BIOPHYSICAL STUDIES OF MOLECULAR RECOGNITION IN PERIPHERAL AND INTEGRAL MEMBRANE PROTEINS

by

XUDONG GUAN

A dissertation submitted to the Graduate Faculty in Chemistry in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2012

This manuscript has been read and accepted for the
Graduate Faculty in Chemistry in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Date	Dr. Ruth E. Stark Chair of Examining Committee
Date	Dr. Maria Tamargo Executive Officer

Dr. Fred Naider

Dr. Ranajeet Ghose

Dr. Marilyn Gunner
Supervision Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

BIOPHYSICAL STUDIES OF MOLECULAR RECOGNITION IN PERIPHERAL AND INTEGRAL MEMBRANE PROTEINS

by

Xudong Guan

Advisor: Professor Ruth E. Stark

The molecular interactions of peripheral and integral membrane proteins with ligands and model biological membranes play important roles in the regulation of human nutrition, cell signaling, and other physiological processes. In the current study, we have used solution- and solid-state nuclear magnetic resonance (NMR) and computational modeling methods to study the interactions of fatty acid-binding proteins (FABPs) with ligands and membrane mimetics, and to examine the conformation of a transmembrane peptide fragment of the Ste2p G protein-coupled receptor (GPCR) in lipids.

Computational docking with NMR-derived restraints has been used to identify critical ligand-protein electrostatic interactions in both an intermediate singly-liganded state and the double-liganded liver fatty acid-binding protein (LFABP). The model structure for a R122L/S124A LFABP mutant reveals that two charged residues are required to ensure that the first oleate ligand adopts an orientation and conformation favoring binding of the second ligand within the protein cavity. A detailed model for the ligand binding process is proposed based on the analysis of these docked structures and solution-state NMR structures determined previously.

A new isotropically tumbling bilayered micelle (bicelle) system composed of dimyristoylphosophatidylcholine (DMPC) and 1,2-di-O-hexyl-sn-glycero-3-phosphocholine (DIOHPC) has been assessed as a membrane-mimetic medium and used to probe molecular interactions with two FABPs. Rat intestinal fatty acid-binding protein (IFABP) and LFABP were titrated with the bicelles, with site-specific changes monitored by chemical shift perturbations. Contrasting transfer mechanisms for fatty acids between FABPs and membranes were differentiated based on their respective chemical shift perturbations, and site-specific information was deduced about the IFABP-membrane interaction.

Additionally, both solution- and solid-state NMR have been utilized to investigate a double transmembrane peptide (TM1TM2) from the Ste2p GPCR in phospholipid environments. A solution-state NMR interaction study of the ^{15}N alanine-labeled peptide with DMPC-sodium docecylsulfate detergent mixtures demonstrated successful reconstitution into a helical conformation, whereas solid-state NMR experiments on TM1TM2 in DMPC multilayers permitted assignment of some signals by residue type.

Acknowledgements

I acknowledge my deepest gratitude to my mentor Professor Ruth E. Stark, who gave me the opportunity to work with these interesting and challenging projects, and guided me through the doctoral study. This dissertation would not have been possible without her tireless help.

I would express my warm and sincere thanks to my committee members, Professor Fred Naider, Professor Ranajeet Ghose, and Professor Marilyn Gunner for their valuable advice and support.

I acknowledge Dr. Cédric Bernard, Dr. Yan He and Dr. Hongwei Wu for their help in my research projects at different stages. I would also thank Dr. Leah Cohen for the collaboration on the solid state NMR study of a transmembrane peptide.

I would like to thank Dr. Hsin Wang for his help and discussion of my research projects and new NMR techniques.

I thank for all my colleagues who worked with me together in Dr. Stark's lab for those enjoyable moments we spent together.

Finally, I thank all my family members, my parents and my wife who always support me.

This dissertation is dedicated to my wife, Hua Qi for her love and support.

Table of Contents

List of Tables

List of Figures

CHAPTER 1 INTRODUCTION

1.1 FATTY ACID-BINDING PROTEINS

Fatty acid-binding proteins (FABPs) are a family of intracellular proteins that bind long-chain fatty acids (FAs) and other hydrophobic ligands. One important function of FABPs is transferring fatty acids between membranes and various intracellular organelles in cells (Figure 1.1)[1]. For example, it is reported that LFABP interacts with and transports ligand to peroxisome proliferator-activated receptor-α (PPARα) and plays a functional role in fatty acid metabolism[2]. Together with the broader family of intracellular lipid-binding proteins (iLBPs), they share a cavity structural feature consisting of 10 antiparallel β-strands and a portal region involving two short α-helices located between the first and second of these β-strands[3-10].

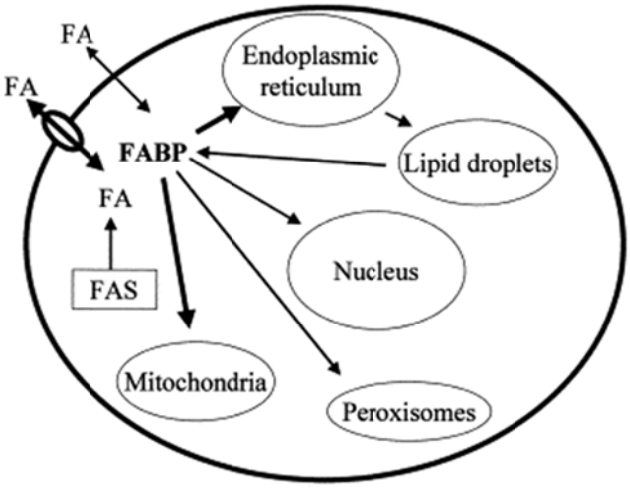

Figure 1.1 Intracellular transport for long-chain fatty acids (Storch and Thumser, 2000[1]). *FA: fatty acid; FABP: fatty acid-binding protein; FAS: fatty acid synthetase.

1.2 BINDING OF FATTY ACIDS WITH RAT LIVER FATTY ACID-BINDING PROTEIN

Rat liver fatty acid-binding protein (LFABP) is a 14-kDa cytosolic protein that is expressed at high levels in the small intestine and the liver[11]. Compared with other fatty acid-binding proteins (FABPs), the ability to bind more than one ligand molecule[12] within its cavity makes LFABP nearly unique in this protein family. The structures of the doubly liganded (holo) protein have been solved by both X-ray and solution NMR, showing that one ligand is bound deeply inside the cavity while another is located close to the surface[4, 6]. Recently a singly-liganded state was observed for LFABP using ^{15}N heteronuclear single quantum coherence spectroscopy (^{15}N HSQC) during the investigation of the ligand binding process by NMR titration[13]. Mutation experiments and computational calculation have suggested that residues arginine 122 and serine 124 play important roles in altering the binding behavior of the protein[14]. In the study of the binding process of R122L/S124A mutant, there was no evidence indicating the presence of a second ligand[14].

To obtain more detailed insight into binding phenomena in LFABP, computational docking with NMR-derived restraints has been employed to study the ligand binding and modeled structures have been generated for both singly-liganded LFABP (intermediate state) and the R122L/S124A LFABP mutant. The structural model of the intermediate state, together with previously solved NMR structures of apo- and holo-LFABP, provide a comprehensive picture of this binding process, testing the contention that the binding site deep inside the protein cavity is occupied first. Comparison of the modeled structures of intermediate states with the R122L/S124A mutant can reveal why the arginine 122 and serine 124 residues are crucial for the binding and provide a rationale for the effect of the mutations on the binding ability of LFABP.

1.3 NEW MEMBRANE MIMETICS AND INTERACTIONS BETWEEN FABPS AND MEMBRANES

Bilayered phospholipid micelles (bicelles) have drawn substantial attention as membrane mimetics for NMR structural studies of membrane-associated amphiphiles and proteins since the 1980s[15]. The ability to maintain planar bilayers makes bicelles better membrane mimetics than micelles, which may better preserve the native conformation of the polypeptide under study. Bicelles can be aligned with respect to the external magnetic field or tumble isotropically in solution, depending on the ratio of long chain to short chain lipids (q). The high-q bicelles have been used widely as aligned media to measure the residual dipolar couplings of soluble proteins for NMR-based structure refinement[15-17], whereas the low-q isotropically tumbling bicelles have made it possible to conduct high resolution solution-state NMR studies of peptides[18] and transmembrane proteins[19-21]. Dimyristoylphosphatidylcholine (DMPC) and dihexanoylphosphatidylcholine (DHPC) are commonly used lipids for bicellar media, but the longterm stability of DMPC-DHPC bicelles has been questioned because their carboxy-ester bonds are susceptible to acid- and base-catalyzed hydrolysis in aqueous solution[16, 22]. Ether-linked lipids were then introduced to enhance bicelle stability over a wide pH range[23].

The mechanism involved in fatty acid transport between membranes and FABPs has been proposed to occur by either direct collisional or diffusional mechanisms[24-27]. A collision-mediated process in which the FABPs and the acceptor membrane are in direct contact occurs for intestinal FABP (IFABP), adipose FABP (AFABP), heart FABP (HFABP), kidney FABP (KFABP), and cellular retinol-binding protein type I (CRBP I)[27], whereas diffusion of ligand through the aqueous phase and subsequent association with the acceptor membrane occurs for liver FABP (LFABP) and cellular retinol-binding protein type II (CRBP II)[24, 27-31].

Herein, we have used NMR methods to investigate the stability, aggregate size, and phospholipid organization for a DMPC-DIOHPC bicelle, a new isotropically tumbling phospholipid system that may serve as a membrane-mimetic medium for studies of iLBPs and other peripheral membrane proteins. The application of the DMPC-DIOHPC bicelle as a membrane mimetic has also been explored by monitoring the interactions with each of two FABPs that have contrasting fatty acid transfer mechanisms. In this work we critically evaluate the formation and promise of a stable isotropically tumbling bicellar complex for high-resolution NMR spectra and structural studies of peripheral membrane proteins. Titrations of LFABP and IFABP with bicelles could be monitored through chemical shift perturbations of the ^{15}N HSQC NMR spectra. These experiments tested whether mechanisms of membrane interactions with LFABP and intestinal fatty acid-binding protein (IFABP) could be differentiated according to the magnitude and spatial distribution of their respective chemical shift perturbations, and whether site-specific information could be obtained about the collisional interactions.

1.4 DOUBLE TRANSMEMBRANE PEPTIDE STUDIED BY SOLID-STATE NMR

G-protein coupled receptors (GPCRs) are a family of integral membrane proteins that share a common structural motif containing 7 transmembrane helices. As important targets for drug design, almost half of modern drugs in the pharmaceutical industry are based on GPCRs[32]. Despite the functional importance of this protein family, molecular-level structural information about GPCRs is severely limited. One fruitful approach to obtain structural information about GPCRs uses peptide fragments that can be reconstituted in a fully functional form[33].

The *Saccharomyces cerevisiae* GPCR, Ste2p, is involved in the regulation of the mating process in *Saccharomyces cerevisiae*. The structures of the first two transmembrane domains, TM1TM2 [Ste2p (G31-T110)] have been solved in LPPG micelles[34] and TFE/water[35] by solution-state NMR. In a complementary fashion, solid-state NMR techniques may allow the structural study of transmembrane peptides and proteins in a lipid bilayer environment[36-44]. In the current study, the reconstitution of the TM1TM2 peptide in dimyristoylphosphatidylcholine-Sodium dodecylsulfate (DMPC-SDS) was evaluated by solution-state NMR. Solid-state NMR was also used for a preliminary study of TM1TM2 in a lipid bilayer, a more native-like environment, to obtain information on the maintenance of its helical structure.

CHAPTER 2 INVESTIGATION OF THE LFABP LIGAND BINDING PROCESS BY NMR TITRATION AND DOCKING METHODS

2.1 INTRODUCTION

Rat liver fatty acid-binding protein (LFABP) is a 14-kDa cytosolic protein expressed abundantly in the small intestine and the liver[11]. Compared with other fatty acid-binding proteins (FABPs), the ability to bind more than one ligand molecule[12] within its cavity makes LFABP nearly unique in the intracellular lipid-binding protein family. Both X-ray and solution NMR structures of the holo protein show that one ligand is bound deeply inside the cavity while the other one is situated close to the surface[4, 6]. Using an NMR titration to investigate the ligand binding process of LFABP, we recently observed a singly-liganded state using ^{15}N heteronuclear single quantum coherence NMR spectroscopy (^{15}N HSQC)[13]. It has also been found that binding sites deep inside the protein cavity are occupied prior to binding sites close to the protein surface. To clarify the description, we refer to the first ligand as the one bound inside the protein cavity and the second ligand as the one lying close to the protein surface.

To understand the binding process of LFABP, both mutations[45, 46] and computational studies have been done to elucidate the driving force behind it. A Multi-Conformational Continuum Electrostatics (MCCE[47, 48]) calculation[14] suggested that R122 and S124 promote the ionization of the carboxylate group of the first ligand. The R122L/S124A mutant of LFABP was constructed and the binding properties were tested in our group[14], revealing that the binding affinity of the first ligand drops significantly after the mutation. In addition to that, there is no evidence indicating the presence of a second ligand in the holo form of the mutant. The mutational study supports the hypothesis that residues arginine 122 and serine 124 play important roles in altering

the binding behavior of the protein[14]. As an interesting residue, arginine is found to be involved in electrostatic interactions with the carboxylate group of fatty acids in structures of several FABP-fatty acid complexes[4, 6, 49, 50].

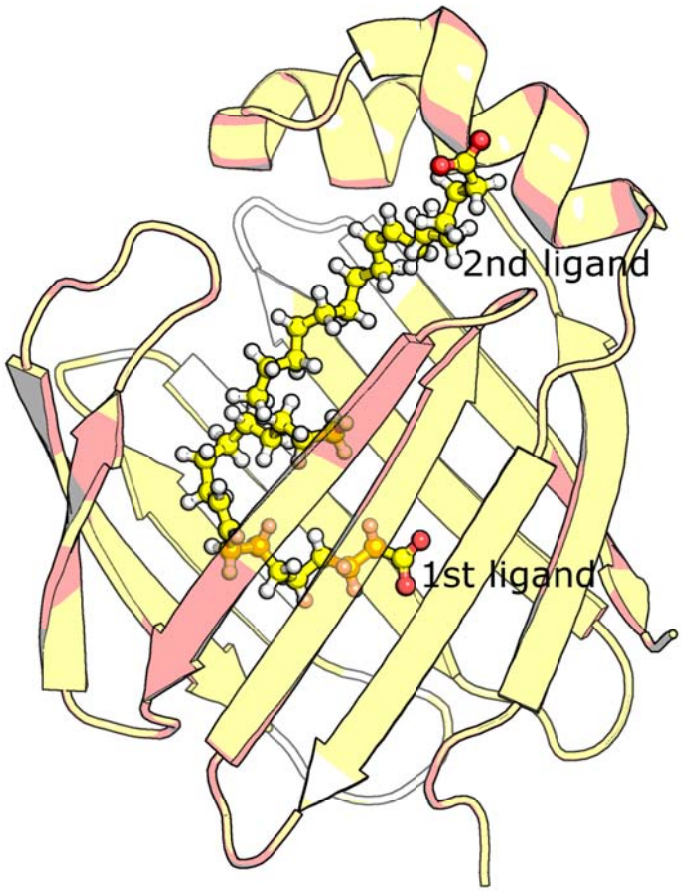

Figure 2.1. NMR-based structure of holo LFABP with two oleates bound[4].

Herein we show that structural models can be generated for both singly-liganded LFABP and the R122L/S124A mutant by computational docking, by incorporating both chemical shift

perturbation (CSP) and intermolecular nuclear Overhauser effect (NOE) restraints. The structural model of the singly-liganded protein reveals the conformation and binding sites in the intermediate state, whereas the structural model of R122L/S124A LFABP shows the change in ligand binding behavior upon mutation. The analysis of these computational models, together with the previously determined apo- and holo-protein structures, provides a comprehensive picture of this binding process. It is also shown why the arginine 122 and serine 124 residues are crucial for binding and how the mutation affects the binding process for R122L/S124A LFABP.

2.2 HIGH AMBIGUITY DRIVEN BIOMOLECULAR DOCKING (HADDOCK)

The docking of wild-type LFABP and the R122L/S124A mutant with oleate was performed using the software HADDOCK2.1 (High Ambiguity Driven biomolecular DOCKing). HADDOCK is a protein-protein or protein-ligand docking approach based on biochemical and/or biophysical interactions.[51] The software was originally designed for protein-protein docking and later expanded to protein-DNA/RNA, protein-peptide and protein-ligand applications. The advantage of this docking approach is that it does not rely solely on the combination of energetics and shape complementarity typical of most docking software. HADDOCK also makes use of biochemical and biophysical information obtained from experimental data, for example, chemical shift perturbation effects measured by solution-state NMR. In addition to that, CNS1.3[52] is loaded as structure calculation software, making it very convenient to load NMR-based restraints obtained from chemical shift perturbation and intermolecular NOE data.

There are three stages comprising the HADDOCK protocol: rigid body energy minimization, semi-flexible simulated annealing, and final refinement with explicit solvent[51, 53]. During each stage of docking, a specific number of structures are generated and sorted by their HADDOCK scores. Then the top HADDOCK-ranked structures are selected and sent to the next stage to continue the docking process. The HADDOCK score used to rank and select solutions during different docking stage is calculated as described by Tomaselli *et al.*[53] as a function of van der Waals energy, electrostatic energy, Ambiguous Interaction Restraints (AIR) energy, buried surface area, and desolvation energy terms at different docking stages.

2.3 NMR DATA BASED RESTRAINTS

As noted above, HADDOCK is able to incorporate some experimental results, especially NMR data, during the docking process. Herein we use both chemical shift perturbations and intermolecular NOEs to define restraints for the docking process.

2.3.1 Chemical shift perturbation

Chemical shift assignments of apo-, holo- and sgl-LFABP (sgl-LFABP refers to singly-liganded LFABP) were obtained from Dr. He's previous work[13]; chemical shift assignments of apo- and holo-R122L/S124A LFABP were made by Dr. Estephan[14]. Chemical shift perturbations at each polypeptide backbone site of wild-type LFABP were obtained from Dr. He's titration work[13] and replotted. Chemical shift perturbation arises from the change in magnetic (chemical) environment of particular residues upon addition of ligands, indicating either the binding interface or global conformational changes[54]. An example of a ^{1}HN and ^{15}N chemical shift perturbation plot comparing the wild-type forms of apo- and sgl-LFABP appears in Figure 2.2.

Chemical shift perturbations for each backbone cross peak were calculated with the following equation:

$\Delta_{HN-N} = ([\boldsymbol{\delta_{HN}(apo)} - \boldsymbol{\delta_{HN}(sgl)}]^{\mathbf{2}} + \left\{\frac{[\boldsymbol{\delta_N(apo)} - \boldsymbol{\delta_N(sgl)}]}{\mathbf{6.5}}\right\}^{\mathbf{2}})^{\mathbf{1/2}}$ as recommended[55]. If the chemical shift perturbations are larger than the sum of the average CSP and standard deviation, those residues are considered to have significant perturbation as indicated by red bars in Figure 2.2. As the chemical shift difference reflects a change in magnetic environment, those residues with large chemical shift perturbations are viewed as likely binding sites of the ligand in this singly-liganded state if they are clustered spatially to form a protein-ligand binding interface.

They are used to define the Ambiguous Interaction Restraints (AIRs) for docking calculations with HADDOCK.

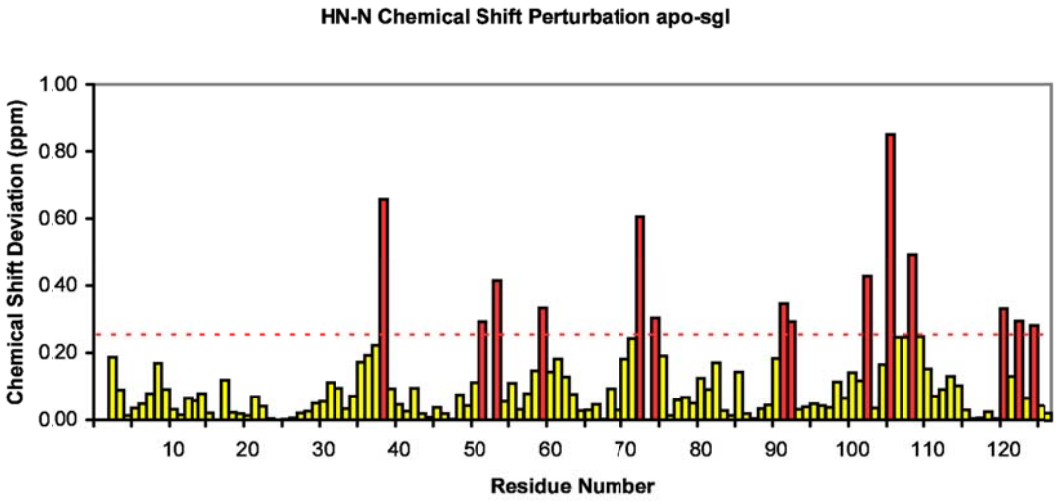

Figure 2.2 Composite $^{1}H–^{15}N$ chemical shift perturbation plotted along the protein sequence for apo- to sgl- states of wild-type LFABP[13]. Average CSP + standard deviation is shown by a red dashed line, and all residues with significant perturbations are colored red.

2.3.2 Intermolecular NOE

Besides chemical shift perturbation data, we also employed intermolecular Nuclear Overhauser Effect (NOE) data in the docking procedure. Those NOE data provide more definitive information about the protein-ligand interface. Intermolecular NOE data were collected with two types of samples for both wild-type and double mutant LFABPs. The samples were prepared with both [U-^{15}N,^{13}C]LFABP and unlabeled oleate and with [U-^{15}N]LFABP and [U-^{13}C]oleate. 3D ^{13}C-filtered NOE[56] measurements were conducted with either a 600 MHz Varian INOVA or 800 MHz Bruker AVANCE II spectrometer. All data were processed with

NMRPipe[57], and processed data were analyzed with NMRViewJ[58] (One Moon Scientific, Inc., Westfield, NJ).

To assign the intermolecular NOE data, chemical shift assignments are needed for both proteins and ligands. Chemical shift assignments of wild-type LFABP in the singly-liganded state and for the holo-R122/S124 LFABP were made by Dr. He and Dr. Estephan,[14] respectively. Chemical shift assignments of oleate in different samples were based on a previous assignment of oleate in pH 7.0 phosphate buffer[59] and titration of LFABP with ^{13}C-labeled long-chain fatty acids, including oleate[60]. The assignment of oleate is illustrated by a ^{13}C HSQC spectrum for the singly-liganded state (Figure 2.3). It can be seen from the spectrum that many proton signals are overlapped with each other in the ^{1}H dimension, which could cause ambiguity when assigning intermolecular NOE data. Although, for example, proton signals of P and Q are overlapped with protons C-F/K-N, fortunately those signals that are overlapped in the ^{1}H dimension have corresponding ^{13}C signals that can be resolved along the ^{13}C dimension.

In order to make more definitive assignments, the intermolecular NOE data were analyzed with the following strategy: 1) examine the ^{13}C dimension of ^{13}C-filtered NOE spectra of the ^{13}C labeled protein sample and ^{13}C labeled oleate samples side by side to identify those signals present in both datasets; 2) assign the signals to protein based on ^{1}H and ^{13}C chemical shifts in the protein-filtered NOE spectrum; 3) assign the signals to oleate based on ^{1}H and ^{13}C chemical shifts in oleate-filtered NOE spectrum. In this way, several intermolecular NOE signals could be assigned unambiguously. For example, Figure 2.4 shows the assignment of one set of intermolecular NOEs for singly-liganded LFABP. ^{13}C-edited NOE experiments for both oleate labeled (A) and protein labeled (B) samples show NOE cross peaks between 1.77 ppm (oleate) and 0.84 ppm (LFABP). The peak with a ^{13}C chemical shift of 29.6 ppm and a ^{1}H chemical shift

of 1.77 ppm can be assigned to carbon and hydrogen G or J of oleate based on spectrum A; the peak with ^{13}C chemical shift of 27.0 ppm and ^{1}H chemical shift of 0.88 ppm can be assigned to C_δ and H_δ of Leucine 50 of the protein based on spectrum B. Thus an intermolecular NOE can be deduced between hydrogen G/J of oleate and H_δ of Leucine 50 in LFABP. A summary of the intermolecular NOE assignments for sgl-LFABP is reported in Table 2.1.

Table 2.1. Intermolecular NOE assignments for sgl-LFABP

Oleate	LFABP
HA	I41.HD11, I109.HG21, I109.HD11
HG/HJ	L50.HD11, L50.HD21, L71.HD11, V83.HG21, T93.HG21
HH/HI	L71.HD11, L71.HG21, V83.HG11, V83.HG21, T93.HG21
HP	L50.HD11, I59.HB
HQ	I59.HD11, L71.HB2, E72.HB2
HR	Y54.HD1, M74.HG2

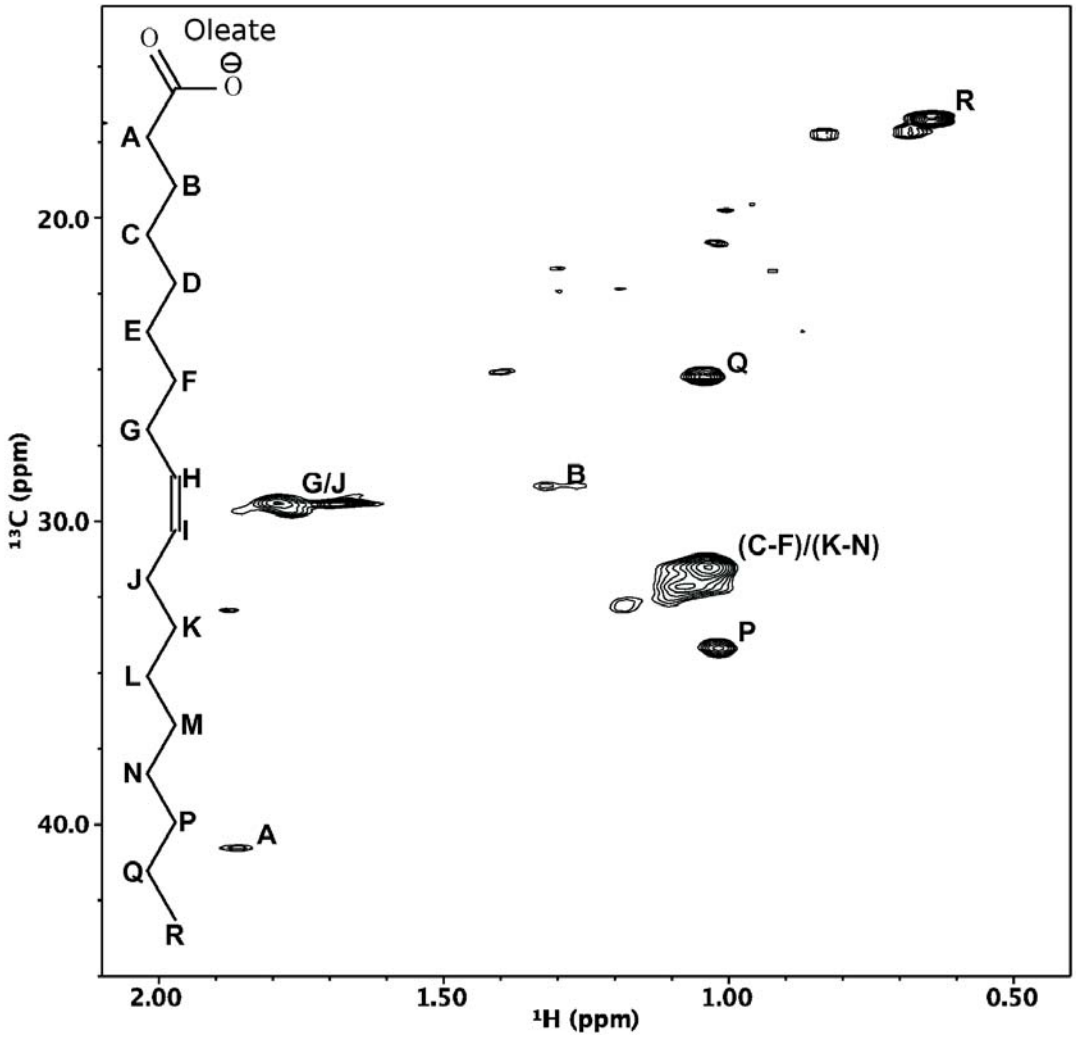

Figure 2.3. Molecular structure and assignment of ^{13}C HSQC spectrum of ^{13}C labeled oleate with unlabeled LFABP in the single-liganded state. (The proton chemical shift of H and I is 5.1 ppm.)

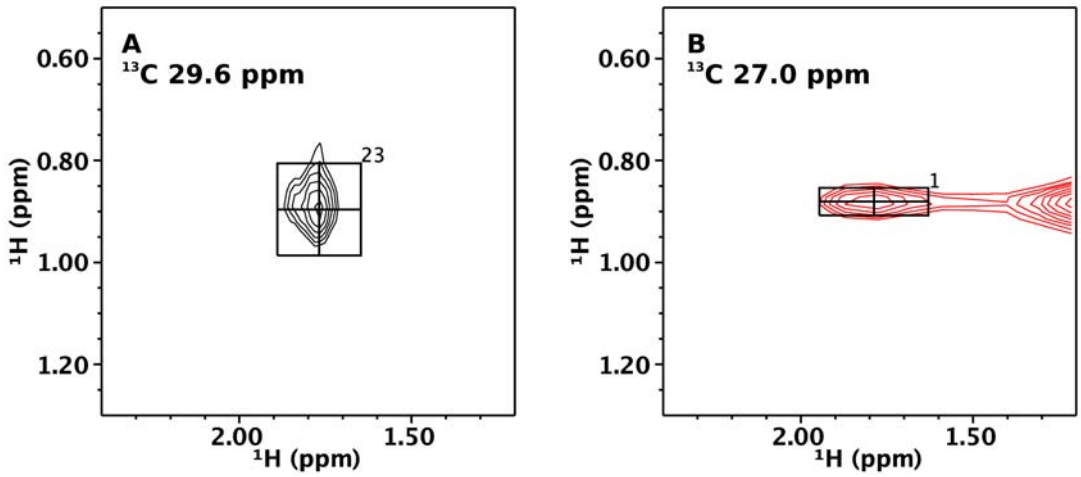

Figure 2.4. Intermolecular NOEs between LFABP and oleate in the singly-liganded state. ^{13}C-filtered NOE spectra of (A) unlabeled LFABP with ^{13}C labeled oleate; (B) ^{13}C labeled LFABP with unlabeled oleate.

2.3.3 Definition of AIRs and unambiguous distance restraints

Ambiguous Interaction Restraints (AIRs) are introduced by HADDOCK as the driving force for the docking process, and defined between residues from the protein/ligand interface with "active" and "passive" residues. Typically the active residues are defined as residues with experimentally observed chemical shift perturbations, while the passive residues refer to residues surrounding these active residues. In our case, active residues of LFABP are defined as those protein sites that are identified from either CSP or intermolecular NOE data in the presence of ligands. Passive residues of LFABP are defined as those residues close to active residues in the three-dimensional structure, which could possibly interact with ligand. The passive residues are chosen from the residues inside the protein cavity that have atoms within a 5 Å range from atoms of the active residues. For oleate the whole molecule was defined as one active residue. An

example of AIRs definition for the binding of singly-liganded LFABP is shown in Figure 2.5. Active residues (red) were defined based on the composite chemical shift perturbation of amide proton and nitrogen (CSP_{HN+N}) nuclei from the apo- to sgl- protein and intermolecular NOE data in sgl-LFABP (Figure 2.4), while passive residues (yellow) are defined as residues surrounding them structurally. As we know the ligand is bound inside the protein cavity in this singly-liganded state, the active and passive residues were filtered so that only those residues inside the cavity were retained in the AIRs definition. In general an upper limit of effective distance is set between any atom *m* of active residue *i* of protein A and any atom *n* of both active and passive residue *k* of protein or ligand B.

$$d_{iAB}^{eff} = \left(\sum_{m_{iA}=1}^{N_{atoms}} \sum_{k=1}^{N_{res}B} \sum_{n_{kB}=1}^{N_{atoms}} \frac{1}{d_{m_{iA}n_{kB}}^{6}} \right)^{(-1/6)}$$

The effective distance is defined using the equation[51] above, where N_{atoms} indicates all atoms of a given residue and N_{res} indicates the sum of active and passive residues for a given protein or ligand. In our case, the ligand is dealt with as a single residue as noted above. Therefore, two types of effective distances are defined: effective distances between every active residue of protein and ligand (protein as A, ligand as B); the effective distance between ligand and all active/passive residues of the protein (ligand as A, protein as B). The restraint for maximum effective distance was set to 4.0 Å for our docking procedures. Beside AIRs, another important driving force we used for docking is unambiguous distance restraints, defined with a maximum distance (carbon-carbon distance) of 8.0 Å between a pair of atoms from protein and ligand based on the intermolecular NOE assignments as discussed in Section 2.3.2.

