

**DYNAMICS OF HUMAN INTESTINAL FATTY ACID
BINDING PROTEIN**

YU BINHAN

(B.Sc., PKU)

**A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

DEPARTMENT OF BIOLOGICAL SCIENCES

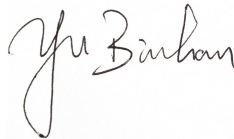
NATIONAL UNIVERSITY OF SINGAPORE

2014

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

A handwritten signature in black ink, reading "Yu Binhan". The signature is written in a cursive style with a large initial 'Y' and 'B'. It is positioned above a horizontal line that is broken at both ends by short dashes.

Yu Binhan

14 Aug 2014

ACKNOWLEDGEMENT

I would like to express my greatest gratitude to my supervisor, Professor Dr. Yang Daiwen for giving me the opportunity to pursue my Ph.D. in his laboratory and for his support, guidance and inspirations throughout the course of this work.

I am also grateful to Dr. Fan Jingsong for the help and valuable discussion in NMR experiments. Special thanks to Associate Professor Dr. Mok Yu-Keung, Henry for the expertise and critical comments in our weekly group meetings. Thanks to Dr. Long Dong for the helpful discussion in the delipidation experiments and to Iman Fabim Hameed for the help and assistance in protein expression and purification in the early stage of this work.

I would like to thank all the colleagues and friends in the structural biology labs for the help and joyful experiences during my study: Dr. Lin Zhi, Dr. Zhang Jingfeng, Dr. Zheng Yu, Liang Chen, Dr. Wang Shujing, Dr. Balakrishna Chandrababu Karthik, Dr. Lim Jack-Wee, Dr. Gao Zhenwei, Liu Xiao, Xiao Tianshu, Lei Cheng, Rusha Chakraborty, Zhou Yang. Special thanks to Lai Chong Cheong for the help and valuable input in molecular cloning experiments.

I would like to thank my parents and family members for the persistent love and support.

Last but not least, I would like to thank National University of Singapore for the financial support during the course of my study.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	I
SUMMARY	VI
LIST OF TABLES	VII
LIST OF FIGURES	VIII
LIST OF ABBREVIATIONS	XII
CHAPTER 1 BACKGROUND AND INTRODUCTION	1
1.1 Overview of fatty acids	4
1.2 Overview of fatty acid binding protein family	5
1.3 Overview of intestinal fatty acid binding protein	11
1.4 Nuclear spin relaxation theory	14
1.4.1 Bloch equations	15
1.4.2 Solomon equations	16
1.4.3 Master equation	17
1.4.4 Spectral density function	19
1.4.5 Chemical exchange effects in NMR	22
CHAPTER 2 SAMPLE PREPARATION AND NMR EXPERIMENTS	26
2.1 Materials and methods	27
2.1.1 Media	27
2.1.2 SDS-polyacrylamide gel (SDS-PAGE) electrophoresis	27
2.1.3 Expression and purification of FABPs	28
2.1.4 Delipidation of FABPs	31

2.2 NMR experiments	35
2.2.1 2D ^{15}N HSQC experiment	36
2.2.2 T_1 experiment	37
2.2.3 $T_{1\rho}$ experiment	37
2.2.4 NOE experiment	38
2.2.5 3D TOCSY-HSQC	38
2.2.6 3D NOESY-HSQC	39
2.2.7 3D HNC0 experiment	39
2.2.8 3D CC(CO)NH experiment	40
2.2.9 3D HCC(CO)NH experiment	40
2.2.10 Ligand titration experiment	41
2.2.11 Transverse relaxation dispersion experiment	41
2.2.12 Chemical exchange saturation transfer experiment	42
2.2.13 H-D exchange experiment	44
2.3 Result and conclusion	44
CHAPTER 3 DATA ANALYSIS AND MODEL SELECTION FOR APO-HIFABP	47
3.1 R_1, $R_{1\rho}$ and NOE experiments	48
3.2 Ligand titration with 1,8-ANS	51
3.3 NH-D₂O exchange experiment	53
3.4 CPMG data analysis	56
3.4.1 Two-state fitting of CPMG data	58
3.4.2 Individual fitting of CPMG data	59
3.4.3 Global fitting of CPMG data	62
3.4.3.1 Grid search aided minimization for parameter space mapping	62
3.4.3.2 Global fitting result for two-state exchange model	67

3.5 CEST data analysis	68
3.6 Combined analysis of CEST and CPMG data	76
3.7 Model selection for residues with exceptional RD and CEST profiles	84
3.7.1 Regional collective motions in local area, faster exchange rate based on three-state exchange model	84
3.7.2 A third exchange process in the presence of the two global motions	87
3.7.3 Four-site exchange for apo-hIFABP	89
3.7.4 Evaluation of the R_{ex} contributed by potential fast exchange process	94
3.8 Final analysis of relaxation data for apo-hIFABP	101
3.8.1 Three-state exchange models	101
3.8.2 Four-state exchange models	103
CHAPTER 4 RESULT AND DISCUSSION	105
4.1 Method validation: robust extraction of kinetic parameters compared to data analysis on single NMR experiment.	106
4.1.1 CPMG data fitting	106
4.1.1.1 Three-site exchange processes approximated by two-site model	106
4.1.1.2 Three-site exchange processes examined by three-site models	109
4.1.2 CEST data fitting	111
4.1.3 Combined CPMG and CEST data analyses	112
4.2 NMR titration experiments	117
4.3 H-D exchange experiments	121
4.4 Four-state equilibrium: coexistence of multiple intermediates of apo-hIFABP in solution	123
4.4.1 Regional motions and a statistical view of proteins in dynamical equilibrium	124
4.4.2 Physical meaning of the four-state equilibrium	129

4.4.2.1 Three-state model I and four-state model I	129
4.4.2.2 Three-state model I and four-state model IV	131
BIBLIOGRAPHY	134

SUMMARY

Biological macromolecules such as proteins exist in dynamic equilibrium in which a lower energy ground state interconverts with one or more “excited states”. Currently, studies on the multistate systems are often limited to three-state exchange models at most due to the high level of complexity increased with the number of the exchanging states. Here we present a millisecond timescale dynamics study of apo-form human intestinal fatty acid binding protein (hIFABP) using ^{15}N NMR CPMG and CEST experiments. A four-state equilibrium, which we believe is the first direct experimental evidence for a fourth state, was observed from relaxation profiles of apo-form human IFABP. In the equilibrium, three intermediates directly interconvert with the major form with exchange rates ranging from 70 s^{-1} to 3400 s^{-1} . Moreover, additional fast exchange processes with k_{ex} in the range of $10^5\sim 10^6\text{ s}^{-1}$ was identified for the protein. The extracted chemical shifts of the minor states indicate that major structural variations of the sparsely populated conformers with respect to the majority of the population occur mainly in the gap region. Based on our observations, we propose that the excited states in equilibrium are irrelevant to the ligand binding process. The FABP-ligand binding should be via a ligand-induced binding mechanism in which the ligand interacts with the native state directly. In the absence of any unfolded states, the three intermediates should represent conformations in the last step of the folding pathway.

LIST OF TABLES

Table 1. Tissue distribution of FABPs.	7
Table 3.4.1 Calculation of v_{cp} from T_{cp}, m and n.	58
Table 3.7.1 F-test for residues in the gap region between three-state model I and four-state model I.	93

LIST OF FIGURES

Figure 1.1. Alignment of amino acid sequences of the members of the FABP family.	6
Figure 1.2 Three-dimensional structures of FABPs.	8
Figure 2.1.1 pET-32a(+) derived vector (pETm).	29
Figure 2.1.2 SDS-PAGE gel of hIFABP.	30
Figure 2.1.3 FPLC profile for hIFABP.	31
Figure 2.1.4 HSQC of holo-form hIFABP, pH8.0 at 302K.	32
Figure 2.1.5 Mono S profiles of hIFABP.	33
Figure 2.1.6 HSQC spectrum of apo-hIFABP purified by mono S column.	33
Figure 2.1.7 HSQC spectrum of apo-hIFABP purified by mono S column after two weeks.	34
Figure 2.1.8 HSQC spectra of FABPs after delipidation.	35
Figure 2.2.1 HSQC spectra of hIFABP at pH 4.6 (green) and pH 7.2 (red).	37
Figure 2.2.2 Pulse scheme for the measurement of relaxation dispersion.	43
Figure 2.3.1 Backbone assignment of hIFABP.	46
Figure 3.1.1 R_1 and $R_{1\rho}$ of apo-hIFABP.	49
Figure 3.1.2 R_2 and NOE of apo-hIFABP.	50

Figure 3.2.1 NMR titration experiments of ^{15}N enriched hiFABP with 1,8-ANS as ligand.	52
Figure 3.2.2 Representative residues in NMR titration experiments.	53
Figure 3.3.1 Representative residues in H-D exchange experiments.	55
Figure 3.3.2 Relative peak intensities of the H-D exchange experiment.	56
Figure 3.4.1 Reference spectrum of CPMG experiments at 500MHz magnetic field.	59
Figure 3.4.2.1 Exchange rates extracted by two-site exchange model.	61
Figure 3.4.2.2 Representative residues in two-state CPMG fitting.	62
Figure 3.4.3.1 Parameter space for a single residue in two-state fitting.	65
Figure 3.4.3.2 Parameter space output projection on p, k_{ex} and d_w dimension.	66
Figure 3.4.3.3 Grid search minimization in global fitting for apo-hIFABP.	66
Figure 3.5.1 Representative residues that show obvious two minor dips.	69
Figure 3.5.2 Residues with obvious chemical exchange effect plotted on ribbon structure of apo-hIFABP.	69
Figure 3.5.3 Representative residues for the three-state model tests.	72
Figure 3.5.4 Kinetic parameters extracted from 3-state individual fitting with CEST data.	73
Figure 3.5.5 Sensitivity test for CEST data to fast and slow exchange processes.	

	74
Figure 3.5.6 Global fitting of CEST data with three-state model I.	76
Figure 3.6.1 Representative residues fitted individually with three-state model I and II by combined analysis of CEST and CPMG data simultaneously.	78
Figure 3.6.2 k_{ex1} distribution of individual fitting by 3-state model I.	79
Figure 3.6.3 Representative residues that could be fitted well only with $k_{ex1} > 2000 \text{ s}^{-1}$.	82
Figure 3.6.4 Gap region of hiFABP.	82
Figure 3.7.1 Three-state model 1 selection for the gap region.	86
Figure 3.7.2 Representative residues fitted with four-state model when the global parameters for the other two exchange processes are fixed.	89
Figure 3.7.3 Fitting residuals for the six four-state models.	90
Figure 3.7.4 Representative residues in the four-state model test.	92
Figure 3.7.5 Residue-specific Model Free Analysis of apo-hIFABP.	98
Figure 3.7.6 Comparison of ΔR_2^{eff} and ΔR_2^{cal} .	99
Figure 3.7.7 Comparison of ΔR_2^{eff} and ΔR_2^{cal} in the presence of ΔR_2^{DD} .	100
Figure 3.8.1 Representative residues in the global fitting with three-state model I.	102
Figure 3.8.2 Representative residues in the global fitting with four-state model I.	103

Figure 4.1.1 Representative residues in the two-state global fitting of the theoretical data.	107
Figure 4.1.2 Residues in the global fitting of the theoretical data by three-state model I.	108
Figure 4.1.3 Representative residues in the global fitting by three-state model II.	110
Figure 4.1.4 Global fitting of synthetic data with combined analysis of CEST and CPMG data.	116
Figure 4.2.1 Chemical shift perturbation of apo-hIFABP upon binding with 1,8-ANS.	118
Figure 4.2.2 Residues with CCSP larger than 0.1 plotted on the ribbon structure of apo-hIFABP.	119
Figure 4.2.3 Combined chemical shift perturbation of 56 residues with CCSP larger than 0.1.	121
Figure 4.3.1 Fast exchanging residues in H-D exchange experiment plotted on ribbon structure of apo-hiFABP.	122
Figure 4.4.1 Chemical shifts comparison between minor states and IDP.	123
Figure 4.4.2 Four-state exchange pathway I for apo-hiFABP.	130
Figure 4.4.3 $\delta\omega$ plotted on the ribbon structure of apo-hiFABP.	131
Figure 4.4.4 Four-state exchange pathway II for apo-hiFABP.	131

LIST OF ABBREVIATIONS

ATP	Adenosine Triphosphate
ANS	1-anilinonaphthalene-8-sulfonic acid
Bis-tris	Bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane
CSP	Chemical shift perturbation
CCSP	Combined chemical shift perturbation
CPMG	Carr-Purcell-Meiboom-Gill
CSA	Chemical shielding anisotropy
DD	Dipole-dipole
DIPSI	Decoupling In the Presence of Scalar Interactions
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetic acid
FA	Fatty Acid
FFA	Free fatty acid
FPLC	Fast performance liquid chromatography
HSQC	Heteronuclear single quantum coherence
IDP	Intrinsically disordered protein
INEPT	Insensitive nuclei enhanced by polarization transfer

IPTG	Isopropyl-beta-D-1-thiogalactopyranoside
kD	kilo Dalton
LB	Luria Bertani
LCFA	Long chain fatty acid
(L/I/A/M/B)FABP binding protein	(Liver/Intestinal/Adipocyte/Muscle/Brain) fatty acid binding protein
MD	Molecular Dynamics
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser enhancement
NOESY	Nuclear Overhauser effect spectroscopy
PAGE	Polyacrylamide gel electrophoresis
ppm	parts per million
PYY	Peptide tyrosine tyrosine
ps/ns/ μ s/ms	picosecond/nanosecond/microsecond/millisecond
RD	Relaxation dispersion
RF	Radio frequency
TOCSY	Total correlation spectroscopy
TPPI	Time proportional phase incrementation

CHAPTER 1

BACKGROUND AND INTRODUCTION

Chapter 1 Background and introduction

Protein functions are largely dependent on their structures. Despite the rich information from high resolution 3D structures obtained by X-ray crystallography and/or Nuclear Magnetic Resonance (NMR), sometimes it is still difficult to appreciate the structure-function relationship of these macromolecules without detailed characterization of their dynamic properties. Benefited from the emerging techniques developed in the last two decades, the importance of the plasticity of proteins to their functions has become increasingly clear. Among many methods used in this area, NMR has the unique capability to provide site-specific information under physiological conditions. A suit of experiments have been developed, covering almost all the functional relevant dynamic motions on time scales ranging from 10^{-12} to 10^5 s. Atomic fluctuations on the picosecond to nanosecond time range, for instance, are estimated by order parameters generated from the Model Free Analysis based on R_1 , R_2 and NOE data. For slower motions on microsecond to millisecond time scale, including protein folding/unfolding, ligand binding, allosteric regulation and enzyme catalysis, the Carr-Purcell-Meiboom-Gill (CPMG) R_2 and spin-lock $R_{1\rho}$ relaxation dispersion experiments are used to extract exchange parameters such as k_{ex} (exchange rate constant between different states) and P_i (relative population of the i state). Besides, ZZ exchange and H-D exchange methods have also been used for dynamic studies on millisecond to second time scale and longer. More recently, the Chemical Exchange Saturation Transfer (CEST) experiment has been proposed for exchanging systems in the range of $20 \sim 200 \text{ s}^{-1}$, further improving both the range and accuracy of dynamic

studies.

Meanwhile, advances in computing have also benefited the studies of protein dynamics, bringing Molecular Dynamic (MD) simulation to the range of microsecond. Ideally, the MD studies can, given proper force field and model, provide detailed information on the trajectories of each atom in a protein molecule, thus generating a complete map of motions over time. Although we are still far from the ultimate goal to “watch” a protein move, a consensus has been reached that protein movements, either regional or global, are collective motions. This simply means that protein conformational fluctuations occur in a concerted manner. Residues that are close to each other with spatial restrictions such as covalent bonds, hydrogen bonds and van der waals interactions, are likely to move in a cooperative way. These collective motions are, however, not necessarily applicable to all the residues across the entire molecule, if there is substantial space between two subunits or regions. In principle, MD simulation is an ideal way to predict and validate the regional collective motions of proteins. However, it is currently not able to provide such a support to this systemic behavior due to the limit of computing power, especially for an average sized protein on the functional related microsecond to second time scale. Although NMR studies generate site-specific nature on the protein system, such information is often considered as a global behavior without further reasoning due to the limited number of data sets (residues) suitable for analysis. Consequently, experimental evidence supporting regional collective motions, especially for proteins in the absence of the binding partners, is still lacking in literature.

1.1 Overview of fatty acids

Fatty acids, usually derived from triglycerides or phospholipids *in vivo*, are carboxylic acids with long aliphatic chains which are either saturated or unsaturated. The fatty acid chains differ by length, often categorized as short to very long. The solubility of fatty acids decreases rapidly as their chain lengths increase. When not attached with other molecules, they are known as “free fatty acids”. Since they are insoluble in aqueous environment, fatty acids are circulated while bound to plasma protein albumin. As a consequence, the level of free fatty acids in blood is limited by the availability of the protein binding site. It has been shown that fatty acids are widely distributed in various forms in different stages in tissues and blood circulation. These organic compounds, either by ingestion or by drawing on triglycerides stored in fat tissues, are important sources of energy for muscular contraction and general metabolism (der Vusse and Reneman 1995) since they yield a large amount of ATP when consumed by mitochondria by beta oxidation. They are also precursors for the biosynthesis of complex membrane lipids as the acyl chain composition of these molecules largely determines the overall membrane structure and function (Pohl et al. 2004)

Humans lack the ability to introduce double bonds in fatty acids beyond carbons 9 and 10, as counted from the carboxylic acid side. These required fatty acids, often cannot be made in sufficient quantity from other substrates in human body, and thus must be obtained from food, are termed as essential fatty acids. Fatty acids absorbed by enterocytes in the jejunum and ileum, are then transported

into the lymphatic system and blood circulation as chylomicrons. Capillary-bound lipoprotein lipase, produced by liver, heart, adipose and other tissues, catalyzes the release of fatty acids from lipoproteins, of which the vast majority is immediately bound to albumin. Apart from diffusion through the phospholipid bilayer suggested by earlier researches, additional ways of distribution are supported by the fact that physiological uptake of fatty acids across cellular membrane is a saturable and inhibitable process (Abumrad, Park, and Park 1984), indicating a protein associated transportation (Pohl et al. 2004).

