Liu, K. Spiegel, M. Dal Peraro, M. L. Klein, K. G. Valentine, A. J. Wand, and W. F. DeGrado. Solution NMR structure of a designed

metalloprotein and complementary molecular dynamics refinement. *Structure*, 16(2):210–215, 2008.

[327] J. Kordel, D. A. Pearlman, and W. J. Chazin. Protein solution structure calculations in solution: solvated molecular dynamics refinement of calbindin d9k. *Journal of Biomolecular NMR*, 10(3):231–243, 1997.

[328] J. P. Linge, M. A. Williams, C. A. Spronk, A. M. Bonvin, and M. Nilges. Refinement of protein structures in explicit solvent. *Proteins*, 50(3):496–506, 2003.

[329] B. Xia, V. Tsui, D. A. Case, H. J. Dyson, and P. E. Wright. Comparison of protein solution structures refined by molecular dynamics simulation in vacuum, with a generalized Born model, and with explicit water. *Journal of Biomolecular NMR*, 22(4):317–331, 2002.

[330] C. Dominguez, R. Boelens, and A. M. Bonvin. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, 125(7):1731–1737, 2003.

[331] J. J. Buffy, T. Hong, S. Yamaguchi, A. J. Waring, R. I. Lehrer, and M. Hong. Solid-state NMR investigation of the depth of insertion of protegrin-1 in lipid bilayers using paramagnetic $Mn^{2+}$. *Biophysical Journal*, 85(4):2363–2373, 2003.

[332] R. Mani, M. Tang, X. Wu, J. J. Buffy, A. J. Waring, M. A. Sherman, and M. Hong. Membrane-bound dimer structure of a beta-hairpin antimicrobial peptide from rotational-echo double-resonance solid-state NMR. *Biochemistry*, 45(27):8341–8349, 2006.

[333] O. Toke, R. D. O'Connor, T. K. Weldeghiorghis, W. L. Maloy, R. W. Glaser, A. S. Ulrich, and J. Schaefer. Structure of KIAGKIA3 aggregates in phospholipid bilayers by solid-state NMR. *Biophysical Journal*, 87(1):675–687, 2004.

[334] C. H. Wu, A. Ramamoorthy, L. M. Gierasch, and S. J. Opella. Simultaneous characterization of the amide $^{1}H$ chemical shift, $^{1}H$-$^{15}N$ dipolar, and $^{15}N$ chemical shift interaction tensors in a peptide bond by three-dimensional solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 117(2):6148–6149, 1995.

[335] M. Tada, M. A. Kirchberger, and A. M. Katz. Phosphorylation of a 22,000-dalton component of the cardiac sarcoplasmic reticulum by adenosine 3':5'-monophosphate-dependent protein kinase. *Journal of Biological Chemistry*, 250(7):2640–2647, 1975.

[336] M. Hoshijima, Y. Ikeda, Y. Iwanaga, S. Minamisawa, M. O. Date, Y. Gu, M. Iwatate, M. Li, L. Wang, J. M. Wilson, Y. Wang, J. Ross Jr, and K. R. Chien. Chronic suppression of heart-failure progression by a pseudophosphorylated mutant of phospholamban via in vivo cardiac rAAV gene delivery. *Nature Medicine*, 8(8):864–871, 2002.

[337] D. M. Kaye, A. Preovolos, T. Marshall, M. Byrne, M. Hoshijima, R. Hajjar, J. A. Mariani, S. Pepe, K. R. Chien, and J. M. Power. Percutaneous cardiac recirculation-mediated gene transfer of an inhibitory phospholamban peptide reverses advanced heart failure in large animals. *Journal of the American College of Cardiology*, 50(3):253–260, 2007.

[338] H. K. Simmerman and L. R. Jones. Phospholamban: protein structure, mechanism of action, and role in cardiac function. *Physiological Reviews*, 78(4):921–47, 1998.

[339] C. B. Karim, T. L. Kirby, Z. Zhang, Y. Nesmelov, and D. D. Thomas. Phospholamban structural dynamics in lipid bilayers probed by a spin label rigidly coupled to the peptide backbone. *Proceedings of the National Academy of Sciences U. S. A.*, 101(40):14437–42, 2004.