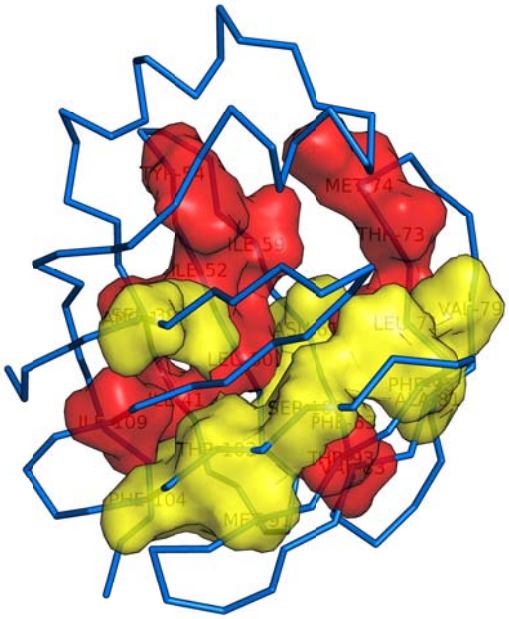

Figure 2.5. Active (red) and passive (yellow) residues of singly-liganded LFABP defined based on CSP and intermolecular NOE data.

2.3.4 Dihedral angle restraints

Dihedral angle restraints were predicted by TALOS+[61] based on the available chemical shift assignments. They were used to preserve the integrity of the secondary structures of protein during the docking procedure.

2.4 DOCKING SETUP

To set up the protein-ligand docking with HADDOCK, we need the molecular coordinates for proteins and ligands, several restraints including AIRs, and definitions of semi-flexible and fully flexible segments.

For molecular coordinates, ensembles of 10 NMR structures of the protein (holo-LFABP (2JU8)[4] with ligand omitted and apo-R122L/S124 LFABP[14]) were used as the starting structures to account for loop flexibility. The oleate coordinates were generated from PRODRG server[62] (http://davapc1.bioch.dundee.ac.uk/prodrg/).

The AIRs and unambiguous distance restraints were defined as discussed in Section 2.3.3. Dihedral angle restraints were predicted by TALOS+[61] using the HN, HA, CA, CB, and N chemical shift assignments to preserve the integrity of secondary structures in the protein. An additional distance restraint with an upper limit of 3.0 Å between the center of mass of protein (C_α only) and oleate (all heavy atoms) was used during the initial rigid body docking stage to make sure that the ligand entered the protein cavity.

To account for conformational changes in the protein, fully flexible segments were defined based on the chemical shift perturbations of the protein backbone (CSP_{HN+N} and CSP_{CA}), while semi-flexible segments were defined to cover all residues defined in AIRs and the fully flexible residues.

The OPLSX nonbonded parameters (slightly modified from the Optimized Potentials for Liquid Simulations (OPLS) force field[63]) with a 8.5 Å distance cut-off from the parallhdg5.4.pro parameter file[64] were used to calculate the nonbonded energies.

During the three stages of the docking process, 4000 structures were generated in the rigid body stage, among which 400 top-scoring HADDOCK structures were refined by following semi-flexible simulated annealing. Finally the 200 top ranking structures from the semi-flexible refinement stage were selected for water refinement, in which the model structures are refined with an explicit solvent layer (8 Å for water)[51].

2.5 Analysis of Docking Results

Model structures calculated after the final water refinement stage were analyzed to choose the best docking results. First the pairwise RMSD (ligand RMSD) matrix over all model structures was generated. Ligand RMSDs were calculated by superimposing the protein backbones first using the PROFIT program (Martin, A. C. R., Profit. http://www.bioinf.org.uk/software/profit).

After the RMSD matrix was generated, it was used to cluster the docking solutions based on the algorithm of Duara.[65] A script searches through the RMSD matrix for groups of structures with a ligand RMSD less than a preset cut-off, placing the group with the maximum number of such structures into cluster1. The same procedure is applied to the remaining structures until the number of structures that are found to fit with each other is less than a pre-set minimum requirement (at least four structures per cluster in our case). Ultimately, more than 50% of the calculated structures must be clustered and each cluster should contain at least 4 structures.

After clustering the structures, all the clusters were ranked based on the average HADDOCK score of their 4 top ranking structures. The best 4 structures were chosen from the cluster with the best average HADDOCK score to represent the docking results. Finally, the quality of the chosen structures was evaluated with PROCHECK[66].

2.6 DOCKING RESULTS

To gain a comprehensive understanding of the binding process of LFABP with fatty acids, docking calculations were conducted for doubly liganded wild type LFABP (wt-LFABP), singly liganded wt-LFABP, and liganded R122L/S124A mutant.

2.6.1 Docking of wt-LFABP with two ligands

Docking of wt-LFABP with two ligands was done to test the feasibility of applying HADDOCK in this LFABP/oleate system, as both NMR and crystal structures of holo-LFABP are available. It was also used to optimize the docking parameters to be used in docking of a single ligand with wt-LFABP and the R122L/S124A mutant.

Docking was set up as discussed in Section 2.4 with AIRs, unambiguous distance restraints, dihedral angle restraints and center of mass restraints. Herein, the two oleate ligands were docked into the protein cavity simultaneously as reported for the docking of liver bile acid binding protein with bile acids[53].

After three docking stages, 200 model structures were generated. 165 of 200 structures were grouped into 10 clusters with a 2.5 Å RMSD cutoff for the two ligands after fitting of the protein backbone. Clusters ranked with the average HADDOCK score of the best 4 structures in the cluster are listed in Table 2.2 together with the ligand RMSDs, van der Waals energy (E_{vdW}), electrostatic energy (E_{elec}) and buried surface area (BSA).

Table 2.2. Cluster statistics for the four best structures from wild-type LFABP docked with two oleate ligands

Cluster#	HADDOCK score	rmsd	N_{struc}	E_{vdwW}[1]	E_{elec}[1]	BSA[1]
clust1	-125.7 ± 0.6	2.1 ± 1.2	78	-45.3 ± 1.0	-242.9 ± 12.9	1700.8 ± 56.1
clust2	-122.0 ± 5.2	1.9 ± 1.1	21	-48.6 ± 2.9	-203.5 ± 22.2	1651.0 ± 69.8
clust5	-114.8 ± 2.0	2.1 ± 1.3	8	-45.1 ± 5.2	-207.7 ± 6.9	1696.5 ± 69.6
clust4	-113.2 ± 4.1	1.5 ± 0.9	15	-45.9 ± 5.1	-212.5 ± 23.7	1619.7 ± 58.5
clust3	-112.5 ± 1.7	1.7 ± 1.0	17	-45.9 ± 2.8	-214.9 ± 12.8	1734.9 ± 28.6
clust7	-107.5 ± 2.5	1.6 ± 1.0	5	-45.6 ± 5.0	-182.9 ± 18.3	1607.7 ± 33.5
clust8	-102.9 ± 8.8	1.9 ± 1.1	5	-38.4 ± 2.7	-208.6 ± 22.6	1655.1 ± 33.4
clust9	-102.7 ± 2.3	1.6 ± 0.9	5	-40.6 ± 1.9	-201.4 ± 16.3	1544.1 ± 37.6
clust6	-100.6 ± 3.0	2.2 ± 1.3	7	-43.0 ± 2.7	-185.9 ± 21.3	1594.8 ± 75.7
clust10	-100.5 ± 5.0	1.8 ± 1.1	4	-44.2 ± 4.8	-166.4 ± 49.0	1605.0 ± 41.8

[1] van der Waals energy (E_{vdW}), electrostatic energy (E_{elec}) and buried surface area (BSA)

It is evident from Table 2.2 that clusters 1 and 2 have similar HADDOCK scores (-125.7 and -122), both of which are much better than other clusters (higher than -115). Comparing the best structures of cluster1 and cluster2, the ligands occupy similar binding sites and adopt similar conformations as shown in Figure 2.6. In both cluster 1 and cluster 2, the first ligand occupies a binding site deep inside the protein cavity and exhibits a 'U' shaped conformation, while the second ligand is bound at a binding site close to the protein surface with an extended conformation. Nonetheless, slight differences in the aliphatic tail orientation of the second ligand can be seen in these structures. Similar results were reported in our previous NMR structural study, which noted two major conformations for the second ligand[4]. This phenomenon is likely to arise from the lack of restraints in the aliphatic tail region of the second ligand, which

may in turn reflect its flexibility. Cluster 1 was used to present the docking results, as it has a greater population and is slightly favored as judged by its lower HADDOCK score.

The reliability of the docking results was checked by comparing the computed structures with results from solution-state NMR. The overlay of the best docking structure with our experimentally determined NMR structure is shown in Figure 2.7. The docked structure aligns very well with the NMR structure, with both ligands bound in the expected positions and adopting similar conformations to those found experimentally. Structural statistics for wild type LFABP docked with two ligands are reported in Table 2.3. Fitting with PROFIT, the best four docking structures in cluster 1 yield a ligand RMSD of 1.7 Å and protein backbone RMSD of 0.7 Å. When compared with the NMR structure, the docking results have a 2.5 Å ligand RMSD and a 1.1 Å protein backbone RMSD, validating this modeling approach as a tool to provide insight into binding sites and ligand conformations in our LFABP–oleate systems. Moreover, the intermolecular energies obtained in the current docking runs are comparable to those reported for an analogous study of chicken liver bile acid binding protein (LBABP) with two molecules of bile acids[53]. LFABP has better values of van der Waals energy (E_{vdW}) and buried surface area (BSA), while LBABP has superior electrostatic energy (E_{elec}), possibly because the bile acids have two more polar -OH groups in addition to the carboxylate present in the oleate ligand of LFABP. Finally, it should be noted that if HADDOCK fitting is attempted without using the unambiguous distance restraints defined from intermolecular NOE data, the conformation of the ligands is poorly defined due to the significant flexibility of oleate. The quality of the model structures is improved in terms of ligand RMSD by including the unambiguous distance restraints.

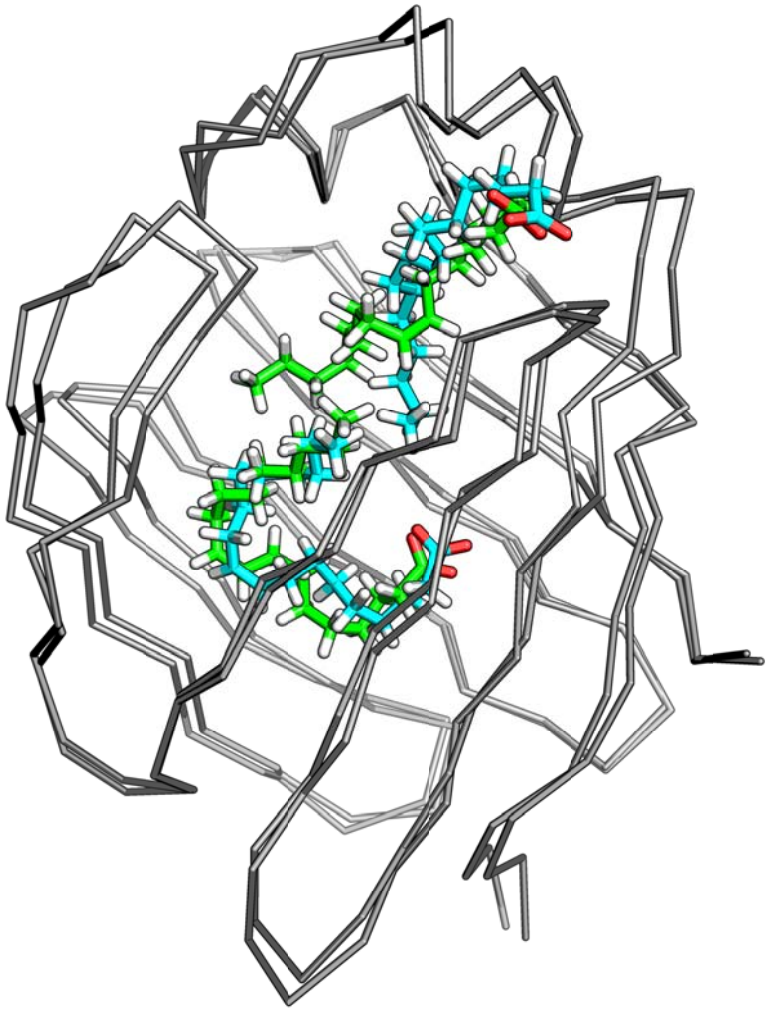

Figure 2.6. Overlay of top ranking structures for cluster 1 (green) and cluster 2 (cyan) of doubly liganded LFABP (dbl-LFABP) docking results.

Table 2.3. Structural statistics for the four best structures of cluster1 from wild-type LFABP docked with two ligands

Ligand RMSD (Å) after fitting protein backbone with respect to		
Best HADDOCK score ranking structure inside cluster	1.72 ±	0.99
NMR structure of dbl-LFABP	2.46 ±	0.12
Backbone RMSD (Å) with respect to		
Flexible interface backbone	0.62 ±	0.24
All backbone	0.70 ±	0.29
NMR structure of dbl-LFABP	1.10 ±	0.32
Number of distance restraints		
Total AIRs	13	
number of AIR violations (>0.3 Å)	1	
Total unambiguous restraints	45	
number of unambiguous restraint violations (>0.3 Å)	0	
Intermolecular energies after water refinement		
E_{vdW} (kcal mol^{-1})	-47.5 ±	1.3
E_{elec} (kcal mol^{-1})	-215.5 ±	25.2
Buried surface area ($Å^2$)	1654 ±	31
RMSD from idealized covalent geometry		
Bonds (Å)	0.004 ±	0.000
Angles (º)	0.55 ±	0.01
Impropers (º)	0.46 ±	0.01
Dihedrals (º)	19.4 ±	1.3
Ramachandran analysis		
Residues in the favored region (%)	89.9	
Residues in the additional allowed region (%)	5.2	
Residues in generously allowed regions (%)	2	
Residues in disallowed regions (%)	2.9	

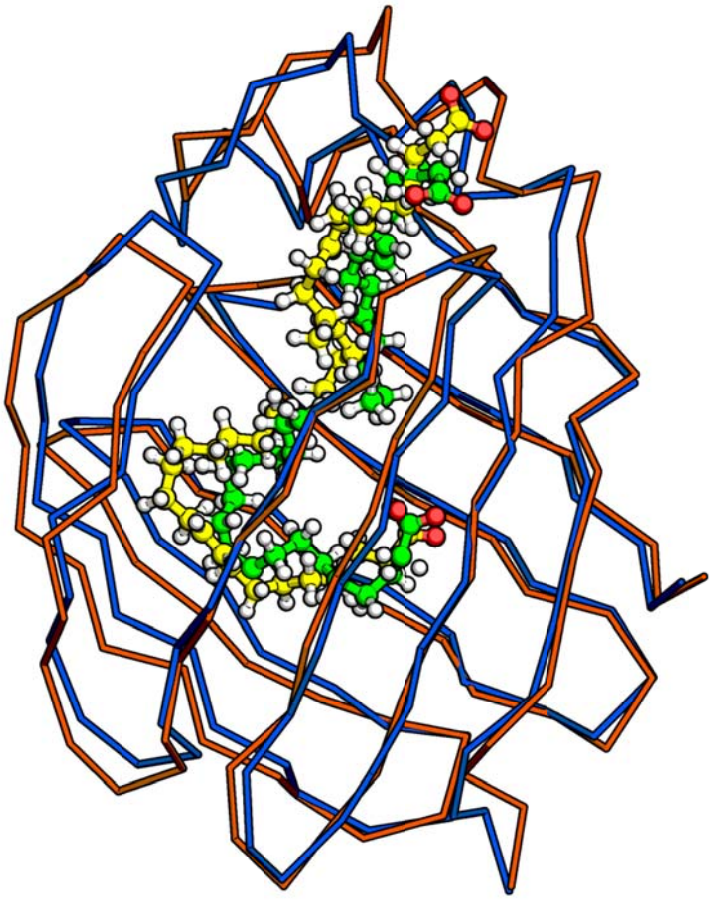

Figure 2.7. Overlay of docking result of doubly liganded LFABP (blue/green) with previously solved NMR structure (orange/yellow)[4].

2.6.2 Docking of wt-LFABP with a single ligand

Docking of wild-type LFABP with one ligand was done similarly to the docking with two ligands described above. AIRs defined based on HN-N chemical shift perturbation from apo- to sgl-LFABP and intermolecular NOEs of sgl-LFABP are listed in Table 2.4 and shown in Figure

2.5. Intermolecular NOEs were also used to define the unambiguous distance restraints as described in Section 2.3.3. Docking results were analyzed as described in Section 2.5. Structural models generated after water refinement were clustered with a 1.8 Å ligand RMSD cutoff and a minimum of four structures per cluster. 135 of 200 structures were grouped into 10 clusters as reported in Table 2.5. Cluster1 was chosen to present the docking results, because it had the best HADDOCK score ranking and was the most populated cluster. The four top HADDOCK-ranked structures of cluster1 are presented in Figure 2.8. The ligands aligned with each other very well among the four structures, displaying a ligand RMSD of 1.3 Å and protein backbone RMSD of 0.7 Å. The structural statistics for single ligand docking are reported in Table 2.6. A ligand RMSD of 3.3 Å was found with respect to the first ligand in the NMR-derived structure of the doubly liganded LFABP.

Table 2.4. AIR definitions based on CSP_{HN-N} and intermolecular NOEs for wild-type LFABP docking with a single ligand

Active		Passive
S39	CSP	N61
I41	NOE	F63
L50	NOE	V79
I52	CSP	A81
Y54	NOE/CSP	M91
I59	NOE	F95
L71	NOE	S100
T73	CSP	F104
M74	NOE	R122
V83	NOE	
T93	NOE	
I109	NOE/CSP	

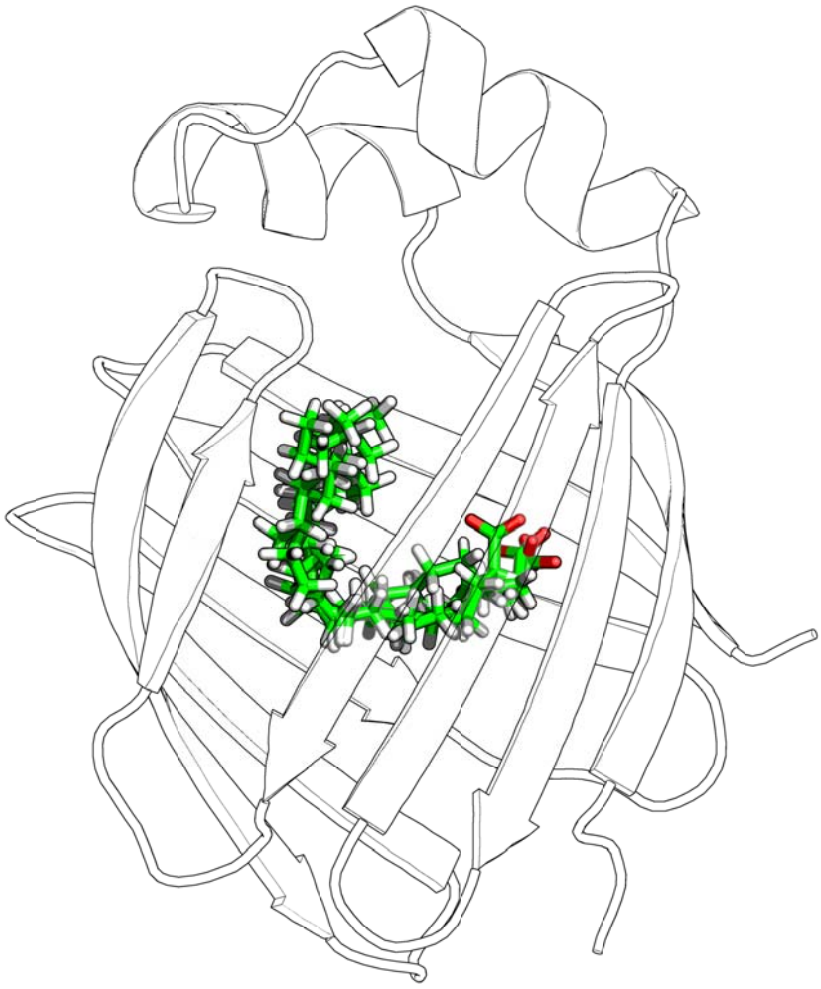

Figure 2.8. Best four structures from sgl-LFABP docking.

Table 2.5. Cluster statistics for the four best structures from wild-type LFABP docked with a single ligand

Cluster#	HADDOCK score	rmsd	N_{struc}	E_{vdW}	E_{elec}	BSA
clust1	-88.4 ± 2.7	1.3 ± 0.7	45	-23.3 ± 2.2	-108.1 ± 5.2	854.8 ± 15.5
clust2	-83.6 ± 0.5	1.2 ± 0.7	35	-20.2 ± 1.2	-107.9 ± 7.8	838.9 ± 7.0
clust6	-80.0 ± 7.3	1.4 ± 0.8	6	-20.5 ± 2.6	-96.6 ± 9.3	860.3 ± 14.3
clust3	-79.7 ± 0.7	1.5 ± 1.0	14	-21.1 ± 2.4	-119.3 ± 12.6	855.5 ± 11.6
clust4	-73.2 ± 1.7	1.4 ± 0.8	11	-22.2 ± 0.6	-115.3 ± 4.3	878.5 ± 21.1
clust7	-71.4 ± 6.0	1.3 ± 0.8	4	-20.2 ± 2.6	-100.8 ± 8.9	856.2 ± 21.3
clust5	-70.1 ± 6.3	1.5 ± 0.9	8	-21.3 ± 3.7	-111.5 ± 6.8	867.1 ± 24.1
clust9	-64.8 ± 2.5	1.4 ± 0.8	4	-21.1 ± 4.4	-92.3 ± 22.0	862.9 ± 19.9
clust8	-63.0 ± 7.7	1.1 ± 0.7	4	-17.6 ± 1.4	-112.0 ± 9.8	839.4 ± 19.2
clust10	-62.4 ± 2.0	1.0 ± 0.6	4	-20.5 ± 1.1	-104.6 ± 3.1	893.4 ± 12.3

Table 2.6. Structural statistics of the four best structures from wild-type LFABP docked with a single ligand

Ligand RMSD (Å) after fitting protein backbone with respect to			
best HADDOCK score ranking structure inside cluster	1.25	±	0.73
first ligand in NMR structure of dbl LFABP	3.34	±	0.47
Backbone RMSD (Å) with respect to			
Mean flexible interface backbone	0.64	±	0.27
All backbone	0.66	±	0.23
NMR structure of dbl LFABP	1.25	±	0.13
Number of distance restraints			
Total AIRs	14		
number of AIR violations (>0.3 Å)	1		
Total unambiguous restraints	20		
number of unambiguous restraint violations(>0.3 Å)	0		
Intermolecular energies after water refinement			
E_{vdW} (kcal mol^{-1})	-23.3	±	2.2
E_{elec} (kcal mol^{-1})	-108.1	±	5.2
Buried surface area ($Å^2$)	854	±	16
RMSD from idealized covalent geometry			
Bonds (Å)	0.004	±	0.000
Angles (º)	0.60	±	0.01
Impropers (º)	0.55	±	0.02
Dihedrals (º)	19.7	±	1.3
Ramachandran analysis			
Residues in the favored region (%)	88.1		
Residues in the additional allowed region (%)	10.4		
Residues in generously allowed regions (%)	0.7		
Residues in disallowed regions (%)	0.9		

2.6.3 Docking of R122L/S124A LFABP with a single ligand

A similar procedure was applied to the docking of the R122L/S124A mutant LFABP with one ligand. AIR definitions are listed in Table 2.7. As we have less information about the binding sites of oleate in the R122L/S124A mutant, all residues that had significant chemical shift perturbation or observed intermolecular NOEs were defined as active residues. Other than AIRs, other parameters for docking were set up in the same way as described for singly-liganded LFABP docking. After docking, 158 of 200 structures were grouped in 10 clusters with a 1.8 Å ligand RMSD cutoff as reported in Table 2.8. The four top HADDOCK-ranked structures in cluster1 were chosen to represent the docking results as done with singly-liganded LFABP. As shown in Figure 2.9, oleate adopts a nearly linear conformation with its carboxylate group close to the portal helical region and its aliphatic tail buried deeply in the protein cavity. Ligands fit with each other very well among the four docked structures, yielding a ligand RMSD of 1.3 Å. Compared with either oleate in the singly-liganded protein or the first ligand of doubly-liganded holo-LFABP, there are clear changes in ligand orientation and conformation for R122L/S124A LFABP. The structural statistics for the docked R122L/S124A LFABP-oleate complex are summarized in Table 2.9.

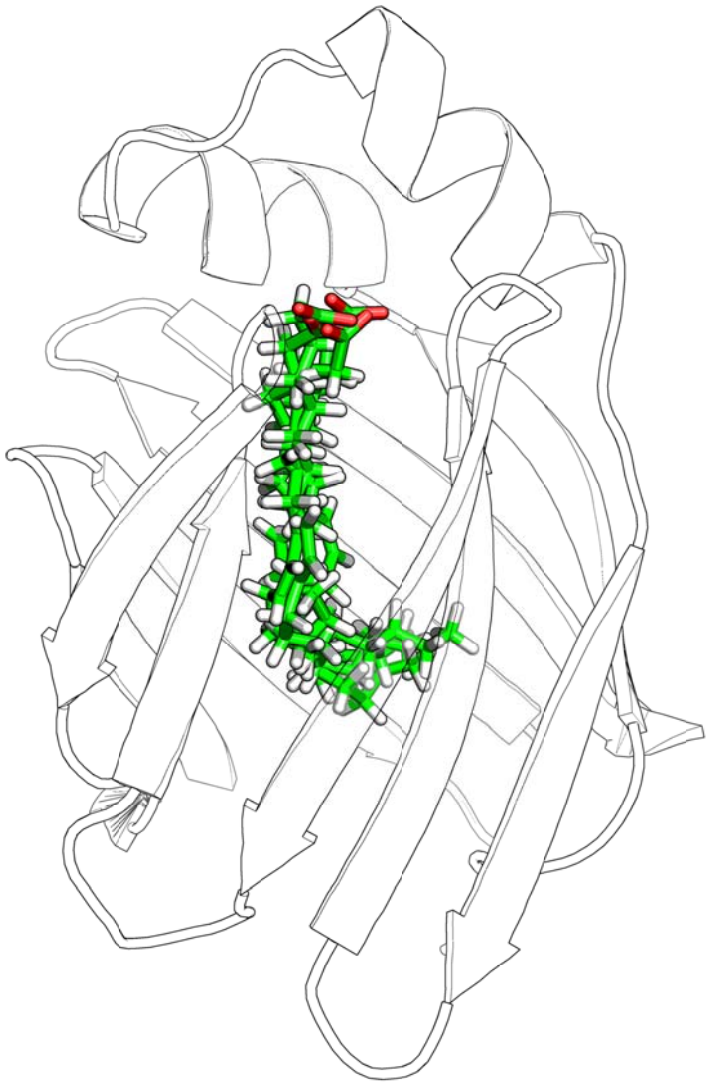

Figure 2.9. Four best docked structures of R122L/S124A mutant with oleate.

Table 2.7. AIR definitions for R122L/S124A LFABP docking with oleate

Active		Passive
V9	NOE	Y7
L24	NOE	Q12
L28	NOE	F15
G32	CSP	L71
S39	NOE	E72
L50	NOE	M74
Y54	CSP	M91
G55	CSP	M113
K57	CSP	Y120
I59	CSP	
F63	CSP	
T75	CSP	
V83	CSP	
T93	CSP/NOE	
F95	CSP	
I98	NOE	
I109	NOE	
N111	CSP	
L122	NOE	
A124	NOE	

Table 2.8. Cluster statistics for the four best structures from R122/S124A LFABP docked with oleate

Cluster#	HADDOCK score	rmsd	N_{struc}	E_{vdW}	E_{elec}	BSA
clust1	-88.7 ± 4.8	1.3 ± 0.8	78	-28.2 ± 2.8	-104.6 ± 11.5	913.8 ± 22
clust3	-73.8 ± 3.7	1.2 ± 0.7	19	-22.6 ± 2.5	-64.1 ± 11.1	879.9 ± 27.9
clust2	-72.7 ± 4.8	1.2 ± 0.7	20	-25 ± 3.5	-77.9 ± 30.5	902.9 ± 16.2
clust4	-70.6 ± 3.3	1.4 ± 0.8	10	-28.2 ± 1.5	-50.6 ± 8.9	900.2 ± 34.5
clust7	-67 ± 11.9	1.4 ± 0.8	5	-21.8 ± 3.1	-84.3 ± 22.8	861.5 ± 34.8
clust5	-65.7 ± 1.7	1.2 ± 0.7	8	-24.8 ± 2.1	-42 ± 24.8	931.3 ± 26.9
clust10	-63.2 ± 6.9	1.3 ± 0.8	4	-22.8 ± 3.6	-69 ± 40.6	877 ± 17.8
clust8	-61.3 ± 7.6	1.2 ± 0.7	4	-25.1 ± 4.1	-62.2 ± 39.6	875.9 ± 46.6
clust9	-59.7 ± 4.7	1.4 ± 0.8	4	-23.5 ± 5.7	-94.5 ± 28.8	854.8 ± 32.9
clust6	-56.2 ± 3.4	1.5 ± 0.9	6	-23.4 ± 3.4	-32 ± 10.5	861.4 ± 69.3

Table 2.9. Structural statistics of the four best structures from R122L/S124A LFABP docked with one ligand

Ligand RMSD (Å) after fitting protein backbone with respect to			
best HADDOCK score ranking structure inside cluster	1.31	±	0.77
Backbone RMSD (Å) with respect to			
Mean flexible interface backbone	0.84	±	0.31
All backbone	0.92	±	0.36
Number of distance restraints			
Total AIRs	25		
number of AIR violations (>0.3 Å)	9		
Total unambiguous restraints	18		
number of unambiguous restraint violations(>0.3 Å)	0		
Intermolecular energies after water refinement			
E_{vdW} (kcal mol^{-1})	-28.2	±	2.8
E_{elec} (kcal mol^{-1})	-104.7	±	11.5
Buried surface area ($Å^2$)	914	±	22
RMSD from idealized covalent geometry			
Bonds (Å)	0.004	±	0.000
Angles (º)	0.84	±	0.03
Impropers (º)	1.09	±	0.02
Dihedrals (º)	21.3	±	1.2
Ramachandran analysis			
Residues in the favored region (%)	86.7		
Residues in the additional allowed region (%)	12.2		
Residues in generously allowed regions (%)	0.2		
Residues in disallowed regions (%)	0.9		

2.6.4 Docking of R122L/S124A LFABP with two ligands

The binding affinity of R122L/S124A LFABP was shown to decrease significantly with the mutation, but the stoichiometry of the R122L/S124A LFABP/oleate complex remained unclear and was deduced indirectly from chemical shift perturbation effects during a stepwise oleate

titration[13]. Both the observation of a single set of signals for oleate in ^{13}C HSQC and the weak binding measured for the first ligand suggest that there is probably no second bound ligand, but exchange averaging of the ligand resonances could undercut this interpretation. The docking of R122L/S124A LFABP with two ligands was undertaken to investigate the capability of the double mutant to accommodate more than one ligand.