1.2 Overview of fatty acid binding protein family

Fatty acid binding proteins (FABPs) were discovered as cytosolic proteins with high non-covalent binding affinity for long chain fatty acids *in vitro*. It has been shown that FABPs could facilitate the transportation of lipid molecules and serve as an intracellular fatty acids reservoir (Hamilton, 2004). FABPs, together with acyl-CoA binding protein, non-specific lipid transfer protein, phosphatidylcholine transfer protein, phosphatidylinositol transfer protein, cellular retinal binding protein and cellular retinoic acid binding protein, form a superfamily of cytosolic non-enzymic proteins of low molecular weight (around 15 KD) and high binding affinity for amphiphilic molecules. At present, nine types of FABP have been isolated from mammals and subsequently named after the tissue where they were first discovered (table 1). FABPs, consisting of 126-134 amino acids, have highly conserved primary structures between species but to a lesser extent among the different types with sequence similarities of 22-73% (fig. 1).

H-FABP	MVDA..FLGT	WKLVD SKNFD	DYMKSLGVGF	ATRQVASMT.	.KPTT IIIEKN	45
B-FABP	MVEA..FCAT	WKL TNSQNF	EYMKALGVGF	ATRQVGNVT.	.KPT VIISQE	45
M-FABP	.SNK..FLGT	WKL VSSNF	DYMKALGVGL	ATRKLGNLA.	.KPT VIISK	45
A-FABP	MCDA..FVGT	WKL VSSNF	DYMKEVGVGF	ATRKVAGMA.	.KPN MIIISVN	45
E-FABP	MATVQOLEGR	WRL VDSKGF	EYMKELGVGI	ALRKMGAMA.	.KPD CIITCD	47
I-FABP	MA...FDST	WKV DRSENYD	KFMEKMGVNI	VKRKLAHDN	LK..LT ITQE	43
L-FABP	MS...FSGK	YQL QSQENFE	AFMKAIGLPE	EL..IQKGD	IKG VSEIVQN	43
I-LBP	MA...FTGK	FEME SEKNYD	EFMKLLGISS	DV..IEKARN	FKI VTEVQQD	43
H-FABP	GDILTLKTHS	TFK NTEIS.F	KLGVFEDET	ADDRKVKSI	TLD.GG KLHV	93
B-FABP	GDKVVIRTLS	TFK NTEIS.F	QLGEEFDETT	ADDRNCKSV	SLD.GD KLHV	93
M-FABP	GDIIITIRTES	TFK NTEIS.F	KLGOEFEEET	ADNRKTKSI	TLQ.RF SLNQ	93
A-FABP	GDVITIKSES	TFK NTEIS.F	ILGOEFDEVT	ADDRKVKSTI	TLD.GG VLHV	93
E-FABP	GKNLTIKTES	TLK TTQFS.C	TLGKFEET	ADGRKTQTV	NFT.DG ALVQ	95
I-FABP	GKNFTVKESS	AFR NIENV.F	ELGVTFNYNL	ADGTELRTW	SLE.GN KLIG	91
L-FABP	GKHFKFTITA	GSK VIQNE.F	TVGEECELET	MTGEKVKTV	QLEG DNKLV	92
I-LBP	GQDFTWSQHY	SGG HMTNKF	TVGKESNIQT	MGGKTFKAT	QME.GG KLHV	92
H-FABP	LQKWD...GQ	ETTL VRELID	G.KLILTLTH	GTAVCTR	TYE KEA	132
B-FABP	IQKWD...GK	ETNF VREIKD	G.KMVMTLTF	GDVVAVRH	YE KA.	131
M-FABP	VQRWN...GK	ETTI KRKLVD	G.KMVAECKM	KGVVCTRI	YE KV.	131
A-FABP	VQRWN...GK	STTI KRKRED	D.KLVVECV	KGVTSTR	YE RA.	131
E-FABP	HQEWD...GK	ESTI TRKLKD	G.KLVVECV	NNVTCTRI	YE KVE	134
I-FABP	KFKRTD.NGN	ELNT VREIIG	D.ELVQTYVY	EGVEAKRI	FK KD.	131
L-FABP	TF....KN.	IKSV TELNG	D.IITNTMTL	GDIVFKRI	SK RI.	126
I-LBP	NF....PN.	YHQ TSEIVG	D.KLVEVSTI	GGVTYER	VSK RLA	127

Figure 1.1. Alignment of amino acid sequences of the members of the FABP family. Abbreviations for FABP type are in table 1. All sequences are for human proteins. Positions of well-conserved amino acids (identical residues present in at least five molecules) are in bold.

Many FABPs are predominantly expressed in a single tissue or cell type, while some of them display broad tissue distributions such as heart-type and epidermal-type FABP. In the meantime, more than one kinds of FABP can be found in a same tissue (e.g., ileal-, intestine- and liver-type FABP in intestine). The lipid-active tissues where the corresponding FABPs are found, such as intestine and liver, often have substantial level of fatty acid flux, suggesting that these FABPs are necessary for intracellular binding and transfer of fatty acids. Consequently, several functions have been proposed for the FABPs: (1) modulation of specific enzymes of lipid metabolic pathways; (2) maintenance of cellular membrane fatty acid levels, serving as a lipid molecule reservoir; (3) regulation of the expression of fatty acid-responsive genes (J. F. Glatz et al.

1995; J. Glatz and Vusse 1996).

In spite of significant variation in primary structures, crystal structures and NMR solved solution structures suggest that they display a highly conserved secondary and tertiary structure for all FABPs. They fold as a slightly elliptical β barrel comprising 10 antiparallel β strands, with two short α helices located between the first and second β strands (Figure 1.2).

Table 1. Tissue distribution of FABPs.

FABP type	Abbreviation	Tissue
Adipocyte	A	Adipose tissue
Brain	B	Brain
Epidermal	E	Skin, brain, lens, capillary, endothelium, retina
Heart	H	Heart, kidney, skeletal muscle, aorta, adrenals, Placenta, brain, testes, ovary, lung, mammary gland, stomach
Ileal	IL	Intestine, ovary, adrenals, stomach
Intestine	I	Intestine, stomach
Liver	L	Liver, intestine, kidney, stomach
Myelin	M	Peripheral nervous system
Testicular	T	Testis

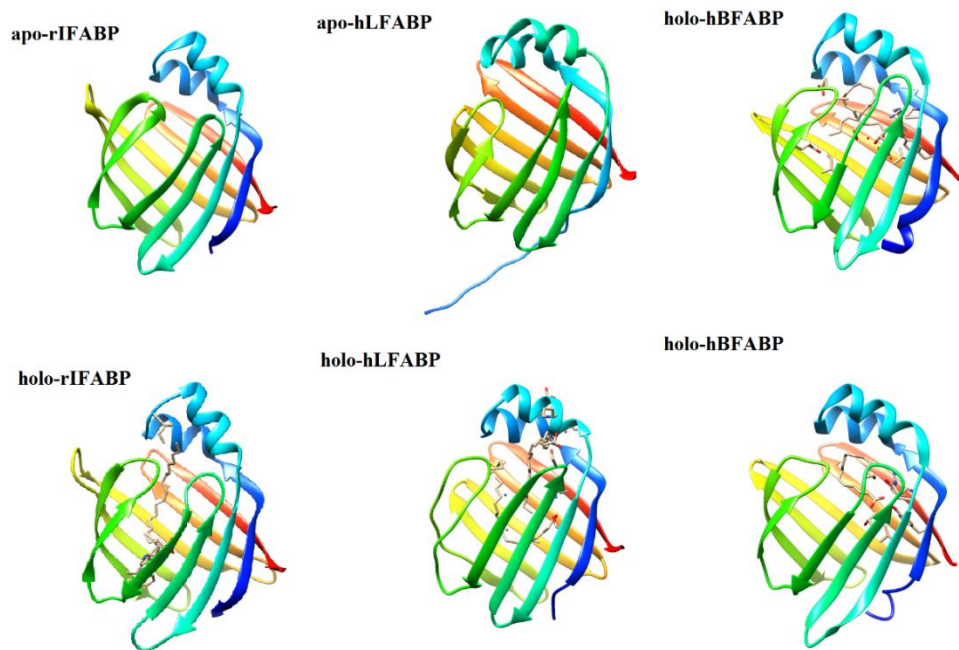


Figure 1.2 Three-dimensional structures of FABPs. Ribbon structures of rat IFABP (apo/holo), human LFABP (apo/holo) and human BFABP (holo) are shown. The PDB entries for them are 1IFB/1ICN, 2F73/1LFQ and 1FF3/1FDQ, respectively.

Structural studies of FABPs show that the β barrel possesses substantial structural stability, as it is basically unaffected by chemical modifications, such as the introduction of bulky fluorescent groups, or site mutagenesis. The β strands are organized into two nearly orthogonal β sheets that form an internalized ligand-binding cavity. The solvent-accessible cavity is capped at the end of the barrel by the helix-turn-helix motif. This helix cap is thus thought to act as a portal region that regulates the ligand entry/exit process. The α -II helix forms long-range interactions with the α -II turn between β strands C and D and is considered a key structural element of the putative portal region. The binding cavity is two to three times larger than the volume of the FA, and the structures reveal ordered water molecules in the cavity that are hydrogen bonded to internal polar residues. Most of the FABPs bind only a single FA, with the carboxylate group oriented toward to the open end of the β barrel.

However, liver-type FABP has the unique property of binding two FAs or other large lipid molecules. Structures derived from X-ray or NMR studies show that the binding cavity possess a volume around three times larger than that of fatty acid molecules, which explains the broad range of hydrophobic molecules they bind, rather than specific-ligand selection.

Static structures from crystallographic analysis show little difference between apo- and holo-FABPs, which is, however, another case from NMR derived solution structures. Sizeable structural variations were observed in the distal half of the α -II helix and the turn between β strands C and D. These two regions are more disordered in the apo- state than those in the holo- state. These changes suggest that the ligand helps to stabilize that portal region by a series of long-range interactions. Additionally, since no obvious openings were observed in both bounded and unbounded FABPs, substantial conformational exchange is required to permit the entry/exit of the ligand to/from the portal region.

Although all FABPs bind fatty acids and other hydrophobic molecules, the binding affinity, ranging from nano- to micromolar, vary significantly among different FABP types. According to functional studies of FABPs, these proteins have been divided into two groups based on the mechanisms of ligand transfer: (1) direct contact and/or collision type; and (2) aqueous-phase diffusion type. The first group of FABPs, including intestinal-type, adipose-type, myelin-type, epidermal-type and heart-type etc., collect and transport their ligands through contact/collision with a cellular membrane; while the other group of FABPs, only liver-type FABP in this case, deliver its ligand to and from membranes

through aqueous-phase diffusion. Consequently, FABPs of the first group are expected to experience intermediate states in which they are in direct physical contact with the target cellular membrane. In fact, studies of several such FABPs have demonstrated direct protein-membrane interactions under physiologically relevant conditions. Hypotheses that suggest targeted interactions of the protein with specific membrane-lipid domains and/or membrane-protein domains in the collisional delivery have also been proposed.

Several studies have analyzed the effects and contributions of electrostatic and hydrophobic interactions to the fatty acid transfer from intestinal fatty acid-binding protein, suggesting the importance of these highly conserved structural elements of the collisional FABP in the ligand transfer process. A study on surface properties of adipocyte-type FABP also shows that the net positive surface electrostatic potential of the helix-turn-helix motif and the amphipathic property of the helix-I are both important for the FABP-membrane interaction. Mutagenesis studies on the structure-function relationship of collisional FABPs also suggest the core position of the α -helical region in determining the ligand delivery mechanism of FABPs. Such a structure-based mechanism demonstrated here, ideally, is anticipated to serve as a generalization for the ligand delivery system of the contact/collisional FABPs.

The precise physiological roles of FABPs and their mechanisms of action remained unclear for several decades after they were identified. As we discuss below, however, the functions of individual members of the FABP family are being elucidated by a combination of gene knockout mouse models, cell culture

studies in which FABP levels are altered, molecule-level analyses of FABP gene regulation (in some cases), and an examination of the physiological correlates to specific human polymorphisms.

It can be seen from the structures obtained with or without ligand that the conformation of the holo form is very similar to that of the apo form, even if the conformation of the bound ligand differs for the distinct FABPs types. Although structural studies have shown that the ligand is well situated in the interior cavity of FABPs, it is still unknown how the ligands get access to the internalized binding site. Several mechanisms were proposed by Cistola and coworkers (Michael E Hodsdon and Cistola 1997) based on careful inspections of the crystal structures and backbone flexibility studies by NMR relaxation experiments as well as molecular dynamic simulations. However, the validity of these assumptions is still to be verified.

1.3 Overview of intestinal fatty acid binding protein

IFABP, coexpressed with LFABP in the small intestine and stomach (Cohn et al. 1992), is found to be most abundant in the proximal intestine with the expression shifted to the distal end of the duodenum (Sacchettini and Hautt 1990; Besnard et al. 2002). Peptide tyrosine tyrosine (PYY), a gut hormone present in endocrine cells in the lower intestine that can be released by the presence of luminal free fatty acids (FFAs), induces the expression of the IFABP gene (Hallden 1997). An *in vivo* study on immunocytochemical localization of FABPs demonstrated the apical cytoplasmic localization in the fasting state and

the redistribution to the entire cytoplasm after fat feeding. This is consistent with the hypothesis that IFABP may contribute to the metabolic compartmentalization of apically presented FAs in the intestine (Alpers et al. 2000).

In spite of the similarities among all the FABPs, tissue-specific expression, different genetic regulation and their distinct FA binding patterns indicate potential possibility of specialized function of each FABP in the dietary lipid uptake and transportation processes. For example, IFABP binds saturated long chain fatty acid (LCFA) with higher affinity than unsaturated FA, while LFABP has greater affinity for unsaturated FA than IFABP (Gary V. Richieri 1994; GV Richieri and Ogata 2000). *In vitro* transfer studies have also demonstrated the fundamentally different mechanisms used by IFABP and LFABP for the FA transfer from and to membranes. Transfer with IFABP occurs via direct contact/collision between the protein and membranes, as opposed to aqueous-phase diffusional transfer for LFABP (Storch 1996; Thumser and Storch 2000). The helices part of the putative portal region has been shown to be determinative for the FA transfer mechanism. The cytosol-facing membranes are considered to be negatively charged surface. Therefore, charge-charge interactions could serve as driving forces for the FA transfer by FABP within the cell. The positive surface electrostatic potential of the helix-turn-helix motif, together with the amphipathic helix-I segment, form an membrane interactive structural element that regulates the transfer process of FA by collisional FABP (Corsico et al. 1998; Córscico, Liou, and Storch 2004; Falomir-Lockhart et al. 2006; LiCata and Bernlohr 1998; Wu et al. 2001).

For more than a decade after the discovery of the first FABP, the putative role of FABPs in the cellular uptake and trafficking of lipid molecules is mostly only indirectly supported. For instance, studies have been generally focused on the binding characteristics of FABPs with different FAs (G V Richieri, Ogata, and Kleinfeld 1995; Wilkinson and Wilton 1987; Ogata 1996; Scapin, Gordon, and Sacchettini 1992; Kurian, Kirk, and Prendergast 1996), examination of the FABPs participation in the FAs transfer process using photo-activated fatty acid probes in cell culture systems (Trigatti, Mangroo, and Gerber 1991; Schmider et al. 1996; Waggoner 1990), effects of FABPs on mitochondrial or microsomal enzyme activities (Peeters, Veerkamp, and Demel 1989; Burnett et al. 1979; Ockner and Manning 1976), correlations between FABP levels and cellular function (Distel, Robinson, and Spiegelman 1992; Hertz and Bernlohr 1998; Haunerland et al. 2013; Veerkamp and van Moerkerk 2013), dietary modulation (Kaikaus et al. 2013; Poirier et al. 1997; Cohn et al. 1992) and many others. More recently, studies of the *Ifabp*^{-/-} mouse provided “direct” evidence for the IFABP function. IFABP gen (*Fabpi*) was disrupted to test the hypothesis that IFABP serves an essential role in the assimilation of dietary fatty acids. The IFABP-knockout mice show alterations in body weight and are hyperinsulinemic. Additionally, the effect of IFABP on the body weight is gender-dependent, male mice are more liable to weight gain in the absence of *Ifabp*. It is possible that the existence of multiple types of FABPs in the same tissue represents a mechanism that ensures functional redundancy. This result clearly indicates that the IFABP is not essential for dietary fat absorption (Vassileva et al. 2000).

Another role of IFABP in lipid assimilation and energy homeostasis has also been proposed by studying the effect of IFABP gene as a possible genetic factor in determining insulin action in the Pima Indian of the southwest America. This native population is found to be associated with abnormal lipid metabolism and an increased incidence of obesity and type II diabetes. A polymorphism at codon 54 of Ifabp was identified that results in Alanine to Threonine substitution. This polymorphism was found to increase the absorption of dietary fat and thus increase the fat oxidation rate which are known to inhibit insulin action (Baier et al. 1995). Moreover, the Ala54Thr IFABP has larger binding affinities for long-chain fatty acids and was found to be associated with insulin resistance and islipidemias in some populations (Pérez-Bravo et al. 2006; Georgopoulos et al. 2007; Pratley et al. 2000). These findings collectively suggests a more specific role of IFABP in cellular trafficking for dietary FAs.