[340] C. B. Karim, Z. Zhang, E. C. Howard, K. D. Torgersen, and D. D. Thomas. Phosphorylation-dependent conformational switch in spin-labeled phospholamban bound to SERCA. *Journal of Molecular Biology*, 358(4):1032–1040, 2006.

[341] Y. E. Nesmelov, C. B. Karim, L. Song, P. G. Fajer, and D. D. Thomas. Rotational dynamics of phospholamban determined by multifrequency electron paramagnetic resonance. *Biophysical Journal*, 93(8):2805–2812, 2007.

[342] E. S. Karp, E. K. Tiburu, S. Abu-Baker, and G. A. Lorigan. The structural properties of the transmembrane segment of the integral membrane protein phospholamban utilizing $^{13}$C CPMAS, $^{2}$H, and REDOR solid-state NMR spectroscopy. *Biochimica et Biophysica Acta - Biomembranes*, 1758(6):772 – 780, 2006.

[343] C. H. Wu, A. Ramamoorthy, and S. J. Opella. High-resolution heteronuclear dipolar solid-state NMR spectroscopy. *Journal of Magnetic Resonance*, 109(2):270–272, 1994.

[344] T. Vorherr, E. Carafoli, A. Wrzosek, and M. Chiesi. Total synthesis and functional properties of the membrane-intrinsic protein phospholamban. *Protein Science*, 2(3):339–347, 1993.

[345] A. A. Nevzorov and S. J. Opella. Selective averaging for high-resolution solid-state NMR spectroscopy of aligned samples. *Journal of Magnetic Resonance*, 185(1):59–70, 2007.

[346] P. L. Gor'kov, E. Y. Chekmenev, C. Li, M. Cotten, J. J. Buffy, N. J. Traaseth, G. Veglia, and W. W. Brey. Using low-E resonators to reduce RF heating in biological samples for static solid-state NMR up to 900 MHz. *Journal of Magnetic Resonance*, 185(1):77–93, 2007.

[347] R. Bertram, T. Asbury, F. Fabiola, J. R. Quine, T. A. Cross, and M. S. Chapman. Atomic refinement with correlated solid-state NMR restraints. *Journal of Magnetic Resonance*, 163(2):300 – 309, 2003.

[348] S. Lambeth, H. Schmid, M. Muenchbach, T. Vorherr, J. Krebs, E. Carafoli, and C. Griesinger. NMR solution structure of phospholamban. *Helvetica Chimica Acta*, 83:2141–2152, 2000.

[349] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.

[350] S. Hiller, R. G. Garces, T. J. Malia, V. Y. Orekhov, M. Colombini, and G. Wagner. Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science*, 321(5893):1206–1210, 2008.

[351] K. A. Baker, C. Tzitzilonis, W. Kwiatkowski, S. Choe, and R. Riek. Conformational dynamics of the KcsA potassium channel governs gating properties. *Nature Structural and Molecular Biology*, 14(11):1089–1095, 11 2007.

[352] M. Bayrhuber, T. Meins, M. Habeck, S. Becker, K. Giller, S. Villinger, C. Vonrhein, C. Griesinger, M. Zweckstetter, and K. Zeth. Structure of the human voltage-dependent anion channel. *Proceedings of the National Academy of Sciences U. S. A.*, 105(40):15370–15375, 2008.

[353] A. Grishaev, V. Tugarinov, L. E. Kay, J. Trewhella, and A. Bax. Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. *Journal of Biomolecular NMR*, 40(2):95–106, 2008.

[354] M. J. Bogusky, G. C. Leo, and S. J. Opella. Comparison of the dynamics of the membrane-bound form of fd coat protein in micelles and in bilayers by solution and solid-state nitrogen-15 nuclear magnetic resonance spectroscopy. *Proteins: Structure, Function, and Bioinformatics*, 4(2):123–130, 1988.