As no distinct peaks could be observed for the putative second ligand, the intermolecular NOEs could not be assigned specifically to the first and second ligands. Therefore no unambiguous distance restraints were used for the docking. The AIRs were defined similarly to the docking of R122L/S124A mutant with one ligand, except that the two ligand molecules were defined ambiguously in AIRs to allow the program to choose the best conformation and binding sites for them. All other setups were the same as before.

A majority of the modeled structures generated from this docking calculation were found to accommodate one ligand inside the protein cavity and leave the other ligand exposed to aqueous solution. Clearly it is unacceptable to have the whole molecule of the long-chain fatty acid remaining in the water. As there is no distinction between the two ligands in the modeling setup, we take the docking results to indicate that R122L/S124A LFABP is not capable of binding more than one ligand.

To exclude the possibility that the HADDOCK program is incapable of handling two ligands without unambiguous restraints for the LFABP-oleate system, the docking of wild-type LFABP with two ligands was repeated using similar settings to those described above for R122L/S124A LFABP. No unambiguous distance restraints were defined using NOEs and only chemical shift perturbations were used to define AIRs, with no distinction made between the first and second

ligand. The docking was successful in terms of putting two ligands into the protein cavity, even though the ligand conformations were not well defined. To refine the conformation of the ligands, intermolecular NOE data are still necessary to be included as unambiguous distance restraints. In fact, it has been shown that HADDOCK is capable of positioning two ligand molecules in the cavity of bile acid binding protein[53]. It has also been reported that docking of LFABP with two ligands was successful using GLIDE[67].

2.7 DISCUSSION

2.7.1 Ligand binding of singly-liganded LFABP

Based on previous NMR titration studies, binding sites were proposed for the singly-liganded state. With additional intermolecular NOE data and using the HADDOCK method, we can confirm that in the singly-liganded state the ligand is situated inside the protein cavity and occupies a similar binding site to the first ligand in doubly liganded holo-LFABP.

As shown in Section 2.6.2, model structures of singly-liganded LFABP were generated successfully. Comparing the oleate in the docked structure of singly-liganded LFABP with the first ligand of holo-LFABP in previously determined NMR structures[4], it is clear that the ligand in the HADDOCK model is bound deeply inside the protein cavity with a similar orientation. The ligand conformations are also in close agreement: the ligand adopts a 'U' shaped conformation as shown in Figure 2.10 in both singly-liganded LFABP and the holo-protein. Despite the similarity, a slight difference in ligand binding can be seen. The difference of ligand RMSD is about 3.4 Å between oleate in singly-liganded LFABP and the first ligand in the holo protein, while the ligand RMSD within the chosen cluster is 1.3 Å (Table 2.6). Compared with the first ligand in the holo NMR structure, oleate is bound less deeply in the singly-liganded state and adopts a more extended conformation. This result is reasonable, since without the presence of the second ligand, extra space is available in the protein cavity for the first ligand to "explore". To enable binding of the second ligand, the first ligand must become somewhat more compact to provide additional space. The final arrangement in holo-LFABP looks as if the first oleate is "pressed" by the second oleate, possibly providing a rationale for the lowered binding affinity of the second ligand.

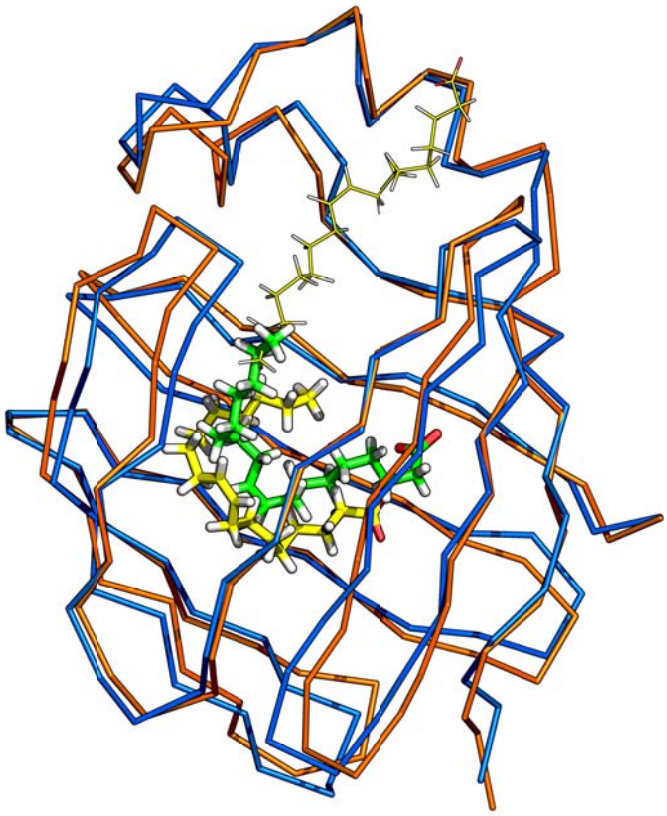

Figure 2.10. Docked structure of wild-type LFABP with single ligand (blue ribbon for protein and green stick for ligand) vs. NMR structure of holo-LFABP (orange ribbon and yellow stick for ligand, with the second ligand is shown as thin stick).

Even though oleate in this intermediate state is bound less deeply in the cavity, there remains space available in the protein cavity for potential binding of an additional ligand. In addition, it has been proposed that binding of the first ligand offers potential favorable hydrophobic contacts for the aliphatic tail of the additional ligand, which may then adopt the specific orientation

observed in the holo-LFABP structures. Contact between the two ligands has been reported in both crystal and NMR structures[4, 6].

2.7.2 Electrostatic interactions in the singly-liganded state

In addition to the ligand conformation, the modeled structures can provide insight into the electrostatic contacts between ligand and protein in the singly-liganded state. As shown in Figure 2.11, the distances between the carboxylate group of the oleate and residues S39, R122 and S124 of wild-type LFABP are short enough (1.7, 1.8 Å) to form hydrogen bonds in the singly-liganded state. Similar electrostatic interactions involving residues S39, R122 and S124 have also been reported in both crystal and NMR structures[4, 6]. It should be noted that although neither R122 nor S124 were defined as active residues in AIRs or involved in unambiguous distance restraints, the program still found them to be in contact with the oleate carboxylate group via electrostatic interactions. As electrostatic interactions are normally stronger than hydrophobic interactions between protein and ligand, it may be proposed that those electrostatic interactions are key determinants of the binding site and orientation of the first ligand. The importance of electrostatic interactions in the ligand binding process has also been implicated by MCCE simulations and supported by site-specific mutagenesis[14]. Specifically, MCCE simulations suggested that residues R122 and S124 promote the ionization of the first ligand. Experimentally, production of the R122L/S124A double mutant showed that the binding affinity drops dramatically for the first ligand in the complex as compared with wild-type LFABP, even though the protein conformations remain similar to each other.[14] The functions of arginine 122 and serine 124 are discussed below in connection with the docking results for R122L/S124A LFABP.

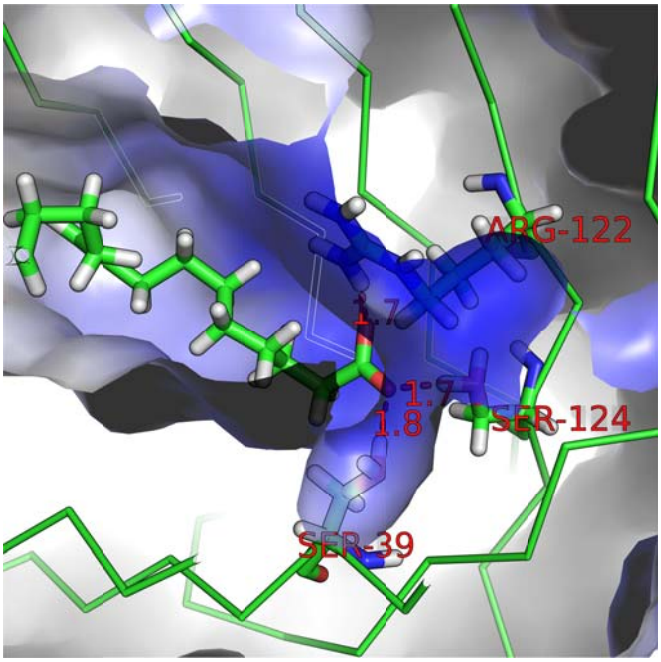

Figure 2.11. Electrostatic interactions between protein and ligand in the singly-liganded state of wt-LFABP. The blue surface represents the positive charge of the protein cavity.

2.7.3 Oleate binding with R122L/S124A LFABP

As mentioned above, the ligand binding properties differ for wild-type LFABP and the R122L/S124A mutant. Although it had been suggested from our prior titration study that the potential binding sites are changed after the two charged and polar residues are replaced by non-polar residues[13], the binding sites and ligand conformation remained provisional without determination of the structure of the protein-ligand complex. With the docked structures of the singly-liganded LFABP and the R122L/S124A complex in hand, we can evaluate how ligand binding is affected structurally by the mutations.

As mentioned in Section 2.6.3 and shown in Figure 2.12, the oleate ligand in the docked R122L/S124A LFABP adopts a near-linear conformation, with its carboxylate group close to the

helices and its aliphatic tail buried deeply inside the protein cavity. It is evident that this orientation of the ligand differs substantially from the first ligand within wild-type LFABP in either holo- or sgl- states (Figure 2.12). The change of orientation upon mutation of the protein supports the hypothesis that orientation of the first ligand in the protein cavity is determined by electrostatic interactions between the carboxylate group of oleate and charged or polar residues R122 and S124 of LFABP. The loss of electrostatic contacts of the carboxylate group in the R122L/S124A mutant compared with wild-type LFABP also corresponds to a less polar protein cavity, which favors hydrophobic interactions. Thus the cavity accommodates the aliphatic tail instead of the carboxylate group.

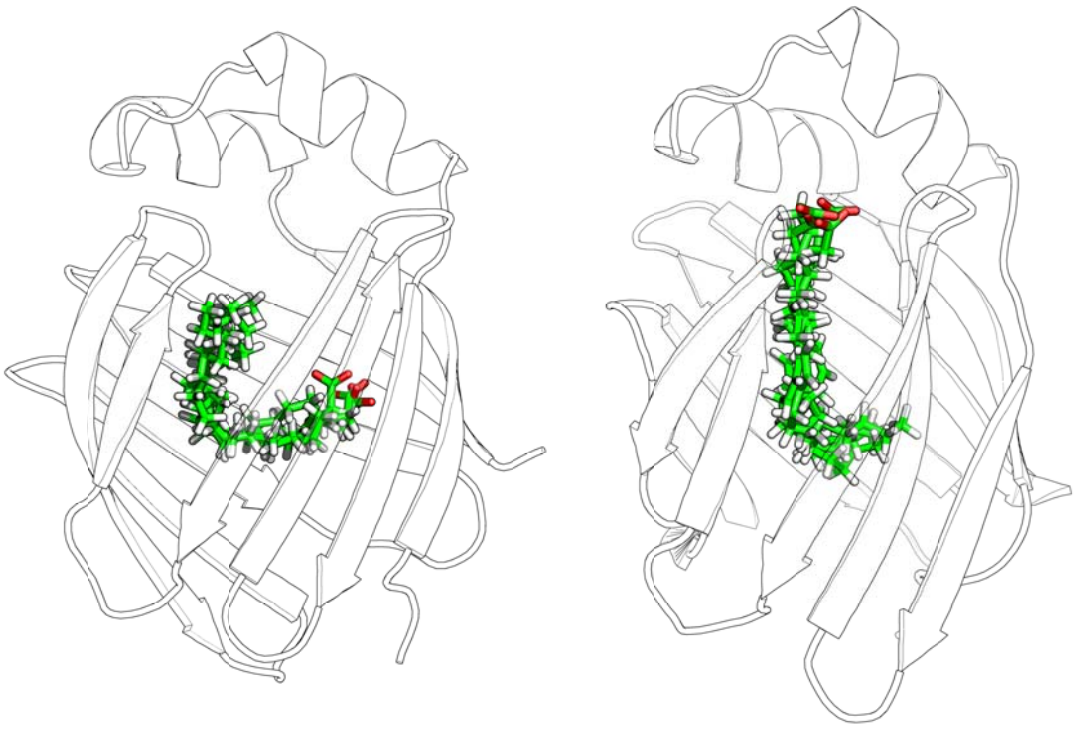

Figure 2.12. Comparison of ligand orientation for singly-liganded LFABP (left) and the R122L/S124A mutant complex (right).

2.7.4 Comparison of the binding process of wt-LFABP and the R122L/S124A mutant

As mentioned above, the conformation and orientation of the first oleate change significantly for singly-liganded LFABP as compared with the R122L/S124A mutant complex. It is clear that the binding properties are altered by the mutation, but to understand these changes and the functions of arginine 122 and serine 124, it was useful to make a close comparison of the modeled structures of wild-type LFABP and the R122L/S122A mutant, with each singly bound ligand.

To evaluate the hypothesis that electrostatic interactions are the key determinants of conformational changes for the ligand in the protein cavity (Section 2.7.2), electrostatic plots of the cavity for both wild-type LFABP and the R122L/S124A mutant with one oleate bound are presented in Figure 2.13. It is clear from the figure that the negatively charged carboxylate group of oleate is situated in a positively charged pocket deep inside the cavity in wild-type LFABP, coinciding with the locations of arginine 122 and serine 124 as pointed out in Section 2.7.2. The replacement of R122 and S124 with non-polar residues in the mutant alters the electrostatic environment of the protein cavity. The positively charged region in the wild-type protein becomes nearly neutral in the R122L/S124A mutant, i.e., the environment is no longer favorable for a negatively charged carboxylate group, but is more amenable to hydrophobic contacts. As a consequence, we expect the negatively charged carboxylate group to be pushed away from the protein cavity and the cavity to be occupied by the aliphatic tail of the oleate. In fact that behavior is seen in the model structure of R122L/S124A mutant. The orientation of the ligand has turned around, with the aliphatic tail in the protein cavity and the carboxylate group staying at the top of the protein cavity, near the positively charged residue lysine 57. These observations further support the hypothesis that electrostatic interactions are the key

determinants of the conformation and orientation of the first ligand. The importance of electrostatic interactions has also been reported in a docking study of bile acid binding protein[53].

Moreover, the conformational change introduced by this electrostatic interaction has a major effect on the potential of the protein to bind more ligands. As seen for the wild-type LFABP (Figure 2.13), there is still room available in the protein cavity after the first ligand is bound (the red outline indicates where the second ligand will be). Besides the available space, the presence of the aliphatic tail of the first ligand reduces the polarity of the protein cavity, making it more suitable for the aliphatic tail of the second ligand. In fact, contact between the aliphatic tails of the two ligands has been deduced from both crystal and NMR structures of holo-LFABP[4, 6]. This reasoning suggests that the binding of the first ligand in wild-type LFABP promotes the binding of the second ligand, in agreement with titration results showing that the ligand binding process is stepwise[4]. When residues R122 and S124 are replaced with nonpolar residues, the less polar protein cavity no longer attracts the carboxylate group, allowing the turn in orientation of the first ligand. In the modeled structure of the R122L/S124A mutant, oleate adopts a near-linear conformation with the carboxylate group at the top of the protein cavity. The new conformation of the first ligand does not facilitate binding of a second ligand: the entrance of an additional ligand through the portal is blocked, and the polar region close to the helices is no longer suitable to accommodate the aliphatic tail of a second ligand. Even though this polar area could provide favorable contact for the carboxylate group of the second ligand, the aliphatic tail would then be exposed to the aqueous solution, an energetically unacceptable arrangement.

The importance of electrostatic interactions was also found in a docking study of LBABP bound to two bile acids[53]. In the modeled structure of this LBABP complex the ligands were positioned in the protein cavity with H-bond/salt bridges formed by charged residues T72/D74/K76 for the

inner ligand and residues T53/R55 for the surface-located ligand. The position and orientation of the bile acids appear to be directed by those electrostatic contacts, similar to the case of the first oleate in the LFABP cavity discussed above.

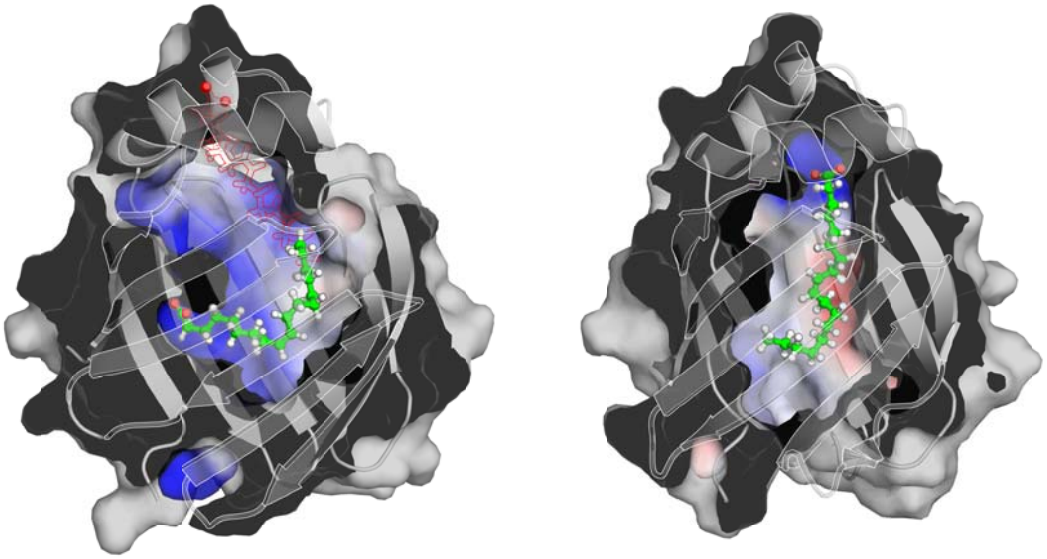

Figure 2.13. Comparison of wild-type LFABP (left) with the R122L/S124A mutant when bound to one oleate molecule. The red outline in the wild-type LFABP structure delineates a potential binding site for the second ligand. (For the protein cavity, blue and red represent positive and negative charged surfaces, respectively.)

2.8 SUMMARY

Modeled structures for singly-liganded LFABP and R122L/S124A LFABP have been generated with a docking calculation that incorporates both NMR chemical shift perturbation and intermolecular NOE data. The model structures provide a better understanding of the binding process of LFABP with long-chain fatty acids. As reported previously, binding is stepwise rather than simultaneous. Based on the analysis of the model structures of wild-type LFABP and R122L/S124A LFABP, the binding process is proposed to occur as follows: the binding of the first ligand is induced mainly by electrostatic interactions between the carboxylate group of oleate and a positively charged pocket containing residues arginine 122 and serine 124. The orientation and conformation of the first ligand are directed by these electrostatic interactions. Furthermore, the binding of the first ligand does not occupy all the space in the protein cavity. The available space together with the more hydrophobic interface provided by the aliphatic tail of the first ligand facilitate the binding of a second ligand with an orientation in which the aliphatic tail is buried in the protein cavity and the carboxylate group is close to the protein surface.

CHAPTER 3 PHOSPHOLIPID ORGANIZATION OF ISOTROPICALLY TUMBLING BICELLE SYSTEMS AND APPLICATION TO PROTEIN MEMBRANE INTERACTION

(A portion of this chapter has been published as "Assessing the size, stability, and utility of isotropically tumbling bicelle systems for structural biology." Wu, H.; Su, K.; Guan, X.; Sublette, M. E.; Stark, R. E., *Biochim. Biophys. Acta* (2010), 1798 (3), 482-8.)

3.1 INTRODUCTION

Bilayered phospholipid micelles (bicelles) have drawn substantial attention as membrane mimetics for NMR structural studies of membrane-associated amphiphiles and proteins since the 1980s[15]. The advantage of bicelles as membrane mimetics over micelles is that they maintain planar bilayers that closely mimic in vivo membranous structures and thus better preserve the native conformation of the polypeptide under study. The phospholipid assemblies with disklike[68, 69] or 'Swiss cheese' morphologies[16, 70] are formed due to the segregation of the two lipid species, with long-chain lipids in the central planar region and short-chain lipids in the curved rim region[68, 69, 71]. Depending on the ratio of long- to short-chain lipids (q), bicelles may tumble isotropically or be aligned with respect to a strong magnetic field. Generally, bicelles will orient in the magnet when q is larger than 3[16, 17] and tumble isotropically when q is less than 1.

Dilute solutions of high-*q* bicelles have been used widely as aligned media to measure the residual dipolar couplings of soluble proteins for NMR-based structure refinement[15-17]. Solid-state investigations[72-75] have typically made use of more concentrated preparations of these bicelles, for which the direction of magnetic alignment can be tuned by adding lanthanide ions[76, 77] that increase the spectral dispersion. In a number of cases the low-*q* isotropically tumbling

bicelles have made it possible to conduct high resolution solution-state NMR studies of peptides[18] and transmembrane proteins[19-21] without compromising their biochemical activity; the small size and rapid tumbling of these assemblies permits acquisition of well-resolved NMR spectra, comparable to those seen in micelles[19]. A bicelle organizational model has also been proposed for mixtures of long- and short-chain phospholipids, e.g. dimyristoylphosphatidylcholine (DMPC) and dihexanoylphospha-tidylcholine (DHPC), respectively. In the bicelle thus produced, the two lipid species tend to segregate, with the long-chain phospholipids located in a central planar region and the short-chain phospholipids concentrated in a curved rim region[68, 69, 71] (Figure 3.2). One interesting feature of this system is that the two bicellar regions experience different magnetic environments and thus display distinct peaks in the ^{31}P NMR spectrum[68].

DMPC and DHPC are commonly used lipids for bicelle media by many investigators. The long term stability of the DMPC-DHPC bicelles has been questioned because their carboxy-ester bonds are susceptible to acid- and base-catalyzed hydrolysis in aqueous solution[16, 22]. Ether-linked lipids such as 1,2-di-O-hexyl-sn-glycero-3-phosphocholine (DIOHPC) and 1,2-di-O-tetradecyl-sn-glycero-3-phosphocholine (DIOMPC) have been introduced to enhance the stability of bicelles because they are more stable than ester-linked lipids DHPC and DMPC to acid- or base-catalyzed hydrolysis over a pH range from 2 to 10[23].

Fatty acid-binding proteins (FABPs) are a family of intracellular proteins which bind long-chain fatty acids (FAs) and other hydrophobic ligands. FABPs act as FA transporters, transferring FAs to and from membranes and playing a role in the regulation of fatty acid metabolism.[1] Despite amino acid sequence homology that ranges between 20 and 70% among FABPs, all of these proteins show a high degree of similarity with respect to their tertiary structures, sharing

structural features including a cavity formed by β-strands and a portal region consisting of two short α-helices located between the first and second of these β-strands[3-10].

It has been proposed that FABPs transfer fatty acids to membranes via either direct collisional or diffusional mechanisms[24-27]. The FABPs and the acceptor membrane are in direct contact for the collision-mediated process, which occurs for intestinal FABP (IFABP), adipose FABP (AFABP), heart FABP (HFABP), kidney FABP (KFABP), and cellular retinol-binding protein type I (CRBP I)[27], whereas diffusion of ligand through the aqueous phase and subsequent association with the acceptor membrane occurs for liver FABP (LFABP) and cellular retinol-binding protein type II (CRBP II)[24, 27-31], as shown in Figure 3.1.

It has been reported that the portal helices are critical for the collisional process of fatty acid transfer, based on mutagenesis studies for HFABP, AFABP and IFABP[28, 78, 79]. Positively charged lysine residues present in various helical regions were proposed to be important for regulating interactions of biological membranes with AFABP and HFABP[28, 79]. Ionic interactions between IFABP and the membrane have been reported to be important by acetylation of surface lysine residues[80]. The α-helical domain has also been shown to be responsible for the diffusion-mediated transfer mechanism of LFABP[81]. Finally, the α-helical domain of IFABP was shown to be crucial for the collisional fatty acid transfer mechanism by a study of engineered α(I)LβIFABP[78, 81].

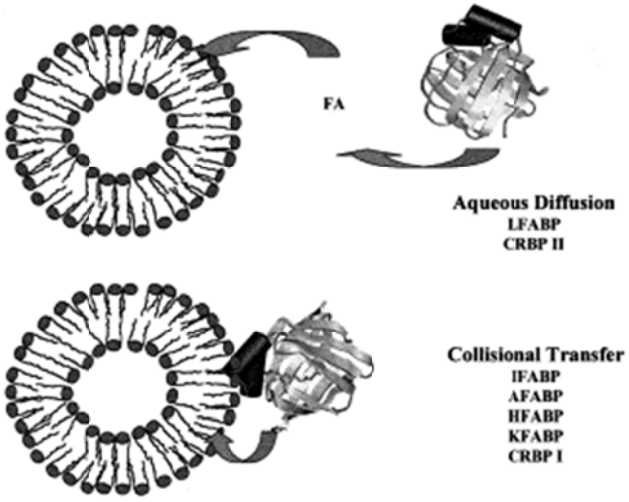

Figure 3.1 Transfer of fatty acids from FABPs to membranes occurs by two different mechanisms (Storch and Thumser 2000[1]).

Herein, we evaluate a new isotropically tumbling phospholipid system as a membrane- mimetic medium for studies of iLBPs and other peripheral membrane proteins. We investigate the stability, aggregate size, and phospholipid organization for a DMPC-DIOHPC bicelle system using NMR methods. The new system is found to be resistant to changes in size and morphology over a wide range of concentrations, temperatures and physiological salt concentrations, and stable enough to serve as a membrane-mimetic medium for NMR structural studies of peripheral membrane proteins. The morphology of the DMPC-DIOHPC bicelle system is investigated and monitored under various conditions by exploiting its well separated ^{31}P NMR peaks, which are observed because of the intrinsic chemical environment difference of lipids in the central planar and rim regions. The application of the DMPC-DIOHPC bicelle as a membrane mimetic is also explored by monitoring the interactions with each of two FABPs. The results support the formation of a stable isotropically tumbling bicellar complex which offers

potential to yield high-resolution NMR spectra for structural studies of peripheral membrane proteins. Titrations of LFABP and IFABP with bicelles were monitored by chemical shift perturbations of the ^{1}H-^{15}N HSQC spectra. Mechanisms of the membrane interactions with LFABP and IFABP were differentiated according to the magnitude and distribution of chemical shift perturbations, and site specific information about the collisional interactions was obtained.

3.2 MATERIALS AND METHODS

3.2.1 Preparation of bicelle samples

1,2-dimyristoyl-sn-glycero-3-phosphocholine (14:0) (DMPC, MW: 677.93), 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (6:0) (DHPC, MW: 453.25), 1,2-di-O-tetradecyl-*sn*-glycero-3-phosphocholine (14:0) (DIOMPC, MW: 649.97) and 1,2-di-O-hexyl-*sn*-glycero-3-phosphocholine (6:0) (DIOHPC, MW: 425.55) were purchased from Avanti Polar Lipids (Alabaster, AL) and used without further purification. Research grade NaCl, KCl, $CaCl_2$ and $MgCl_2$ were purchased from Fisher Scientific (Fair Lawn, NJ).

We compared properties of three bicellar systems, DMPC-DHPC, DIOMPC-DHPC, and DMPC-DIOHPC. To prepare a DMPC-DIOHPC bicelle stock solution with c_L = 25% (w/v of total lipid), the appropriate amount of DMPC was suspended in deionized water and then vortexed at room temperature to form a slurry. An aqueous DIOHPC stock solution (0.5g/ml in water) was mixed with the DMPC slurry to achieve q = 0.5, and water was added to maintain an overall lipid concentration of c_L = 25% (w/v). The samples were vortexed, centrifuged and vortexed again, then heated briefly at 37 °C in a water bath and cooled down to 0 °C in ice water. This protocol was repeated for at least 5 cycles until a clear, homogeneous bicelle solution was obtained. DIOMPC-DHPC and DMPC-DHPC bicelles were handled in the same fashion as the DMPC-DIOHPC mixtures.

Samples with q = 0.5 and a series of c_L values were prepared for NMR studies by adding appropriate amounts of the 25% (w/v) bicelle stock solution and D_2O, then diluting with water to a final volume of 500 µl and 10% D_2O. Selected samples contained various salts (NaCl, KCl, $CaCl_2$ and $MgCl_2$) that were added to yield final concentrations of 50, 100, 150, and 200 mM.

3.2.2 Preparation of Intestinal and Liver Fatty Acid-binding Proteins

Rat intestinal fatty acid-binding protein (IFABP) and rat liver fatty acid binding protein (LFABP) were expressed and purified as described previously[13, 82]. ^{15}N-enriched apo-IFABP and apo-LFABP solutions were prepared in 50 mM NaH_2PO_4, 100 mM NaCl, 5μM EDTA, 0.02% NaN_3, and 5% D_2O at pH 7.0; a 25% (w/v) DMPC-DIOHPC stock bicelle solution in pure water was added directly to the protein solution to achieve a final bicelle concentration of 0.25~2.0% in a total NMR sample volume of 500 μl.

3.2.3 Nuclear Magnetic Resonance Spectroscopy

The ^{31}P NMR data were acquired on Varian spectrometers at either the College of Staten Island or the City College of New York: a ^{UNITY}INOVA 600 instrument (Palo Alto, CA) equipped with an IDQG probe and operating at ^{1}H and ^{31}P frequencies of 599.497 and 242.856 MHz, respectively, or a VNMRS 600 spectrometer equipped with an IDQG probe and operating at 599.761 MHz. 2D ^{1}H-^{15}N Heteronuclear Single Quantum Correlation NMR Spectroscopy (HSQC)[83], 3D ^{15}N-edited 3D Total Correlation SpectroscopY (TOCSY-HSQC) and ^{15}N-edited 3D Nuclear Overhauser Effect SpectroscopY (NOESY-HSQC)[84] data were acquired on either a 600 MHz or 700 MHz Bruker Avance spectrometer equipped with a cryoprobe at the New York Structural Biology Center. One-dimensional ^{31}P spectra were recorded at various temperatures using a proton-decoupled single pulse experiment with 32 – 1024 scans for total lipid concentrations of c_L = 15% to 0.125%. The spectra were recorded with 4096 complex points and a sweep width of 4882.8Hz (~20 ppm), producing a digital resolution of < 0.01 ppm/point or

2.39 Hz/point. A solution of 85% H_3PO_4 was used as an external reference corresponding to a ^{31}P chemical shift of 0 ppm. The data were processed with Varian VnmrJ software. Two-dimensional ^{1}H-^{15}N HSQC spectra were recorded with 512x128 complex points; 3D TOCSY-HSQC and NOESY-HSQC were recorded with 512x32x96 complex points. The ^{1}H carrier frequency was set at the water peak (referenced to 4.629 ppm at 37 °C). 2D and 3D data were processed with NMRpipe[57] and analyzed with NMRViewJ[58].