1.4 Nuclear spin relaxation theory

When a nuclei ensemble is to be put in a static magnetic field and undisturbed for a certain period of time, the system will reach its thermal equilibrium state in which the populations are given by the Boltzmann distribution. At this point, if an r.f. field is allowed to present shortly, the thermal equilibrium will first deviate from and then return to its original state where equilibrium is regained. The process by which the equilibrium is regained through interaction of the spin system with the thermal molecular environment, is called spin relaxation. Relaxation process can be divided into two groups. Spin-lattice relaxation, also called longitudinal relaxation, is concerned with the change of the spin

population going back to its Boltzmann distribution. Spin-spin relaxation, on the other hand, is about the spin motions through which decay of coherence occurs. This process is also known as the transverse relaxation. The spin-lattice and spin-spin relaxation evolves independently of each other and is affected by different experimental factors. In the following sections, general aspects of spin relaxation in liquids that are concerned with protein dynamics studies will be briefly reviewed. Fundamental theories on relaxation processes had been covered in previous work and reviews (Ernst, Bodenhausen, and Wokanu 1987; Abragam, A. 1961; Redfield 1965; Wangsness and Bloch 1953; Cavanagh et al. 2007; McConnell, J. 1987). Theoretical formalism describing relaxation is complicated and is usually replaced by the product operator formalism in practice. Detailed descriptions of the derivation of the relaxation equations are presented elsewhere (Goldman 1988; Abragam, A. 1961; McConnell, J. 1987).

1.4.1 Bloch equations

The evolution of an isolated spin can be described by the Bloch equations (Wangsness and Bloch 1953):

$$\begin{aligned}
 \frac{dM_z(t)}{dt} &= \gamma(M(t) \times B(t))_z - R_1(M_z(t) - M_0) \\
 \frac{dM_x(t)}{dt} &= \gamma(M(t) \times B(t))_x - R_2 M_x(t) \\
 \frac{dM_y(t)}{dt} &= \gamma(M(t) \times B(t))_y - R_2 M_y(t)
 \end{aligned}
 \tag{1.1}$$

Where $M(t)$ and $B(t)$ are the magnetization vector and the field strength at time

t, respectively; R_1 and R_2 are the longitudinal and transverse relaxation rates, respectively; γ represents for the gyromagnetic ratio of the spin. The Bloch formalism provides quantitative description through which the two relaxation rates can be measured experimentally.

1.4.2 Solomon equations

Longitudinal relaxation for interacting spins can be treated by considering transition rates of spins between different energy levels (Bloembergen, Purcell, and Pound 1948). By deriving the population of each energy state, the magnetizations of two spins are given:

$$\begin{aligned}\frac{d\Delta I_z(t)}{dt} &= -\rho_I \Delta I_z(t) - \sigma_{IS} \Delta S_z(t) \\ \frac{d\Delta S_z(t)}{dt} &= -\rho_S \Delta S_z(t) - \sigma_{IS} \Delta I_z(t)\end{aligned}\quad (1.2)$$

In which, ρ_I and ρ_S are the auto-relaxation rate constants (equivalent to longitudinal relaxation rate in the Bloch equations) for spin I and S, respectively. σ_{IS} is the cross-relaxation rate for exchange of magnetization between the two spins. The equations could be extended to multiple interacting spins:

$$\begin{aligned}\frac{d\Delta I_{kz}(t)}{dt} &= -\rho_K \Delta I_{Kz}(t) - \sum_{k \neq j} \sigma_{kj} \Delta I_{jz}(t) \\ \rho_K &= \sum_{k \neq j} \rho_k\end{aligned}\quad (1.3)$$

This equation describes the relaxation of the k th spin by all other spins in the absence of interference effects. Written in matrix form, we have:

$$\frac{d\Delta M_z(t)}{dt} = -R\Delta M_z(t) \quad (1.4)$$

In which R is a N by N matrix with elements defined as $R_{kk}=\rho_k$, $R_{kj}=\sigma_{kj}$; $\Delta M_z(t)$ is an N by 1 column vector and $\Delta M_k(t)=\Delta I_{kz}(t)$. By matrix manipulations, the Solomon equations have the formal solution as:

$$\begin{aligned} a_{II}(t) &= a_{SS}(t) = 0.5 \exp[-(\rho - \sigma)t][1 + \exp(-2\sigma t)] \\ a_{IS}(t) &= a_{SI}(t) = -0.5 \exp[-(\rho - \sigma)t][1 + \exp(-2\sigma t)] \end{aligned} \quad (1.5)$$

The Solomon equations are crucial for the study of NOE for biological macromolecules.

1.4.3 Master Equation

In practical applications, the most useful approach is a microscopic semi-classical theory of spin relaxation developed by Bloch, Wangsness and Redfield (BWR) (Wangsness and Bloch 1953; Redfield 1965). The Hamiltonian of the spin system is described as the sum of a deterministic quantum-mechanical Hamiltonian and a stochastic Hamiltonian:

$$H(t) = H_{\text{det}}(t) + H_1(t) = H_0 + H_{rf}(t) + H_1(t) \quad (1.6)$$

in which, H_0 is the Zeeman and scalar coupling Hamiltonians, $H_{rf}(t)$ is the Hamiltonian for applied rf fields. The deterministic Hamiltonian applies only on the spin system itself while the stochastic Hamiltonian couples the spin system to the lattice. The time-independent Hamiltonian H_0 , which under the effects of the time-dependent Hamiltonians $H_{rf}(t)$ and $H_1(t)$, could be removed

by transforming the Liouville-von Neumann equation into the interaction frame. In the absence of rf field, the stochastic Hamiltonians and density operator in the interaction frame are given by:

$$\begin{aligned} H^T(t) &= \exp(iH_0t)H_1(t)\exp(-iH_0t) \\ \sigma^T(t) &= \exp(iH_0t)\sigma_1(t)\exp(-iH_0t) \end{aligned} \quad (1.7)$$

And the final result is:

$$\frac{d\sigma^T(t)}{dt} = -i[H_1^T(t), \sigma^T(t)] \quad (1.8)$$

Solve the equation by successive approximations to second order and we have:

$$\frac{\sigma^T(t)}{dt} = -i[H_1^T(t), \sigma^T(0)] - \int_0^t d\tau [H_1^T(t), [H_1^T(t-\tau), \sigma^T(0)]] \quad (1.9)$$

The above equation represents the evolution of the density operator for a specific realization of $H_1(t)$. The equation for any macroscopic sample can be obtained by performing ensemble average on both side of the above equation under certain assumptions that: $H_1^T(t)$ of the ensemble is zero; $H_1^T(t)$ and $\sigma^T(t)$ are not correlated; the characteristic correlation time, τ_c , is much shorter than t and $1/R$. The final result will be:

$$\frac{d\sigma^T(t)}{dt} = - \int_0^t d\tau [H_1^T(t), [H_1^T(t-\tau), \sigma^T(0)]] \quad (1.10)$$

By secular approximation ($\omega_{\alpha\alpha} = \omega_{\beta\beta}$), the matrix formalism of the master equation will be (Ernst, Bodenhausen, and Wokanu 1987):

$$\frac{d\sigma_{\alpha\alpha'}^T(t)}{dt} = \sum_{\beta, \beta'} R_{\alpha\alpha'\beta\beta'} (\sigma_{\beta\beta'}^T(t) - \sigma_{0\beta\beta'}) \quad (1.11)$$

In which

$$R_{\alpha\alpha'\beta\beta'} = 0.5 [J_{\alpha\beta\alpha'\beta'}(\omega_{\beta'\alpha'}) + J_{\alpha\beta\alpha'\beta'}(\omega_{\alpha\beta}) - \delta_{\alpha'\beta'} \sum J_{\gamma\beta\gamma\alpha}(\omega_{\gamma\alpha}) - \delta_{\alpha'\beta'} \sum J_{\gamma\alpha'\gamma\beta'}(\omega_{\alpha'\gamma})]$$

$$J_{\alpha\alpha'\beta\beta'} = 2 \int_0^{\infty} \langle \alpha | H_1(t) | \alpha' \rangle \langle \beta | H_1(t+\tau) | \beta' \rangle \exp(-i\omega\tau) d\tau$$

1.4.4 Spectral density function

The power spectral density (PSD) function describes how the power of a signal or time series is distributed over the different frequencies, in which the Fourier transform is often used. For a protein sample, the PSD function measures the contribution to reorientational dynamics of the molecule from motions with frequency components from ω to $\omega + \delta\omega$.

The rotational Brownian motions of molecules in solution differ, and hence do the resultant oscillating magnetic fields. For a rigid spherical molecule in liquid the auto-spectral density function takes the form of:

$$j(\omega) = c_0^2 \operatorname{Re} \left\{ \int_{-\infty}^{\infty} C_{00}^2(\tau) \exp\{-i\omega\tau\} d\tau \right\} \quad (1.12)$$

In which $C_{00}^2(\tau)$ is the reorientational correlation function:

$$C_{00}^2(\tau) = \overline{Y_2^0[\Omega(t)] Y_2^0[\Omega(t+\tau)]} \quad (1.13)$$

For isotropic rotational diffusion of a rigid molecule, the above correlation

function is:

$$C_{00}^2(\tau) = 0.2 \exp\left(\frac{-\tau}{\tau_c}\right) \quad (1.14)$$

where τ_c is the characteristic correlation time of the molecule and is a measurement of the average time for the molecule to rotate by one radian. The correlation time for biological macromolecules in solution is of the order of nanoseconds and is affected by certain factors such as molecular size, solvent viscosity and temperature. The resultant spectral density function therefore is:

$$J(\omega) = \frac{0.4\tau_c}{1 + \omega^2\tau_c^2} \quad (1.14)$$

An important application of the spectral density function called Model Free analysis pioneered by Lipari-Szabo (Lipari and Szabo 1982b; Lipari and Szabo 1982a) and extended later by Clore and coworkers (Clore et al. 1990) in which the corresponding function is given as:

$$j(\omega) = 0.4\overline{c_0^2} \left[\frac{S^2\tau_c}{1 + \omega^2\tau_c^2} + \frac{(1-S^2)\tau_e}{1 + \omega^2\tau_e^2} \right]$$

$$\tau^{-1} = \tau_c^{-1} + \tau_e^{-1}$$

$$S^2 = \left[\overline{c_0^2} \right]^{-1} \sum_{q=-2}^2 \left| \overline{c_0 Y_2^q(\Omega)} \right|^2 \quad (1.15)$$

in which, τ_e is the effective correlation time for the internal motions, S^2 is the square of the generalized order parameter that defines the amplitude of the intramolecular motion and satisfies the inequality that: $0 \leq S^2 \leq 1$. The overbar indicates an ensemble average performed over the equilibrium distribution of orientations Ω in the molecular reference frame. $Y_2^q(\Omega)$ is the modified second order spherical harmonic functions while Ω is the time-dependent polar angle

defining the orientation of the internuclear unit vector for dipolar interactions or the symmetry axis of the chemical shielding tensor for CSA interactions. Larger S^2 indicates increased rigidity so that $S^2=0$ or $S^2=1$ represents complete flexible or rigid in terms of internal motions for that bond vector. The above equation has been used extensively in analyzing relaxation data of proteins (M E Hodsdon and Cistola 1997; Stone et al. 1992; Xia, Deng, and Levy 2013; Osborne et al. 2001).

In a typical Model Free analysis (J. Chen, Brooks, and Wright 2004), T_1 , T_2 and steady state heteronuclear NOE are used collectively in extracting the desired parameters such as the generalized order parameter and the rotational/effective correlation time. The three observables are expressed as:

$$\begin{aligned}
 R_1 &= \frac{d^2}{4} [3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_N + \omega_H)] + \frac{c^2}{3} J(\omega_N) \\
 R_2 &= \frac{d^2}{8} [4J(0) + 3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_N + \omega_H)] \\
 &+ \frac{c^2}{18} [4J(0) + 3J(\omega_N)] + R_{ex} \\
 NOE &= 1 + \frac{d^2 \gamma_H}{4R_1 \gamma_N} [6J(\omega_N + \omega_H) - J(\omega_H - \omega_N)] \\
 d &= (\mu_0 h \gamma_H \gamma_N / (8\pi^2)) \langle r_{NH}^{-3} \rangle \\
 c &= (\omega_N / \sqrt{3}) \Delta\sigma
 \end{aligned} \tag{1.16}$$

in which μ_0 is the permeability of the free space, γ_N and γ_H are the gyromagnetic ratio of ^{15}N and ^1H , respectively, h is the Plank's constant, r_{NH} is the length of the NH bond, $\Delta\sigma$ is the chemical shift anisotropy of ^{15}N and is the difference

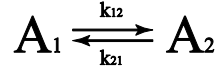
between the parallel and perpendicular components of the assumed axially symmetric chemical shift tensor(Abragam, A. 1961; Cavanagh et al. 2007).

In cases where the overall rotational diffusion is anisotropic, an alternative approach of the model-free analysis was proposed by Schurr and coworkers(Schurr, Babcock, and Fujimoto 1994). In this treatment, a localized correlation time is allocated to each residue instead of a global overall correlation time, hence generating residue-specific results. The application of this approach has been proved helpful to flexible molecule such as partially unfolded proteins(Yang et al. 1997; Ran et al. 2003).

1.4.5 Chemical exchange effects in NMR

The exchange of a nucleus between environments due to chemical reactions or conformational rearrangement has certain effects on its magnetization. By observing the magnetization evolution of the nucleus in magnetically distinct sites, NMR spectroscopy offers a unique way to monitor such exchange processes. The chemical shift timescale is defined by the frequency differences of two exchanging resonances. If the exchange rate is slow ($k \ll \Delta\omega$), two signals are observed for nuclei in the two sites; if the exchange rate is fast ($k \gg \Delta\omega$), a single resonance is observed at the population-weighted average chemical shift of the nucleus in the related two sites; for intermediate exchange ($k \approx \Delta\omega$), a broadened signal will be observed.

For a simple two-site exchange process:



in which k_{12} and k_{21} represent for the reaction or exchange rate constant for the forward and reverse reaction, respectively.

In the absence of chemical exchange, the magnetization of a nucleus follows the description of the Bloch equation (Wangsness and Bloch 1953):

$$\frac{d}{dt} M_j(t) = \gamma(1 - \sigma_j) M_j(t) \times B(t) - R_j [M_j(t) - M_{j0}] \quad (1.17)$$

in which σ is the chemical shielding constant, R is the relaxation matrix and M is a column vector defines the initial magnetization of all species.

In the presence of chemical exchange, the modified Bloch equation that describes the system are given as (McConnell 1958):

$$\frac{d}{dt} M_j(t) = \gamma(1 - \sigma_j) M_j(t) \times B(t) - R_j [M_j(t) - M_{j0}] + \sum_r K_{jr} M_r(t) \quad (1.18)$$

In which K is the kinetic matrix related to the exchange rate constants. In magnetic equilibrium, the magnetization is proportional to the instant concentration of the corresponding species:

$$M_{j0}(t) = M_0 \frac{[A_j](t)}{\sum_k [A_k]} \quad (1.19)$$

In the rotating frame, the longitudinal and transverse magnetization evolve independently of each other:

$$\begin{aligned}
\frac{d}{dt}M_j^+(t) &= (i\Omega - R_{2j})M_j^+ + \sum_r K_{jr}M_r^+ \\
\frac{d}{dt}M_{jz}(t) &= -R_{1j}(M_{jz} - M_{j0}(t)) + \sum_r K_{jr}M_{rz}
\end{aligned} \tag{1.20}$$

In which $M_j^+ = M_{jx} + iM_{jy}$ and $\Omega_j = -\gamma(1 - \sigma_j)B_0$ is the chemical shift frequency. The above can be conveniently written to matrix form:

$$\begin{aligned}
\frac{d}{dt}M^+(t) &= L^+M^+(t) \\
\frac{d}{dt}M_z(t) &= L \left[M_z(t) - M_0(t) \right] + KM_0(t) \\
L^+ &= i\Omega - \Lambda + K \\
L &= -R + K
\end{aligned} \tag{1.21}$$

When in equilibrium, the last term of the second equation in [1.21] becomes zero. The kinetic matrix L^+ and L include the effect to magnetization evolution from precession, relaxation and chemical exchange process. The element of diagonal matrix Ω corresponds to chemical shift frequencies of the related species. The transverse relaxation matrix Λ is also diagonal, its elements are defined as $\lambda_{ij} = \sigma_{ij}R_{2j}$. The cross relaxation rates of nuclei between different species are represented by the off-diagonal elements. Therefore, for a two-site exchange system, we have:

$$\frac{d}{dt}M^+ = \begin{bmatrix} -(R_{2a}^0 + k_a) & kb \\ ka & -(R_{2b}^0 + k_b + i\delta\omega) \end{bmatrix} M_t \tag{1.22}$$

The modified Bloch equation for a three-site exchanging system is slightly cumbersome and is given as:

$$\dot{M}_j^+ = \begin{bmatrix} -k_{12} - k_{13} - R_{2(1)} + i\Omega_1 & k_{21} & k_{31} \\ k_{12} & -k_{21} - k_{23} - R_{2(2)} + i\Omega_2 & k_{32} \\ k_{13} & k_{23} & -k_{31} - k_{32} - R_{2(3)} + i\Omega_3 \end{bmatrix} M_j^+ \quad (1.23)$$

Equation 1.22 could be solved analytically(Palmer et al. 2001):

$$\dot{M}_j^+ = \begin{bmatrix} a_{AA}(t) & a_{AB}(t) \\ a_{BA}(t) & a_{BB}(t) \end{bmatrix} M_j^+ \quad (1.24)$$

In which

$$\begin{aligned} a_{AA}(t) &= \frac{(tt_1 + tt_2)}{2} \\ a_{BB}(t) &= \frac{(tt_3 + tt_4)}{2} \\ tt_1 &= \left(1 - \frac{-i\Delta\omega + R_{2A}^0 - R_{2B}^0 + k_{ex}(p_B - p_A)}{\lambda_+ - \lambda_-} \right) \exp(-\lambda_- t) \\ tt_2 &= \left(1 + \frac{-i\Delta\omega + R_{2A}^0 - R_{2B}^0 + k_{ex}(p_B - p_A)}{\lambda_+ - \lambda_-} \right) \exp(-\lambda_+ t) \\ tt_3 &= \left(1 + \frac{-i\Delta\omega + R_{2A}^0 - R_{2B}^0 + k_{ex}(p_B - p_A)}{\lambda_+ - \lambda_-} \right) \exp(-\lambda_- t) \\ tt_4 &= \left(1 - \frac{-i\Delta\omega + R_{2A}^0 - R_{2B}^0 + k_{ex}(p_B - p_A)}{\lambda_+ - \lambda_-} \right) \exp(-\lambda_+ t) \\ a_{AB}(t) &= \frac{k_{ex}p_A}{\lambda_+ - \lambda_-} [\exp(-\lambda_- t) - \exp(-\lambda_+ t)] \\ a_{BA}(t) &= \frac{k_{ex}p_B}{\lambda_+ - \lambda_-} [\exp(-\lambda_- t) - \exp(-\lambda_+ t)] \end{aligned} \quad (1.25)$$

and

$$\begin{aligned} \lambda_{\pm} &= 0.5 \left\{ -i\Omega_A - i\Omega_B + R_{2A}^0 + R_{2B}^0 + k_{ex} \right. \\ &\quad \left. \pm \left[(-i\Delta\omega + R_{2A}^0 + R_{2B}^0 + k_{ex}(p_B - p_A))^2 + 4p_A p_B k_{ex}^2 \right]^{1/2} \right\} \end{aligned} \quad (1.26)$$

CHAPTER 2

SAMPLE PREPARATION AND NMR EXPERIMENTS

Chapter 2 Sample preparation and NMR experiments

2.1 Materials and methods

In this section, protein expression and purification process will be briefly described.