[355] S. J. Opella, F. M. Marassi, J. J. Gesell, A. P. Valente, Y. Kim, M. Oblatt-Montal, and M. Montal. Structures of the M2 channel-lining segments from nicotinic acetylcholine and nmda receptors by NMR spectroscopy. *Nature Structural Biology*, 6(4):374–9, 1999.

[356] K. Werner, I. Lehner, H. Dhiman, C. Richter, C. Glaubitz, H. Schwalbe, J. Klein-Seetharaman, and H. Khorana. Combined solid state and solution NMR studies of $\alpha, \epsilon - {}^{15}$N labeled bovine rhodopsin. *Journal of Biomolecular NMR*, 37:303–312, 2007. 10.1007/s10858-007-9143-0.

[357] E. M. Kelly, Z. Hou, J. Bossuyt, D. M. Bers, and S. L. Robia. Phospholamban oligomerization, quaternary structure, and sarco(endo)plasmic reticulum calcium ATPase binding measured by fluorescence resonance energy transfer in living cells. *The Journal of Biological Chemistry*, 283(18):12202–12211, 2008.

[358] O. C. Andronesi, S. Becker, K. Seidel, H. Heise, H. S. Young, and M. Baldus. Determination of membrane protein structure and dynamics by magic-angle-spinning solid-state NMR spectroscopy. *Journal of the American Chemical Society*, 127(37):12965–12974, 2005.

[359] J. P. Schmitt, M. Kamisago, M. Asahi, G. H. Li, F. Ahmad, U. Mende, E. G. Kranias, D. H. MacLennan, J. G. Seidman, and C. E. Seidman. Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science*, 299(5611):1410–3, 2003.

[360] K. Haghighi, F. Kolokathis, A. O. Gramolini, J. R. Waggoner, L. Pater, R. A. Lynch, G. C. Fan, D. Tsiapras, R. R. Parekh, G. W. Dorn 2nd, D. H. MacLennan, D. T. Kremastinos, and E. G. Kranias. A mutation in the human phospholamban gene, deleting arginine 14, results in lethal, hereditary cardiomyopathy. *Proceedings of the National Academy of Sciences U. S. A.*, 103(5):1388–1393, 2006.

[361] S. Abu-Baker, J. X. Lu, S. Chu, K. K. Shetty, P. L. Gor'kov, and G. A. Lorigan. The structural topology of wild-type phospholamban in oriented lipid bilayers using $^{15}$N solid-state NMR spectroscopy. *Protein Science*, 2007.

[362] S. L. Robia, N. C. Flohr, and D. D. Thomas. Phospholamban pentamer quaternary conformation determined by in-gel fluorescence anisotropy. *Biochemistry*, 44(11):4302–4311, 2005.

[363] C. Toyoshima, M. Asahi, Y. Sugita, R. Khanna, T. Tsuda, and D. H. MacLennan. Modeling of the inhibitory interaction of phospholamban with the $Ca^{2+}$ ATPase. *Proceedings of the National Academy of Sciences U. S. A.*, 100(2):467–72, 2003.

[364] F. M. Marassi and S. J. Opella. Simultaneous assignment and structure determination of a membrane protein from NMR orientational restraints. *Protein Science*, 12(3):403–411, 2003.

[365] F. M. Marassi, C. Ma, H. Gratkowski, S. K. Straus, K. Strebel, M. Oblatt-Montal, M. Montal, and S. J. Opella. Correlation of the structural and functional domains in the membrane protein Vpu from HIV-1. *Proceedings of the National Academy of Sciences U. S. A.*, 96(25):14336–41, 1999.

[366] C. Tian, P. F. Gao, L. H. Pinto, R. A. Lamb, and T. A. Cross. Initial structural and dynamic characteriaztion of the M2 protein transmembrane and amphipathic helices in lipid bilayerws. *Protein Science*, 12(11):2597–2605, 2003.

[367] G. Drin, J. Casella, R. Gautier, T. Boehmer, and T. U. Schwartzand B. Antonny. A general amphipathic alpha-helical motif for sensing membrane curvature. *Nature Structural and Molecular Biology*, 14(2):138–146, 02 2007.