3.2.4 Modeling of Bicelle Size

The bicelle organizational models proposed previously[17, 85] were used to relate the phospholipid ratio in the aggregate to its diameter, allowing us to assess the robustness of the DMPC-DIOHPC bicelle under various experimental conditions. Figure 3.2 shows a schematic representation of this bicelle system, where the long-chain phospholipids occupy the central planar bilayer region and the short-chain phospholipids are present in the curved micelle-like rim[68, 76]. Straightforward geometrical considerations provide a relationship between the molar ratio of the two components and the dimensions of the bicelle assembly:

q = [DMPC] / [DIOHPC] = (area of central plane) / (area of rim) = $R^2/\{(\pi R + 2r)r\}$

(1), where R is the radius of the central planar region and r is the radius of the rim area. The thickness of the DMPC bilayer h has been estimated as about 4 nm[86]. As the radius of the DHPC micelles has been estimated to be 2 nm[76], we use this value as an approximation for the ether-linked DIOHPC case.

The model was refined by accounting for both the partitioning of free-state DHPC monomers ($[DHPC]_{free}$) and DHPC molecules in the bicelle rim, and the distinctive head group areas of the

differently located phospholipids[87]. These considerations lead to an effective q value for our DMPC-DIOHPC bicelle:

q_{eff} = [DMPC]/([DIOHPC]$_{total}$ – [DIOHPC]$_{free}$) = [DMPC] / [DIOHPC]$_{bound}$

= a^{-1}(the area of central plane)/(the area of rim)

$$= R^2/\{a(\pi R + 2r)r\} \quad (2),$$

where a is the ratio of DMPC to DIOHPC (or DHPC) head group areas in planar bilayers and rim micelles, respectively. The head group area of bilayered DMPC has been reported as 0.60 nm^2 [88], whereas for micellar DHPC it has been reported variously as 0.66 nm^2 [89] and 1.02 nm^2 [90]. In conformance with other researchers[68, 91], our calculations use the latter value and thus an a value of 0.6, making the assumption that DHPC and DIOHPC have the same head group area.

Finally, equation (2) yields a bicelle diameter as follows:

$$D_{\text{bicelle}} = 2(R + \text{r}), \text{ where } R = aq_{eff}\left[\pi + \left(\pi^2 + 8/aq_{eff}\right)^{1/2}\right]/2. \quad (3).$$

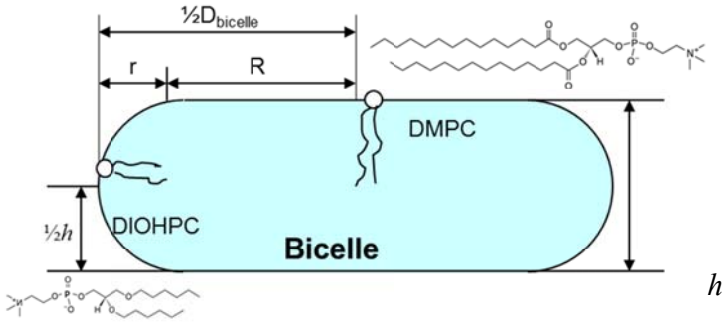

Figure 3.2 Schematic representation of DMPC-DIOHPC (or DMPC-DHPC) bicelles.

3.3 MORPHOLOGY STUDY OF THE DMPC-DIOHPC BICELLE

3.3.1 ^{31}P NMR spectra of bicelle systems

The chemical environment differences between bilayer and rim regions produce two peaks in the ^{31}P NMR spectrum of the q = 0.5 DMPC-DHPC bicelle system[68], which allow us to assess the morphology of the bicelle under various conditions. For the isotropically tumbling low-q DMPC-DHPC bicelle system, the spectrum typically shows the overlap of the two peaks (Figure 3.3 (b)), which presents an obstacle to deducing bicelle size from peak integrations[68]. Figure 3.3 shows the ^{31}P NMR spectra of four types of bicelle systems: (a) DMPC-DIOHPC, (b) DMPC-DHPC, (c) DIOMPC-DHPC, and (d) a 1:1 mixture of (a) and (c). Peaks in the ^{31}P NMR spectra were assigned based on the integrals of their respective peak areas. A summary of the chemical shift assignments appears in Table 3.1. It should be noted that the chemical shifts of a given lipid in the different bicelle systems are consistent with each other. For example, chemical shifts of DMPC remain between -0.160 to -0.156 ppm in samples (a), (b) and (d), indicating that DMPC exists in the bilayer region in each of these three bicelle samples. Analogously, DHPC should be located in the rim area of the bicelle.

The chemical shift difference between DMPC and DHPC is mainly induced by the segregation of phospholipids in planar bilayer or rim regions as reported[68]. For the ether-linked lipids, DIOMPC and DIOHPC, an additional contribution to the chemical shift difference is introduced by the different chemical environment introduced by replacing ester bonds with ether bonds in either of the lipids. It is also observed that the chemical shift difference between DIOMPC and DIOHPC is consistent and similar to the difference between DMPC and DHPC, about 0.08 ppm, implying that ether-linked lipids DIOMPC and DIOHPC exist in similar aggregation states to ester-linked lipids DMPC and DHPC in the bicelle systems, respectively. In the bicelle systems,

the short-chain lipids DHPC and DIOHPC are always present in the rim region and long-chain lipids DMPC and DIOMPC are always present in the central planar bilayer region.

As mentioned above, the peaks of DMPC-DHPC (b) overlap with each other, while the bicelles with ether-linked lipids, (a) and (c), display improved spectral separation. The largest separation is observed for the DMPC-DIOHPC bicelle system (c), which combines the effects of differences in aggregation states and chemical structures. The observed separation indicates that the difference between chemical environments of long-chain lipid and short-chain lipid was amplified by using the ether-linked DIOHPC to replace ester-linked DHPC. The large separation of peaks in the DMPC-DIOHPC bicelle system enables us to obtain accurate integrals for bicelle size calculations. It is also evident in Figure 3.3 (d) that the chemical shifts of the lipids are retained in the mixed sample of DMPC-DHIOHPC (a) and DIOMPC-DHPC (c).

Table 3.1 ^{31}P Chemical shift analysis of phospholipids in different bicelle systems

Chemical Shifts in ^{31}P NMR (ppm)	δ_{DIOHPC}	δ_{DIOMPC}	δ_{DHPC}	δ_{DMPC}	Difference between rim and bilayer region		Difference between ester- and ether-linked lipids	
					$\Delta\delta_{DHPC-DMPC}$	$\Delta\delta_{DIOHPC-DIOMPC}$	$\Delta\delta_{DIOMPC-DMPC}$	$\Delta\delta_{DIOHPC-DHPC}$
(a) DMPC-DIOHPC	0.187			-0.159	0.078-0.088	0.077-0.082	0.263-0.270	0.259-0.270
(b) DMPC-DHPC			-0.081	-0.160				
(c) DIOMPC-DHPC		0.110	-0.072					
(d) Mixture*	0.189	0.107	-0.073	-0.156	~0.08		~0.26/0.27	

*1:1 Mixture of samples (a) and (c)

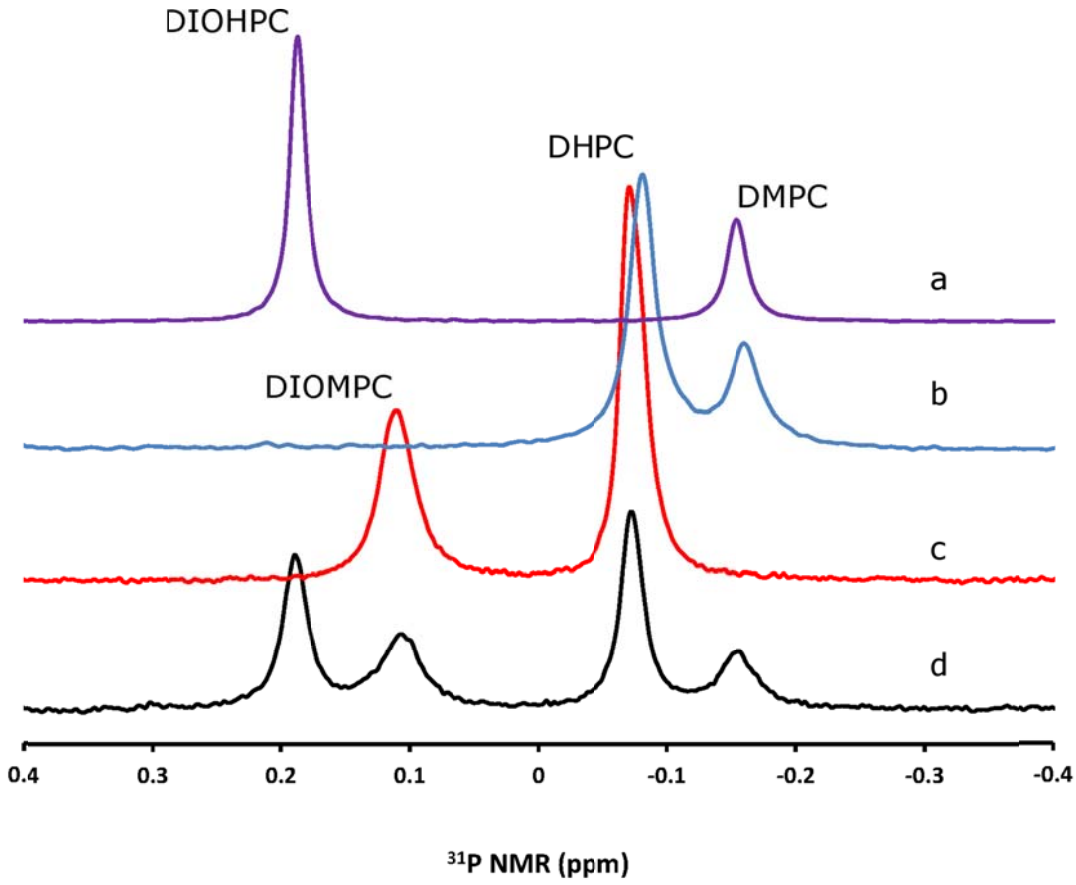

Figure 3.3 Proton decoupled ^{31}P NMR spectra of (a) DMPC-DIOHPC, (b) DMPC-DHPC, (c) DIOMPC-DHPC, and (d) a mixture of samples (a) and (c) in a 1:1 ratio. All bicelle samples had values of q = 0.5, c_L = 5% (w/v) and were dissolved in deionized water to which 10% D_2O was added for NMR spectroscopy conducted at 37 ^oC.

3.3.2 Long term stability of the DMPC-DIOHPC bicelle system

As mentioned above, the use of DMPC and DHPC has raised issues about long term stability due to the susceptibility to acid- and base-catalyzed hydrolysis of their carboxy-ester bonds[16, 22]. Ether-linked lipids DIOHPC and DIOMPC have been used to replace the less stable ester-linked lipids DHPC and DMPC to improve the stability of bicelles against hydrolysis over a wide pH range[23]. For the DMPC-DIOHPC bicelle, the ether-linked DIOHPC is used to replace the ester-linked DHPC, focusing on overcoming the hydrolysis of lipids in the more aqueous-accessible rim region of the phospholipid aggregate. The ester-linked DMPC molecules are expected to be less water accessible, as they are packed more tightly in the central bilayer area compared with the DIOHPC in the rim and are shielded from water by the DIOHPC. It has also been suggested that DHPC experiences more hydrolysis than DMPC in bicelle solutions[16]. As the more aqueous-accessible rim region of the DMPC-DHPC bicelle is replaced by DIOHPC, which is resistant to chemical breakdown, we expect the hydrolysis of lipids in the DMPC-DIOHPC bicelle to be retarded compared with the DMPC-DHPC bicelle. Successive NMR examination of the q = 0.5, c_L= 2% w/v and q = 0.5 c_L = 8% w/v DMPC-DIOHPC systems (3 months at room temperature, 2 years in a freezer at -20^{o}C) gave ^{31}P spectra that were unchanged and free from degradation products. Thus their chemical stability is at least comparable to the DMPC-DHPC bicelle system under similar conditions[16]. In our hands, hydrolysis products were observed by ^{31}P NMR for some DMPC-DHPC bicelle samples within a short time (~ 1 week). In addition to the improvement in chemical stability of DMPC-DIOHPC compared with the DMPC-DHPC bicelle system, using DMPC instead of ether-linked DIOMPC saves about 10 times the cost in chemicals.

3.3.3 Concentration of short-chain lipid in solution

The q_{eff} and aggregate size for dilute bicelle solutions are affected by the partition of short-chain lipids between monomeric and bicellar locations, properties that have been addressed for the DMPC-DHPC bicelle system by determining the monomer concentration of DHPC by ^{31}P NMR analysis[68]. A similar approach is adopted here for the DMPC-DIOHPC bicelle system. The observed chemical shift of DIOHPC can be expressed using the following equation[68].

$$\delta_{obs} = (\delta_{free} - \delta_{bicelle}) \frac{[DIOHPC]_{free}}{[DIOHPC]_{total}} + \delta_{bicelle}$$

where $[DIOHPC]_{total}$ and $[DIOHPC]_{free}$ are the total concentration and monomer concentration of DIOHPC, and δ is the chemical shift. The chemical shifts of DMPC and DIOHPC as a function of total lipid concentration are presented in Table 3.2. The chemical shift of DIOHPC remains constant in highly dilute solutions (0.25 and 0.125%), so that value may be taken as the chemical shift for the DIOHPC monomer, δ_{free} = 0.373 ppm. As δ_{free} , $\delta_{bicelle}$ and $[DIOHPC]_{free}$ should be constant, a linear relationship between δ_{free} and $1/[DIOHPC]_{total}$ is expected, as verified in Figure 3.4. Based on the slope and intercept of the regression line, $\delta_{bicelle}$ and $[DIOHPC]_{free}$ can be deduced as 0.166 ppm and ~6 mM.

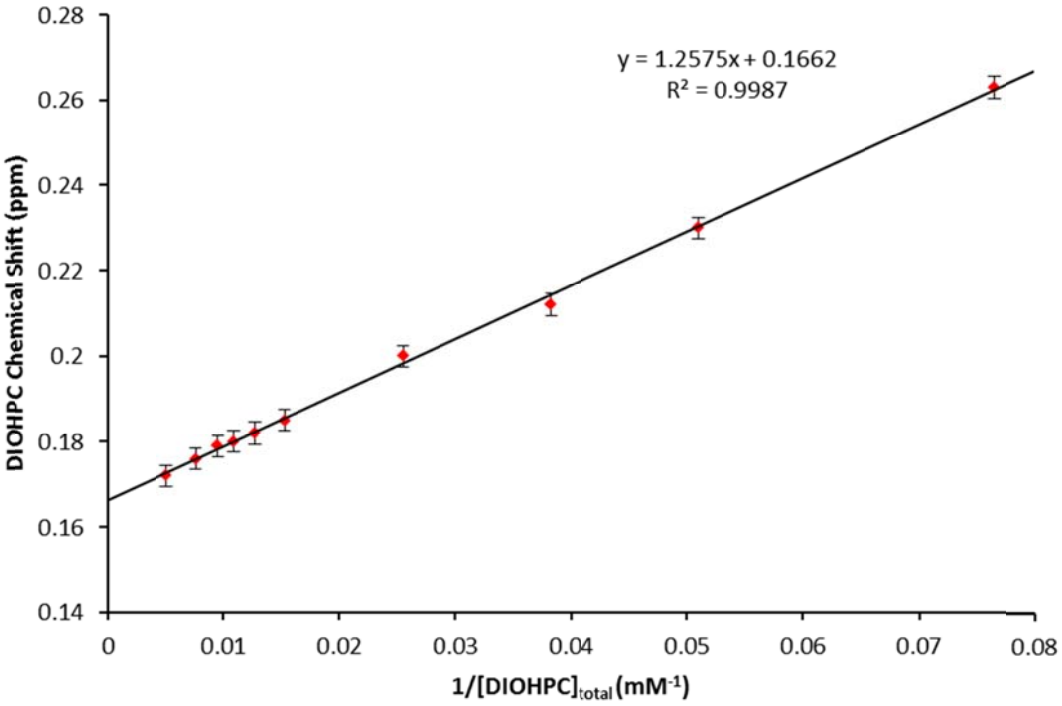

Figure 3.4 Variation of chemical shift as a function of the inverse of total DIOHPC concentration ($1/[DIOHPC]_{total}$). Errors were estimated as the line width at half height of the ^{31}P signal.

Table 3.2 Average ^{31}P chemical shift of lipids in a DMPC-DIOHPC bicelle as a function of total lipid concentration

c_L (%)	Total lipids (mM)	DMPC (mM)	DMPC δ (ppm)	DIOHPC (mM)	DIOHPC δ (ppm)
15	294.3	98.1	-0.155	196.2	0.172
10	196.2	65.4	-0.156	130.8	0.176
8	157.0	52.32	-0.155	104.64	0.179
7	137.3	45.78	-0.155	91.56	0.18
6	117.7	39.24	-0.156	78.48	0.182
5	98.1	32.7	-0.155	65.4	0.185
3	58.86	19.62	-0.156	39.24	0.2
2	39.24	13.08	-0.156	26.16	0.212
1.5	29.43	9.81	-0.158	19.62	0.23
1	19.62	6.54	-0.163	13.08	0.263
0.25	4.905	1.635		3.27	0.373
0.125	2.453	0.8175		1.635	0.373

3.3.4 Variation of $[DIOHPC]_{free}$ with temperature and total lipid concentration

To evaluate whether the bicelle size is maintained over a range of temperatures that might be used in structural studies of biomolecules, the temperature dependence of $[DIOHPC]_{free}$ was evaluated by additional ^{31}P NMR measurements made at 30 °C, 41 °C and 46 °C. The modest changes in ^{31}P chemical shift separations (no more than 0.025 ppm) argued for complete segregation of DIOHPC at the rim rather than miscibility in the planar region[92-94], allowing us to determine the concentration of free short-chain lipid and bicelle size via Eqs. (2) and (3). The

concentration of short-chain phospholipid monomers remained essentially constant at 5.7–6.5 mM over this temperature range (data not shown), in accord with prior reports for DHPC. For instance, dynamic light scattering has been used to show that the variation of hydrodynamic radii with c_L for q = 0.5 bicelles is invariant to temperature, with a value of $[DHPC]_{free}$ at about 5 mM in the range of 25–37 °C[91]. Analogous temperature invariance of small angle neutron scattering (SANS) profiles for q = 0.5 bicelles has been reported in the temperature range of 10–40 °C[69].

It was also of interest to establish the range of total lipid concentration over which the bicelle size is retained. Subtracting $[DIOHPC]_{free}$ from $[DIOHPC]_{total}$ and calculating the effective q from Eq. (2), we find that q_{eff} remains essentially constant at 0.5 when $c_L \geqslant 98$ mM (5% (w/v)). By comparison, c_L must exceed 130 mM (7% (w/v)) for the DMPC-DHPC bicelle system in order to keep the q_{eff} value constant at 0.5[68]. This result demonstrates an augmented versatility for bicellar media containing an ether-linked short-chain phospholipid, since it becomes possible to work with smaller amounts of lipid and either peptide or protein while still maintaining a small, rapidly tumbling bicellar structure and high-resolution NMR spectra.

3.3.5 Influence of salts on DMPC-DIOHPC isotropic bicelle size

Losonczi and Prestegard first reported the beneficial effects of ionic strength on the stability and lifetime of phospholipid bicelle preparations[85]. It has also been reported that the size and magnetic field alignment properties of high-q bicelles are affected by both monovalent and divalent ions[86, 95]. Nonetheless, it has been reported that trivalent lanthanide ions (e.g. Er^{3+}, Yb^{3+}, Tm^{3+}, and Eu^{3+}) do not influence the bicelle size despite their ability to change the orientation of liquid crystalline aggregates in the magnetic field[76, 77]. To evaluate the effects of ionic strength

in our system, values of q were measured for DMPC-DIOHPC bicelle samples that contained four commonly used physiological salts (NaCl, KCl, $CaCl_2$, and $MgCl_2$), based on the ^{31}P NMR peak integrals observed at different temperatures. The results are summarized in Table 3.3. The largest perturbation of bicelle size occurred upon addition of up to 200 mM NaCl to the c_L = 2% bicelle at 25 ^oC. The q value increased from 0.47 to 0.54, which corresponds to an increment in bicelle diameter from 9.2 to 9.8 nm (6.5%). Considering the error involved in the measurements, no significant influence on bicelle size was observed for these salts present at concentrations up to 200 mM and at temperatures from 25 to 46 ^oC.

Table 3.3 Values of lipid ratio (q) measured for DMPC-DIOHPC bicelles at different temperatures and salt concentrations[a]

Salt concentration (mM)	KCl[b]			NaCl[c]			$CaCl_2$[b]			$MgCl_2$[b]		
	q at T (°C)			q at T (°C)			q at T (°C)			q at T (°C)		
	25	37	46	25	37	46	25	37	46	25	37	46
0	0.46	0.46	0.46	0.47	0.47	0.47	0.46	0.47	0.48	0.46	0.46	0.47
50	0.49	0.46	0.48	0.50	0.46	0.49	0.48	0.48	0.47	0.48	0.48	0.48
100	0.48	0.48	0.49	0.53	0.45	0.51	0.48	0.48	0.48	0.47	0.47	0.48
150	[d]			0.54	0.45	0.52	[d]			[d]		
200	0.48	0.48	0.49	0.54	0.47	0.51	0.46	0.47	0.47	0.48	0.48	0.48

[a] The measured q value is derived from the integrals for DMPC and DIOHPC peaks in ^{31}P NMR spectra; errors of 2% were estimated from duplicate measurements.

[b] The mixture was prepared with c_L = 10% (w/v), q = 0.5, 10% D_2O.

[c] The mixture was prepared with c_L = 2% (w/v), q = 0.5, 10% D_2O.

[d] Not measured.

3.4 PROTEIN-MEMBRANE INTERACTION PROBED WITH DMPC-DIOHPC BICELLES

As mentioned above, DMPC-DHPC bicelles have been used widely as membrane- mimetic media[15]. The new DMPC-DIOHPC bicelle system was evaluated with LFABP and IFABP peripheral membrane proteins in terms of its utility as a model membrane for structural biology research. The interaction between FABPs and the bicelle was monitored by ^{1}H-^{15}N HSQC of the proteins. For example, an overlay of ^{1}H-^{15}N HSQC spectra for apo-IFABP alone and with 2% w/v DMPC-DIOHPC bicelles is shown in Figure 3.5. Both spectra exhibit sharp resonances, excellent spectral resolution, and the high degree of chemical shift dispersion typical of β-sheet structures. The high quality of the spectrum of IFABP with bicelle, and a solution that resisted precipitation for 45 days, indicate that the protein tumbles rapidly and does not form large aggregates. Chemical shift changes for particular resonances reflect the interactions between protein and bicelle, to be discussed further below. A single NH resonance appears for each observed backbone site, consistent with rapid exchange between bound and unbound states.

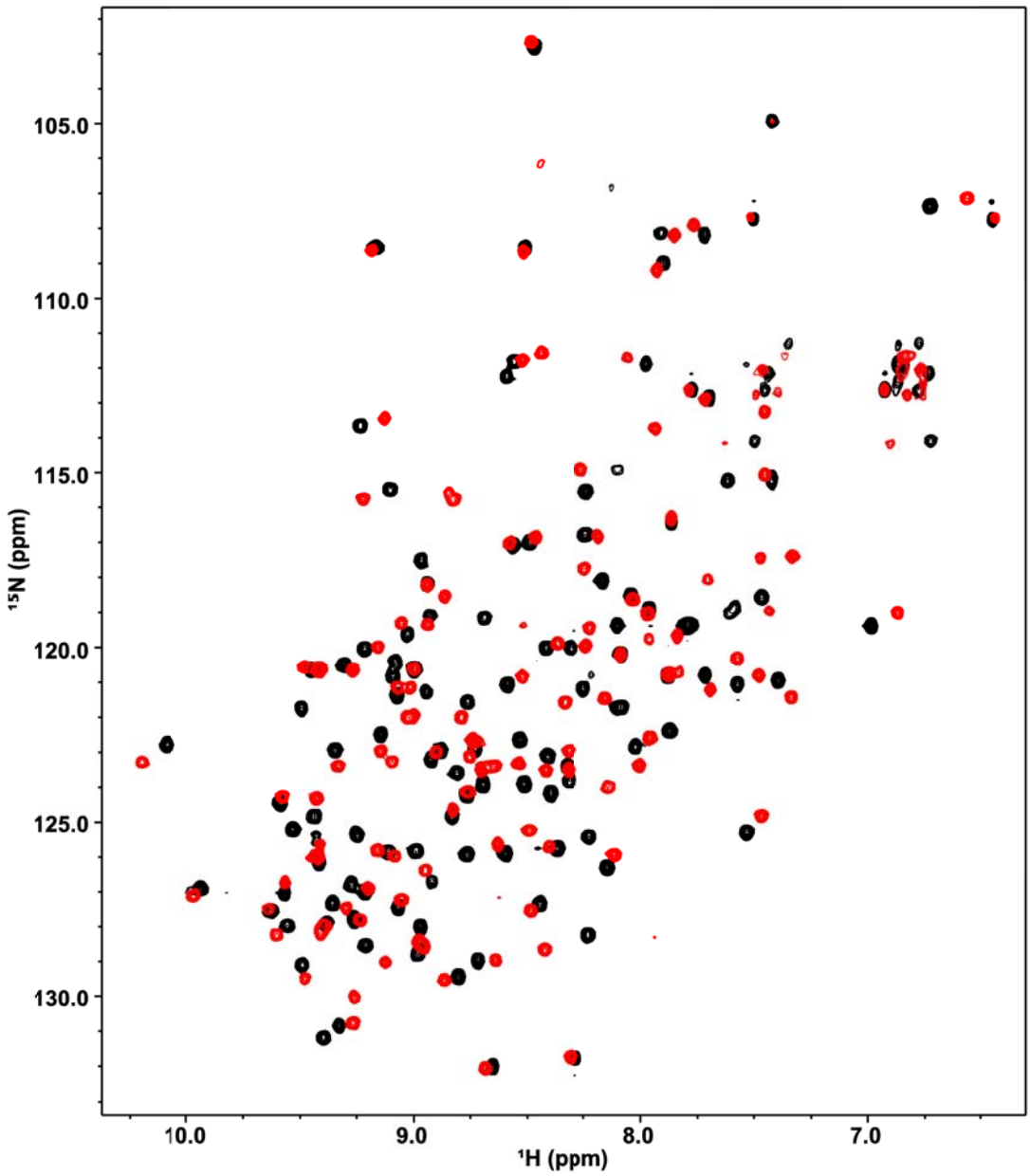

Figure 3.5 An overlay of the amide regions of ^{1}H-^{15}N HSQC NMR spectra for ~0.2 mM IFABP in 50mM phosphate, 100mM NaCl, 5 µM EDTA, 0.02% NaN_3, and 5% D_2O at pH 7.0, representing samples with (red) and without (black) a DMPC-DIOHPC bicelle solution (c_L = 2% w/v, q = 0.5).

3.4.1 Assignment of LFABP and IFABP NMR spectra

To obtain more detailed information about the interaction between FABPs and bicelles, site-specific chemical shift assignments are needed for the ^{1}H-^{15}N HSQC spectra of both LFABP and IFABP. The ^{1}H-^{15}N HSQC spectra of apo-IFABP and apo-LFABP were assigned based on published reports[4, 96, 97] and finally confirmed by checking the spin connectivities observed in 3D HSQC-TOCSY and NOESY-HSQC experiments. The spectra of the protein-bicelle assemblies were assigned by following spectra at the titration points visually, and then adjusted if necessary based on the connections found in 3D TOCSY-HSQC and NOESY-HSQC spectra. For example, Figure 3.6 illustrates confirmation of the assignments for residues 7 to 11 of apo-IFABP by NOE connectivities between residues i-1 and i in NOESY-HSQC and TOCSY-HSQC spectra.

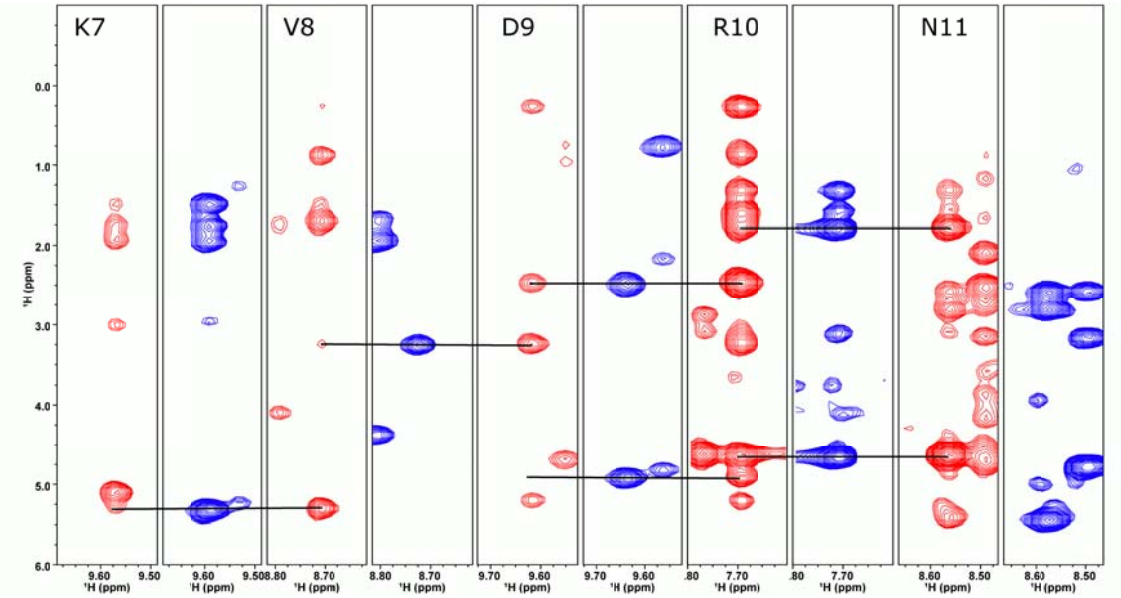

Figure 3.6 NOESY-HSQC (red) and TOCSY-HSQC (blue) spectra of apo-IFABP. The lines crossing the strips indicate connections between residues i-1 and i.

3.4.2 LFABP and IFABP titrations with a DMPC-DIOHPC bicelle

To probe the interaction between FABPs and membranes, both apo-LFABP and apo-IFABP were titrated stepwise with a 25% (w/v) DMPC-DIOHPC bicelle stock solution to achieve bicelle concentrations c_L from 0.25% to 2.25% in the NMR samples. The titrations were monitored by following the chemical shift changes in ^{1}H-^{15}N HSQC spectra of each protein. Figure 3.7 shows the shift of peaks in ^{1}H-^{15}N HSQC spectra of apo-LFABP upon addition of the DMPC-DIOHPC bicelle.

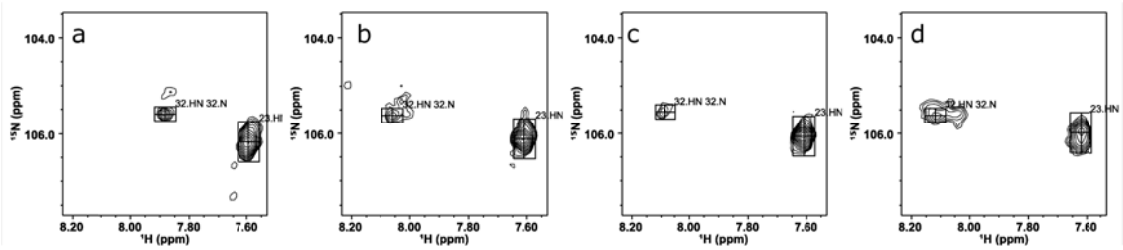

Figure 3.7 Titration of apo-LFABP with a DMPC-DIOHPC bicelle monitored by ^{1}H-^{15}N HSQC: (a) apo-LFABP, (b) with 0.25% bicelle, (c) with 0.5% bicelle, and (d) with 2% bicelle. It is clear from the spectra that the backbone resonance of residue 32 shifted while the resonance for residue 23 remained in the same position.