2.1.1 Media

LB media

LB media was used for initial tests of the expression system and production of the unlabeled protein. The preparation of the LB media and LB agar was done according to manufacturer's instructions.

M9 media

M9 media was prepared by mixing the following ingredients: 1 g NH_4Cl , 1-4 g Glucose, 0.1 ml CaCl_2 (1M), 2 ml MgSO_4 (1M), 100 ml 10x M9 stock (64 g $\text{Na}_2\text{HPO}_4\cdot 7\text{H}_2\text{O}$, 15 g KH_2PO_4 , 2.5 g NaCl), 0.5-1 ml ampicillin (0.1g/L) and 873 ml deionized water. ^{15}N enriched NH_4Cl and/or ^{13}C enriched were used for isotope-labeled protein.

All media were autoclaved at 121 °C for 20 min prior to usage.

2.1.2 SDS-polyacrylamide gel (SDS-PAGE) electrophoresis

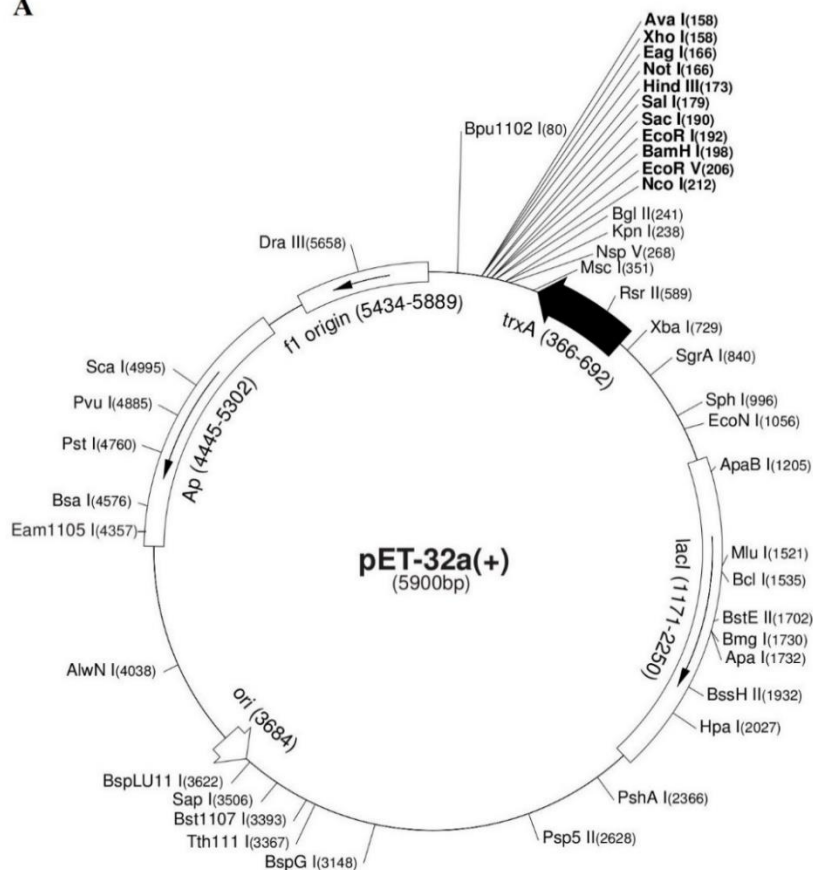
Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was used in protein expression and purification.

The 12.5% and 15% SDS-PAGE were prepared according to the standard recipes. All protein samples in loading buffer were water-boiled for 3-5 min (10 min for whole-cell) before subjected to the gel well. The electrophoresis buffer was prepared by diluting 9 g Tris base, 43.2 g Glycine and 7.5 ml 20% SDS into deionized water to a volume of 1.5 liters. After electrophoresis, the gel was stained (25% (v/v) isopropanol, 10% (v/v) acetic acid, 0.1% (w/v) comassie blue R) for 20-30 min and then de-stained (10% (v/v) acetic acid) to remove the background color(Lin 2006).

2.1.3 Expression and purification of FABPs

Expression vector pET-M (Figure 2.1.1) with cDNA coding of the human intestinal fatty acid binding protein was transformed into *E.coli* expression strain-BL21 (DE3). A single colony was inoculated into 10 ml LB media with 100 µg/ml Ampicillin (amp+) for overnight culture at 37 °C shaking at the speed of 200 rpm. The bacteria was then inoculated into 1 L LB (amp+) and grown until the OD₆₀₀ nm reached 0.6-0.8. The overexpression of hIFABP was then catalyzed by isopropyl-beta-D-1-thiogalactopyranoside (IPTG) with 0.5-1 mM final concentration at 20 °C overnight with 200 rpm shaking. Protein expression was monitored by SDS-PAGE (Figure 2.1.2) throughout the process. The *E. coli* cells were then harvested by centrifugation at 6000g for 25 min.

A



B

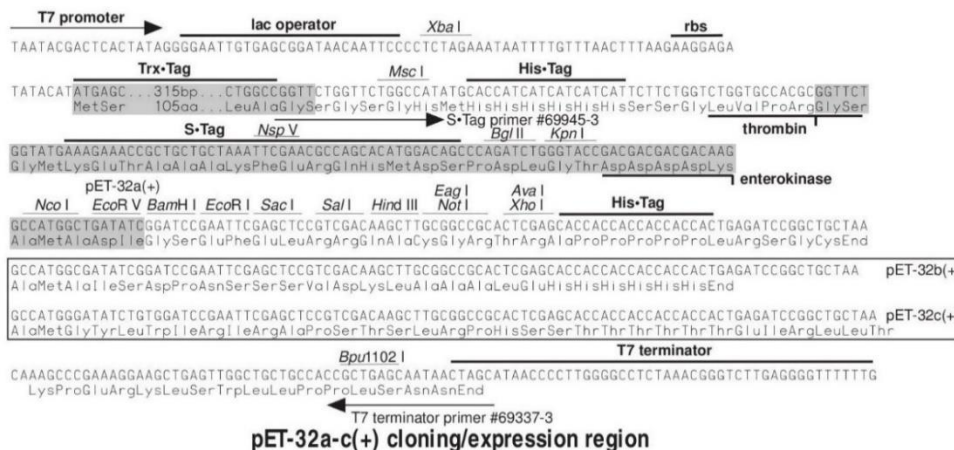


Figure 2.1.1 pET-32a(+) derived vector (pETm). (A) The circular diagram of pET-32a(+) plasmid. (B) The multiple cloning and expression region of pET-32a(+). Region in grey box are removed in pETm.

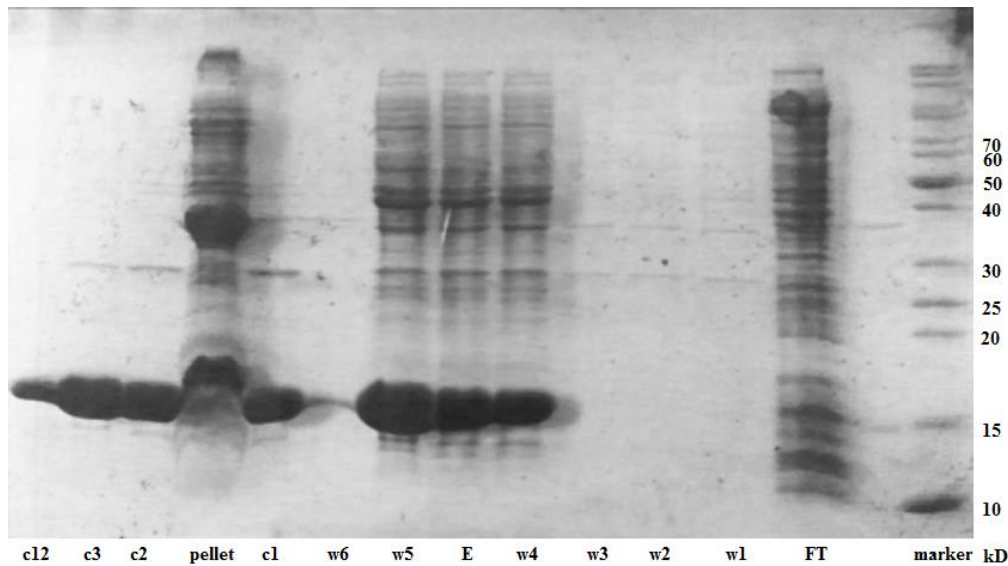


Figure 2.1.2 SDS-PAGE gel of hIFABP. FT is the flow through after binding with the nickel column. w_i is the i_{th} wash. E is the elution. C stands for collected portion after the nickel column. This process was repeated twice for better purity since hIFABP expression level was very high.

The cell pellet was re-suspended in 30 ml lysis buffer (20 mM phosphate, 300 mM NaCl, 5 mM imidazole, pH 8.0) and sonicated on ice for 15 min (1 second on, 2 seconds off; 45 min in all). After centrifuging at 12,000g for 60 min, the supernatant was collected and subjected to 6 ml Ni-NTA beads pre-equilibrated with lysis buffer and was incubated at 4 °C with mild shaking for 1 hour. The resin was then gradient-washed with 65 ml (50ml, 1mM imidazole + 15ml, 10 mM imidazole) washing buffer (20 mM phosphate, 300 mM NaCl, pH 8.0). The hIFABP with His-tag was finally eluted with 15 ml elution buffer (20 mM phosphate, 300 mM NaCl, 500 mM imidazole, pH 8.0). The protein solution with imidazole was then subjected to dialysis against the lysis buffer (20 mM phosphate, 50 mM NaCl, pH 8.0). The N terminal His-tag was then cleaved by thrombin (optional). The protein solution was then concentrated to 4 ml at 3000-4000g and subjected to fast protein liquid chromatography (FPLC) (Figure 2.1.3) equipped with pg-75 column. After gel filtration, the sample was then

applied to ion exchange chromatography equipped with mono S column for further purification (optional).

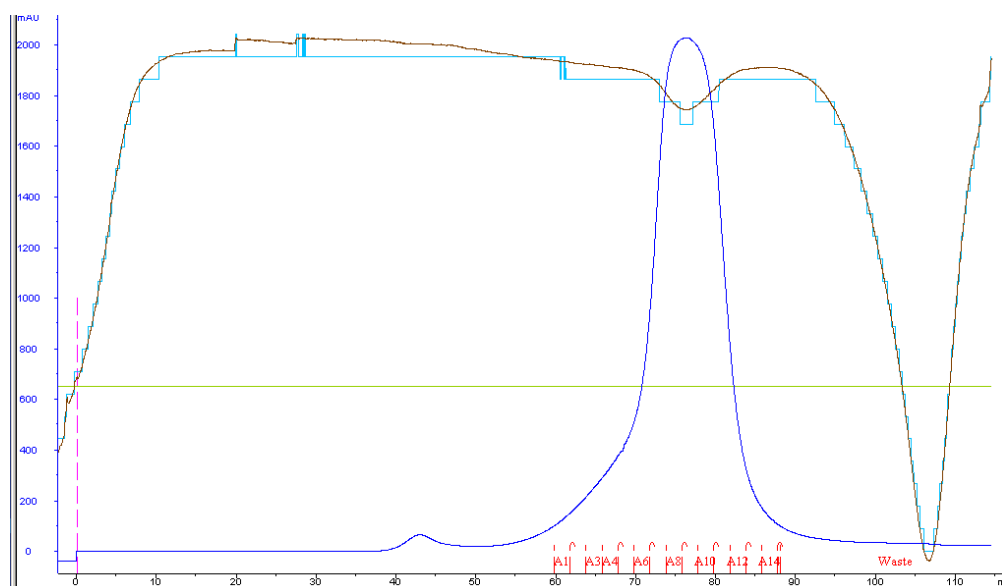


Figure 2.1.3 FPLC profile for hIFABP. Buffer condition used was 20 mM sodium phosphate, 50 mM sodium chloride, pH 7.0.

The protocols for expression and purification of isotope-labelled rat-iFABP and hLFABP were the same as stated above except that ^{15}N enriched NH_4Cl and/or ^{13}C enriched glucose were used for isotope-labeled protein. The amount of glucose used was optimized to 1-2 g/L.

2.1.4 Delipidation of FABPs

After purification, the FABP samples were subjected to delipidation to remove lipid ligands introduced during expression (Figure 2.1.4). Three methods, including lipidex affinity chromatography (Glatz & Veerkamp, 1983a; Glatz & Veerkamp, 1983b; Lassen *et al.*, 1995), charcoal treatment (R. F. Chen 1967; Spector, John, and Fletcher 1969) and solvent extraction (Cham and Knowles 1976) were tested for the delipidation reaction.

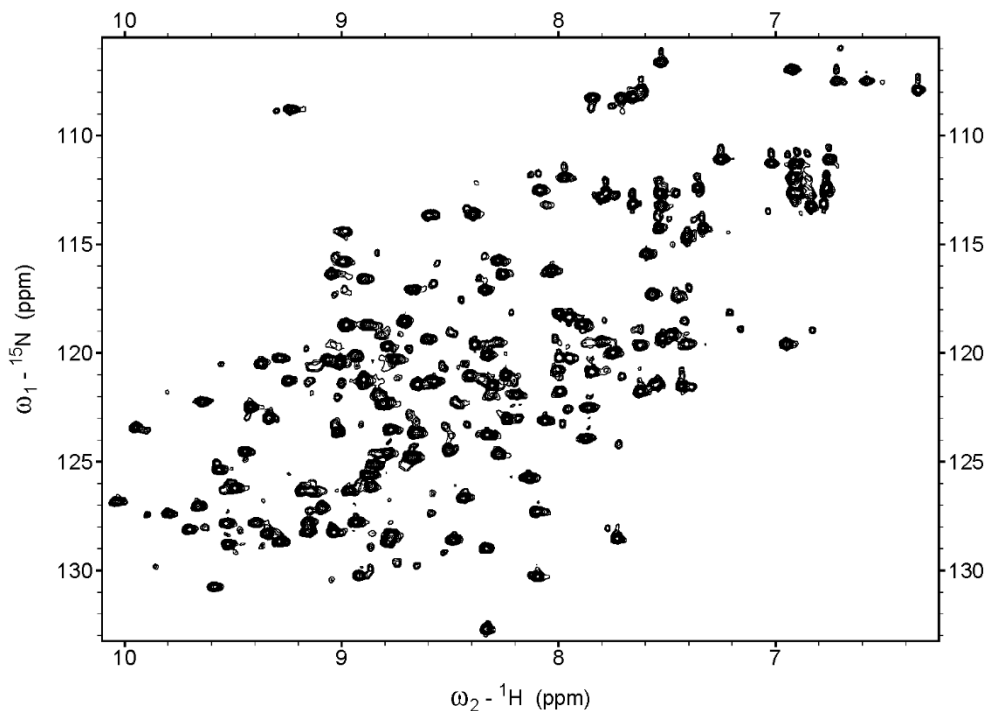


Figure 2.1.4 HSQC of holo-form hIFABP, pH8.0 at 302K. Weak accompanying peaks are very obvious. The spectrum was recorded on a 800 MHz spectrometer.

Among the three methods, solvent extraction with a mixture of 1-butanol and di-isopropyl ether in a 40:60 (v/v) ratio worked best (Figure 2.1.6) in both sample recovery rate and lipid residual value.

In addition, during purification, we found that the cation ion exchange mono S column could separate apo-form FABPs from the holo-forms. The mono S profiles of FABP indicate that there are several kinds of FABPs that show significant differences in surface charge. The major peak in the mono S column could be further separated by reducing buffer gradient (Figure 2.1.5).

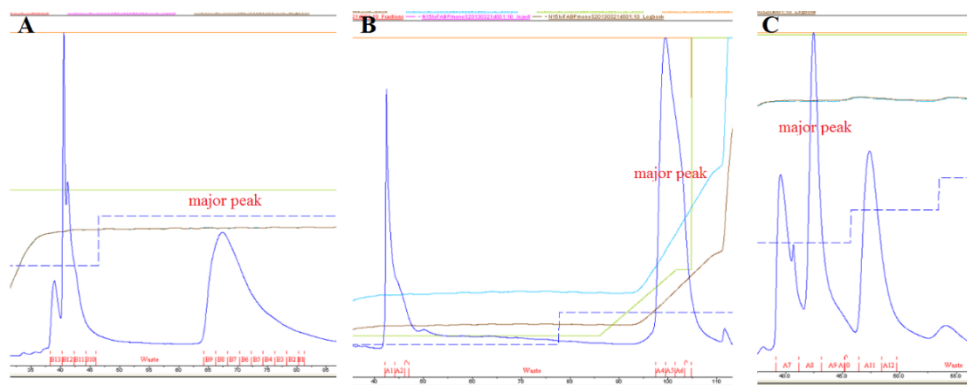


Figure 2.1.5 Mono S profiles of hIFABP. Buffer A: 20 mM NaPO₄, 35 mM NaCl, pH 5.7; buffer B: 20 mM NaPO₄, 3 M NaCl, pH 6.1; (A) Washed with B/A~0.15. (B) B/A ~0.12 (C) B/A~0.08. Different separations of the major peak were achieved by change washing buffer gradient (B/A).