[368] R. B. Cornell and S. G. Taneva. Amphipathic helices as mediators of the membrane interaction of amphitropic proteins, and as modulators of bilayer physical properties. *Current Protein & Peptide Science*, 7(6):539–552, 2006.

[369] B. M. Fung, A. K. Khitrin, and K. Ermolaev. An improved broadband decoupling sequence for liquid crystals and solids. *Journal of Magnetic Resonance*, 142(1):97–101, 2000.

[370] E. Vinogradov, P. K. Madhu, and S. Vega. High-resolution proton solid-state NMR spectroscopy by phase-modulated lee-goldberg experiment. *Chemical Physics Letters*, 314:443–50, 1999.

[371] B. M. Fung, K. Ermolaev, and Y. Yu. $^{13}$C NMR of liquid crystals with different proton homonuclear dipolar decoupling methods. *Journal of Magnetic Resonance*, 138:28–35, 1999.

[372] A. Bax, G. W. Vuister, S. Grzesiek, F. Delaglio, A. C. Wang, R. Tschudin, and G. Zhu. Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods in Enzymology*, 239:79–105, 1994.

[373] J. S. Hu, S. Grzesiek, and A. Bax. Two-dimensional NMR methods for determining $\chi$1 angles of aromatic residues in proteins from three-bond JC'Cg and JNCgcouplings. *Journal of the American Chemical Society*, 119(12):1803–1804, 1997.

[374] A. Senes, I. Ubarretxena-Belandia, and D. M. Engelman. The calpha–h...o hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proceedings of the National Academy of Sciences U. S. A.*, 98(16):9056–9061, 2001.

[375] P. P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen Der Physik*, 369:253–287, 1921.

[376] G. J. Martyna, D. J. Tobias, and M. L. Klein. Constant pressure molecular dynamics algorithms. *Journal of Chemical Physics*, 101(5):4177–4189, 1994.

[377] S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks. Constant pressure molecular dynamics simulation: The langevin piston method. *Journal of Chemical Physics*, 103(11):4613–4622, 1995.

[378] R. Verardi, L. Shi, N. J. Traaseth, N. Walsh, and G. Veglia. Structural topology of phospholamban pentamer in lipid bilayers by a hybrid solution and solid-state NMR method. *Proceedings of the National Academy of Sciences U. S. A.*, 108(22):9101–9106, 2011.

[379] R. L. Cornea, L. R. Jones, J. M. Autry, and D. D. Thomas. Mutation and phosphorylation change the oligomeric structure of phospholamban in lipid bilayers. *Biochemistry*, 36(10):2960–7, 1997.

[380] P. Boknik, C. Unkel, U. Kirchhefer, U. Kleideiter, O. Klein-Wiele, J. Knapp, B. Linck, H. Luss, F. U. Muller, W. Schmitz, U. Vahlensieck, N. Zimmermann, L. R. Jones, and J. Neumann. Regional expression of phospholamban in the human heart. *Cardiovascular Research*, 43(1):67–76, 1999.

[381] I. T. Arkin, P. D. Adams, A. T. Brunger, S. O. Smith, and D. M. Engelman. Structural perspectives of phospholamban, a helical transmembrane pentamer. *Annual Review of Biophysics and Biomolecular Structure*, 26:157–79, 1997.

[382] R. J. Kovacs, M. T. Nelson, H. K. Simmerman, and L. R. Jones. Phospholamban forms $Ca^{2+}$-selective channels in lipid bilayers. *The Journal of Biological Chemistry*, 263(34):18364–18368, 1988.

[383] L. Becucci, A. Cembran, C. B. Karim, D. D. Thomas, R. Guidelli, J. Gao, and G. Veglia. On the function of pentameric phospholamban: ion channel or storage form? *Biophysical Journal*, 96(10):L60–2, 2009.

[384] S. L. Robia, K. S. Campbell, E. M. Kelly, Z. Hou, D. L. Winters, and D. D. Thomas. Forster transfer recovery reveals that phospholamban exchanges slowly from pentamers but rapidly from the SERCA regulatory complex. *Circulation Research*, 101(11):1123–1129, 2007.