Chemical shift perturbations of apo-LFABP with the DMPC-DIOHPC bicelle

Figure 3.8 shows the chemical shift perturbations for apo-LFABP by the DMPC-DIOHPC bicelle, displayed as a function of protein sequence. Chemical shift perturbations for each backbone cross peak were calculated with the following equation:

$\Delta_{HN-N} = ([\delta_{HN}(apo) - \delta_{HN}(sgl)]^2 + \left\{\frac{[\delta_N(apo) - \delta_N(sgl)]}{6.5}\right\}^2)^{1/2}$ as recommended[55]. The residues with chemical shift perturbation larger than the sum of average perturbation and standard deviation are considered to be significantly perturbed and are indicated as red bars in the plot.

The average chemical shift perturbation of apo-LFABP with bicelle is 0.056 ppm; the maximum perturbation of 0.331 ppm occurs for residue V38, which is much smaller than the chemical shift perturbations observed for LFABP binding with oleate with average perturbation of 0.203 ppm and maximum perturbation of 1.129 ppm[13]. The residues with significant perturbations are F15, I29, G32, V38, S39, G45, H60, T74, K78, V79, V83, V92, F95, T102 and K121. Figure 3.9 maps the residues with significant perturbation onto a previously solved NMR structure[4]. Based on this mapping, the highly perturbed residues are found to be distributed all over the protein structure. Thus, no preferred interaction region can be identified for LFABP with the DMPC-DIOHPC bicelle.

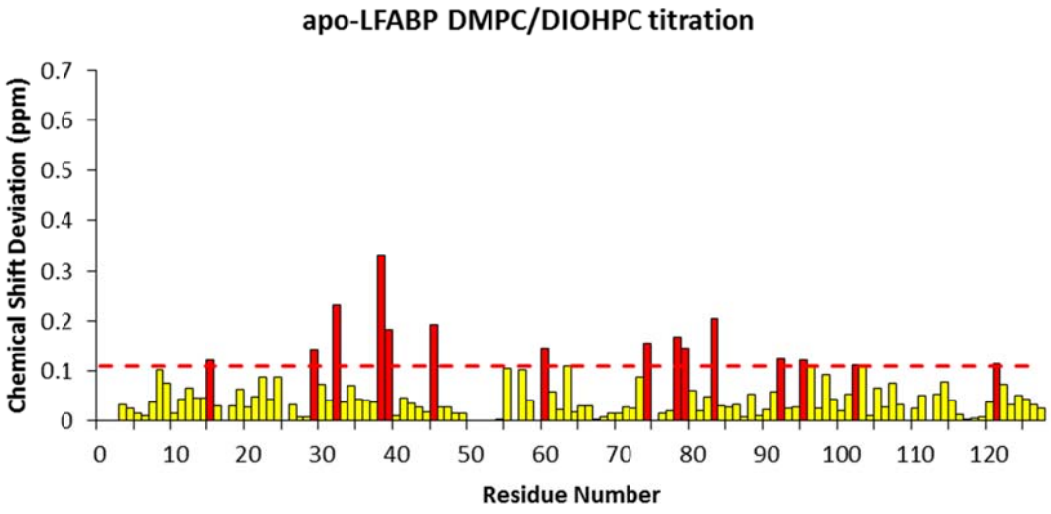

Figure 3.8 Chemical shift perturbations (CSP) of apo-LFABP upon addition of the DMPC-DIOHPC bicelle (c_L = 2.0%, q = 0.5). Average CSP + standard deviation is shown by a red dashed line, and all residues with significant perturbations are colored red.

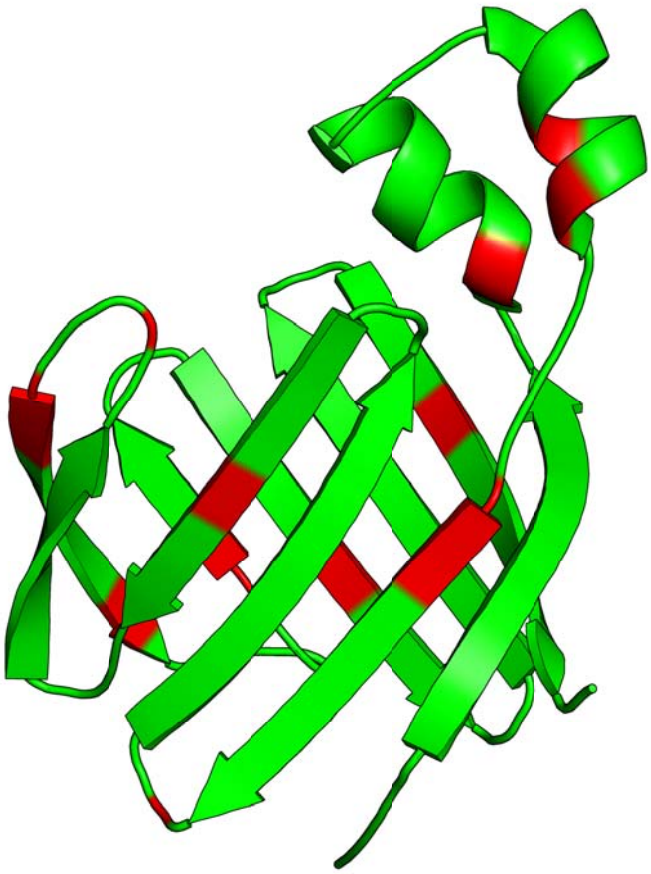

Figure 3.9 Mapping of significantly perturbed residues onto the LFABP structure (residues with perturbations greater than the mean plus one standard deviation are shown in red).

Chemical shift perturbations of apo-IFABP with the DMPC-DIOHPC bicelle

Analogously, the chemical shift perturbations of apo-IFABP upon addition of the DMPC-DIOHPC bicelle are plotted in Figure 3.10 as a function of protein sequence. Residues K27,

K29, G31, H33, D34, N35, K37, E51, F55, A73, and K125 were found to have significant perturbation. The average chemical shift perturbation in this case is 0.111 ppm; the maximum perturbation of 0.595 ppm is observed for residue K29. The mapping of the chemical shift perturbations onto the crystal structure of IFABP[98] is summarized in Figure 3.11. Those residues with significant perturbation are located either in the helical region (K27, K29, G31, H33, D34, and N35) or have close spatial contact with the helices (K37, E51, F55, and A73), especially the second helix. The only exception is K125, which seems to be unrelated to the helical region but is located spatially near the portal.

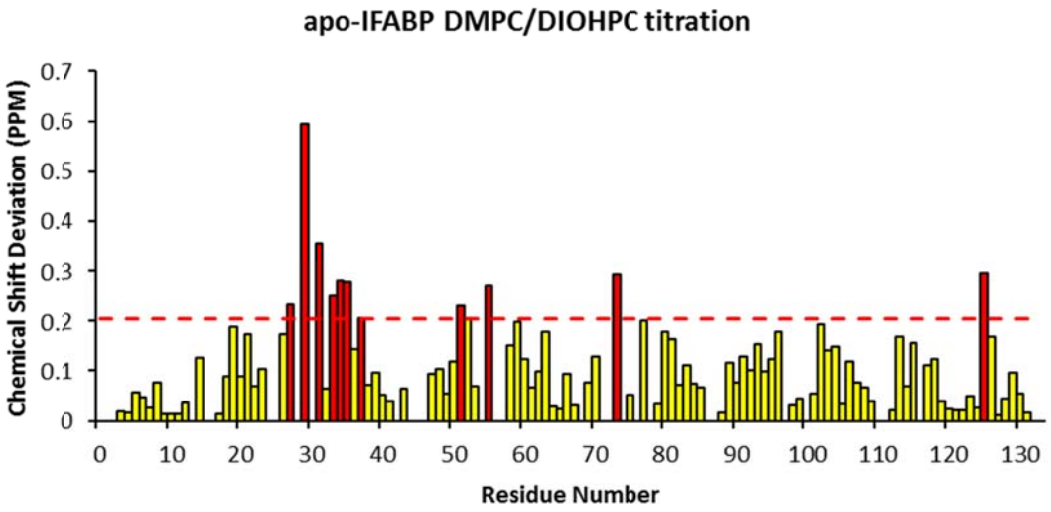

Figure 3.10 Chemical shift perturbation of apo-IFABP titrated with a DMPC-DIOHPC bicelle. Average CSP + standard deviation is shown by the red dashed line, and all residues with significant perturbations are colored red.

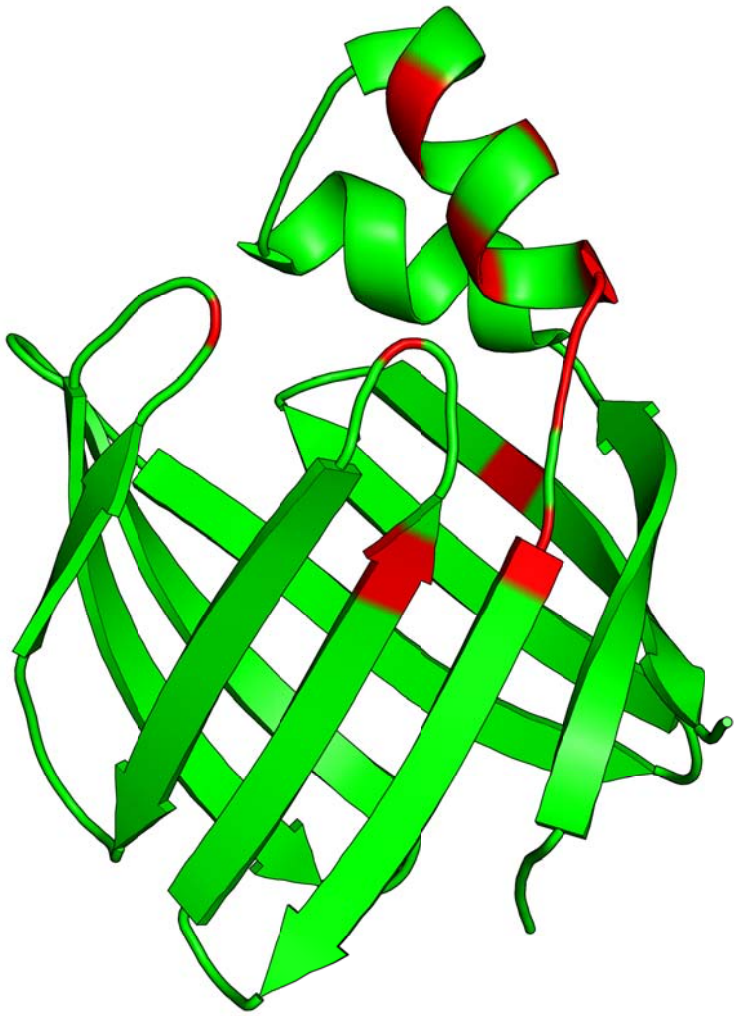

Figure 3.11 Mapping of chemical shift perturbation on IFABP. Residues with significant perturbations are colored red.

3.4.3 Contrasting mechanisms of protein-membrane interactions for FABPs

As described above, the significantly perturbed residues of apo-LFABP upon addition of the DMPC-DIOHPC bicelle are distributed all over the structure of the protein, which indicates a

more global interaction of LFABP with the model membrane rather than a site-specific interaction. In contrast, the residues of apo-IFABP that display significant chemical shift perturbations upon bicelle addition are located primarily in either the helices themselves or in close contact with the helices, indicating that the interaction between IFABP and the membrane is more specifically related to the helical region, especially the second helix. It is notable that positively charged residues lysine 27 and 29 in the second helix have significant perturbations upon adding bicelles, especially K29 which is the most perturbed residue (0.595 ppm). Our hypothesis is that these residues could to be important in regulating interactions between IFABP and membranes, as proposed previously for AFABP and HFABP[28, 79].

The contrasting patterns of these interactions of LFABP and IFABP with the bicelle reflect the different mechanisms established for FABP-membrane interactions[30]. IFABP obtains fatty acids from the membrane by a direct collisional mechanism, whereas LFABP interacts with the membrane and transfers fatty acids through a diffusion-driven mechanism[1, 30]. The failure to observe specific interactions between LFABP and bicelles is in accord with the diffusional mechanism of LFABP-mediated fatty acid transfer to membranes. Conversely, the observation of specific bicelle interactions located in the region of the IFABP helices supports the collisional mechanism of protein-membrane interactions, which is more site specific than the diffusional mechanism. Moreover, the alpha helices of IFABP have been implicated in the collisional mechanism[78, 99], validating the observation of chemical shift perturbations of IFABP with the DMPC-DIOHPC bicelle specifically in the helical region (Figure 3.11).

Additionally, the difference in amplitude of the bicelle-associated chemical shift perturbations between LFABP and IFABP supports their contrasting ligand transfer mechanisms. The average chemical shift perturbation for IFABP is about twice of that for LFABP (0.111 ppm vs. 0.056

ppm). The larger average chemical shift perturbation suggests stronger interactions between IFABP and the bicelle as compared with LFABP, since the observed resonance positions are a weighted average between bound and unbound protein states. It is reasonable as well to expect the stronger interaction of IFABP to be associated with its collisional membrane interaction mechanism and the mild interaction of LFABP to reflect the diffusional mechanism of fatty acid transfer to membranes.

3.4.4 Possible protein-lipid binding

It is essential to establish that our observations of chemical shift perturbations may be attributed to interactions between FABPs and the DMPC-DIOHPC bicelle, and to exclude the possibility of simply binding DIOHPC or DMPC monomers in either protein cavity. This issue was addressed by independent experiments using DMPC and DHPC monomers: no chemical shift perturbation was found for LFABP with added DMPC and only small nonspecific perturbations were observed for DHPC[82]. In our titration of IFABP with the DMPC-DIOHPC bicelle, there are almost no chemical shift perturbations at the first titration point, even though the amount of DIOHPC monomer would correspond to about 10 equivalents of protein. This negative observation demonstrates that there is no interaction between lipids and protein until the amount of bicelle reaches a certain threshold. As mentioned above, the short chain lipids exist mainly as monomer or micelle in a very dilute bicelle solution. In other words, the interaction between protein and monomers is negligible.

Figure 3.12 illustrates this argument by monitoring of the changes in HSQC spectra for residue T77 of IFABP. When the bicelle concentration is low (0.25%), the peak moves very little (less

than 0.01ppm) (spectrum b). Substantial changes in chemical shift (~0.2 ppm) are observed when the amount of bicelle exceeds 0.5% (spectrum c). With further increases in bicelle concentration, a stable bicelle assembly forms in addition to DHPC monomers or micelles and significant chemical shift perturbations are attributable to the interaction of IFABP with bicelle.

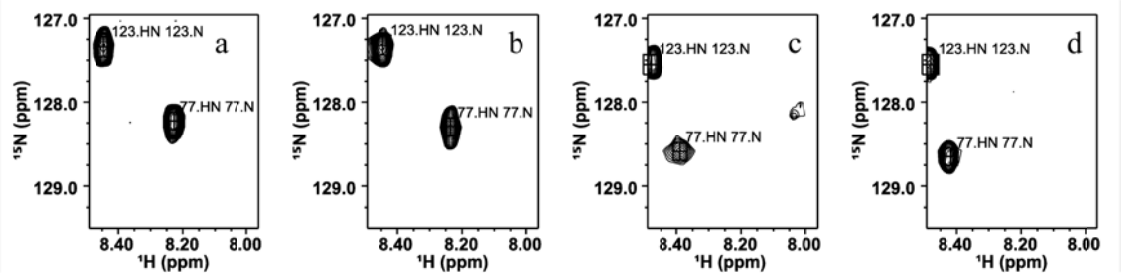

Figure 3.12 Titration of apo-IFABP with a DMPC-DIOHPC bicelle monitored by ^{1}H-^{15}N HSQC: (a) apo-LFABP, (b) with 0.25% bicelle, (c) with 0.5% bicelle, and (d) with 2% bicelle. Negligible backbone chemical shift perturbation is observed for residue T77 with the 0.25% bicelle and a large perturbation with the 0.5% bicelle.

3.5 SUMMARY

A new DMPC-DIOHPC isotropic bicelle system has been evaluated as a membrane- mimetic medium for biomolecular NMR and related structural biology research. Compared with other similar systems, DMPC-DIOHPC isotropically tumbling bicelles exhibit improved hydrolytic stability and superior ^{31}P NMR peak separation, allowing more accurate measurements of integrated signal intensities and associated sizes. Moreover, the new system extends the range of experimental conditions over which constant size and aggregate organization can be maintained, including overall phospholipid concentrations down to c_L of 5%, temperatures from 25 to 46 ^{o}C, and the presence of physiologically relevant salts at concentration up to 200 mM. The q = 0.5 DMPC-DIOHPC bicelle system shows potential as a versatile medium for high-resolution NMR studies of peripheral membrane proteins. The bicelle interactions of intestinal and liver fatty acid-binding proteins that accomplish ligand transfer via different mechanisms have been differentiated according to the spatial distribution of their significantly perturbed residues and the overall amplitude of their chemical shift perturbations. In addition, the interaction between IFABP and membrane via the collisional mechanism is proposed to be localized in the helical regions of the protein based on the chemical shift perturbation results, a result that supports the functional importance proposed for this portal region in AFABP and HFABP[28, 79].

CHAPTER 4 SOLID STATE NMR STUDY OF A TRANSMEMBRANE PEPTIDE

4.1 INTRODUCTION

G-protein coupled receptors

G-protein coupled receptors (GPCRs) are a family of integral membrane proteins which share the common structural motif of 7 transmembrane (TM) helices, illustrated in Figure 4.1 for the *Saccharomyces cerevisiae* GPCR (Ste2p)[100]. GPCRs are involved in cell signaling and signal transduction for taste, smell, and light stimulation as well as many other cell signaling processes that are mediated by protein-ligand interactions[101]. As important targets for drug design, almost half of modern drugs in the pharmaceutical industry target the GPCR proteins[32]. Despite the functional importance of GPCRs, structural information on this protein family is rather limited because of the low protein expression levels and difficulty in generating crystals for X-ray crystallography for membrane proteins[102]. Since the first crystal structure of rhodopsin was solved in 2000[103], only a few additional GPCR crystal structures have been solved: the β_1-adrenergic G-protein-coupled receptor[104, 105], the G protein-coupled chemokine receptor CXCR4[106], the human β_2-adrenergic receptor[107-109], and the agonist-bound human adenosine A_{2A} receptor[110, 111]. There has also appeared one structure determination for a seven-helix transmembrane protein by solution NMR in a micelle[112]. Compared with the number of GPCRs predicted in the human genome, the structural information on GPCRs is still severely limited.

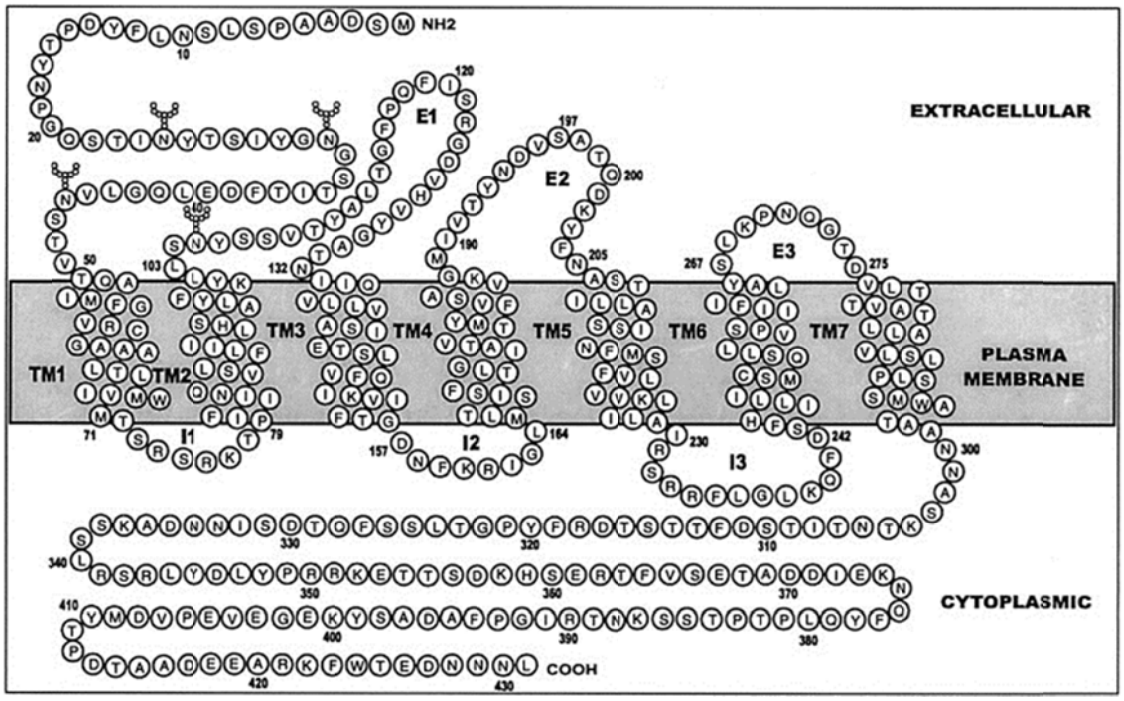

Figure 4.1 Cartoon of the yeast-mating factor G protein-coupled receptor, Ste2p.[100]

The folding of integral membrane proteins has been proposed to be a two-step process: (1) the insertion of polypeptide into the membrane and formation of helical secondary structures; (2) the association of helices into tertiary structures[113-115]. The highly helical structures of both single and double TMs from GPCRs have been observed. Receptor reconstitutions of coexpressed fragments of Ste2p (the *Saccharomyces cerevisiae* GPCR) split at cytoplasmic or extracellular loops have shown that the fragments can assemble and transduce a signal as a functional receptor[33]. The ability of transmembrane fragments to find each other in a membrane mimetic medium has also been demonstrated by an interaction study between the transmembrane domains of the human adenosine A_{2A} receptor using a combination of Forster resonance energy transfer (FRET) measurements and circular dichroism (CD) spectroscopy[116]. The data support the two-

step folding process mentioned above and argue that structural study of separated fragments can provide valid information for the intact GPCRs. GPCR fragments have been used widely in structural studies due to the difficulty of obtaining stable preparations of the full-length GPCRs that are suitable for high-resolution structure determination.

The *Saccharomyces cerevisiae* GPCR Ste2p plays a role in the regulation of the mating process in *Saccharomyces cerevisiae* by interacting with the pheromone polypeptide secreted from cells of the opposite mating type. Fragments of Ste2p have been studied for many years to obtain structural information on Ste2p, and helical structures have been reported for synthetic single-transmembrane domains[117-120]. Moreover, NMR structures of first two transmembrane domains, TM1TM2 [Ste2p (G31-T110)] have been solved in LPPG micelles[34] and TFE/water[35] by solution-state NMR.

Solid-state NMR applications to membrane proteins

Solution-state NMR has developed into a standard method for protein structure determination[121] since the structure of proteinase inhibitor IIA from bull seminal plasma[122] was determined. Recently developed solid-state NMR techniques make it possible to obtain structural information for peptides, amyloid fibrils, and membrane proteins[36-39]. The spectral resolution and sensitivity of solid-state NMR spectra have been improved using a combination of high magnetic field strength and magic angle spinning (MAS)[123] to average out the line broadening effects from chemical shift anisotropy (CSA) and strong dipolar coupling [36, 37, 39].

The development of new solid-state NMR techniques together with different isotopic labeling schemes e.g., reverse labeling, make it possible to conduct structural and dynamic studies of

transmembrane proteins in a lipid bilayer environment. For example, partial site-specific assignments have been reported for the solid-state NMR spectra of uniformly [^{13}C, ^{15}N]-labeled light-harvesting complex 1, a 160 kDa integral membrane protein[43] and uniformly [^{13}C, ^{15}N]-labeled cytochrome bo_3, a 144 kDa integral membrane protein[124]. Besides assignment of the NMR spectra, solid-state NMR has also been applied to determine protein structures. The high resolution structure of microcrystalline GB1 was determined[125] and refined with chemical shift tensor information,[126, 127] and a solid-state NMR-based structure was reported for ubiquitin in a uniformly [^{13}C, ^{15}N]-labeled microcrystalline sample[128]. Structural studies of transmembrane proteins with solid-state NMR have also been reported for the seven-helix integral membrane proton pump proteorhodopsin[44] and the Ca^{2+}-ATPase-bound phospholamban protein in lipid bilayers[42].

Sample preparation of membrane proteins for solid-state NMR

A crucial aspect of successful solid-state NMR studies of membrane proteins is sample preparation. The procedure for sample preparation can be divided into two stages: reconstitution of protein in lipids and sample precipitation. The transmembrane protein is first reconstituted in a lipid/detergent mixture in solution. Then the detergent is removed, customarily by either dialysis or using Bio-beads. Dialysis has been reported for solid-state NMR studies of DsbB[41] and cytochrome bo_3[124]. Bio-beads were used for sample preparation with proteorhodopsin[44]. Finally the protein precipitates together with the lipids and is collected after ultracentrifugation to remove excess water. In addition, a solid-state NMR study has appeared for the transmembrane protein cytochrome P450 (CYP) 3A4 reconstituted in nanodiscs, which yield narrow line widths

(~0.5 ppm) and high sensitivity in 2D spectra[40]. The interaction between protein and lipids in nanodisc has also been reported[129].

As mentioned above, the structures of TM1TM2 [Ste2p (G31-T110)] have been solved in both a LPPG micelle[34] and TFE/water[35]. Herein we describe a preliminary solid-state NMR study of TM1TM2 carried out in a native-like lipid bilayer environment, to obtain some structural information.

4.2 MATERIALS AND METHODS

1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (14:0) (DMPC, MW: 677.93) and 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (6:0) (DHPC, MW: 453.25) were purchased from Avanti Polar Lipids (Alabaster, AL) and used without further purification. Sodium dodecylsulfate (SDS) was purchased from Fisher Scientific (Fair Lawn, NJ). [U-^{15}N, ^{13}C] glutamine was purchased from Isotec-Sigma-Aldrich (Miamisburg, OH). TM1TM2 peptides with different isotope labeling schemes were provided by Ms. Leah Cohen (laboratory of Prof. Fred Naider, College of Staten Island). The expression and purification of TM1TM2 have been described previously[35].

4.2.1 Sample preparation

Preparation of TM1TM2 in lipid bilayers for solid-state NMR was conducted by first reconstituting the peptide in either phospholipid bicelles or a lipid-detergent mixture; then the short-chain lipid or detergent was removed by dialysis against a phosphate buffer. The assemblies of peptide with phospholipids were precipitated from the solution and collected with an ultracentrifuge operating at 10,000g. Finally the sample was packed into a 3.2 mm thin wall rotor (36 μL) for solid-state NMR experiments conducted with magic angle spinning. For the samples prepared with isotropically tumbling bicelles, appropriate amounts of TM1TM2, DMPC (dimyristoylphosphatidylcholine), and DHPC (dihexanoylphosphatidylcholine) were dissolved in a 1:1 mixture of chloroform and methanol. The solvent was evaporated under a stream of nitrogen gas to form a film in a small glass vial. Finally the sample was hydrated with phosphate buffer in preparation for dialysis. For samples prepared with a lipid-detergent mixture, TM1TM2 was dissolved in SDS solution with brief sonication and a hydrated DMPC slurry was

added. A clear solution was formed by vortexing in preparation for dialysis. The lipid-to-peptide molar ratio was 70 to 80 in all samples.

4.2.2 NMR spectroscopy

The solid-state NMR data were all collected on either a Varian VNMRS 600 instrument (Palo Alto, CA) equipped with a BioMAS ^{1}H-^{13}C-^{15}N 3.2 mm probe and operating at a ^{1}H frequency of 599.497 MHz at the City College of New York or a 900 MHz Bruker Avance III spectrometer equipped with an HXY 3.2 mm probe at the New York Structural Biology Center. All spectra were collected with a magic angle spinning rate of 13.3 kHz unless otherwise specified. The ^{13}C chemical shifts were referenced directly to the methylene carbon of adamantane at 40.48 ppm,[130] and ^{15}N chemical shifts were referenced to 0 ppm based on a calculation from the frequency ratios of ^{15}N and ^{13}C, following IUPAC guidelines[131] (see Appendix I). The sample temperature under MAS conditions was calibrated using ^{207}Pb NMR of a $Pb(NO_3)_2$ standard[132, 133] by following the protocol described in Appendix II[134]. TPPM decoupling[135] at strengths of 100 or 80 kHz was used for 600 MHz and 900 MHz spectrometers, respectively. Experimental conditions were optimized with either a [U-^{15}N, ^{13}C] glutamine or a [U-^{15}N, ^{13}C] GB1 sample for 1D and 2D experiments. Solution state ^{1}H-^{15}N HSQC spectra were acquired on a Varian VNMRS 600 instrument (Palo Alto, CA) equipped with an IDQG probe and operating at ^{1}H frequency of 599.497 MHz at The City College of New York. 2D data were processed with NMRpipe[57] and analyzed with NMRViewJ[58] (One Moon Scientific Inc.).

Cross polarization and magic angle spinning

Cross polarization (CP) provides high sensitivity for rare nuclear spins such as ^{13}C and ^{15}N by polarization transfer from proton spins[136] (pulse sequence in Figure 4.2), and magic angle spinning (MAS) improves the resolution of solid-state NMR spectra by eliminating broadening effects from chemical shift anisotropy and dipolar coupling. A ^{13}C CP spectrum of [U-^{15}N, ^{13}C]-glutamine appears in Figure 4.3(b), displaying 5 carbon signals corresponding to the 5 magnetically inequivalent carbon atoms in the molecule. The ^{13}C resonances are well separated and have narrow linewidths typical of crystalline solids (~120 Hz line width at half height).

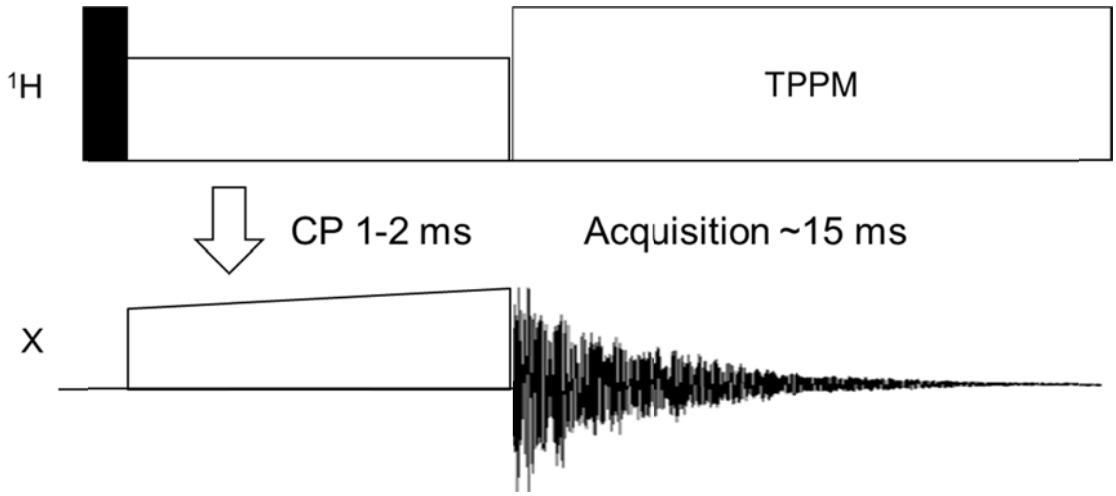

Figure 4.2 The cross-polarization (CP) pulse sequence for a CPMAS experiment.