The major peak separated by mono S column produced very good HSQC spectra (Figure 2.1.6). The minor NMR peaks attributed to the ligand-bound forms were not observed in HSQC, indicating complete separation of the apo-form from the holo-forms. Compared to the three methods reported previously in literature, the cation exchange column was proved the best in preparing

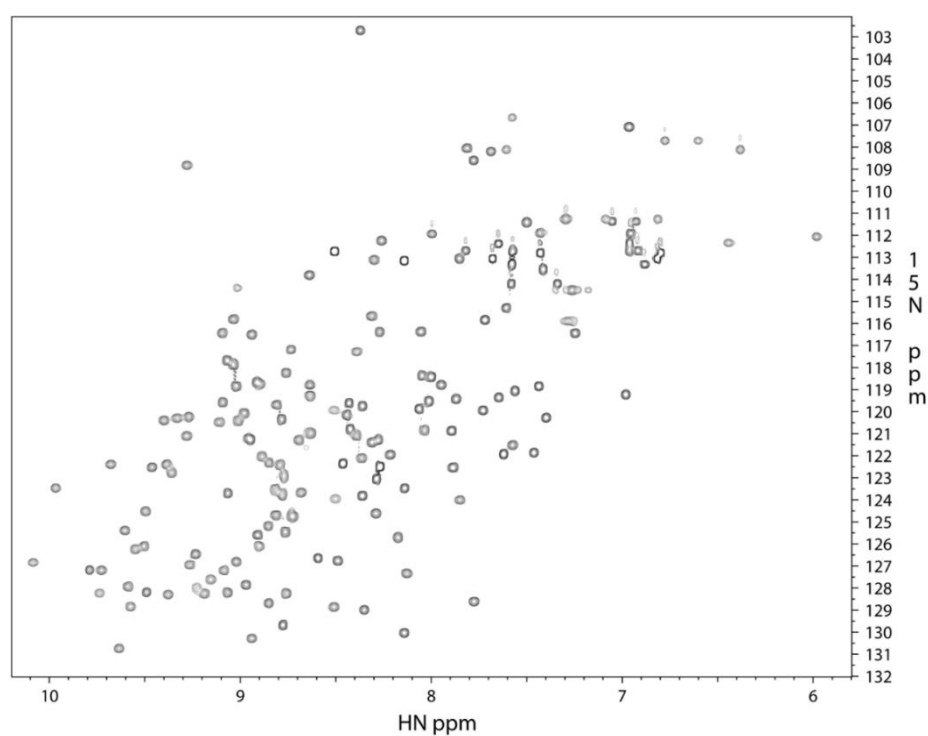


Figure 2.1.6 HSQC spectrum of apo-hIFABP purified by mono S column. The spectrum was recorded on a 800 MHz spectrometer at pH 5.75, 302K.

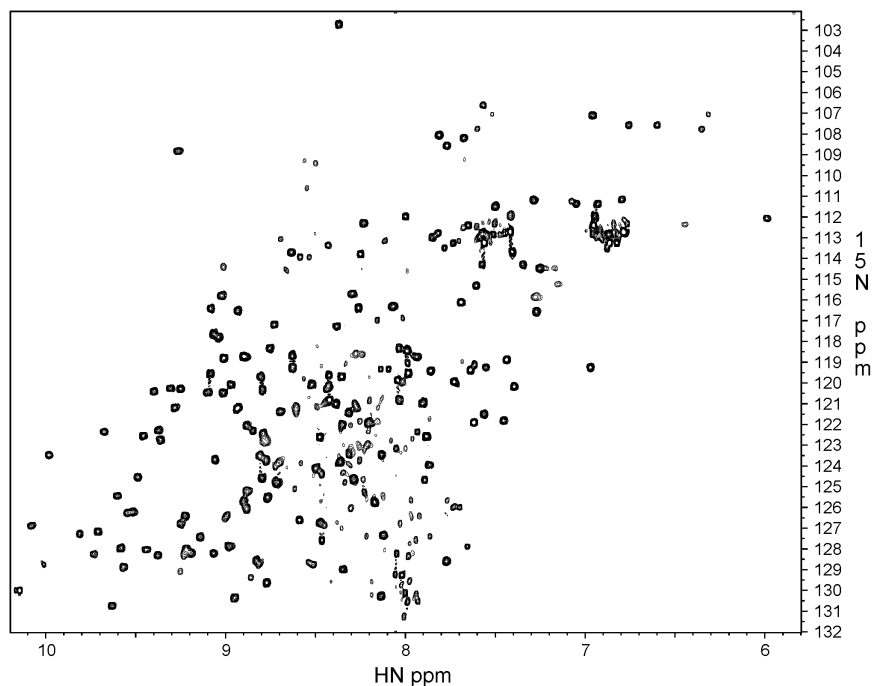


Figure 2.1.7 HSQC spectrum of apo-hIFABP purified by mono S column after two weeks. The spectrum was recorded on a 800 MHz spectrometer at pH 5.75, 302K. Obvious degradation of hIFABP was observed.

the apo-form FABPs. However, the sample obtained with mono S was prone to degradation. The HSQC spectra recorded two weeks after the mono S column showed obvious peaks of the degraded protein, indicating significant reduction in sample stability (Figure 2.1.7).

Moreover, the mono S column is not suitable for mass production of the apo-form FABPs due to the low binding capacity compared to other methods. Therefore, solvent extraction method was used for the preparation of the defatted FABPs.

Delipidation reactions by solvent extraction were performed at room temperature in 15 ml eppendorf tubes. Five ml protein solution each containing

0.5 mg of ethylenediamine tetraacetate (EDTA) (to prevent interaction with possible contaminating traces of peroxide) were added to 10ml of organic solvent consisting of DIPE and 1-butanol (60:40, v/v). The tubes were then put on a shaker with speed at 30 rpm for 1 hour. After that, the aqueous phase was separated from the organic solvent. This process repeated twice in the same manner. Thorough solvent exchange was then performed by dialysis or centrifugation to remove trace amount of the organic component. Finally, protein samples were prepared (~2.5 mM ^{13}C and/or ^{15}N -labeled hIFABP/rIFABP/hLFABP, 20 mM phosphate, 50 mM NaCl, 1 mM EDTA, pH 7.0, 95% H_2O , 5% D_2O and 0.01% NaN_3) for NMR experiments (Figure 2.1.8).

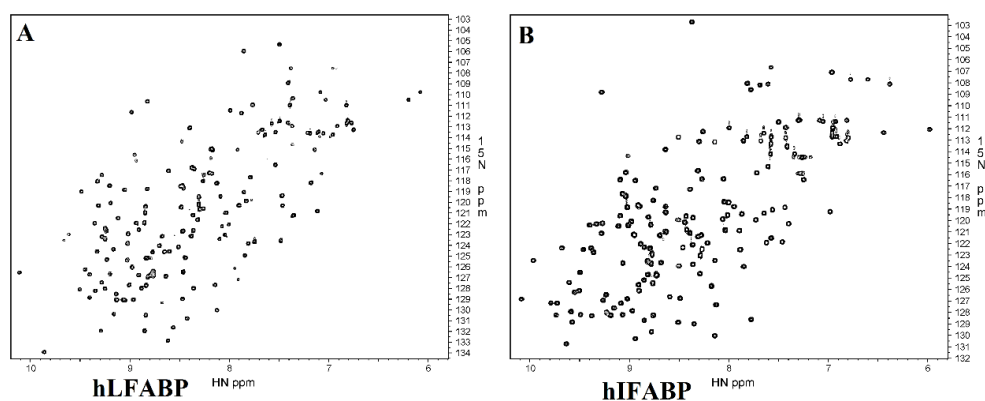


Figure 2.1.8 HSQC spectra of FABPs after delipidation. (A) HSQC of ^{15}N apo-hLFABP. (B) HSQC of apo-hIFABP. The spectra were recorded at 302 K, pH 7.0 in 18.8 T static magnetic field.

2.2 NMR experiments

The three dimensional structure of human intestinal fatty acid binding protein in the absence of ligand has been reported previously. In this work, we focus on

dynamic properties of apo-form hIFABP. The 4D experiments commonly used for structure determination are thus not necessary. Instead, a set of 2D and 3D experiments were performed for sequential backbone assignment. NMR spectra were recorded on a 800 MHz (Bruker) spectrometer equipped with cryo-probe for ^{15}N and/or ^{13}C -enriched hIFABP and hLFABP at 307K. All spectra were processed with nmrPipe software package (Delaglio et al. 1995) on linux platform. Further analysis and backbone sequential assignment was done using Sparky software(Kneller, n.d.).

2.2.1 2D ^{15}N HSQC experiment

Two-dimensional ^1H - ^{15}N HSQC spectra with sensitivity improvement (Palmer et al. 1991; Kay, Keifer, and Saarinen 1992; Schleucher et al. 1994) were recorded for hIFABP, rIFABP and hLFABP on a regular basis to check sample quality and stability before and after delipidation. Sample conditions, including pH values, salt concentration, temperature, were tested thoroughly for better spectrum resolution and stability (Figure 2.2.1).

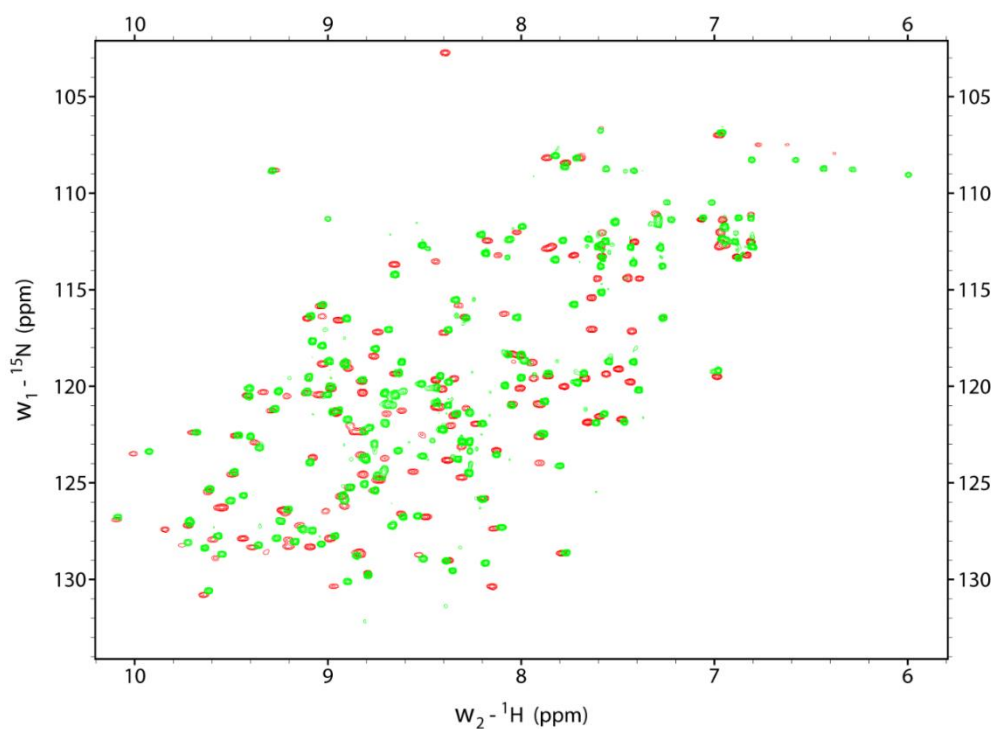


Figure 2.2.1 HSQC spectra of hIFABP at pH 4.6 (green) and pH 7.2 (red). The spectra were recorded at 800MHz spectrometer at 302K.

2.2.2 T₁ experiment

A series of 2D experiments were performed to detect the longitudinal relaxation rate of hIFABP. The delays used were 0.01 s, 0.15 s, 0.3 s, 0.45 s, 0.6 s, 0.85 s and 1.1 s. Values of R₁ were extracted according to:

$$\frac{I_t}{I_0} = \exp(-R_i t)$$

In which I_{t(0)} is the peak intensity at the delay t (0), R_i is the inverse of the relaxation time (i=1 for longitudinal relaxation, i=2 for transverse relaxation).

2.2.3 T_{1rho} experiment

Seven 2D spectra were recorded for detecting the R_{1rho} of the apo-hIFABP.

Delays were set at 0.001 s, 0.01 s, 0.02 s, 0.03 s, 0.04 s, 0.06 s and 0.08 s, respectively. The spin lock field used was 1560 Hz. Values of R_{1rho} were extracted in a manner same with R_1 . R_2 values were calculated according to:

$$R_2 = \frac{R_{1rho} - R_1 \cos^2 \left(\arctan \left(\frac{\Delta\omega}{\omega_1} \right) \right)}{\sin^2 \left(\arctan \left(\frac{\Delta\omega}{\omega_1} \right) \right)}$$

In which $\Delta\omega$ and ω_1 are the resonance offset and spin-lock field strength, respectively.

2.2.4 NOE experiment

The backbone ^1H - ^{15}N heteronuclear Nuclear Overhauser Enhancement (NOE) provides information about the motion of individual N-H bond vectors. Two different spectra were recorded in an interleaved manner with and without proton saturation. The proton saturation time and recycle delay were both set at 3 s. Values of steady-state NOEs are established from the ratio of peak intensities according to:

$$NOE = \frac{I_{saturated}}{I_{equilibrium}} - 1$$

2.2.5 3D TOCSY-HSQC

^{15}N -3D TOCSY-HSQC experiment correlates all side chain hydrogens with the ^1HN and ^{15}N of that residue. Magnetization is transferred between all $^1\text{H-X}$

spins and then to the neighbouring ^{15}N and finally back to ^1H for detection. DIPSI2 sequence was used for homonuclear Hartman-Hahn transfer. ^1H -X correlation was achieved with sensitivity improved double inept transfer. The spectrum consisted of $1280 \times 92 \times 144$ complex points with carrier frequencies set to 4.7 ppm, 118.3 ppm and 4.7 ppm, respectively. Sweep widths for each dimension were 11161 Hz, 2108 Hz and 8000 Hz with four scans, respectively.

2.2.6 3D NOESY-HSQC

^{15}N -3D NOESY-HSQC correlates the NH group of that residue to all the hydrogens nearby using the NOE. Magnetization is transferred between all ^1H -X spins and back to NH for final detection. Sensitivity improved double inept transfer was used for ^1H -X correlation. Echo-Antiecho-TPPI gradient selection was used for phase sensitive acquisition with decoupling. The spectra recorded with four scans consisted of $1280 \times 90 \times 168$ complex points for F1 (HN), F2 (^{15}N) and F3 (^1H) dimension, respectively. The corresponding carrier frequencies and the sweep widths for each dimension are 4.7 ppm, 118.3 ppm 4.7 ppm, and 11161 Hz, 2108 Hz, 9601 Hz, respectively.

2.2.7 3D HNC0 experiment

In HNC0 experiment, magnetization is passed from ^1H to ^{15}N and then selectively to the carbonyl ^{13}C via the ^{15}N - ^{13}C CO J-coupling. Magnetization is then transferred back to ^1HN via ^{15}N for detection. In each NH strip, there is one CO peak visible which belongs to the *i*-1 residue. This experiment

correlates the NH and ^{15}N chemical shift of one residue to the CO chemical shift of the preceding residue, providing valuable information on sequential connectivity. 1280, 96, 64 points were collected for F1 dimension (^1H), F2 dimension (^{13}C), and F3 dimension (^{15}N), respectively. The corresponding spectrum width for the three dimensions is 11161 Hz, 2027 Hz and 2415 Hz, respectively. The carrier frequencies were set at 4.7ppm, 173.3ppm and 118.6ppm for each dimension.

2.2.8 3D CC(CO)NH experiment

In CC(CO)NH experiment, magnetization is transferred from the side-chain hydrogen to their attached ^{13}C . After that, by isotropic ^{13}C mixing, the magnetization is transferred between side chain carbons and then to the carbonyl carbon which goes back to the amide nitrogen and finally the amide hydrogen for detection. This experiment correlates the HN and ^{15}N resonances of one residue with the side chain ^{13}C atoms of the preceding residue, providing crucial information on backbone sequence. 1280 points were collected for F1 dimension (^1H) with sweep width of 11161 Hz. 96 points were collected for F2 dimension (^{15}N) with sweep width of 2027 Hz. For F3 dimension (^{13}C), 160 points were collected with the sweep width of 12468 Hz. The carrier frequencies for the three dimensions were set to 4.7 ppm, 118.6 ppm and 39ppm, respectively.

2.2.9 3D HCC(CO)NH experiment

In HCC(CO)NH experiment, the magnetization transfer is similar to that of the CC(CO)NH experiment. The difference is that the resulting three-dimensional spectrum consists of one nitrogen and two hydrogen dimensions. This experiment correlates the side chain hydrogens of one residue with the ^1HN and ^{15}N of the following residue. The spectra were recorded with $1280 \times 96 \times 128$ complex points and the sweep widths were 11161 Hz, 2027 Hz and 5200 Hz, respectively. The carrier frequencies of the three dimensions were set at 4.7 ppm, 118.6 ppm and 4.7 ppm respectively.

2.2.10 Ligand titration experiment

NMR titration experiment was performed on ^{15}N hIAFBP with 1,8-ANS as ligand. The protein concentration was estimated to be $\sim 0.37\text{mM}$. Concentrated ANS solution was added into protein sample stepwise to reach the final molar ratios (ANS/hIFABP): 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 2.0, 3.0 and 4.0. A total number of twelve HSQC spectra were recorded on 800 MHz NMR at 302K.