[385] K. N. Ha, L. R. Masterson, Z. Hou, R. Verardi, N. Walsh, G. Veglia, and S. L. Robia. Lethal arg9cys phospholamban mutation hinders $Ca^{2+}$-ATPase regulation and phosphorylation by protein kinase A. *Proceedings of the National Academy of Sciences U. S. A.*, 108(7):2735–2740, 2011.

[386] L. E. Kay, M. Ikura, R. Tschudin, and A. Bax. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *Journal of Magnetic Resonance*, 89(3):496–514, 1990.

[387] S. Grzesiek and A. Bax. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *Journal of Magnetic Resonance*, 114(16):6291–6293, 1992.

[388] T. A. Cross, M. Sharma, M. Yi, and H. X. Zhou. Influence of solubilizing environments on membrane protein structures. *Trends in Biochemical Sciences*, 2010.

[389] S. Abu-Baker, J. X. Lu, S. Chu, K. K. Shetty, P. L. Gor'kov, and G. A. Lorigan. The structural topology of wild-type phospholamban in oriented lipid bilayers using $^{15}N$ solid-state NMR spectroscopy. *Protein Science*, 16(11):2345–2349, 2007.

[390] R. L. Cornea and D. D. Thomas. Effects of membrane thickness on the molecular dynamics and enzymatic activity of reconstituted Ca-ATPase. *Biochemistry*, 33(10):2912–2920, 1994.

[391] R. J. Bick, L. M. Buja, W. B. Van Winkle, and G. E. Taffet. Membrane asymmetry in isolated canine cardiac sarcoplasmic reticulum: comparison with skeletal muscle sarcoplasmic reticulum. *The Journal of Membrane Biology*, 164(2):169–175, 1998.

[392] W. Liu, J. Z. Fei, T. Kawakami, and S. O. Smith. Structural constraints on the transmembrane and juxtamembrane regions of the phospholamban pentamer in membrane bilayers: Gln29 and Leu52. *Biochimica et Biophysica Acta*, 1768(12):2971–2978, 2007.

[393] S. I. O'Donoghue and M. Nilges. *Calculation of Symmetric Oligomer Structures from NMR Data*, volume 17 of *Structure computation and dynamics in protein NMR*, chapter 4, pages 131–158. Kluwer Academic/Plenum Publishers, New York, NY, 1st edition, 1999.

[394] I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall 3rd, J. Snoeyink, J. S. Richardson, and D. C. Richardson. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research*, 35(Web Server issue):W375–83, 2007.

[395] C. B. Karim, J. D. Stamm, J. Karim, and L. R. Thomas Jones D. D. Cysteine reactivity and oligomeric structure of phospholamban and its mutants. *Biochemistry*, 37:12074–12081, 1998.

[396] N. Bocquet, H. Nury, M. Baaden, C. Le Poupon, J. P. Changeux, M. Delarue, and P. J. Corringer. X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature*, 457(7225):111–114, 2009.

[397] R. J. Hilf and R. Dutzler. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature*, 457(7225):115–118, 2009.

[398] C. Maffeo and A. Aksimentiev. Structure, dynamics, and ion conductance of the phospholamban pentamer. *Biophysical Journal*, 96(12):4853–4865, 2009.

[399] T. Kim, J. Lee, and W. Im. Molecular dynamics studies on structure and dynamics of phospholamban monomer and pentamer in membranes. *Proteins*, 2008.

[400] X. M. Zhang, Y. Kimura, and M. Inui. Effects of phospholipids on the oligomeric state of phospholamban of the cardiac sarcoplasmic reticulum. *Circulation Journal*, 69(9):1116–1123, 2005.

[401] E. Hughes, J. C. Clayton, and D. A. Middleton. Cytoplasmic residues of phospholamban interact with membrane surfaces in the presence of SERCA: a new role for phospholipids in the regulation of cardiac calcium cycling? *Biochimica et Biophysica Acta*, 1788(2):559–566, 2009.