Band selective double-cross polarization (DCP)

Band selective double-cross polarization (DCP) with SPECIFIC polarization transfer from ^{15}N to $^{13}C_{\alpha}$ and $^{13}C'$[137-139], namely NCA and NCO experiments, were set up and optimized with [U-^{15}N,

^{13}C] glutamine. The DCP pulse sequence in Figure 4.4 shows a polarization transfer from ^{1}H to ^{15}N via CP in the first step, and polarization transfer from ^{15}N to ^{13}C via SPECIFIC CP in the second step. For example, Figure 4.3 shows optimized NCA (a) and ^{13}C CP spectra (b) of [U-^{15}N, ^{13}C] glutamine. It can be seen from the spectra that all ^{13}C signals show up in the ^{13}C CP spectrum (b), while only the CA signal appears in the NCA spectrum (a) upon band selective DCP transfer. The signal intensity of CA in the optimized NCA spectrum is about 40% of the signal intensity in the corresponding ^{13}C CP spectrum. Similar experiments conducted with a microcrystalline GB1 protein sample give ~65% DCP efficiency (data not shown). The lower DCP efficiency of glutamine is attributable to the slow relaxation rate. The NCA experiment was set up with a MAS rate of 13.3 kHz. Radiofrequency (rf) field strengths for SPECIFIC CP are 46.6 kHz in the ^{15}N channel and 33.3 kHz in the ^{13}C channel, namely 3.5 and 2.5 times the MAS rate, including a 10% ramp for the ^{13}C channel and 100 kHz CW decoupling on the ^{1}H channel during SPECIFIC CP. The rf field strengths for SPECIFIC CP in ^{15}N and ^{13}C channel must be tuned carefully to achieve optimal spectral sensitivity.

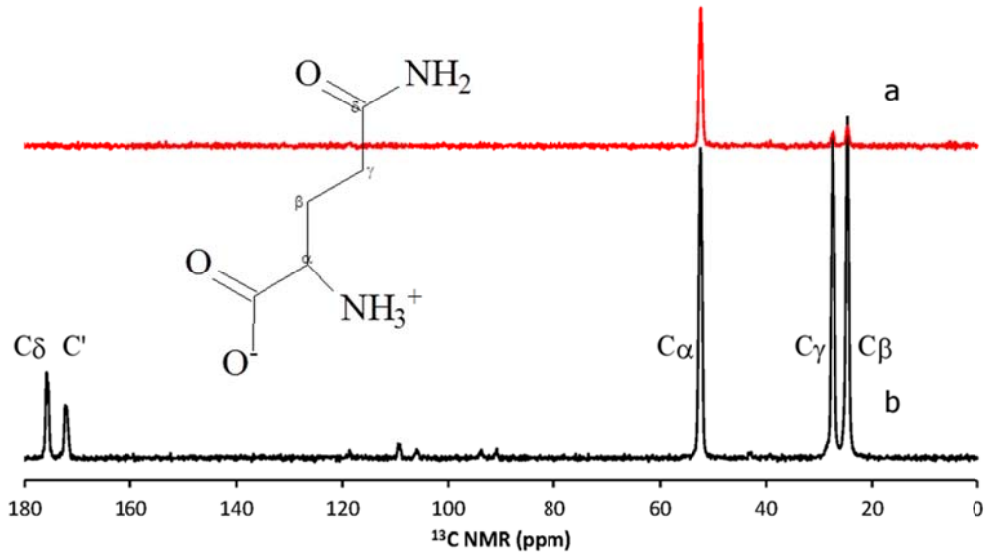

Figure 4.3 1D NCA with SPECIFIC CP transfer from N to CA (a) and ^{13}C CP (b) spectra of U-[^{15}N, ^{13}C]-glutamine. Only the CA resonance is presented in the NCA spectrum.

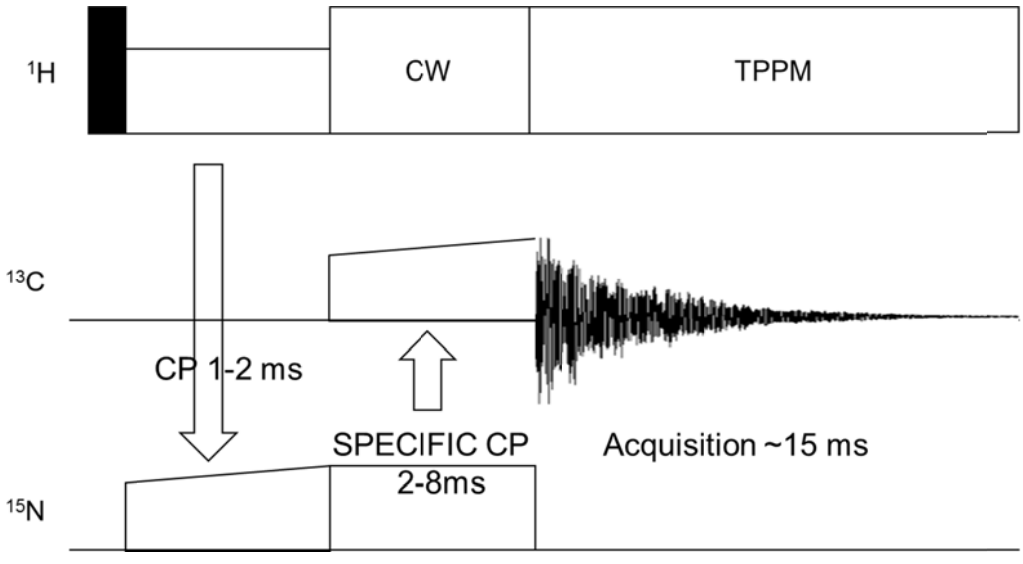

Figure 4.4 Pulse sequence used for DCP type experiments. The magnetization transfers from proton to ^{15}N via the first CP step, followed by the SPECIFIC CP transfer from ^{15}N to either CA or CO.

Dipolar-Assisted Rotational Resonance

Dipolar-Assisted Rotational resonance (DARR) is a ^{13}C-^{13}C recoupling method under MAS that utilizes the ^{13}C-^{1}H dipolar interaction[140]. The DARR pulse sequence in Figure 4.5 shows that ^{13}C-^{13}C recoupling is achieved by irradiation on the proton channel with the ^{1}H rf field (ω_1) set to equal the MAS rate (ω_r) during the DARR mixing time[140, 141]. The DARR technique provides information about the interactions between ^{13}C pairs that are close in space. With longer mixing times, it is possible to achieve an upper distance limit of 6 Å[142, 143]. Figure 4.6. shows a ^{13}C-^{13}C correlation spectrum of [U-^{15}N, ^{13}C] glutamine with 10 ms DARR mixing, where the correlations between ^{13}C pairs are indicated by the lines connecting the cross peaks. For example, correlations of C_α to all other carbons (C_β, C_γ, C_δ, and C') in the molecule can be observed in the spectrum. Even the correlation between C' and C_δ shows up as a weak cross peak around 175 ppm.

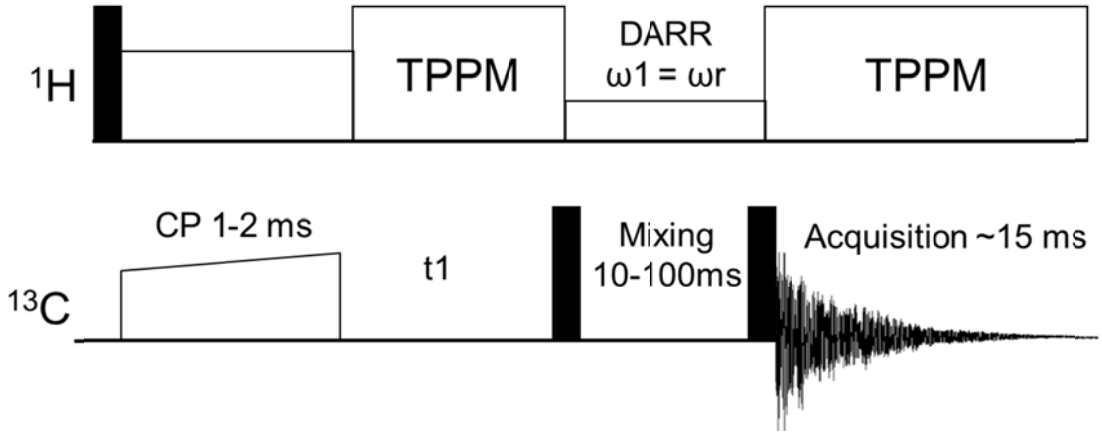

Figure 4.5 DARR pulse sequence.

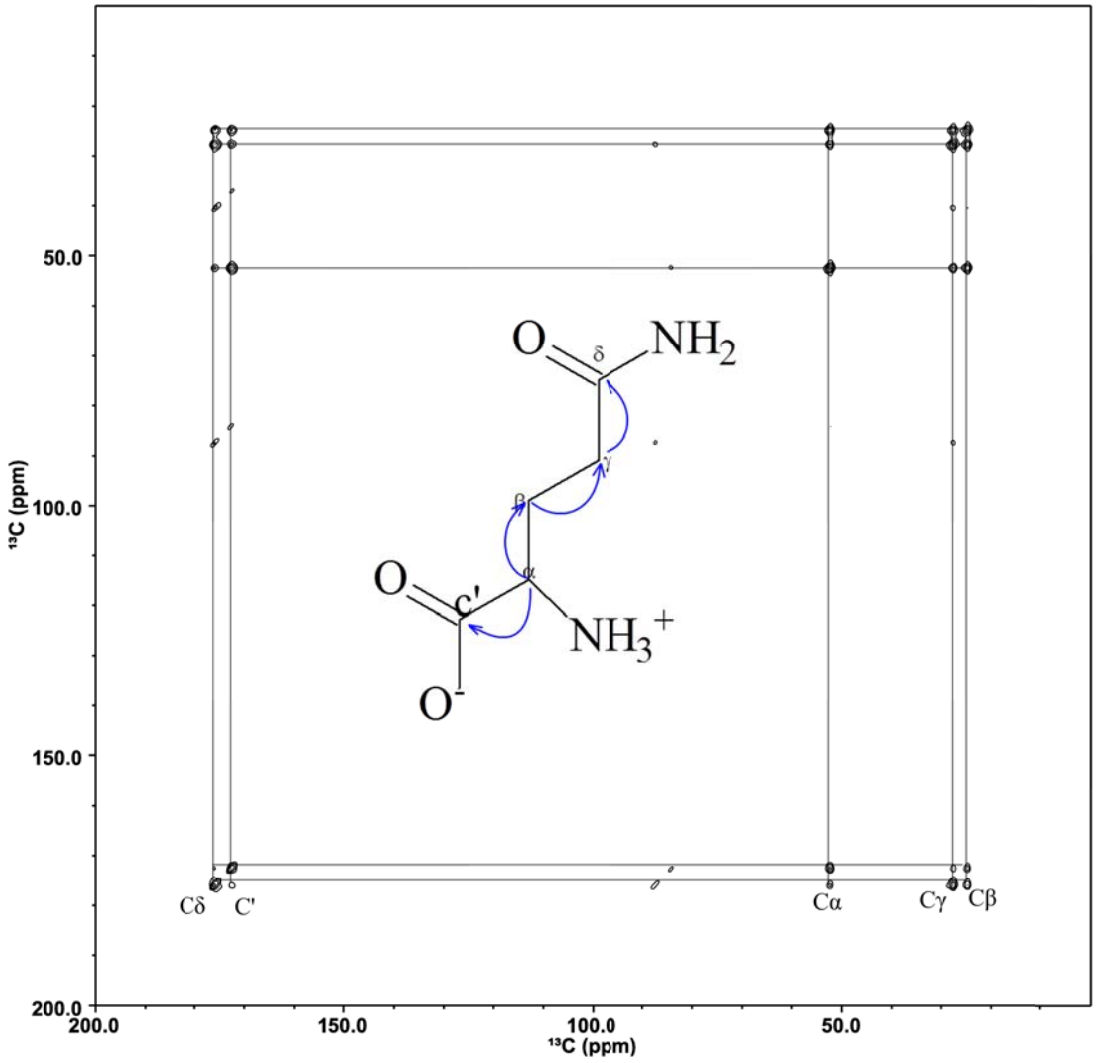

Figure 4.6 ^{13}C-^{13}C correlation spectrum of [U-^{15}N, ^{13}C]-glutamine with 10 ms DARR mixing.

NCACX and NCOCX

The double CP experiments NCA and NCO can be expanded to 2D and 3D experiments with the addition of evolution and mixing times. Figure 4.7 shows an overlay of (N)CACX and (N)COCX spectra for [U-^{15}N, ^{13}C]-glutamine. The magnetization transfer from HN to N to CA or CO (C_δ in glutamine) occurs first via DCP; then the magnetization transfer to other ^{13}C atoms

proceeds via DARR mixing. For example, the magnetization transfers from N to CA first, then from CA to C', CB and CG (weaker signals due to longer distance from CA). The arrows indicate the transfers of magnetization. All ^{13}C atoms can be assigned based on the analysis of NCACX and NCOCX spectra.

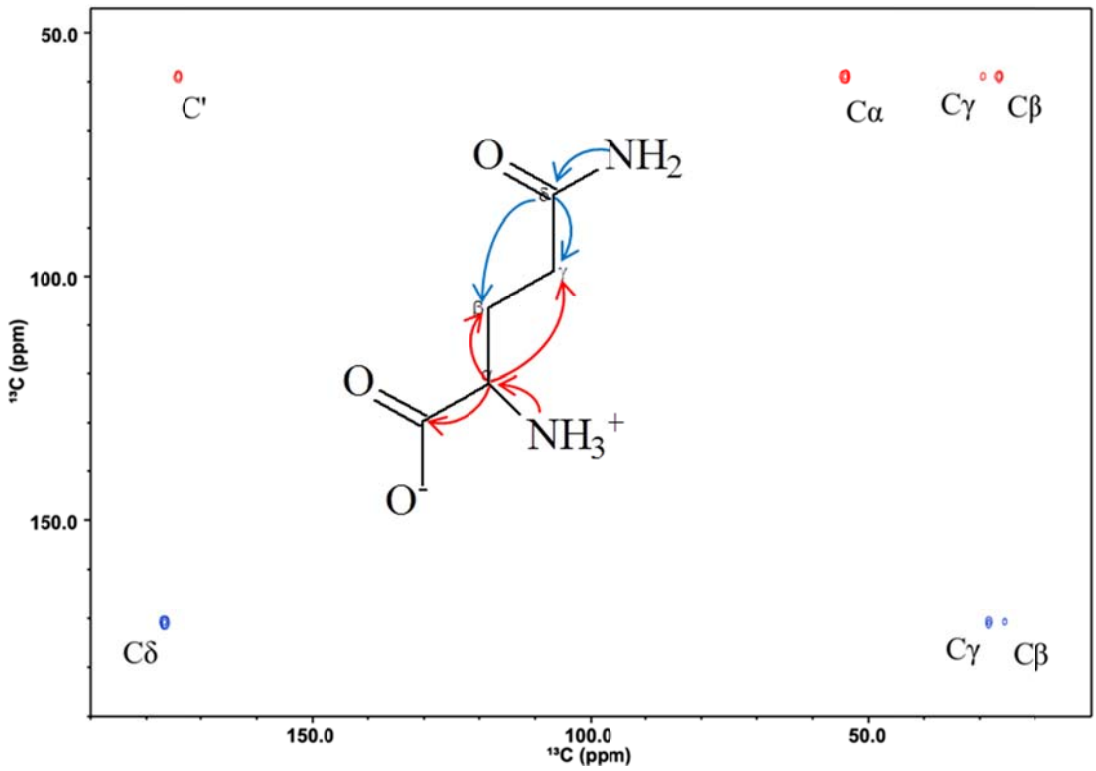

Figure 4.7 (N)CACX (red) and (N)COCX (blue) spectra of [U-^{15}N, ^{13}C]-glutamine. The red arrows indicate the transfer of magnetization in (N)CACX, while the blue arrows show the magnetization transfer for (N)COCX experiments.

4.3 RESULTS AND DISCUSSION

As mentioned above, the peptide was reconstituted in either lipid bicelles or a DMPC/SDS mixture. For the sample reconstituted in bicelles, the mass percentage of peptide in the final precipitate was low, probably because of the difficulty of removing DHPC by dialysis. About 1 mg of peptide is contained in the samples (30-35 mg total) that were packed into rotors. The small amount of TM1TM2 results in low sensitivity for the solid state NMR experiments, making it impractical to collect multidimensional spectra. For the peptide reconstituted in DMPC/SDS, about 3.5 mg of TM1TM2 in 35 mg of total sample can be packed into the 3.2 mm thin wall rotor. The data discussed here are collected with those samples unless otherwise specified.

4.3.1 Peptide-DMPC interactions

To prepare a sample that will be suitable for solid state NMR, it is necessary to establish that the peptide is reconstituted in the lipid-detergent mixture and folds into a well-defined helical structure. The interaction between DMPC and the TM1TM2 peptide was investigated by monitoring the ^{15}N HSQC spectra for ^{15}N alanine-labeled TM1TM2 in SDS solution upon addition of a DMPC slurry. The peptide-SDS solution remains optically clear when the DMPC slurry is added, indicating that DMPC has interacted with SDS and has the potential to interact with the TM1TM2 peptide. It has been reported that TM1TM2 is soluble in SDS and maintains its helical structure as assessed by circular dichroism (CD)[35]. It has also been shown that TM1TM2 in SDS displays reasonable resolution in the solution state ^{15}N HSQC NMR spectrum, but the sample does not last long enough to collect 3D or 4D experiments for assignments and

structural studies[35]. Herein we use SDS to dissolve TM1TM2 for the initial step of sample preparation, which does not require long term stability.

^{15}N HSQC spectra were collected with ^{15}N alanine-labeled TM1TM2 at 49.0 °C. Five isolated peaks appeared in the spectra corresponding to five ^{15}N-labeled alanine residues in the peptide; the signals were assigned by comparison with published assignments for TM1TM2 in LPPG micelles[34]. Figure 4.8 clearly shows that the peaks shift when even a small amount of DMPC is added into the system (1 mg, a DMPC-to-peptide ratio of ~17), indicating that the chemical environment has changed for the peptide. Such peptide chemical shift perturbations confirm that DMPC participates in peptide-SDS assemblies and moreover, that it interacts with the TM1TM2 peptide. Furthermore, the observation of interactions even with modest amounts of DMPC indicates a preference of the peptide for the phospholipid as compared with the SDS detergent. When the amount of DMPC added to the solution is higher, the chemical shifts of the alanines become closer to the peptide signals in LPPG[34]. Since the NMR structure of TM1TM2 has already been determined in LPPG[34], it is reasonable to argue that upon interaction with DMPC the peptide has folded into helices, adopting a similar conformation to what was observed in LPPG. In addition to the change in chemical shifts, a decrease in peptide peak intensity with increasing DMPC is evident from Figure 4.8. The decrease in peak intensity is likely to be caused by line broadening that accompanies an increase in size of the TM1TM2/DMPC/SDS assemblies because of participation of DMPC.

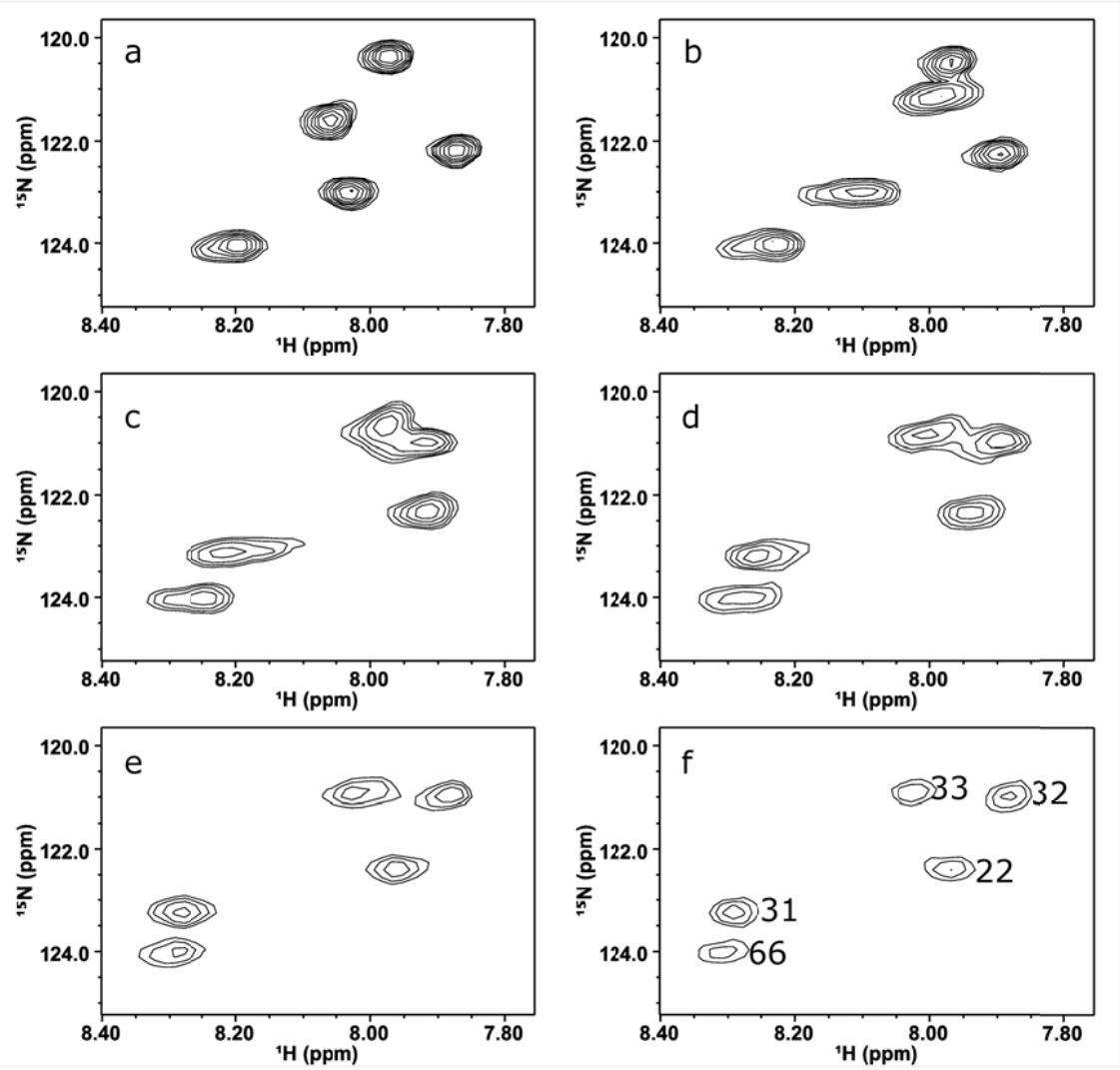

Figure 4.8 ^{1}H-^{15}N HSQC spectra of alanine-labeled TM1TM2 in SDS with varying amounts of added DMPC: (a) 0 mg; (b) 1 mg; (c) 2 mg; (d) 3mg; (e) 4 mg; (f) 5 mg.

Besides the interaction of the peptide with DMPC, the stability of TM1TM2/DMPC/SDS assemblies was monitored by acquiring ^{15}N HSQC spectra periodically for 9 days at room temperature to make sure the peptide was not denatured during the dialysis. As shown in Figure 4.9, no chemical shift changes were observed after the SDS-dissolved peptide was incubated at

room temperature for 9 days with a DMPC slurry added. Both the resolution and the peak intensity were well maintained for more than a week in room temperature, which means that the peptide was not denatured in the DMPC/SDS mixture. The stability of TM1TM2 in the this mixed assembly implies that the peptide is sequestered from SDS by a surrounding layer of DMPC, because SDS alone tends to denature TM1TM2 within a week[35].

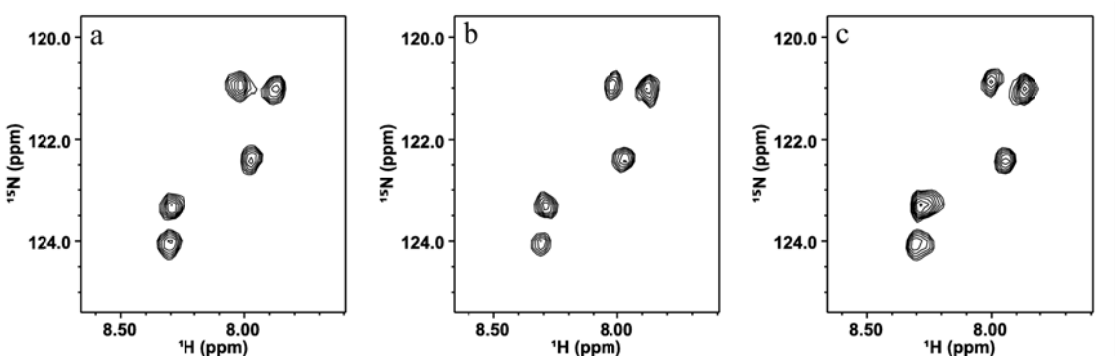

Figure 4.9 ^{15}N HSQC spectra of SDS-dissolved TM1TM2 in DMPC/SDS after (a) 2 days, (b) 4 days, and (c) 9 days. No chemical shift change or peak intensity decay was observed.

4.3.2 1D ^{13}C and ^{15}N NMR of TM1TM2

^{13}C and ^{15}N CPMAS spectra of TM1TM2 in DMPC lipids are presented in Figure 4.10 and Figure 4.11, respectively. In the ^{13}C spectrum, different regions may be identified as carbonyls (170-180 ppm), aromatics (115-140 ppm), C_α (48-65 ppm) and other aliphatic carbons (10-48 ppm). The ^{13}C signal at ~160 ppm belongs to C_ζ of tyrosine. In the ^{15}N spectrum, the main backbone resonances span the range 110-130 ppm, with maximum signal intensity at 120 ppm, approximately where the backbone residues resonate. Besides the backbone signal, the signals of the arginine side-chain nitrogens appear at 71 and 84 ppm, and the signal of the lysine side-chain nitrogen resonates at 32 ppm.

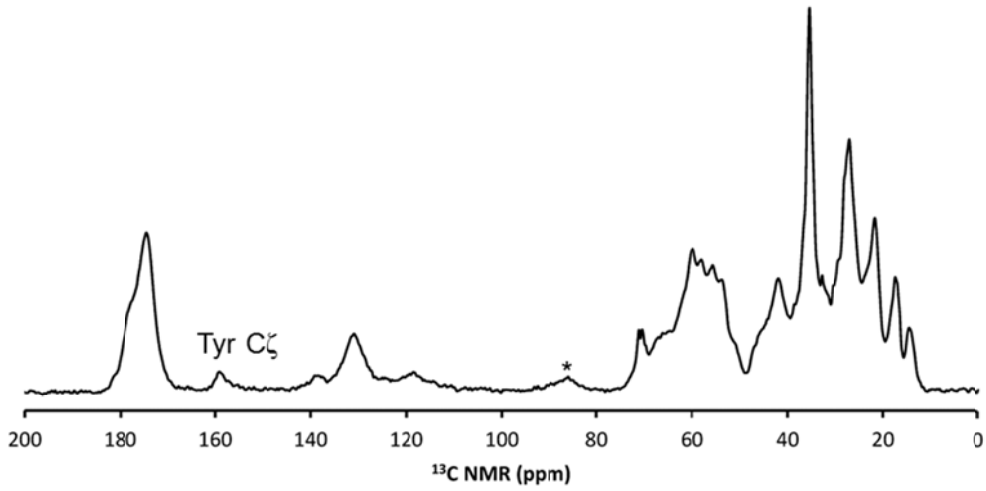

Figure 4.10 ^{13}C CP spectrum of TM1TM2 in lipids at -20 ^{o}C, processed with 50 Hz line broadening. *: spinning side band of C'.

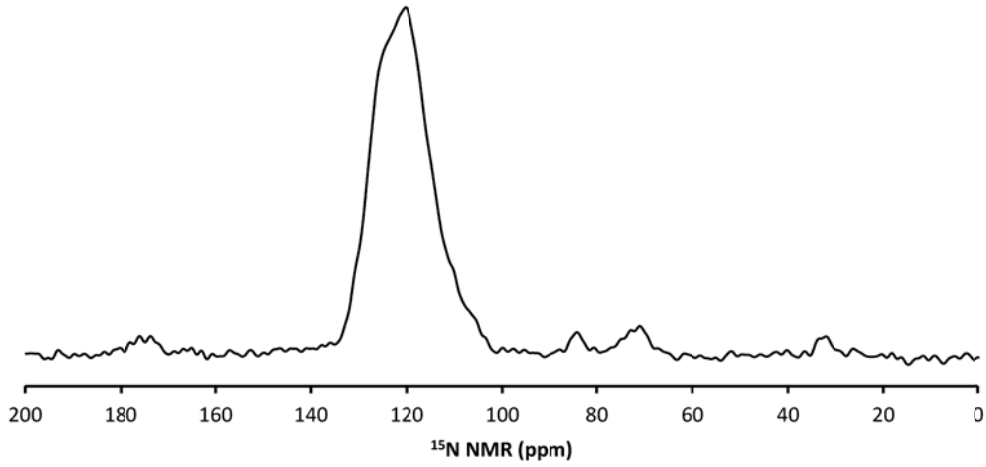

Figure 4.11 ^{15}N CP spectrum of TM1TM2 in DMPC, processed with 50 Hz line broadening.

4.3.3 2D ^{13}C-^{13}C and ^{13}C-^{15}N correlation experiments

Further characterization of TM1TM2 in lipids was carried out using a ^{13}C-^{13}C correlation spectrum with DARR mixing. The experiment was set up at -20 °C, 13.3 kHz MAS and using a 15 ms DARR mixing time. The spectrum of [U-^{15}N, ^{13}C] TM1TM2 in DMPC is shown in Figure 4.12. Some signals may be assigned tentatively in terms of their residue types, such as C_α-C_β correlations of alanine, valine, isoleucine, serine, threonine, and the C_α-C' correlation of glycine. The resolution of the spectrum is too limited for further assignment of the signals. In addition to the ^{13}C-^{13}C correlation spectrum, the 2D ^{15}N-^{13}C correlation spectra, namely N(CA)CX and N(CO)CX, were collected. Due to extensive signal overlap, the resonances could not be assigned. The 2D N(CA)CX spectrum is shown in Figure 4.13 as an example: most of signals overlap with each other in the NCA and N(CA)CO regions. Only a few signals that are distant spectroscopically from the overlapped region (indicated by a dashed circle in Figure 4.13) can be resolved, possibly corresponding to the C_β resonances of serine and threonine. Based on the analysis of these ^{13}C-^{13}C and ^{15}N-^{13}C correlation spectra, the solid state NMR resolution for [U-^{15}N, ^{13}C] TM1TM2 in lipid bilayer media remains too limited to achieve full signal assignment. The disappointing resolution of the spectra may be attributed to the conformational heterogeneity and flexibility of the peptide conformation in these lipid-rich model membrane protein systems.