2.2.11 Transverse relaxation dispersion experiment

Relaxation dispersion spectra were recorded on a Bruker 800 MHz spectrometer at 307 K using the pulse scheme shown in figure 2.2.2. A constant time delay ($T_{\text{CP}} = 0.05\text{ s}$) was used with a series of CPMG field strengths (40, 80, 120, 160, 200, 240, 280, 320, 400, 480, 560, 640, 800, 960 Hz). The experiment (with $\nu_{\text{CP}} = 120\text{ Hz}$) was repeated twice for the estimation of experimental

uncertainties. The ^{15}N and ^1H carrier frequency were set at 118.6 ppm and 4.7 ppm, respectively. The R_2^{eff} values obtained from these measurements were corrected with the R_1 and R_2 values according to:

$$R_2^{eff} = R_{2\text{exp}}^{eff} + \frac{(R_2 - R_1)v_{CP}(n - m)}{2v_{1N}^0(2n - m)}$$

In which R_1 and R_2 are the longitudinal and transverse relaxation rate, respectively; $R_{2\text{exp}}^{eff}$ is the directly measured effective relaxation rate; $n-m$ and m are the repetitive elements of the refocusing pulses.

2.2.12 Chemical exchange saturation transfer experiment

^{15}N CEST experiments (Bouvignies and Kay 2012) were performed with two weak radiofrequency (rf) fields of 15 and 30 Hz. For each rf field, 55 2D ^1H - ^{15}N HSQC spectra were acquired with a series of ^{15}N carrier frequencies ranging from 105.5 to 132.5 ppm at a spacing of 0.5 ppm. Each 2D data set comprised 1280×200 complex points in the ^1H and ^{15}N dimensions and was recorded with 2 scans, an inter-scan delay of 1.4 s and a saturation time (T_{EX}) of 0.5 s. Reference spectra were also recorded with similar parameters except that $T_{\text{EX}} = 0$ s.

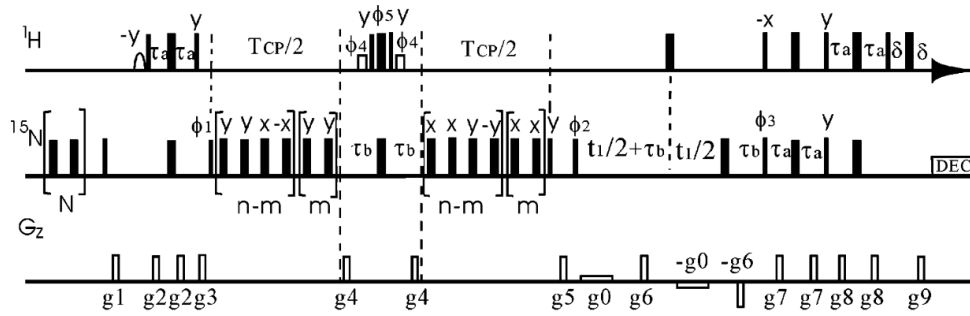


Figure 2.2.2 Pulse scheme for the measurement of relaxation dispersion. All narrow (wide) bars represent 90° (180°) rectangular pulses, applied with phase x , unless indicated otherwise. The first shaped ^1H pulse is a 1.4 ms sinc 90° pulse, while the open rectangles denote 1.6 ms ^1H 90° pulses. The ^{15}N pulses before the first 90° pulse and those pulses from pulse ϕ_1 to the point just before pulse ϕ_2 are employed with a field strength of 5.6 kHz. The delays used are $\tau_a = 2.3$ ms, $\tau_b = 2.68$ ms, $\delta = 1.1$ ms, $T_{\text{CP}} = 50$ ms $= 2[(n - m)(4\tau_{\text{cp}}) + m(2\tau_{\text{cp}})]$, where τ_{cp} is the delay between the centers of two successive 180° pulses. $N = 4(1 + n_{\text{max}}) - [4(n - m) + 2m]$, where n_{max} is the maximum of $n - m$ and is set to 12. The durations (ms) and strengths (G/cm) of sine-shaped gradients are: $g_1 = (1, 5)$, $g_2 = (1, 15)$, $g_3 = (2, 22.5)$, $g_4 = (0.5, 20)$, $g_5 = (1, -10)$, $g_6 = (1, 20)$, $g_7 = (1, 25)$, $g_8 = (1, 15)$, $g_9 = (1, 4.05)$. Weak bipolar gradients g_0 (1.5 G/cm) are used during the t_1 period. The phase cycling used is: $\phi_1 = x, -x$; $\phi_2 = 4(y), 4(-y)$; $\phi_3 = x$; $\phi_4 = 2(x), 2(-x)$; $\phi_5 = 2(-x), 2(x)$; $\text{rec} = 2(x, -x), 2(-x, x)$.

2.2.13 H-D exchange experiment

H₂O-D₂O exchange experiments were performed with ¹⁵N-¹³C-enriched hIFABP on Bruker 800MHz spectrometer at 307 K. The sample was freeze-dried overnight to remove H₂O and dissolved in D₂O immediately before recording. Each 2D spectrum was consist of 1024×100 complex points in the ¹H and ¹⁵N dimensions and was recorded with 2 scans. The carrier frequencies were set to 4.7 ppm and 118.6 ppm with sweep width of 12820 Hz and 2027 Hz for ¹H and ¹⁵N dimension, respectively. Since the H₂O-D₂O exchange started, each spectrum was recorded at 373 s, 467 s, 559 s, 652 s, 869 s, 1081 s, 1347 s, 1610 s, 1875 s, 2137 s, 2401 s, 2671 s, 2936 s, 3319 s, 3707 s, 4090 s, 4476 s, 4859 s, 5245 s, 5629 s, 6015 s, 6400 s, 6785 s, 7170 s, 7556 s, 7940 s, 9868 s, 12196 s, 14081 s, 18083 s and 20024 s, respectively. A total number of 31 2D spectra were obtained in the same manner.

2.3 Result and conclusion

In this section, hIFABP, hLFABP and rIFABP were expressed and purified with the protocol described in chapter 2.1 and 2.2. The unlabeled protein was used to test experimental conditions. To remove the indigenous lipids introduced by the expression system, three methods reported previously and the cation exchange mono S column method were tested for hI/rI/hLFABP. Although the mono S column produces the apo-form FABPs with the highest purity, the samples suffer from serious degradation problem within a short period. Therefore, the solvent extraction method was used for mass production of the

apo-form FABPs.

^{15}N and ^{15}N - ^{13}C labeled samples were prepared and defatted with the solvent extraction method to obtain the apo-form FABPs for NMR experiments. Buffer conditions for NMR samples, including pH (4.6 to 8.0), salt concentration (20 mM sodium/potassium phosphate, 20-300 mM NaCl/KCl) and temperature (302 K and 307 K) were tested thoroughly. The final sample condition used for NMR experiments was: ~2.5 mM FABPs, 20 mM Na_2PO_4 , 50 mM NaCl, 1 mM EDTA, pH 7.0, 95% H_2O , 5% D_2O and 0.01% NaN_3 . All NMR samples were protected by N_2 and put in 4 °C for storage.

For the backbone sequential assignment of hIFABP, 134 peaks were assigned (Figure 2.3.1), including 4 residues in the HIS-TAG (V-4, G-3, T-2, Q-1) and two sidechain ^{15}N ($\text{W6N}_{\epsilon 1}$ and $\text{W82N}_{\epsilon 1}$), residue N13, A73 and D74 are missing in the spectra due to fast NH- H_2O exchange. Series of 2D experiments were performed to measure the dynamic properties of FABPs. Spectra were processed and used to extract peak intensities of the corresponding residues for systematic data analysis.

CHAPTER 3

DATA ANALYSIS AND MODEL SELECTION FOR APO-HIFABP

Chapter 3 Data analysis and model selection for apo-hIFABP

Spectra were processed with nmrPipe software package as mentioned in section 2.1.5. All calculations were done with in-house written scripts executed in Matlab software (“MATLAB 8.1 and Statistics Toolbox 8.1” 2013) unless stated otherwise.

3.1 R_1 , $R_{1\rho}$ and NOE experiments

Peak intensities were extracted and subjected to nonlinear least square fittings with the corresponding time delays as described in section 2.1.5.2 and 2.1.5.3. The resultant longitudinal relaxation rates range from 0.75 to 1.40 s⁻¹ with the mean value 0.99 s⁻¹ (Figure 3.1.1). The $R_{1\rho}$ ranges from 5.21 to 33.72 s⁻¹ with the mean value 14.65 s⁻¹ (Figure 3.1.2). For the NOE, the minimum, maximum and average values are 0.31, 0.85 and 0.78, respectively (Figure 3.1.3). To exclude the possible contribution to the transverse relaxation rate due to chemical exchange processes, R_2 values (Figure 3.1.4) were calculated from R_1 , $R_{1\rho}$ and NOE of the corresponding residues according to:

$$R_2 = \frac{R_{1\rho} - R_1 \cos^2\left(\arctan\left(\frac{\Delta\omega}{\omega_1}\right)\right)}{\sin^2\left(\arctan\left(\frac{\Delta\omega}{\omega_1}\right)\right)} \quad (3.1)$$

In which $\Delta\omega$ and ω_1 are the resonance offset and spin-lock field strength, respectively. The use of the spin-lock field of 1562.5 Hz helped to suppress the additional R_2 from chemical exchanges within the same range.

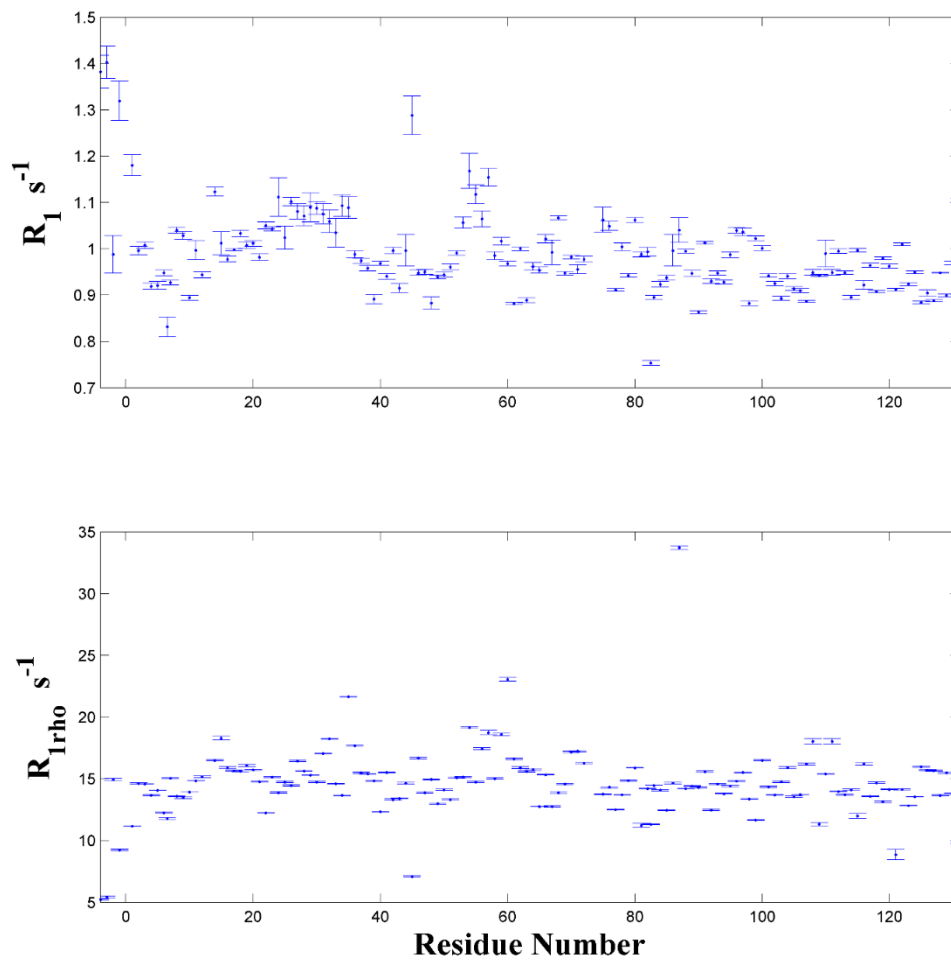


Figure 3.1.1 R_1 and $R_{1\rho}$ of apo-hIFABP. The corresponding errors are represented with the error bar. Three residues, N13, A73 and D74 are missing in the spectrum.

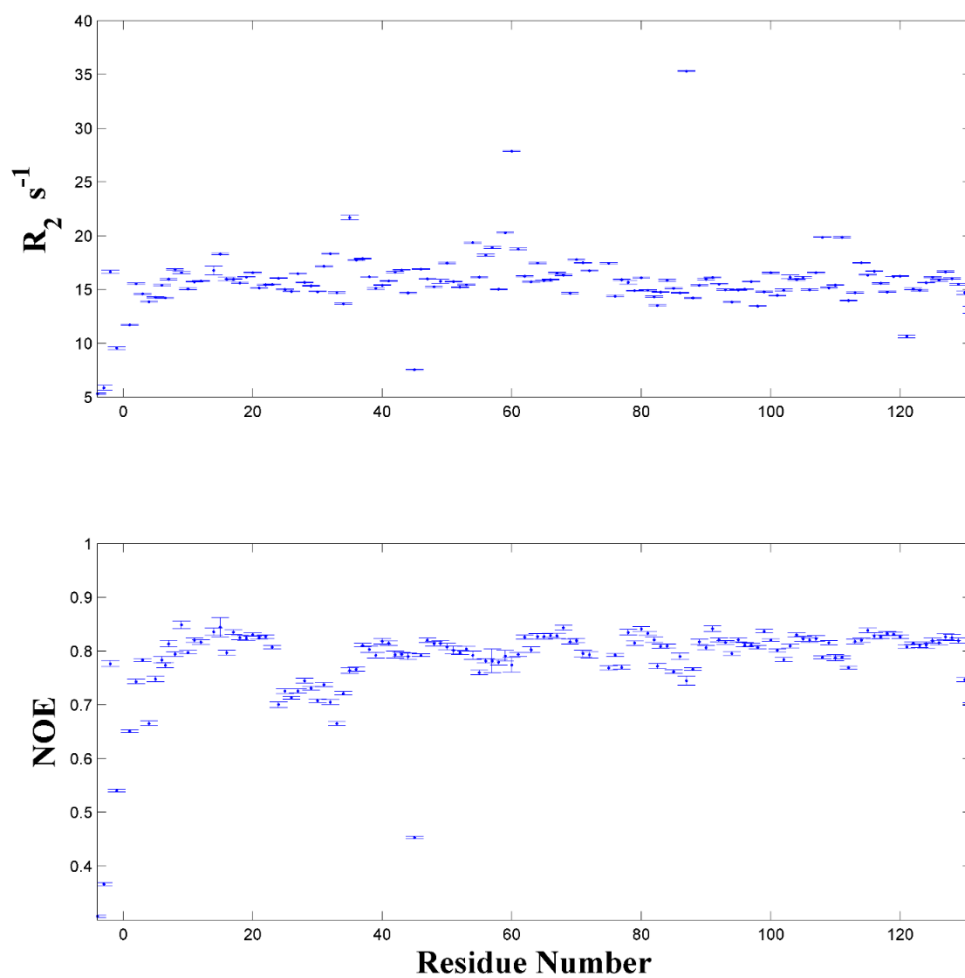


Figure 3.1.2 R_2 and NOE of apo-hIFABP. The corresponding errors are represented with the error bar. Three residues, N13, A73 and D74 are missing in the spectrum.

3.2 Ligand titration with 1,8-ANS

Chemical shift perturbation (CSP), or also known as chemical shift mapping or complexation-induced changes in chemical shift (CIS), is a very simple experiment to study the binding process of a protein to its ligands. In this study, concentrated 1,8-ANS was added stepwise to the ^{15}N -labeled apo-hIFABP to the final molar ratios ranging from 0 to 4.0 as described in section 2.1.10. A 2D ^{15}N HSQC was recorded at each stage of binding throughout the titration experiment. The chemical shift changes of each residue upon ligand binding were monitored and used to track the binding process by identifying residues with significant CSP. Combined chemical shift perturbation (CCSP) (Tochio et al. 2000; Piserchio et al. 2002) were calculated according to:

$$\Delta_{ppm}(CCSP) = \sqrt{(\Delta\delta_{HN})^2 + (\Delta\delta_N \times \alpha_N)^2} \quad (3.2)$$

In which $\Delta\delta_{HN}$ and $\Delta\delta_N$ are the differences of ^1H and ^{15}N chemical shifts in the absence and presence of ligands, respectively. α_N is a scaling factor and has a value of 0.17.