[402] S. Chu, S. Abu-Baker, J. Lu, and G. A. Lorigan. $^{15}$N solid-state NMR spectroscopic studies on phospholamban at its phosphorylated form at Ser-16 in aligned phospholipid bilayers. *Biochimica et Biophysica Acta*, 1798(3):312–317, 2010.

[403] A. M. Slovic, J. D. Lear, and W. F. DeGrado. De novo design of a pentameric coiled-coil: decoding the motif for tetramer versus pentamer formation in water-soluble phospholamban. *The Journal of Peptide Research*, 65(3):312–321, 2005.

[404] D. Bray and T. Duke. Conformational spread: the propagation of allosteric states in large multiprotein complexes. *Annual Review of Biophysics and Biomolecular Structure*, 33:53–73, 2004.

[405] J. Zhai, A. G. Schmidt, B. D. Hoit, Y. Kimura, D. H. MacLennan, and E. G. Kranias. Cardiac-specific overexpression of a superinhibitory pentameric phospholamban mutant enhances inhibition of cardiac function in vivo. *The Journal of Biological Chemistry*, 275(14):10538–10544, 2000.

[406] G. Chu, L. Li, Y. Sato, J. M. Harrer, V. J. Kadambi, B. D. Hoit, D. M. Bers, and E. G. Kranias. Pentameric assembly of phospholamban facilitates inhibition of cardiac function in vivo. *The Journal of Biological Chemistry*, 273(50):33674–33680, 1998.

[407] S. Abu-Baker and G. A. Lorigan. Phospholamban and its phosphorylated form interact differently with lipid bilayers: a $^{31}$P, $^{2}$H, and $^{13}$C solid-state NMR spectroscopic study. *Biochemistry*, 45(44):13312–13322, 2006.

[408] M. Gustavsson, N. J. Traaseth, C. B. Karim, E. L. Lockamy, D. D. Thomas, and G. Veglia. Lipid-mediated folding/unfolding of phospholamban as a regulatory mechanism for the sarcoplasmic reticulum $Ca^{2+}$-ATPase. *Journal of Molecular Biology*, 2011.

[409] C. Fernandez, C. Hilty, G. Wider, and K. Wuthrich. Lipid-protein interactions in DHPC micelles containing the integral membrane protein OmpX investigated by NMR spectroscopy. *Proceedings of the National Academy of Sciences U. S. A.*, 99(21):13533–13537, 2002.

[410] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes. $^{1}$H, $^{13}$C and $^{15}$N chemical shift referencing in biomolecular NMR. *Journal of Biomolecular NMR*, 6(2):135–40, 1995.

[411] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. NMRPipe: A multidimensional spectral processing system based on unix pipes. *Journal of Biomolecular NMR*, 6:277–293, 1995.

[412] K. Takegoshi, S. Nakamura, and T. Terao. $^{13}$C-$^{1}$H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chemical Physics Letters*, 344:631, 2001.

[413] S. Jo, J. B. Lim, J. B. Klauda, and W. Im. CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes. *Biophysical Journal*, 97(1):50–58, 2009.

[414] A. D. MacKerell Jr, M. Feig, and C. L. Brooks 3rd. Improved treatment of the protein backbone in empirical force fields. *Journal of the American Chemical Society*, 126(3):698–699, 2004.

[415] E. A. Dolan, R. M. Venable, R. W. Pastor, and B. R. Brooks. Simulations of membranes and other interfacial systems using P21 and Pc periodic boundary conditions. *Biophysical Journal*, 82(5):2317–2325, 2002.

[416] T. Darden, D. York, and L. Pedersen. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089–10092, 1993.

[417] O. S. Smart, J. M. Goodfellow, and B. A. Wallace. The pore dimensions of gramicidin A. *Biophysical Journal*, 65(6):2455–2460, 1993.

[418] S. O. Smith, T. Kawakami, W. Liu, M. Ziliox, and S. Aimoto. Helical structure of phospholamban in membrane bilayers. *Journal of Molecular Biology*, 313(5):1139–1148, 2001.