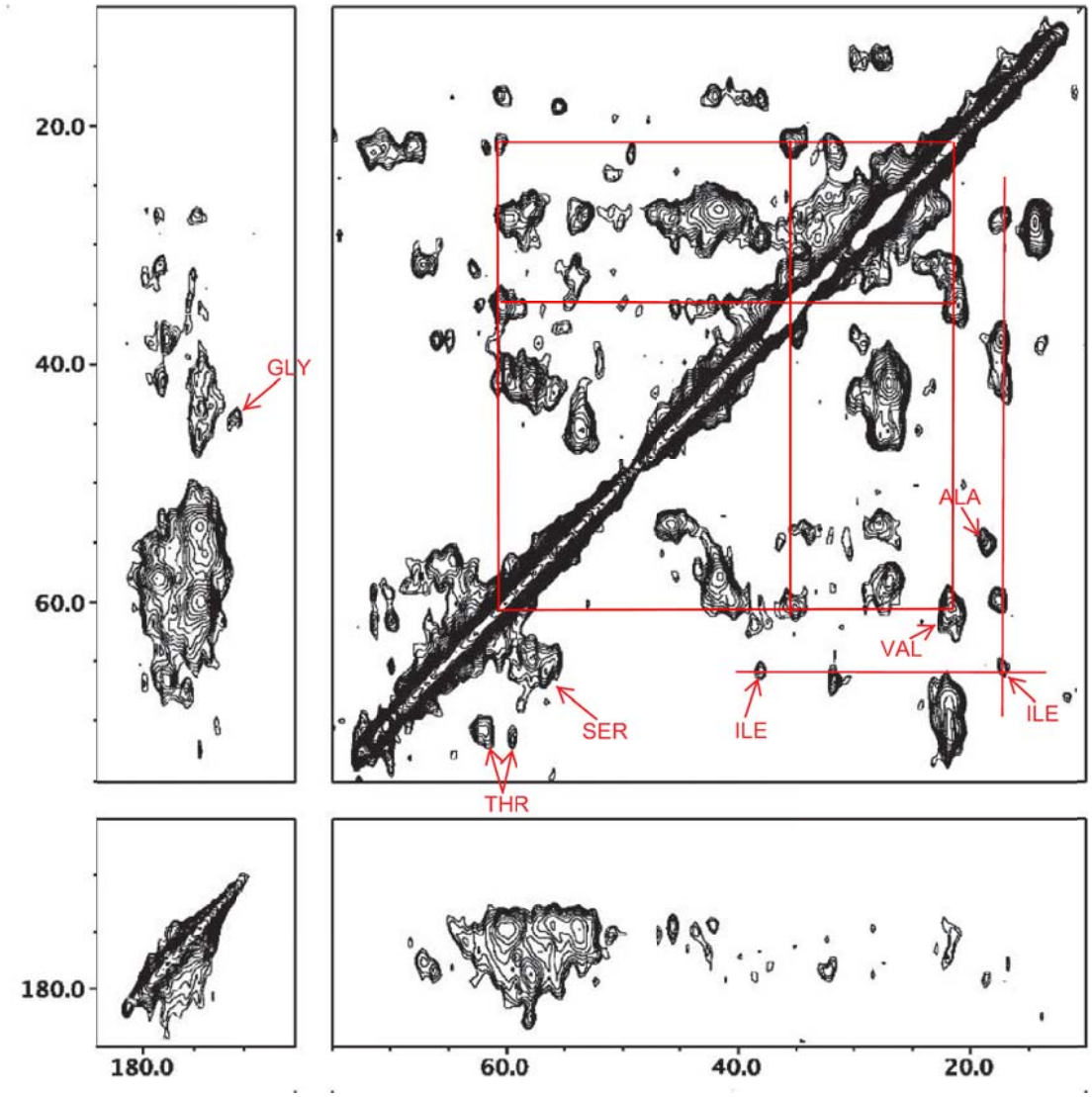

Figure 4.12 ^{13}C-^{13}C correlation spectrum with DARR mixing for [U-^{15}N,^{13}C]-TM1TM2 in DMPC. Assignments of residue types are labeled on the spectrum.

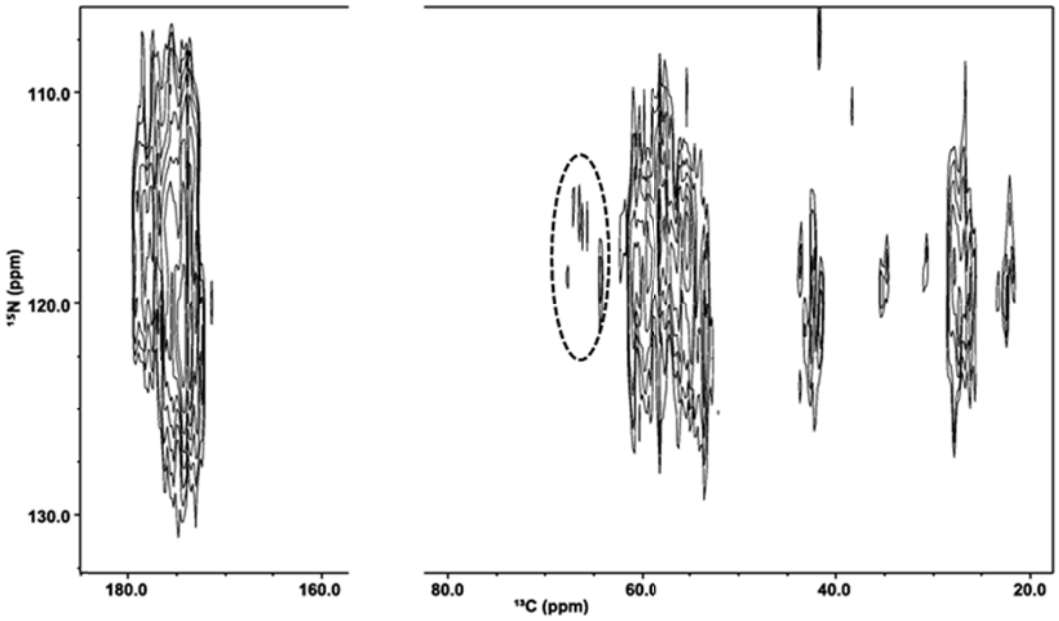

Figure 4.13 2D N(CA)CX spectrum of TM1TM2 in lipids.

4.4 SUMMARY

A preliminary study of TM1TM2 structure was conducted with solid state NMR. The peptide was reconstituted successfully into a helical conformation with DMPC/SDS in solution, as evaluated by solution state NMR. The study of a ^{15}N alanine-labeled peptide with DMPC/SDS hints that TM1TM2 folds into helical structures with a conformation similar to that determined in LPPG micelles[34] based on the similarity of HSQC spectra. Although well-resolved NMR spectra and reasonable sample stability could be achieved for ^{15}N alanine-labeled TM1TM2 in DMPC/SDS, addition of sufficient DMPC to form a good membrane mimetic also results in formation of large aggregates and unacceptable line broadening, especially for a uniformly labeled sample.

Solid state NMR experiments of TM1TM2 in DMPC multilayers show that some signals can be assigned by residue type, but further site-specific assignment was limited by the resolution of the spectra with [U-^{15}N, ^{13}C] TM1TM2. The low resolution could be due to heterogeneity of the prepared sample and/or the intrinsic mobility of the peptide. The sample preparation needs further improvement to make feasible a solid state NMR study of this peptide with lipids. Possible measures to consider are as follows: (1) try other lipids instead of DMPC; (2) use Bio-beads to remove the detergent used during reconstitution; (3) reconstitute the peptide in a nanodisc; (4) use reverse labeling or selective labeling to simplify and assign spectra.

CHAPTER 5 CONCLUSIONS

Both X-ray and NMR structures of holo-LFABP show that two ligands are bound within the protein cavity [4, 6]. A singly-liganded state was observed for the LFABP binding with oleate and the binding process has been proposed to be stepwise rather than simultaneous based on an NMR titration study [13]. The singly-liganded LFABP and R122L/S124A mutant has been investigated with a docking calculation that incorporates both NMR chemical shift perturbations and intermolecular NOE data. Modeled structures for singly-liganded LFABP and R122L/S124A LFABP have been generated successfully with the docking methodology. The structural model of singly-liganded LFABP provides a better understanding of the ligand binding process of LFABP in comparison with previously determined apo- and holo-protein structures[4]. The different binding behavior of the wild-type LFABP and R122L/S124A mutant indicates the importance of electrostatic interactions between the protein cavity and oleate in determining ligand orientation and conformation. Based on the analysis of these NMR and modeled structures, binding is proposed to occur as follows: an electrostatic interaction between the carboxylate group of oleate and a positively charged pocket containing residues arginine 122 and serine 124 initiates the binding of the first ligand in a specific orientation, leaving space available in the protein cavity; the available space together with the more hydrophobic interface provided by the aliphatic tail of the first ligand facilitates the binding of a second ligand with an orientation in which the aliphatic tail is buried in the protein cavity and the carboxylate group is close to the protein surface. Electrostatic interactions between residues R122, S124 and oleate are found to be crucial in determining the orientation and conformation for the first ligand, which then favors the binding of a second ligand.

A new DMPC-DIOHPC isotropic bicelle system has been evaluated as a membrane- mimetic medium for biomolecular NMR and related structural biology research. Compared with other similar systems, DMPC-DIOHPC isotropically tumbling bicelles exhibit improved hydrolytic stability and superior ^{31}P NMR peak separation, allowing more accurate measurements of integrated signal intensities and associated sizes. Moreover, the new system extends the range of experimental conditions over which constant size and aggregate organization can be maintained, including overall phospholipid concentrations down to c_L of 5%, temperatures from 25 to 46 ^{o}C, and presence of physiologically relevant salts at concentrations up to 200 mM. The q = 0.5 DMPC-DIOHPC bicelle system shows potential as a versatile medium for high-resolution NMR studies of peripheral membrane proteins, as illustrated by liver and intestinal FABPs. The bicelle interactions of intestinal and liver fatty acid-binding proteins that accomplish ligand transfer via different mechanisms have been differentiated according to the spatial distribution of their significantly perturbed residues and the overall amplitude of their chemical shift perturbations. In particular, the interaction between IFABP and membranes via a collisional mechanism is proposed to be localized in the helical regions of the protein based on the chemical shift perturbation results, a result that supports the functional importance inferred for this portal region in AFABP and HFABP [28, 79].

Preliminary structural characterization was carried out for the double transmembrane peptide TM1TM2 in membrane-mimetic media by solution- and solid-state NMR. The reconstitution of the peptide into a DMPC-SDS lipid-detergent mixture has been evaluated by solution-state NMR using ^{15}N HSQC. The study of an ^{15}N alanine-labeled peptide with DMPC-SDS indicates a preference of peptide to associate with DMPC and the adoption of a helical conformation resembling the structure reported in LPPG micelles[34]. The solid-state NMR experiments in

DMPC multibilayers show that some signals can be assigned based on their residue types in a [U-^{15}N, ^{13}C] TM1TM2 sample. Further assignment was limited by signal overlap and limited resolution of the spectra, which may be due to sample heterogeneity and/or mobility of the peptide in this lipid-rich mixture. Redesign of the sample preparation protocol would be required for further solid state NMR study of GPCR peptide fragments in membrane-mimetic media.

APPENDIX I A PRACTICAL CHEMICAL SHIFT REFERENCING PROCEDURE FOR SOLID-STATE NMR

Chemical shift referencing can be achieved by internal referencing, external referencing, substitution methods, or referencing via direct measurement of the absolute frequency of the field-frequency/lock signal (normally from a ^{2}H signal in the solvent)[131]. All of these methods are used for solution-state NMR and the last one is adopted in all commercially available spectrometers. That is why chemical shifts are always correct when the ^{2}H solvent signal is locked properly. The chemical shift for X nuclei is defined by the following equation:

$$\delta_{X,sample} = \frac{(\upsilon_{X,sample} - \upsilon_{X,reference})}{(\upsilon_{X,reference}/10^6)} \quad [131] \qquad (1)$$

Provided that the value $\upsilon_{reference}$ is determined, the chemical shift of the sample can be referenced.

Modern NMR spectrometers include field-frequency locking and frequency synthesizers, which provide interrelated frequencies for all nuclei through a master lock frequency. The frequencies are related by the Ξ value determined with the equation below:

$$\Xi/\% = 100\left(\upsilon_X^{obs} \Big/ \upsilon_{TMS}^{obs}\right) \quad [131] \qquad (2)$$

The values of Ξ have been summarized in a set of IUPAC recommendations [131], so that 0 ppm frequencies υ_X^{obs} for all nuclei can be calculated when the frequency of tetramethylsilane (TMS) is known, in another words when the proton spectrum has been referenced. The υ_X^{obs} in equation (2) is $\upsilon_{X,reference}$ in equation (1). The lock system works in such a way that an additional field

is added on top of the external magnetic field to make sure the proton spectrum is referenced, and then the frequency of TMS is deduced. Spectra of all other nuclei can be referenced from the 0 ppm frequency calculated by the Ξ value.

For solid-state NMR, chemical shift referencing is normally carried out by using the substitution method, which means spectra of the sample and reference compound are recorded individually. Nonetheless, it is difficult to find standard compounds for some nuclear species. If several different nuclei will be examined in a given sample, it must be swapped out with standard compounds several times. Moreover, some standard compounds have broad signals, introducing errors in chemical shift referencing. Herein, we developed a practical and convenient chemical shift referencing procedure for MAS solid-state NMR, which borrows the idea of a field-frequency lock from solution-state NMR.

As mentioned above, all nuclei can be referenced correctly with Ξ values, if the proton spectrum is referenced with a proper lock setup. As this method is based on the Ξ values, if a spectrum of any nucleus (not necessarily ^{1}H) is referenced, the other nuclei may be referenced by a calculation, normally done automatically with a macro set of instructions by the manufacturer. Our goal in the current instance was to find an alternative way to do the field/frequency locking. This "pseudo locking" is done by adjusting 'Z0' according to the signal of the methylene group of adamantane, which is normally used for referencing solid-state ^{13}C NMR spectra.

The data collected in a Varian VNMRS 600 instrument are used to illustrate the procedure. To determine the relationship between the chemical shift difference and Z0 value, ^{13}C NMR spectra of adamantane were collected at different Z0 values. The changes in chemical shift were plotted

vs. Z0 value as shown in Figure A.1. An equation that describes the relation between Z0 and the chemical shift change was found as shown below by linear fitting:

$$Z0 = -11.493 * \Delta_{CS}\,(Hz) + 5.2 \qquad (3).$$

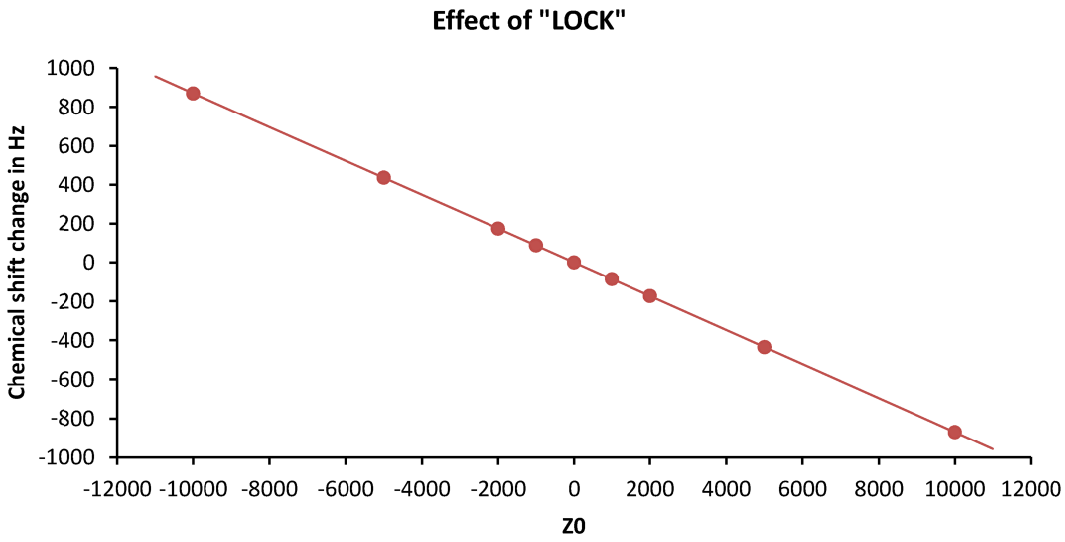

Figure A.1 Chemical shift change of adamantine CH_2 with various Z0 values.

Based on equation (3), the Z0 value can be calculated from the difference between the recorded methylene chemical shift and the reported value [130]. In this way, ^{13}C NMR can be referenced by the 'pseudo lock' with calculated Z0. For other nuclei, the chemical shift can be referenced based on calculation from equations (1) and (2). For the Agilent/Varian system, the calculation is done with the macro 'setref'. A macro is written for Z0 calculation as follows:

```
"res
r1=-11.49281*(38.48*sfrq-cr)+5.20311
r2=Z0+r1
Z0=r2"
```

The procedure proposed here is easy to set up and convenient to use for referencing chemical shifts in MAS solid-state NMR without disturbing any system parameter. The advantage of this method is that all nuclei can be referenced with one adamantane sample, which is also used for referencing in ^{13}C NMR and shimming to optimize the lineshape (3 Hz line width at half height) for MAS solid-state NMR. The reported precision of ±0.03 ppm for ^{13}C chemical shifts that is achieved easily with external adamantane referencing [130], was confirmed by the independent measurements in the current work. Referencing with Ξ values can be achieved to a precision of 0.01 ppm for nuclei with high γ values (commonly observed nuclei such as ^{1}H, ^{13}C, ^{31}P) and 0.5 ppm for nuclei with very low γ values [131]. The 'pseudo lock' described herein does not account for the field drift as is done by the lock system in solution-state NMR; such effects are normally not a problem for broad solid-state NMR signals if the magnet is reasonably stable (<5 Hz drift per day), but precautions need to be taken for experiments that last more than a week.

APPENDIX II TEMPERATURE CALIBRATION PROTOCOL FOR MAS NMR PROBES AT ARBITRARY SPINNING SPEEDS

(Adapted from "A General Protocol for Temperature Calibration of MAS NMR Probes at Arbitrary Spinning Speeds" X. Guan and R.E. Stark, *Solid State Nucl Magn Reson* (2010), 38, 74-76.)

Temperature control is essential to many physical studies of thermosensitive samples such as biomolecules and engineered materials. In both solution- and solid-state NMR, the disparities between the set temperature and the true sample temperature inside the probe are well known. Thus, careful calibration is required to obtain the desired value or range of sample temperatures for a particular investigation. The ^{207}Pb NMR resonance of lead nitrate ($Pb(NO_3)_2$) has become the accepted thermometer for calibration of sample temperature in magic-angle spinning (MAS) probes[133, 144, 145]. If the MAS rate is varied, however, the heating effects due to friction between the bearings, the variable temperature (VT) gas, and the spinning rotor complicate the relationship between frictional heating and heat diffusion. Thus to achieve comprehensive temperature calibration, an empirical approach may be preferred to an analytical formalism for description of the sample temperature as a function of both set temperature and MAS rate.

To test this idea, we used a Varian VNMRS system equipped with a BioMAS probe (3.2 mm, 3.0 - 13.3 kHz) and a FastMAS probe (1.6 mm, 15 - 35 kHz), each operating in a 14.1 T magnet. Both probes supply VT gas and bearing gas separately to achieve temperature regulation. The VT gas flow rate was set to 40 liters per minute using an FTS low-temperature source running at -70 ºC, whereas the bearing and drive gas were both operated at ambient temperature. For the

BioMAS probe, the drive gas pressure was set between 3 and 13 psi, whereas the bearing gas was varied between 11 and 20 psi. For the FastMAS probe, the drive gas was varied between 3 and 40 psi, the bearing gas between 5 and 15 psi. The set temperature range was between -30.0 and 45.0 ºC, controlled to within ±0.2 ºC. The true temperature at each spinning condition was calculated based on the linear relationship between the chemical shift of lead nitrate and the sample temperature: $\delta(T) = -3714.6$ ppm + (0.760 ppm/K) * T, where T is the temperature in degrees Kelvin[144]. The chemical shifts of ^{207}Pb were referenced to 0 ppm based on a calculation from frequency ratios of ^{207}Pb and ^{13}C, as recommended by IUPAC guidelines[131]. The ^{13}C spectrum was referenced directly to the methylene carbon of adamantane at 38.48 ppm[130]. Using this method, the chemical shifts of lead nitrate were determined more conveniently and accurately than by analyzing the powder lineshape of a static sample to derive the isotropic chemical shift under ambient sample conditions[132].

To illustrate the procedure, we focus mainly on trials with the FastMAS probe. As expected, excellent linear correlations ($R^2 \geqq 0.9998$) were observed between sample temperature (T_s) and set temperature (T_o) at each MAS rate (ω_r) for both the BioMAS and FastMAS probes (Figure A.2) and as reported previously[133, 144, 146]. Each of these relationships may be expressed as

$$T_s(T_o,\omega_r) = a(\omega_r) * T_o + b(\omega_r) \qquad (1).$$

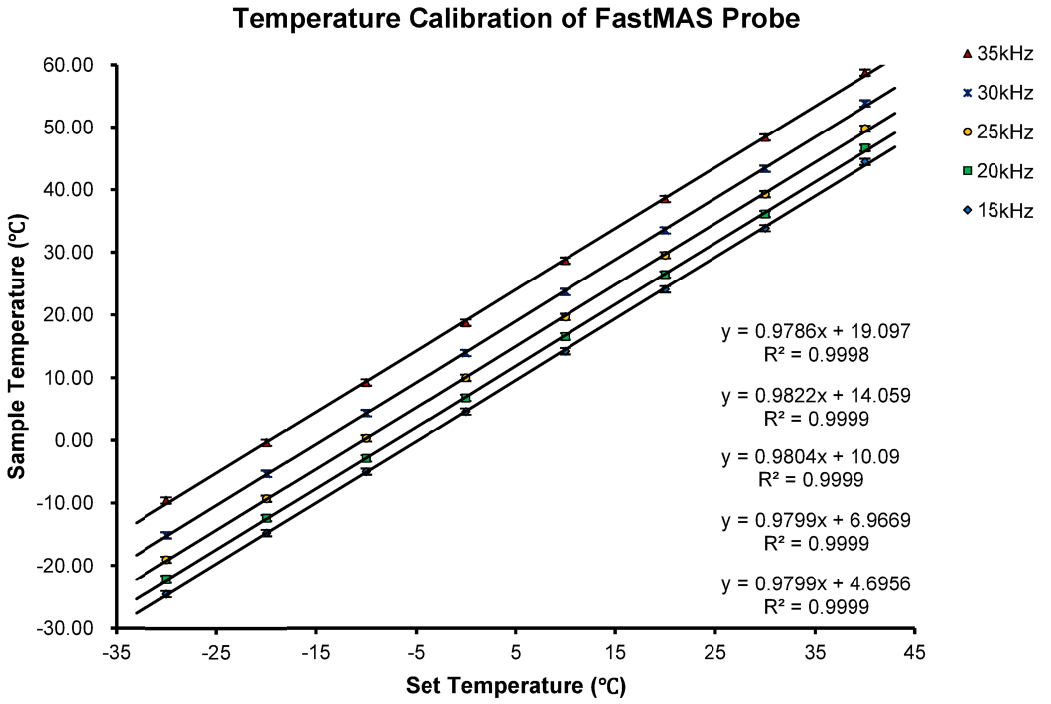

Figure A.2 Sample temperature (T_s) vs. set temperature (T_o) at different MAS rates (ω_r) for a BioMAS probe. Similar results are obtained for a FastMAS probe (data not shown). Error limits were estimated from the linewidth of the lead nitrate signal in the ^{207}Pb NMR spectrum.

To find the sample temperature at any given set temperature (T_o) and MAS rate (ω_r), it is necessary to incorporate the dependence of the MAS rate into Eq. 1. This is possible by expressing $a(\omega_r)$ and $b(\omega_r)$ as functions of MAS rate (ω_r). Considering the slopes $a(\omega_r)$ first, the values are found to be nearly identical at different spin rates 0.9786 - 0.9822 shown in Figure A.2 for the FastMAS probe ; 0.9692 - 0.9736 for the BioMAS probe (not shown), also reported previously for a 2.5-mm Bruker probe[146]. By using degrees Celsius as the unit of temperature, the effect of the tiny differences among slopes is minimized. Thus Eq. 1 may be simplified by using an average value of $a(\omega_r)$. Turning to the intercepts $b(\omega_r)$, Figure A.3 shows an exponential dependence on the MAS rate:

$b(\omega_r) = b_0 + A * \exp(\omega_r / r)$ (2), where A and r are constants.

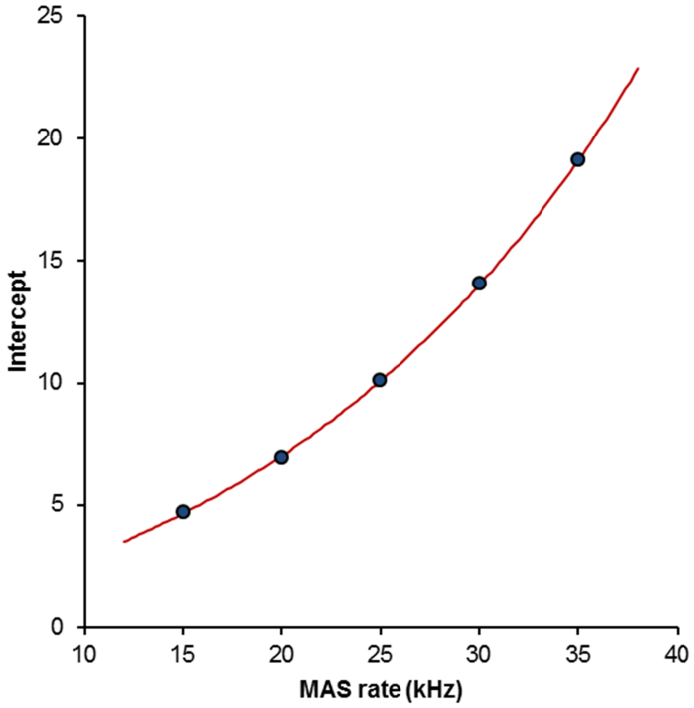

Figure A.3 Dependence of the intercept $b(\omega_r)$ on MAS rate ω_r for the FastMAS probe.

Neither step of this procedure involves more than three adjustable parameters, so the coefficients $a(\omega_r)$ and $b(\omega_r)$ should be better determined than with the 5-paramter fit described in Ref.. Thus the temperature is specified completely for the FastMAS probe using the following coefficients:

$T_s(f) = 0.98 * T_o + 3.79\ ^{o}C * \exp(\omega_r/19.6\ kHz) - 3.49\ ^{o}C$ (3).

Using an analogous procedure, the equation for the BioMAS probe is found to be

$T_s(b) = 0.97 * T_o + 1.34\ ºC * \exp(\omega_r/7.53\ kHz) - 0.77\ ºC$ (4).

In practice, these results show, for instance, that a set temperature of -10 ºC for a delicate sample spinning at 10 kHz corresponds to -5.4 ºC in the BioMAS probe and -7.0 ºC in the fastMAS probe, respectively.

Eqs. (3) and (4) add to the researcher's physical insight by revealing the independent contributions of the set temperature (T_o) and MAS rate (ω_r) to the sample temperature (T_s). In the absence of MAS ($\omega_r = 0$), Eq (3) becomes $T_s(b) = 0.98 * T_o + 0.30$ ºC, resembling a typical equation used for temperature calibration of liquid-state NMR probes that are not equipped for MAS. Conversely, the heating effect arising from MAS is reflected only in the intercept $b(\omega_r)$. Although the frictional heating should be proportional to the square of the MAS rate (ω_r), the heat generated is not absorbed completely by the rotor, but instead diffuses to some extent with the VT and bearing gases. As a consequence, $b(\omega_r)$ need not be proportional to the square of the MAS rate (ω_r); in our trials an exponential growth provided the best empirical fit.

Because the separate bearing and VT gas streams are directed at the drive tip and the central sample-containing part of the rotor, respectively, we expected that the main source of the frictional heating would involve the rotor and the VT gas. This hypothesis was confirmed for the BioMAS probe: using an 8.0-kHz spin rate and temperature setting of 22 °C, we measured an essentially negligible 0.1 °C change in temperature when the bearing gas pressure was varied between 13.4 and 19.1 psi. This negative result suggests that even in challenging spinning situations that require raising the bearing gas pressure by as much as 2-3 psi (e.g., thin-walled rotors, heterogeneous materials), its impact on frictional heating may be neglected. However,

the flow rate and pressure of the VT gas should be kept constant. For probes that use the bearing gas to regulate temperature, the operating pressure should be maintained at a constant value.

Finally, the reliability of the sample temperatures (T_s) predicted from Eqs. (3) and (4) was assessed by comparing the measured sample temperatures for both BioMAS and FastMAS probes. Figure A.4 shows an excellent correlation between the predicted and measured values, with slopes close to 1 (0.9985, 1.0023) and intercepts near 0 (0.0114, 0.0131). The maximum deviations in our data sets were 0.56 ºC for the BioMAS probe and 0.79 ºC for the FastMAS probe, respectively, acceptable for most practical applications and superior to deviations of ±2 K reported about a decade ago[146]. In addition to consistency, the excellent accuracy of our protocol is illustrated by the finding that under ambient conditions (room temperature of 20 ºC, no spinning), the predicted temperatures are 19.97 and 19.90 ºC for BioMAS and Fast MAS probes, respectively.

In summary, these results demonstrate the feasibility of calibrating NMR sample temperatures for any MAS probe at spinning rates up to 35 kHz using $Pb(NO_3)_2$, using a method that is straightforward, reliable, and physically insightful. The true sample temperature can be predicted accurately with an empirical equation based on set temperature and MAS rate, allowing us to adjust the sample temperature as desired for spectroscopic experiments and (bio)chemical materials of interest.

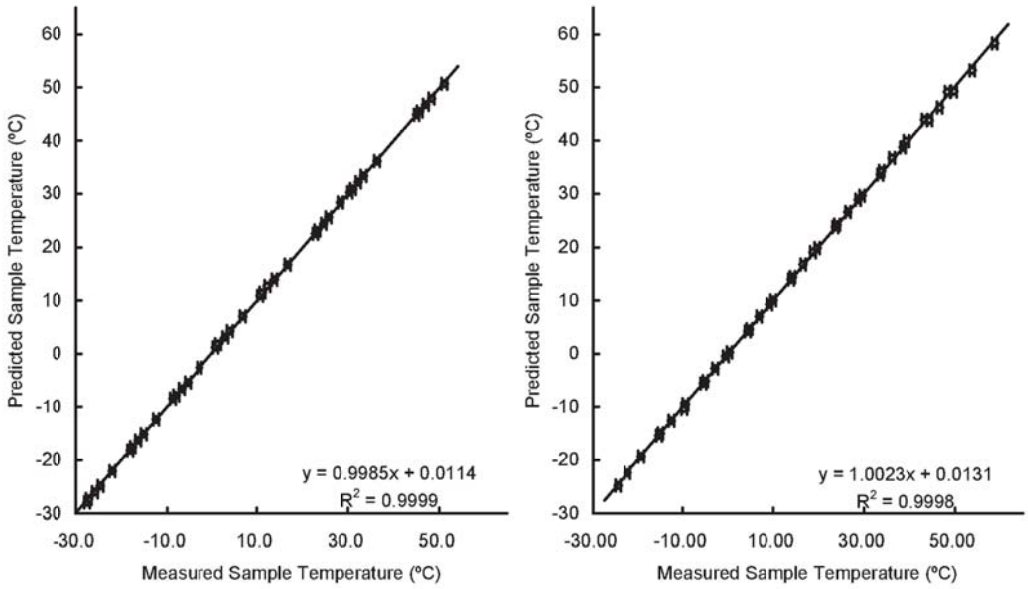

Figure A.4 Predicted sample temperature vs. measured sample temperature for BioMAS (a) and FastMAS (b) probes. Error limits were estimated from the linewidth of the lead nitrate signal in the ^{207}Pb NMR spectrum.

REFERENCES

1. Storch, J., and Thumser, A. E. A. (2000) The fatty acid transport function of fatty acid-binding proteins, *Bba-Mol Cell Biol L 1486*, 28-44.

2. Hostetler, H. A., McIntosh, A. L., Atshaves, B. P., Storey, S. M., Payne, H. R., Kier, A. B., and Schroeder, F. (2009) *L-FABP directly interacts with PPARalpha in cultured primary hepatocytes*, Vol. 50.

3. Zhang, F., Lücke, C., Baier, L. J., Sacchettini, J. C., and Hamilton, J. A. (2003) Solution Structure of Human Intestinal Fatty Acid Binding Protein with a Naturally-Occurring Single Amino Acid Substitution (A54T) that Is Associated with Altered Lipid Metabolism†,‡, *Biochemistry 42*, 7339-7347.

4. He, Y., Yang, X., Wang, H., Estephan, R., Francis, F., Kodukula, S., Storch, J., and Stark, R. E. (2007) Solution-state molecular structure of apo and oleate-liganded liver fatty acid-binding protein, *Biochemistry 46*, 12543-12556.