As shown in figure 3.2.1, a large number of residues were perturbed upon interaction of 1,8-ANS with hIFABP. The chemical shift walking direction indicates that the binding of 1,8-ANS to hIFABP could be described as a two-stage interaction. As can be seen in figure 3.2.2, in stage one (molar ratio 0-1), the chemical shift of most residues move to one direction. When the ligands become excessive (molar ratio between 1 and 4), the chemical shift walking changed to another direction. This result agrees well with the 1:1 stoichiometry of IFABP. In stage one, the ligand binds to the internalized binding site of

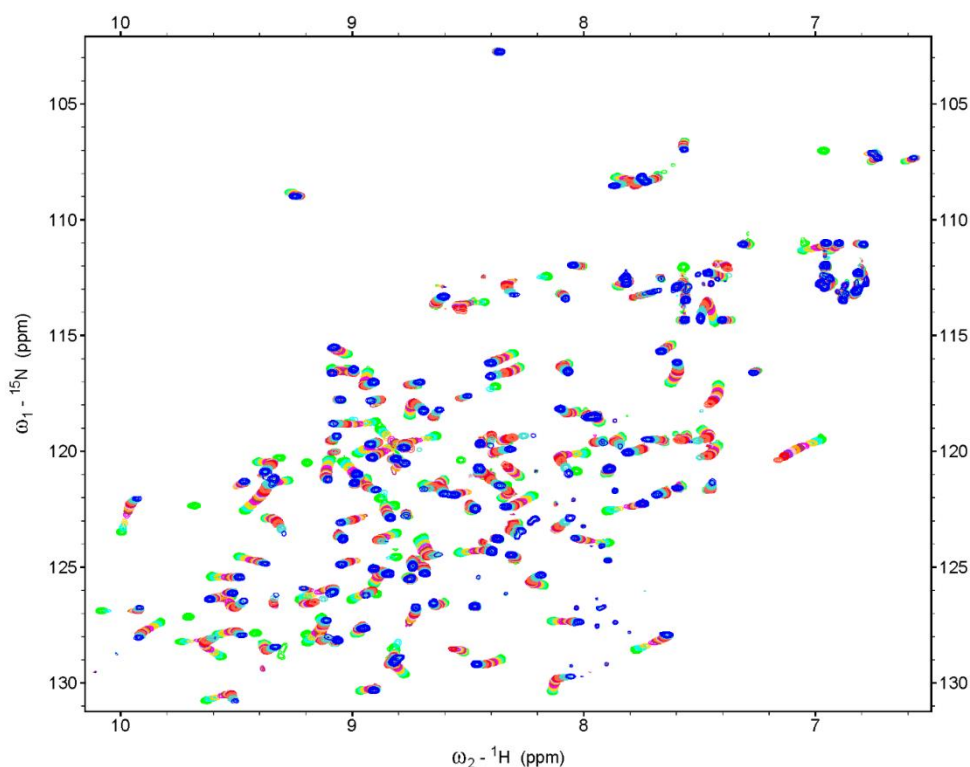


Figure 3.2.1 NMR titration experiments of ^{15}N enriched hIFABP with 1,8-ANS as ligand. A large number of residues were perturbed upon ligand binding. The protein concentration was estimated to be ~ 0.37 mM; the final molar ratios (ANS/hIFABP) were: 0 (green), 0.2 (cyan), 0.4 (gold), 0.6 (magenta), 0.8 (maroon), 1.0 (orange), 1.2 (pink), 1.4 (purple), 1.6 (red), 2.0 (tomato), 3.0 (turquoise) and 4.0 (blue).

hIFABP, the corresponding chemical shift walking is caused by the specific binding. In stage two, the chemical shift walking changes its direction on a residue basis and it should be due to the additional weak binding of ligands onto the surface of the protein. The changes of the chemical shift walking direction were not observed within the stoichiometry limit. Besides, the changes did not become obvious when the ligand to protein ratio was below 1.4. This shows that the affinity of the non-specific association is much lower than the binding inside the β -cavity. In addition, degradation was observed in the titration experiment (additional weak peaks) along with decrease of the peak intensities, showing that the protein is not stable in the presence of large amount of 1,8-ANS and the

non-specific binding on the surface of the protein might cause random aggregation that accounted for the loss of signal intensities.

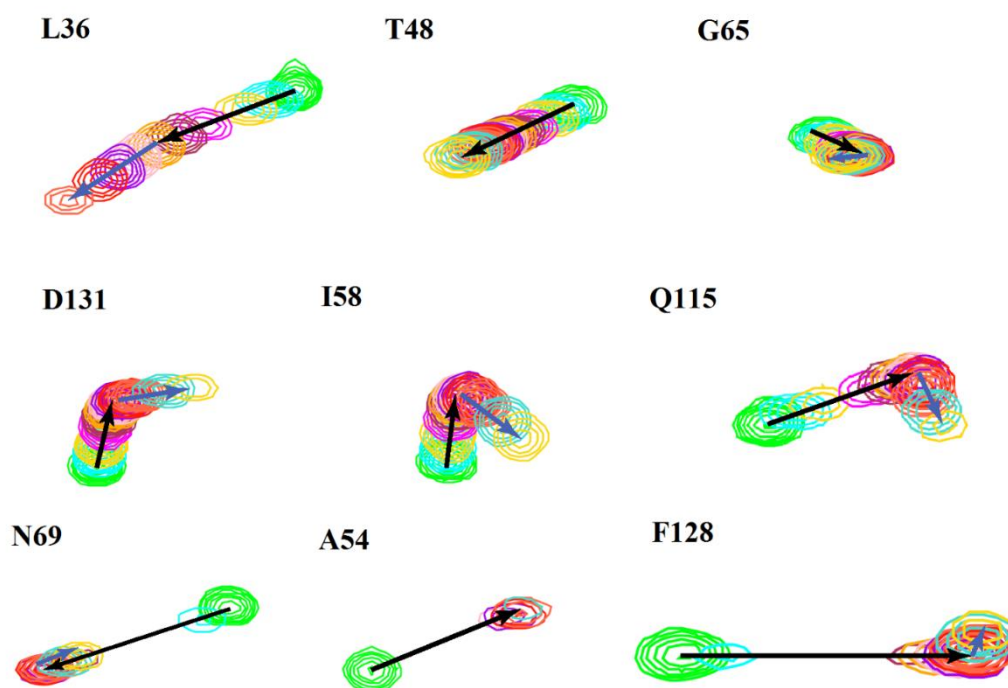


Figure 3.2.2 Representative residues in NMR titration experiments. The chemical shift walking direction due to ligand binding is indicated by the black arrow; the minor change of the direction due to additional weak association is indicated by the blue arrow.

3.3 NH-D₂O exchange experiment

The H-D exchange experiments were performed with ¹⁵N-¹³C labeled apo-hIFABP sample with a concentration ~2.5 mM, pH7.0 at 307 K as described in section 2.1.3. Peak intensities were extracted and subjected to nonlinear least square fittings according to the mono-exponential equation:

$$I_t = Ae^{-k_{prot}t} + C \quad (3.3)$$

In which I_t is the peak intensity at time t , A is the peak intensity at time 0, C is

the correction factor which shifts the curve along y axis, k_{prot} is the apparent exchange rate for the H-D exchange process. The measured exchange rates (k_{prot}) were later used to compare with calculated random chain rates (k_{rc}) for quantitative evaluation of the structural protection or slowing factors, which is defined as:

$$P = \frac{k_{rc}}{k_{prot}} \quad (3.4)$$

The reference exchange rate for a specific residue type in a random coil, k_{rc} is given by (Bai et al. 1993):

$$\begin{aligned} k_{rc} &= k_A + k_B + k_W \\ &= k_{A,ref} (A_L \times A_R) [D^+] + k_{B,ref} (B_L \times B_R) [OD^-] + k_{W,ref} (B_L \times B_R) \end{aligned} \quad (3.5)$$

In which k_A , k_B and k_W represent the exchange rate constant catalyzed by acid, base and water, respectively. $k_{A,ref}$, $k_{B,ref}$ and $k_{W,ref}$ are the standard rate constants for the pertinent alanine reference peptide catalyzed by acid, base and water, respectively. A_L , A_R and B_L , B_R are the respective side chain-specific acid and base factors. D^+ and OD^- are the concentration of the corresponding ions in D_2O in the experiments. The pD value was calculated from pH values according to:

$$pD = pH_{D_2O} + 0.4 \quad (3.6)$$

In which pH_{D_2O} is the pH value measured in D_2O and has a value of 6.98 for the hIFABP sample.

The predicted first-order rate constant for acid, base and water catalyzed

exchanges are given by:

$$\begin{aligned}
 k_A &= e^{k_{A,ref} + A_L + A_R - PD} \\
 k_B &= e^{k_{B,ref} + B_L + B_R + \log_{10}^{k_{D_2O}/10}(-PD)} \\
 k_W &= e^{k_{W,ref} + B_L + B_R}
 \end{aligned}
 \tag{3.7}$$

The temperature correction are given by:

$$k_{rc}(T) = k_{rc}(293)e^{(-Ea[1/T-1/293]/R)}
 \tag{3.8}$$

Where Ea is the activation energy with a value of 14, 17 and 19 Kcal/mol for the acid, base and water catalyzed reactions, respectively; T is the experimental temperature in Kelvin, R is the universal gas constant with a value of 1.987×10^{-3} Kcal/mol. Equation 3.8 was used to correct the three rate constants in equation 3.7 to calculate the exchange rate of a specific residue in a random coil at experimental temperature T.

The peak intensity of each residue was extracted and fitted to equation 3.3 (Figure 3.3.1), the resultant k_{prot} was then used to extract the protection factor of the corresponding residue.

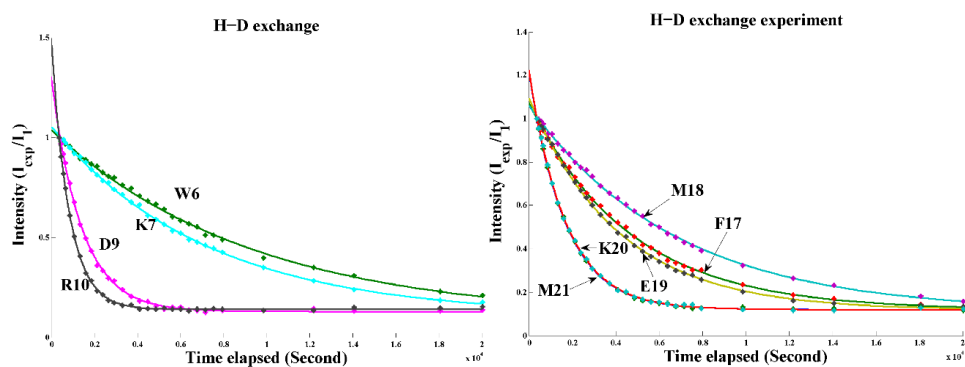


Figure 3.3.1 Representative residues in H-D exchange experiments. Colored dots are the experimental data, solid lines are the best fits.

The peak intensities of the first H-D exchange experiments were analyzed to estimate the fast exchange residues in apo-hIFABP (Figure 3.3.2). As can be seen in the picture, the H-D exchange processes of some residues had been already finished by the time the first 2D spectrum was recorded, indicating high solvent accessibility of these area.

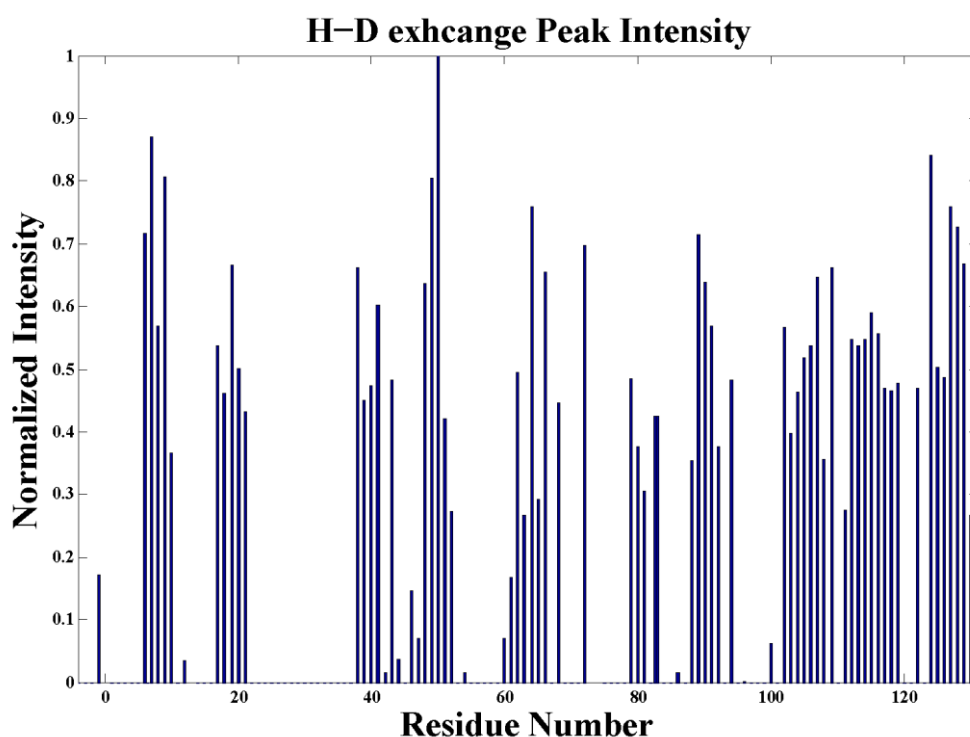


Figure 3.3.2 Relative peak intensities of the H-D exchange experiment. The ratios are calculated as the peak heights of the first spectrum of the H-D exchange experiment divided by the peak heights of the HSQC spectrum followed by normalization.

3.4 CPMG data analysis

The transverse relaxation dispersion experiments were done at two static magnetic fields as described in section 2.1.1. The peak intensities of the two sets of spectra were extracted and converted to relaxation rates and subjected to

fittings at the same time to reduce the degree of freedom of the system. Usually, relatively low radio frequency (rf) field power is used during the CPMG pulse trains in the dispersion experiments to preserve NMR probes (especially cryoprobes). In this case, substantial systematic errors could be introduced to the apparent relaxation rates due to imperfect refocusing CPMG pulses originating from off-resonance effects, rf field inhomogeneity, and miscalibrated pulses parameters. For this reason, the measured transverse relaxation rates were corrected according to (Long, Liu, and Yang 2008):

$$R_2^{eff} = R_{2\text{exp}}^{eff} + \frac{(R_2 - R_1)v_{cp}(n - m)}{2\nu_{1N}^0(2n - m)} \quad (3.9)$$

In which R_2 and R_1 are the transverse and longitudinal relaxation rates of the ^{15}N spin, respectively; $v_{cp}=1/(2\tau_{cp})$; $n-m$ and m are the numbers of the first and second repetitive elements during one $T_{cp}/2$ period, respectively. The relation of v_{cp} to T_{cp} , n and m is given in table 3.4.1. The second term in equation 3.9 represents the correction to relaxation difference of the magnetization evolving during the refocusing pulses along the x and y axes in the CPMG sequence.

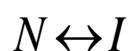
v_{cp} (Hz)	n-m	m
$2/T_{cp}$	0	1
$4/T_{cp}$	1	0
$6/T_{cp}$	1	1
$8/T_{cp}$	2	0
$10/T_{cp}$	2	1
$12/T_{cp}$	3	0
$14/T_{cp}$	3	1
$16/T_{cp}$	4	0
$18/T_{cp}$	4	1
$20/T_{cp}$	5	0
$22/T_{cp}$	5	1
$24/T_{cp}$	6	0
$26/T_{cp}$	6	1
$28/T_{cp}$	7	0

30/T _{cp}	7	1
32/T _{cp}	8	0
34/T _{cp}	8	1
36/T _{cp}	9	0
38/T _{cp}	9	1
40/T _{cp}	10	0
42/T _{cp}	10	1
44/T _{cp}	11	0
46/T _{cp}	11	1
48/T _{cp}	12	0
50/T _{cp}	12	1

Table 3.4.1 Calculation of v_{cp} from T_{cp} , m and n . T_{cp} is the constant time delay used in the CPMG experiments with a value of 50 ms.

3.4.1 Two-state fitting of CPMG data

Conventionally, CPMG data was examined by the two-site exchange model first to estimate the exchange process:



In which N and I stand for the major form and minor form, respectively.

Among all the residues that have been assigned, 108 residues were selected for the two-state CPMG fitting, excluding residues that overlap in both the 500MHz and 800MHz spectra (G44 and G86) and residues that are well separated in 800 MHz field but partially overlap in the 500 MHz (V8 and E43, K37 and A124, Y117 and E120, M18 and K88, I108 and D111) (Figure 3.4.1).

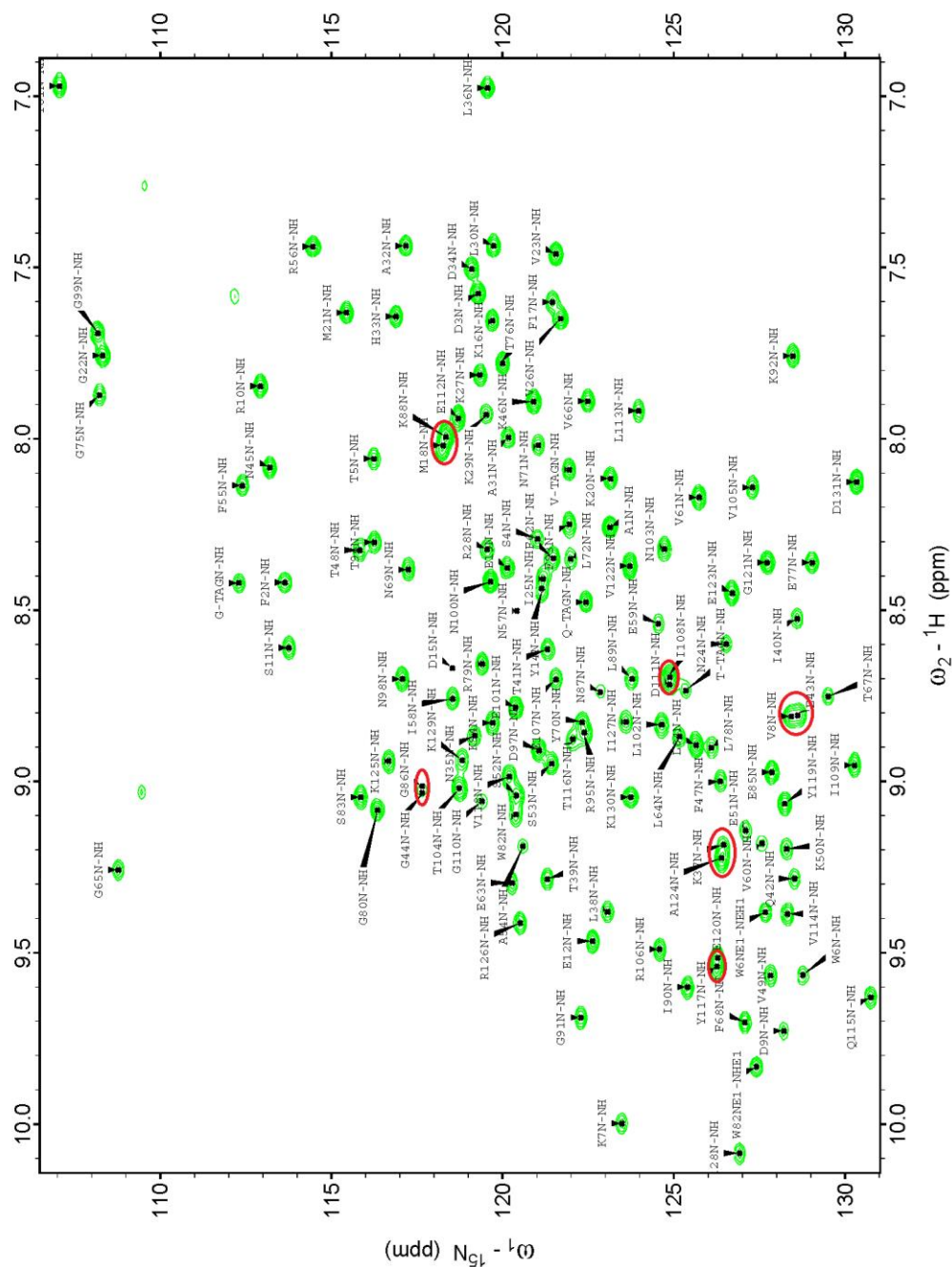


Figure 3.4.1 Reference spectrum of CPMG experiments at 500MHz magnetic field. Overlapping residues are indicated by red circles. Except G44 and G86, other highlighted residues are well separated in 800MHz field but partially overlap at the lower field and are excluded in the CPMG data analysis.