[419] Z. Ahmed, D. G. Reid, A. Watts, and D. A. Middleton. A solid-state NMR study of the phospholamban transmembrane domain: local structure and interactions with $Ca^{2+}$-ATPase. *Biochimica et Biophysica Acta*, 1468(1-2):187–198, 2000.

[420] D. A. Middleton, Z. Ahmed, C. Glaubitz, and A. Watts. REDOR NMR on a hydrophobic peptide in oriented membranes. *Journal of Magnetic Resonance*, 147(2):366–370, 2000.

[421] E. Hughes and D. A. Middleton. Solid-state NMR reveals structural changes in phospholamban accompanying the functional regulation of $Ca^{2+}$-ATPase. *Journal of Biological Chemistry*, 278(23):20835–42, 2003.

[422] L. R. Masterson, L. Shi, E. Metcalfe, J. Gao, S. S. Taylor, and G. Veglia. Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy. *Proceedings of the National Academy of Sciences U. S. A.*, 108(17):6969–6974, 2011.

[423] A. Malmendal, S. Halpain, and W. J. Chazin. Nascent structure in the kinase anchoring domain of microtubule-associated protein 2. *Biochemistry and Biophysic Research Communication*, 301(1):136–42, 2003.

[424] T. D. Goddard and D. G. Kneller. Sparky 3, 1999.

[425] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by $^{15}N$ NMR relaxation. *Biochemistry*, 33(19):5984–6003, 1994.

[426] N. Tjandra, P. Wingfield, S. Stahl, and A. Bax. Anisotropic rotational diffusion of perdeuterated HIV protease from $^{15}N$ NMR relaxation measurements at two magnetic fields. *Journal of Biomolecular NMR*, 8(3):273–284, 1996.

[427] C. Wang, M. Rance, and A. G. Palmer 3rd. Mapping chemical exchange in proteins with MW $>$ 50 kD. *Journal of the American Chemical Society*, 125(30):8968–8969, 2003.

[428] S. A. Tatulian, L. R. Jones, L. G. Reddy, D. L. Stokes, and L. K. Tamm. Secondary structure and orientation of phospholamban reconstituted in supported bilayers from polarized attenuated total reflection FTIR spectroscopy. *Biochemistry*, 34(13):4448–56, 1995.

[429] M. K. Fenwick and R. E. Oswald. NMR spectroscopy of the ligand-binding core of ionotropic glutamate receptor 2 bound to 5-substituted willardiine partial agonists. *Journal of Molecular Biology*, 378(3):673–685, 2008.

[430] B. Ma and R. Nussinov. Enzyme dynamics point to stepwise conformational selection in catalysis. *Current Opinion in Chemical Biology*, 14(5):652–659, 2010.

[431] S. S. Taylor and E. Radzio-Andzelm. Protein kinase inhibition: natural and synthetic variations on a theme. *Current Opinion in Chemical Biology*, 1(2):219–226, 1997.

[432] R. A. Engh and D. Bossemeyer. Structural aspects of protein kinase control-role of conformational flexibility. *Pharmacology & Therapeutics*, 93(2-3):99–111, 2002.

[433] J. Zhang, F. J. Adrian, W. Jahnke, S. W. Cowan-Jacob, A. G. Li, R. E. Iacob, T. Sim, J. Powers, C. Dierks, F. Sun, G. R. Guo, Q. Ding, B. Okram, Y. Choi, A. Wojciechowski, X. Deng, G. Liu, G. Fendrich, A. Strauss, N. Vajpai, S. Grzesiek, T. Tuntland, Y. Liu, B. Bursulaya, M. Azam, P. W. Manley, J. R. Engen, G. Q. Daley, M. Warmuth, and N. S. Gray. Targeting Bcr-Abl by combining allosteric with ATP-binding-site inhibitors. *Nature*, 463(7280):501–506, 2010.

[434] L. Yang, G. Song, A. Carriquiry, and R. L. Jernigan. Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes. *Structure*, 16(2):321–330, 2008.