5. Hellberg, K., Grimsrud, P. A., Kruse, A. C., Banaszak, L. J., Ohlendorf, D. H., and Bernlohr, D. A. (2010) X-ray crystallographic analysis of adipocyte fatty acid binding protein (aP2) modified with 4-hydroxy-2-nonenal, *Protein Sci 19*, 1480-1489.

6. Thompson, J., Winter, N., Terwey, D., Bratt, J., and Banaszak, L. (1997) The crystal structure of the liver fatty acid-binding protein - A complex with two bound oleates, *Journal of Biological Chemistry 272*, 7140-7150.

7. Xu, Y., Long, D., and Yang, D. (2007) Rapid data collection for protein structure determination by NMR spectroscopy, *J Am Chem Soc 129*, 7722-7723.

8. Sacchettini, J. C., Gordon, J. I., and Banaszak, L. J. (1989) Crystal structure of rat intestinal fatty-acid-binding protein. Refinement and analysis of the Escherichia coli-derived protein with bound palmitate, *J Mol Biol 208*, 327-339.

9. Young, A. C. M., Scapin, G., Kromminga, A., Patel, S. B., Veerkamp, J. H., and Sacchettini, J. C. (1994) Structural Studies on Human Muscle Fatty-Acid-Binding Protein at 1.4-Angstrom Resolution - Binding Interactions with 3 C18 Fatty-Acids, *Structure 2*, 523-534.

10. Hodsdon, M. E., and Cistola, D. P. (1997) Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by 15N NMR relaxation and 1H exchange, *Biochemistry 36*, 2278-2290.

11. Gordon, J. I., Elshourbagy, N., Lowe, J. B., Liao, W. S., Alpers, D. H., and Taylor, J. M. (1985) Tissue specific expression and developmental regulation of two genes coding for rat fatty acid binding proteins, *The Journal of Biological Chemistry 260*, 1995-1998.

12. Nemecz, G., Jefferson, J. R., and Schroeder, F. (1991) Polyene Fatty-Acid Interactions with Recombinant Intestinal and Liver Fatty Acid-Binding Proteins - Spectroscopic Studies, *Journal of Biological Chemistry 266*, 17112-17123.

13. He, Y., Estephan, R., Yang, X., Vela, A., Wang, H., Bernard, C., and Stark, R. E. (2011) A Nuclear Magnetic Resonance-Based Structural Rationale for Contrasting Stoichiometry and Ligand Binding Site(s) in Fatty Acid-Binding Proteins, *Biochemistry 50*, 1283-1295.

14. Estephan, R. (2005) Effect of R122L/S124A and K31L mutations on molecular structure and function in rat liver fatty acid-binding protein., *Ph.D. Dissertation, City University of New York, New York.*

15. Prosser, R. S., Evanics, F., Kitevski, J. L., and Al-Abdul-Wahid, M. S. (2006) Current applications of bicelles in NMR studies of membrane-associated amphiphiles and proteins, *Biochemistry 45*, 8453-8465.

16. Ottiger, M., and Bax, A. (1998) Characterization of magnetically oriented phospholipid micelles for measurement of dipolar couplings in macromolecules, *J Biomol NMR 12*, 361-372.

17. Whiles, J. A., Deems, R., Vold, R. R., and Dennis, E. A. (2002) Bicelles in structure-function studies of membrane-associated proteins, *Bioorg Chem 30*, 431-442.

18. Vold, R. R., Prosser, R. S., and Deese, A. J. (1997) Isotropic solutions of phospholipid bicelles: a new membrane mimetic for high-resolution NMR studies of polypeptides, *Journal of biomolecular NMR 9*, 329-335.

19. Poget, S. F., and Girvin, M. E. (2007) Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better, *Biochimica et biophysica acta 1768*, 3098-3106.

20. Poget, S. F., Cahill, S. M., and Girvin, M. E. (2007) Isotropic bicelles stabilize the functional form of a small multidrug-resistance pump for NMR structural studies, *Journal of the American Chemical Society 129*, 2432-2433.

21. Lee, D., Walter, K. F., Bruckner, A. K., Hilty, C., Becker, S., and Griesinger, C. (2008) Bilayer in small bicelles revealed by lipid-protein interactions using NMR spectroscopy, *Journal of the American Chemical Society 130*, 13822-13823.

22. Grit, M., and Crommelin, D. J. A. (1993) Chemical stability of liposomes: implications for their physical stability, *Chemistry and Physics of Lipids 64*, 3-18.

23. Ottiger, M., and Bax, A. (1999) Bicelle-based liquid crystals for NMR-measurement of dipolar couplings at acidic and basic pH values, *Journal of Biomolecular NMR 13*, 187-191.

24. Hsu, K. T., and Storch, J. (1996) Fatty acid transfer from liver and intestinal fatty acid-binding proteins to membranes occurs by different mechanisms, *Journal of Biological Chemistry 271*, 13317-13323.

25. Thumser, A. E., Tsai, J., and Storch, J. (2001) Collision-mediated transfer of long-chain fatty acids by neural tissue fatty acid-binding proteins (FABP): studies with fluorescent analogs, *J Mol Neurosci 16*, 143-150; discussion 151-147.

26. Kurz, M., Brachvogel, V., Matter, H., Stengelin, S., Thuring, H., and Kramer, W. (2003) Insight into the bile acid transportation system: The human ileal lipid-binding protein-cholyltaurine complex and its comparison with homologous structures, *Proteins 50*, 312-328.

27. Kim, H. K., and Storch, J. (1992) Mechanism of free fatty acid transfer from rat heart fatty acid-binding protein to phospholipid membranes. Evidence for a collisional process, *Journal of Biological Chemistry 267*, 20051-20056.

28. Herr, F. M., Aronson, J., and Storch, J. (1996) Role of Portal Region Lysine Residues in Electrostatic Interactions between Heart Fatty Acid Binding Protein and Phospholipid Membranes†, *Biochemistry 35*, 1296-1303.

29. Gericke, A., Smith, E. R., Moore, D. J., Mendelsohn, R., and Storch, J. (1997) Adipocyte fatty acid-binding protein: Interaction with phospholipid membranes and thermal stability studied by FTIR spectroscopy, *Biochemistry 36*, 8311-8317.

30. Thumser, A. E. A., and Storch, J. (2000) Liver and intestinal fatty acid-binding proteins obtain fatty acids from phospholipid membranes by different mechanisms, *J Lipid Res 41*, 647-656.

31. Wootan, M. G., Bernlohr, D. A., and Storch, J. (1993) Mechanism of fluorescent fatty acid transfer from adipocyte fatty acid binding protein to membranes, *Biochemistry 32*, 8622-8627.

32. Hopkins, A. L., and Groom, C. R. (2002) The druggable genome, *Nat Rev Drug Discov 1*, 727-730.

33. Martin, N. P., Leavitt, L. M., Sommers, C. M., and Dumont, M. E. (1999) Assembly of G protein-coupled receptors from fragments: identification of functional receptors with discontinuities in each of the loops connecting transmembrane segments, *Biochemistry 38*, 682-695.

34. Neumoin, A., Cohen, L. S., Arshava, B., Tantry, S., Becker, J. M., Zerbe, O., and Naider, F. (2009) Structure of a double transmembrane fragment of a G-protein-coupled receptor in micelles, *Biophys J 96*, 3187-3196.

35. Cohen, L. S. (2010) Structure Determination of a Double Transmembrane Fragment of the G protein-coupled Receptor Ste2p in Membrane Mimetic Environments.

36. McDermott, A. (2009) Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR, *Annu Rev Biophys 38*, 385-403.

37. Tycko, R. (2001) Biomolecular solid state NMR: advances in structural methodology and applications to peptide and protein fibrils, *Annu Rev Phys Chem 52*, 575-606.

38. Baldus, M. (2006) Solid-state NMR spectroscopy: molecular structure and organization at the atomic level, *Angew Chem Int Ed Engl 45*, 1186-1188.

39. Luca, S., Heise, H., and Baldus, M. (2003) High-resolution solid-state NMR applied to polypeptides and membrane proteins, *Acc Chem Res 36*, 858-865.

40. Kijac, A. Z., Li, Y., Sligar, S. G., and Rienstra, C. M. (2007) Magic-angle spinning solid-state NMR spectroscopy of nanodisc-embedded human CYP3A4, *Biochemistry 46*, 13696-13703.

41. Li, Y., Berthold, D. A., Gennis, R. B., and Rienstra, C. M. (2008) Chemical shift assignment of the transmembrane helices of DsbB, a 20-kDa integral membrane enzyme, by 3D magic-angle spinning NMR spectroscopy, *Protein science : a publication of the Protein Society 17*, 199-204.

42. Seidel, K., Andronesi, O. C., Krebs, J., Griesinger, C., Young, H. S., Becker, S., and Baldus, M. (2008) Structural characterization of Ca(2+)-ATPase-bound phospholamban in lipid bilayers by solid-state nuclear magnetic resonance (NMR) spectroscopy, *Biochemistry 47*, 4369-4376.

43. Huang, L., and McDermott, A. E. (2008) Partial site-specific assignment of a uniformly (13)C, (15)N enriched membrane protein, light-harvesting complex 1 (LH1), by solid state NMR, *Biochimica et biophysica acta 1777*, 1098-1108.

44. Shi, L. C., Ahmed, M. A. M., Zhang, W. R., Whited, G., Brown, L. S., and Ladizhansky, V. (2009) Three-Dimensional Solid-State NMR Study of a Seven-Helical Integral Membrane Proton Pump-Structural Insights, *Journal of Molecular Biology 386*, 1078-1093.

45. Thumser, A. E., Evans, C., Worrall, A. F., and Wilton, D. C. (1994) Effect on ligand binding of arginine mutations in recombinant rat liver fatty acid-binding protein, *Biochem J 297 (Pt 1)*, 103-107.

46. Thumser, A. E., Voysey, J., and Wilton, D. C. (1996) Mutations of recombinant rat liver fatty acid-binding protein at residues 102 and 122 alter its structural integrity and affinity for physiological ligands, *Biochem J 314 (Pt 3)*, 943-949.

47. Alexov, E. G., and Gunner, M. R. (1997) Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties, *Biophys J 72*, 2075-2093.

48. Alexov, E. G., and Gunner, M. R. (1999) Calculated protein and proton motions coupled to electron transfer: electron transfer from QA- to QB in bacterial photosynthetic reaction centers, *Biochemistry 38*, 8253-8270.

49. Hodsdon, M. E., Ponder, J. W., and Cistola, D. P. (1996) The NMR solution structure of intestinal fatty acid-binding protein complexed with palmitate: application of a novel distance geometry algorithm, *Journal of Molecular Biology 264*, 585-602.

50. Xu, Z. H., Bernlohr, D. A., and Banaszak, L. J. (1993) The Adipocyte Lipid-Binding Protein at 1.6-a Resolution - Crystal-Structures of the Apoprotein and with Bound Saturated and Unsaturated Fatty-Acids, *Journal of Biological Chemistry 268*, 7874-7884.

51. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information, *J Am Chem Soc 125*, 1731-1737.

52. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr D Biol Crystallogr 54*, 905-921.

53. Tomaselli, S., Ragona, L., Zetta, L., Assfalg, M., Ferranti, P., Longhi, R., Bonvin, A. M., and Molinari, H. (2007) NMR-based modeling and binding studies of a ternary complex between chicken liver bile acid binding protein and bile acids, *Proteins 69*, 177-191.

54. Medek, A., Hajduk, P. J., Mack, J., and Fesik, S. W. (2000) The use of differential chemical shifts for determining the binding site location and orientation of protein-bound ligands, *Journal of the American Chemical Society 122*, 1241-1242.

55. Mulder, F. A., Schipper, D., Bott, R., and Boelens, R. (1999) Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins, *J Mol Biol 292*, 111-123.

56. Zwahlen, C., Legault, P., Vincent, S. J. F., Greenblatt, J., Konrat, R., and Kay, L. E. (1997) Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: Application to a bacteriophage lambda N-peptide/boxB RNA complex, *Journal of the American Chemical Society 119*, 6711-6721.

57. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *Journal of biomolecular NMR 6*, 277-293.

58. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol Biol 278*, 313-352.

59. He, Y. (2000) NMR Studyies of Rat Liver Fatty Acid-Binding Protein (LFABP), *Ph.D. Dissertation, City University of New York, New York.*

60. Wang, H., He, Y., Kroenke, C. D., Kodukula, S., Storch, J., Palmer, A. G., and Stark, R. E. (2002) Titration and exchange studies of liver fatty acid-binding protein with 13C-labeled long-chain fatty acids, *Biochemistry 41*, 5453-5461.

61. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, *J Biomol NMR 44*, 213-223.

62. Schuttelkopf, A. W., and van Aalten, D. M. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta Crystallogr D Biol Crystallogr 60*, 1355-1363.

63. Jorgensen, W. L., and Tiradorives, J. (1988) The Opls Potential Functions for Proteins - Energy Minimizations for Crystals of Cyclic-Peptides and Crambin, *Journal of the American Chemical Society 110*, 1657-1666.

64. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M., and Nilges, M. (2003) Refinement of protein structures in explicit solvent, *Proteins 50*, 496-506.

65. Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W. F., and Mark, A. E. (1999) Peptide Folding: When Simulation Meets Experiment, *Angewandte Chemie International Edition 38*, 236-240.

66. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *Journal of Biomolecular NMR 8*, 477-486.

67. Chuang, S., Velkov, T., Horne, J., Porter, C. J., and Scanlon, M. J. (2008) Characterization of the drug binding specificity of rat liver fatty acid binding protein, *J Med Chem 51*, 3755-3764.

68. Glover, K. J., Whiles, J. A., Wu, G., Yu, N., Deems, R., Struppe, J. O., Stark, R. E., Komives, E. A., and Vold, R. R. (2001) Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules, *Biophys J 81*, 2163-2171.

69. Luchette, P. A., Vetman, T. N., Prosser, R. S., Hancock, R. E., Nieh, M. P., Glinka, C. J., Krueger, S., and Katsaras, J. (2001) Morphology of fast-tumbling bicelles: a small angle neutron scattering and NMR study, *Biochimica et biophysica acta 1513*, 83-94.

70. Gaemers, S., and Bax, A. (2001) Morphology of three lyotropic liquid crystalline biological NMR media studied by translational diffusion anisotropy, *Journal of the American Chemical Society 123*, 12343-12352.

71. Vold, R. R., and Prosser, R. S. (1996) Magnetically oriented phospholipid bilayered micelles for structural studies of polypeptides. Does the ideal bicelle exist?, *J Magn Reson Ser B 113*, 267-271.

72. Durr, U. H., Waskell, L., and Ramamoorthy, A. (2007) The cytochromes P450 and b5 and their reductases--promising targets for structural studies by advanced solid-state NMR spectroscopy, *Biochimica et biophysica acta 1768*, 3235-3259.

73. Durr, U. H., Yamamoto, K., Im, S. C., Waskell, L., and Ramamoorthy, A. (2007) Solid-state NMR reveals structural and dynamical properties of a membrane-anchored electron-carrier protein, cytochrome b5, *Journal of the American Chemical Society 129*, 6670-6671.

74. Smith, P. E., Brender, J. R., and Ramamoorthy, A. (2009) Induction of negative curvature as a mechanism of cell toxicity by amyloidogenic peptides: the case of islet amyloid polypeptide, *Journal of the American Chemical Society 131*, 4470-4478.

75. Xu, J., Durr, U. H., Im, S. C., Gan, Z., Waskell, L., and Ramamoorthy, A. (2008) Bicelle-enabled structural studies on a membrane-associated cytochrome B5 by solid-state MAS NMR spectroscopy, *Angew Chem Int Ed Engl 47*, 7864-7867.

76. Prosser, R. S., Hunt, S. A., DiNatale, J. A., and Vold, R. R. (1996) Magnetically aligned membrane model systems with positive order parameter: Switching the sign of S-zz with paramagnetic ions, *Journal of the American Chemical Society 118*, 269-270.

77. Prosser, R. S., Volkov, V. B., and Shiyanovskaya, I. V. (1998) Novel chelate-induced magnetic alignment of biological membranes, *Biophys J 75*, 2163-2169.

78. Franchini, G. R., Storch, J., and Corsico, B. (2008) The integrity of the alpha-helical domain of intestinal fatty acid binding protein is essential for the collision-mediated transfer of fatty acids to phospholipid membranes, *Biochim Biophys Acta 1781*, 192-199.

79. Liou, H. L., and Storch, J. (2001) Role of surface lysine residues of adipocyte fatty acid-binding protein in fatty acid transfer to phospholipid vesicles, *Biochemistry 40*, 6475-6485.

80. Corsico, B., Franchini, G. R., Hsu, K. T., and Storch, J. (2005) Fatty acid transfer from intestinal fatty acid binding protein to membranes: electrostatic and hydrophobic interactions, *J Lipid Res 46*, 1765-1772.

81. Corsico, B., Liou, H. L., and Storch, J. (2004) The alpha-helical domain of liver fatty acid binding protein is responsible for the diffusion-mediated transfer of fatty acids to phospholipid membranes, *Biochemistry 43*, 3600-3607.

82. Francis, F. (2007) Solution-State NMR Studies of Fatty Acid Binding Proteins (FABs), *Ph.D. Dissertation, City University of New York, New York.*

83. Mandal, P. K., and Majumdar, A. (2004) A comprehensive discussion of HSQC and HMQC pulse sequences, *Concepts in Magnetic Resonance Part A 20A*, 1-23.

84. Cavanagh, J., Fairbrother, W. J., Arthur G. Palmer, I., Rance, M., and Skelton, N. J. (2007) Protein NMR Spectroscopy (Second Edition): Principles and Practice.

85. Losonczi, J. A., and Prestegard, J. H. (1998) Improved dilute bicelle solutions for high-resolution NMR of biological macromolecules, *Journal of biomolecular NMR 12*, 447-451.

86. Raffard, G., Steinbruckner, S., Arnold, A., Davis, J. H., and Dufourc, E. J. (2000) Temperature−Composition Diagram of Dimyristoylphosphatidylcholine−Dicaproylphosphatidylcholine "Bicelles" Self-Orienting in the Magnetic Field. A Solid State 2H and 31P NMR Study, *Langmuir 16*, 7655-7662.

87. Dufourc, E. J., Aussenac, F., and Lavigne, B. (2005) Toward bicelle stability with ether-linked phospholipids: Temperature, composition, and hydration diagrams by H-2 and P-31 solid-state NMR, *Langmuir 21*, 7129-7135.

88. Nagle, J. F., and Tristram-Nagle, S. (2000) Structure of lipid bilayers, *Biochimica et biophysica acta 1469*, 159-195.

89. Tausk, R. J. M., Karmiggelt, J., Oudshoorn, C., and Overbeek, J. T. G. (1974) Physical chemical studies of short-chain lecithin homologues. I. : Influence of the chain length of the fatty acid ester and of electrolytes on the critical micelle concentration, *Biophysical Chemistry 1*, 175-183.

90. Lin, T. L., Chen, S. H., Gabriel, N. E., and Roberts, M. F. (1986) Use of Small-Angle Neutron-Scattering to Determine the Structure and Interaction of Dihexanoylphosphatidylcholine Micelles, *Journal of the American Chemical Society 108*, 3499-3507.

91. van Dam, L., Karlsson, G., and Edwards, K. (2004) Direct observation and characterization of DMPC/DHPC aggregates under conditions relevant for biological solution NMR, *Biochimica et biophysica acta 1664*, 241-256.

92. Soong, R., and Macdonald, P. M. (2009) Water diffusion in bicelles and the mixed bicelle model, *Langmuir : the ACS journal of surfaces and colloids 25*, 380-390.

93. Triba, M. N., Warschawski, D. E., and Devaux, P. F. (2005) Reinvestigation by phosphorus NMR of lipid distribution in bicelles, *Biophys J 88*, 1887-1901.

94. Yamamoto, K., Soong, R., and Ramamoorthy, A. (2009) Comprehensive analysis of lipid dynamics variation with lipid composition and hydration of bicelles using nuclear magnetic resonance (NMR) spectroscopy, *Langmuir : the ACS journal of surfaces and colloids 25*, 7010-7018.

95. Arnold, A., Labrot, T., Oda, R., and Dufourc, E. J. (2002) Cation modulation of bicelle size and magnetic alignment as revealed by solid-state NMR and electron microscopy, *Biophys J 83*, 2667-2680.

96. Wang, H., He, Y., Hsu, K. T., Magliocca, J. F., Storch, J., and Stark, R. E. (1998) H-1, N-15 and C-13 resonance assignments and secondary structure of apo liver fatty acid-binding protein, *Journal of Biomolecular Nmr 12*, 197-199.

97. Hodsdon, M. E., Toner, J. J., and Cistola, D. P. (1995) 1H, 13C and 15N assignments and chemical shift-derived secondary structure of intestinal fatty acid-binding protein, *Journal of Biomolecular NMR 6*, 198-210.

98. Ropson, I. J., Yowler, B. C., Dalessio, P. M., Banaszak, L., and Thompson, J. (2000) Properties and Crystal Structure of a [beta]-Barrel Folding Mutant, *Biophys J 78*, 1551-1560.

99. Corsico, B., Cistola, D. P., Frieden, C., and Storch, J. (1998) The helical domain of intestinal fatty acid binding protein is critical for collisional transfer of fatty acids to phospholipid membranes, *Proc Natl Acad Sci U S A 95*, 12174-12178.

100. Arevalo, E., Estephan, R., Madeo, J., Arshava, B., Dumont, M., Becker, J. M., and Naider, F. (2003) Biosynthesis and biophysical analysis of domains of a yeast G protein-coupled receptor, *Biopolymers 71*, 516-531.

101. Gether, U. (2000) Uncovering molecular mechanisms involved in activation of G protein-coupled receptors, *Endocrine Reviews 21*, 90-113.

102. Carpenter, E. P., Beis, K., Cameron, A. D., and Iwata, S. (2008) Overcoming the challenges of membrane protein crystallography, *Curr Opin Struct Biol 18*, 581-586.

103. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor, *Science 289*, 739-745.

104. Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G., Tate, C. G., and Schertler, G. F. (2008) Structure of a beta1-adrenergic G-protein-coupled receptor, *Nature 454*, 486-491.

105. Warne, T., Moukhametzianov, R., Baker, J. G., Nehme, R., Edwards, P. C., Leslie, A. G., Schertler, G. F., and Tate, C. G. (2011) The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor, *Nature 469*, 241-244.

106. Wu, B., Chien, E. Y., Mol, C. D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F. C., Hamel, D. J., Kuhn, P., Handel, T. M., Cherezov, V., and Stevens, R. C. (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists, *Science 330*, 1066-1071.

107. Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Kuhn, P., Weis, W. I., Kobilka, B. K., and Stevens, R. C. (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor, *Science 318*, 1258-1265.

108. Bokoch, M. P., Zou, Y., Rasmussen, S. G., Liu, C. W., Nygaard, R., Rosenbaum, D. M., Fung, J. J., Choi, H. J., Thian, F. S., Kobilka, T. S., Puglisi, J. D., Weis, W. I., Pardo, L., Prosser, R. S., Mueller, L., and Kobilka, B. K. (2010) Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor, *Nature 463*, 108-112.

109. Rasmussen, S. G., DeVree, B. T., Zou, Y., Kruse, A. C., Chung, K. Y., Kobilka, T. S., Thian, F. S., Chae, P. S., Pardon, E., Calinski, D., Mathiesen, J. M., Shah, S. T., Lyons, J. A., Caffrey, M., Gellman, S. H., Steyaert, J., Skiniotis, G., Weis, W. I., Sunahara, R. K., and Kobilka, B. K. (2011) Crystal structure of the beta2 adrenergic receptor-Gs protein complex, *Nature 477*, 549-555.

110. Xu, F., Wu, H., Katritch, V., Han, G. W., Jacobson, K. A., Gao, Z. G., Cherezov, V., and Stevens, R. C. (2011) Structure of an agonist-bound human A2A adenosine receptor, *Science 332*, 322-327.

111. Lebon, G., Warne, T., Edwards, P. C., Bennett, K., Langmead, C. J., Leslie, A. G., and Tate, C. G. (2011) Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation, *Nature 474*, 521-525.

112. Gautier, A., Mott, H. R., Bostock, M. J., Kirkpatrick, J. P., and Nietlispach, D. (2010) Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy, *Nat Struct Mol Biol 17*, 768-774.

113. Popot, J. L., and Engelman, D. M. (2000) Helical membrane protein folding, stability, and evolution, *Annual Review of Biochemistry 69*, 881-922.

114. Fiedler, S., Broecker, J., and Keller, S. (2010) Protein folding in membranes, *Cellular and Molecular Life Sciences 67*, 1779-1798.

115. White, S. H., and Wimley, W. C. (1999) Membrane protein folding and stability: physical principles, *Annual review of biophysics and biomolecular structure 28*, 319-365.

116. Thevenin, D., and Lazarova, T. (2008) Stable interactions between the transmembrane domains of the adenosine A2A receptor, *Protein Science : a publication of the Protein Society 17*, 1188-1199.

117. Neumoin, A., Arshava, B., Becker, J., Zerbe, O., and Naider, F. (2007) NMR studies in dodecylphosphocholine of a fragment containing the seventh transmembrane helix of a G-protein-coupled receptor from Saccharomyces cerevisiae, *Biophys J 93*, 467-482.

118. Estephan, R., Englander, J., Arshava, B., Samples, K. L., Becker, J. M., and Naider, F. (2005) Biosynthesis and NMR analysis of a 73-residue domain of a Saccharomyces cerevisiae G protein-coupled receptor, *Biochemistry 44*, 11795-11810.

119. Arshava, B., Taran, I., Xie, H., Becker, J. M., and Naider, F. (2002) High resolution NMR analysis of the seven transmembrane domains of a heptahelical receptor in organic-aqueous medium, *Biopolymers 64*, 161-176.

120. Xie, H., Ding, F. X., Schreiber, D., Eng, G., Liu, S. F., Arshava, B., Arevalo, E., Becker, J. M., and Naider, F. (2000) Synthesis and biophysical analysis of transmembrane domains of a Saccharomyces cerevisiae G protein-coupled receptor, *Biochemistry 39*, 15462-15474.

121. Wuthrich, K. (2001) The way to NMR structures of proteins, *Nat Struct Biol 8*, 923-925.

122. Williamson, M. P., Havel, T. F., and Wuthrich, K. (1985) Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry, *Journal of Molecular Biology 182*, 295-315.

123. Schmidt-Rohr, K., and Spiess, W. (1994) Multidimensional Solid-State NMR and Polymers.

124. Frericks, H. L., Zhou, D. H., Yap, L. L., Gennis, R. B., and Rienstra, C. M. (2006) Magic-angle spinning solid-state NMR of a 144 kDa membrane protein complex: E. coli cytochrome bo3 oxidase, *Journal of biomolecular NMR 36*, 55-71.

125. Nieuwkoop, A. J., Wylie, B. J., Franks, W. T., Shah, G. J., and Rienstra, C. M. (2009) Atomic resolution protein structure determination by three-dimensional transferred echo double resonance solid-state nuclear magnetic resonance spectroscopy, *J Chem Phys 131*, 095101.

126. Wylie, B. J., Schwieters, C. D., Oldfield, E., and Rienstra, C. M. (2009) Protein Structure Refinement Using 13Cα Chemical Shift Tensors, *Journal of the American Chemical Society 131*, 985-992.

127. Wylie, B. J., Sperling, L. J., Nieuwkoop, A. J., Franks, W. T., Oldfield, E., and Rienstra, C. M. (2011) Ultrahigh resolution protein structures using NMR chemical shift tensors, *Proceedings of the National Academy of Sciences of the United States of America*.

128. Manolikas, T., Herrmann, T., and Meier, B. H. (2008) Protein structure determination from 13C spin-diffusion solid-state NMR spectroscopy, *Journal of the American Chemical Society 130*, 3959-3966.

129. Kijac, A., Shih, A. Y., Nieuwkoop, A. J., Schulten, K., Sligar, S. G., and Rienstra, C. M. (2010) Lipid-protein correlations in nanoscale phospholipid bilayers determined by solid-state nuclear magnetic resonance, *Biochemistry 49*, 9190-9198.

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

131. Harris, R. K., Becker, E. D., Cabral de Menezes, S. M., Goodfellow, R., and Granger, P. (2002) NMR Nomenclature: Nuclear Spin Properties and Conventions for Chemical Shifts: IUPAC Recommendations 2001, *Solid State Nuclear Magnetic Resonance 22*, 458-483.

132. Neue, G., Dybowski, C., Smith, M. L., Hepp, M. A., and Perry, D. L. (1996) Determination of 207Pb2+ chemical shift tensors from precise powder lineshape analysis, *Solid State Nuclear Magnetic Resonance 6*, 241-250.

133. Bielecki, A., and Burum, D. P. (1995) Temperature Dependence of 207Pb MAS Spectra of Solid Lead Nitrate. An Accurate, Sensitive Thermometer for Variable-Temperature MAS, *Journal of Magnetic Resonance, Series A 116*, 215-220.

134. Guan, X., and Stark, R. E. (2010) A general protocol for temperature calibration of MAS NMR probes at arbitrary spinning speeds, *Solid State Nucl Magn Reson 38*, 74-76.

135. Bennett, A. E., Rienstra, C. M., Auger, M., Lakshmi, K. V., and Griffin, R. G. (1995) *Heteronuclear decoupling in rotating solids*, Vol. 103, AIP.

136. Hartmann, S. R., and Hahn, E. L. (1962) Nuclear Double Resonance in the Rotating Frame, *Physical Review 128*, 2042.

137. Griffin, R. G., Rienstra, C. M., Hohwy, M., and Hong, M. (2000) 2D and 3D N-15-C-13-C-13 NMR chemical shift correlation spectroscopy of solids: Assignment of MAS spectra of peptides, *Journal of the American Chemical Society 122*, 10979-10990.

138. Baldus, M., Petkova, A. T., Herzfeld, J., and Griffin, R. G. (1998) Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems, *Mol Phys 95*, 1197-1207.

139. Baldus, M., Geurts, D. G., Hediger, S., and Meier, B. H. (1996) Efficient15N–13C Polarization Transfer by Adiabatic-Passage Hartmann–Hahn Cross Polarization, *Journal of Magnetic Resonance, Series A 118*, 140-144.

140. Takegoshi, K., Nakamura, S., and Terao, T. (2001) 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR, *Chemical Physics Letters 344*, 631-637.

141. Oas, T. G., Griffin, R. G., and Levitt, M. H. (1988) *Rotary resonance recoupling of dipolar interactions in solid-state nuclear magnetic resonance spectroscopy*, Vol. 89, The Journal of Chemical Physics.

142. Yang, J., Paramasivam, S., Marulanda, D., Cataldi, M., Tasayco, M. L., and Polenova, T. (2007) Magic angle spinning NMR spectroscopy of thioredoxin reassemblies, *Magn Reson Chem 45 Suppl 1*, S73-83.

143. Crocker, E., Patel, A. B., Eilers, M., Jayaraman, S., Getmanova, E., Reeves, P. J., Ziliox, M., Khorana, H. G., Sheves, M., and Smith, S. O. (2004) Dipolar assisted rotational resonance NMR of tryptophan and tyrosine in rhodopsin, *Journal of biomolecular NMR 29*, 11-20.

144. Neue, G., and Dybowski, C. (1997) Determining temperature in a magic-angle spinning probe using the temperature dependence of the isotropic chemical shift of lead nitrate, *Solid State Nucl Magn Reson 7*, 333-336.

145. Takahashi, T., Kawashima, H., Sugisawa, H., and Baba, T. (1999) 207Pb chemical shift thermometer at high temperature for magic angle spinning experiments, *Solid State Nucl Magn Reson 15*, 119-123.

146. Langer, B., Schnell, I. I., Spiess, H. W., and Grimmer, A. R. (1999) Temperature calibration under ultrafast MAS conditions, *J Magn Reson 138*, 182-186.