3.4.2 Individual fitting of CPMG data

The CPMG data was first examined with the two-site model by nonlinear least square fittings according to the Carver & Richards equation (Carver 1972; Davis,

Perlman, and London 1994):

$$R_{2(\frac{1}{\tau_{cp}})} = \frac{1}{2} \left(R_a + R_b + k_{ex} - \frac{1}{\tau_{cp}} [D_+ \cosh(\eta_+) - D_- \cos(\eta_-)] \right) \quad (3.10)$$

in which

$$D_{\pm} = \frac{1}{2} \left[\pm 1 + \frac{\psi + 2\Delta\omega^2}{(\psi^2 + \zeta^2)^{1/2}} \right]$$

$$\eta_{\pm} = \frac{\tau_{cp}}{\sqrt{2}} \left[\pm \psi + (\psi^2 + \zeta^2)^{1/2} \right]^{1/2}$$

$$\psi = (R_a - R_b - p_a k_{ex} + p_b k_{ex})^2 - \Delta\omega^2 + 4p_a p_b k_{ex}^2$$

$$\zeta = 2\Delta\omega(R_a - R_b - p_a k_{ex} + p_b k_{ex})$$

$$k_{ex} = k_a + k_b = \frac{k_a}{p_b} = \frac{k_b}{p_a}$$

In the analysis, residues could be explained very well by the two-state model. The resultant fractional populations for the minor state (p_1) range from 2% to 10% while the exchange rates (k_{ex}) distribute between 200 to 3500 s⁻¹ (Figure 3.4.2.1 and Figure 3.4.2.2). The extracted kinetic parameters were then used as initial values in the subsequent global fitting.

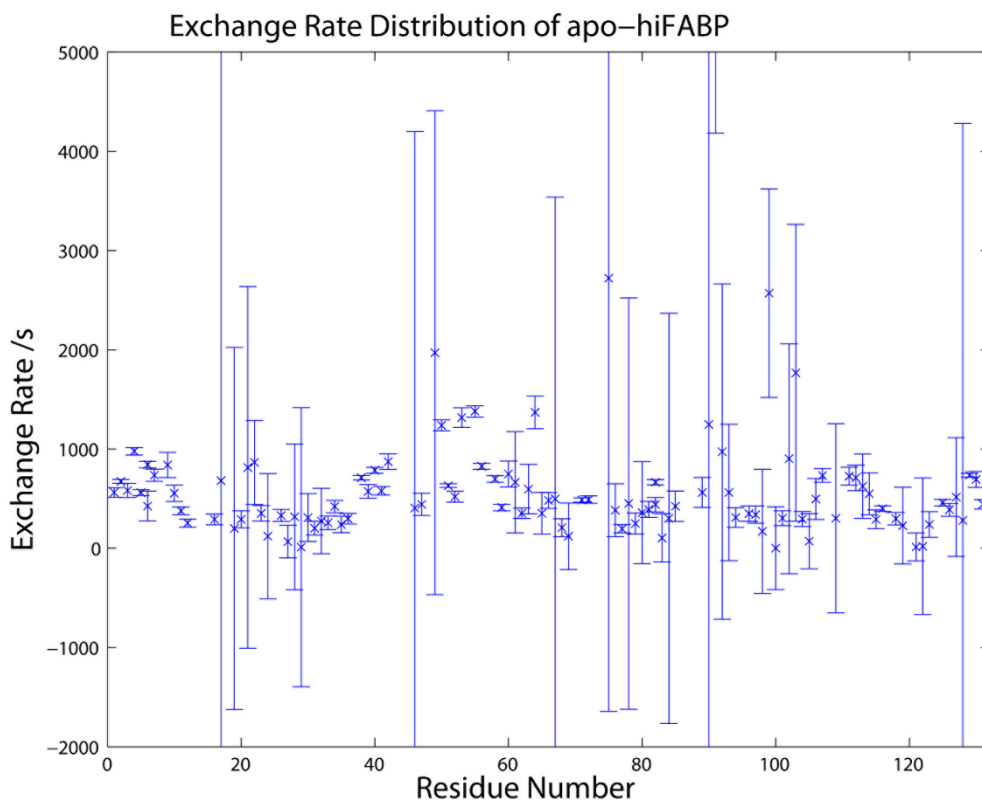
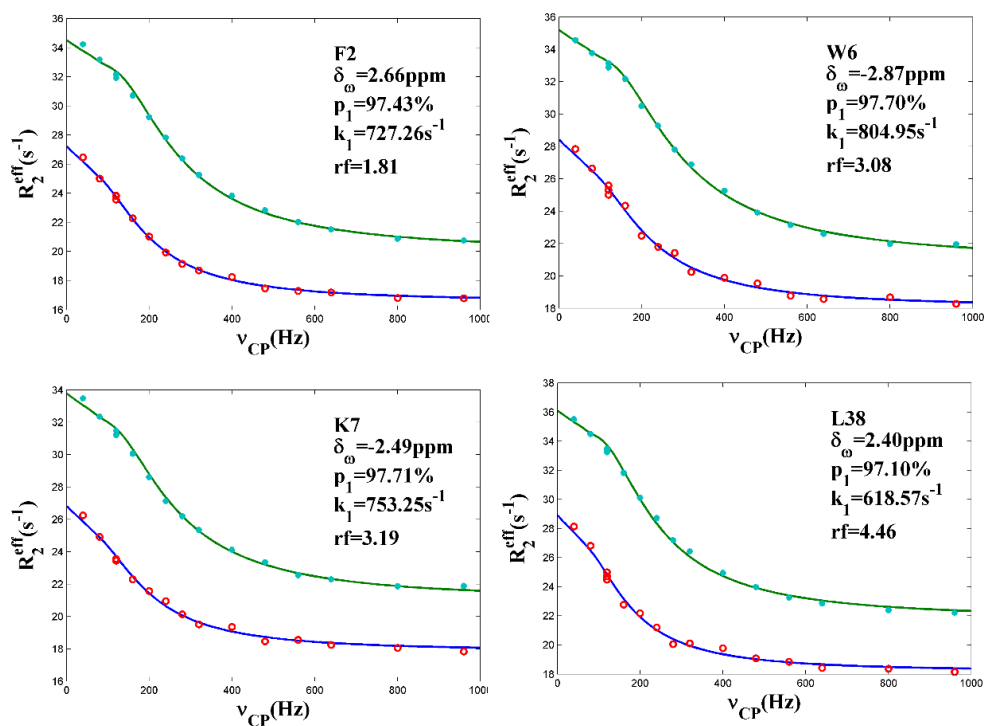


Figure 3.4.2.1 Exchange rates extracted by two-site exchange model. Errors are represented by error bar. Residues with exceptionally large errors usually have relatively small R_{ex} ($< 2 \text{ s}^{-1}$).



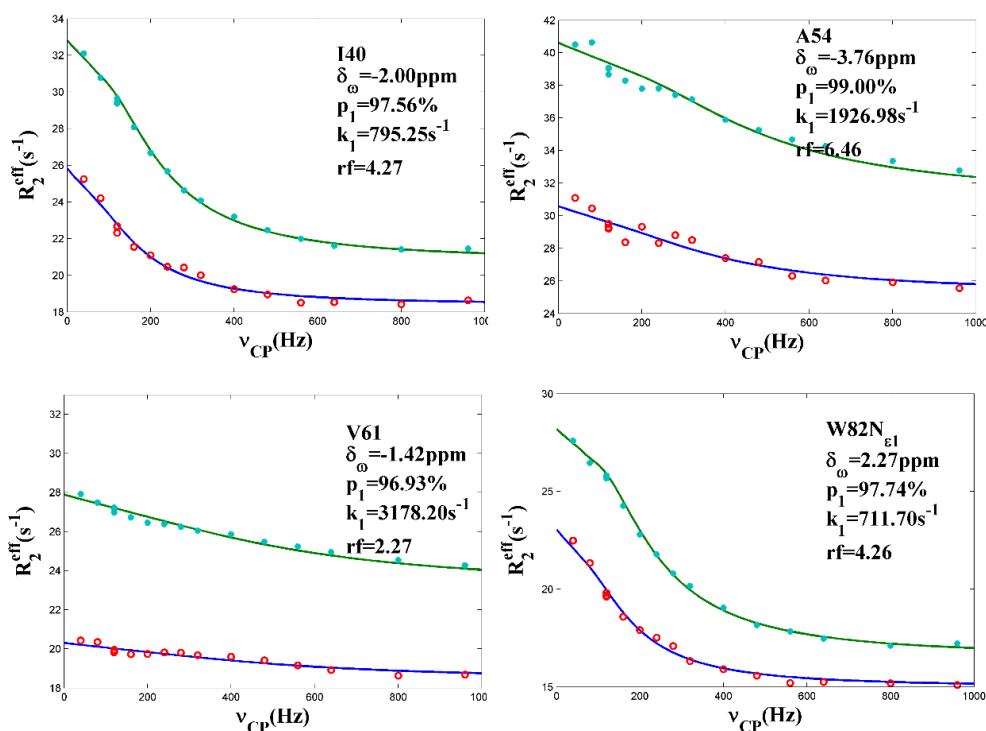


Figure 3.4.2.2 Representative residues in two-state CPMG fitting. The boundaries set for the fitting were: $0.90 \leq p_N \leq 0.99$, $300 \text{ s}^{-1} \leq k_{\text{ex}} \leq 4000 \text{ s}^{-1}$. δ_{ω} was not limited.

3.4.3 Global fitting of CPMG data

Theoretically, if a protein undergoes conformational rearrangement between two different sites, the fractional populations and the exchange rate should be the same for every residue. By analyzing data of different residues simultaneously, the degree of freedom is reduced compared to that of the individual fittings. Thus, by global fitting, one extracts a set of global parameters that accurately describe the overall chemical exchange process.

3.4.3.1 Grid search aided minimization for parameter space mapping

As described in section 3.4.2, the two-state fitting of CPMG data could be flexible since the chemical shift of each state is not limited, the change of fractional population and (or) exchange rate could be compensated by the change of the chemical shifts, thus generating multiple local minima in the parameter hyper surface. In order to extract kinetic parameters more accurately, grid search aided minimization was used for the parameter space mapping in the global fitting process.

In the parameter space, each dimension represents one fitting parameter. For a single residue, there are five parameters in the individual two-site fitting: fractional population of the major state p_N , exchange rate between the two states k_{ex} , the chemical shift difference between the two states δ_ω , and the intrinsic transverse relaxation rates of the ^{15}N spin in the two static fields R_{21} and R_{22} . The corresponding parameter space hence consists of a total number of six dimensions, five for the fitting parameters and one for the target function (fitting residual). As shown in figure 3.4.3.1, multiple local minima were observed along the p_N dimension while k_{ex} ranges from 0 to around 1000 s^{-1} . Notably, fitting residuals in these local minima are very close in value and thus are very hard to differentiate in the fitting. This plot was generated by applying grid search minimization of a theoretical data set of one residue which undergoes two-site chemical exchange. During minimization, the δ_ω and $R_{21(2)}$ were set the same as the input value, which means these three dimensions were excluded from the minimization. Consequently, more local minima are expected in real fittings, especially in the presence of random and systematic errors.

To evaluate the parameter space map output, m-n dimensions were projected onto n dimensions with logical output (Figure 3.4.3.2). In which m+n is the number of parameter space, n is the dimensions that represent the parameters through which the output is evaluated or visualized. The number of each cross point in the observing surface is calculated by $num_{ture} = d_1 \times d_2 \times d_3 \dots \times d_m$. In which d_i is the number of positive output in the i_{th} dimension. The fractional population of the cross point that has an output smaller than the predefined threshold is calculated by:

$$P_s = (d_1 \times d_2 \times d_3 \dots \times d_m) / (D_1 \times D_2 \times D_3 \dots \times D_m)$$

In which D_i is the number of output in the i_{th} dimension, m is the number of dimensions to be projected. The larger the P_s is, the more tolerable the chi square is to the change of the other parameters. If P_s is 0, the corresponding condition (e.g., a specific p_N and k_{ex}) cannot explain the CPMG data.

The target function was defined as:

$$\chi^2 = \sum (re_{1i} - rin_{1i})^2 + \sum (re_{2i} - rin_{2i})^2$$

$$\chi^2 = \sum \frac{(re_{1i} - rin_{1i})^2}{std_{1i}} + \sum \frac{(re_{2i} - rin_{2i})^2}{std_{2i}}$$

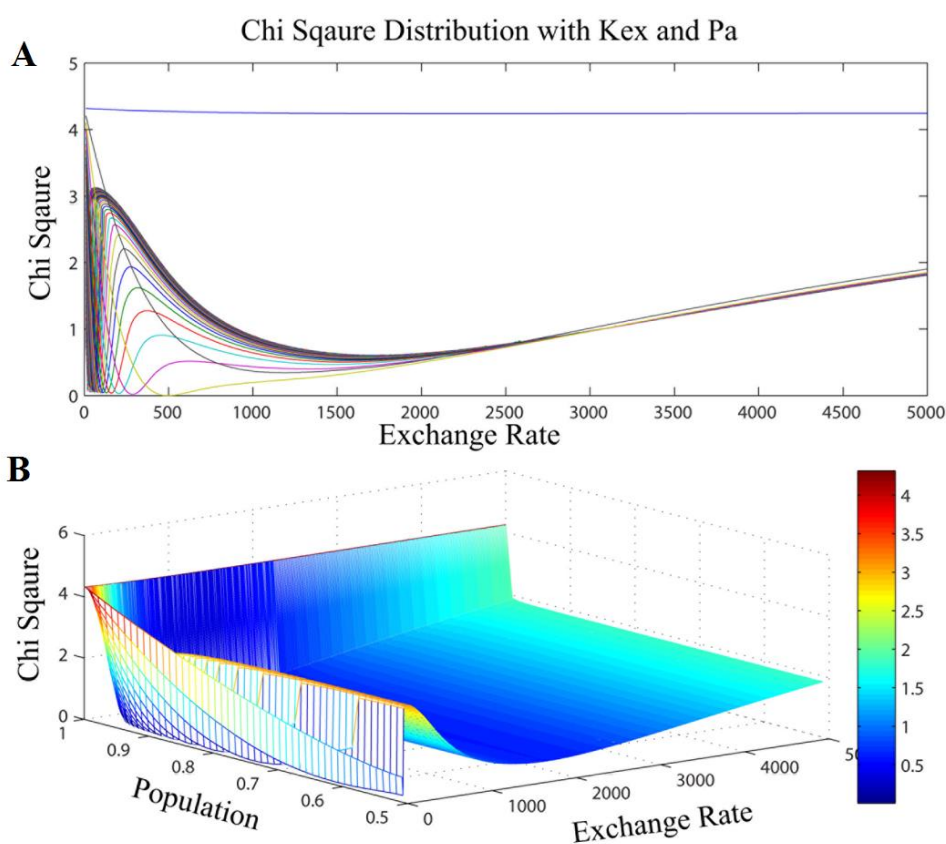
$$\chi^2 = \sum \frac{(re_{1i} - rin_{1i})^2}{rex_{1i}} + \sum \frac{(re_{2i} - rin_{2i})^2}{rex_{2i}}$$

$$rex_{1i} = R_{2i}^{eff} (v_{cp}=40Hz, \omega_H=500MHz) - R_{2i}^{eff} (v_{cp}=960Hz, \omega_H=500MHz)$$

$$rex_{2i} = R_{2i}^{eff} (v_{cp}=40Hz, \omega_H=800MHz) - R_{2i}^{eff} (v_{cp}=960Hz, \omega_H=800MHz)$$

respectively to examine the effect of weighing factors. In which re_{1i} , re_{2i} and

rin_{1i} and rin_{2i} are the calculated and experimental transverse relaxation rates in the 500 and 800 MHz fields, respectively. std_{1i} and std_{2i} are the standard deviation of the i_{th} residue in the two static fields while rex represents the difference of the apparent transverse relaxation rate when v_{cp} in the CPMG experiments is set at minimum and maximum.



$K_{ex}=500 /s$, $d\omega=1000$ rad/s, $P_a=0.98$

Figure 3.4.3.1 Parameter space for a single residue in two-state fitting. The theoretical data was generated based on a set of input value: $p_N=0.98$, $\delta_\omega=1000$ rad/s, $k_{ex}=500$ s⁻¹. (A) Multiple local minima along the k_{ex} dimension. (B) Parameter hyper surface projected onto two dimensions, k_{ex} and p_N .

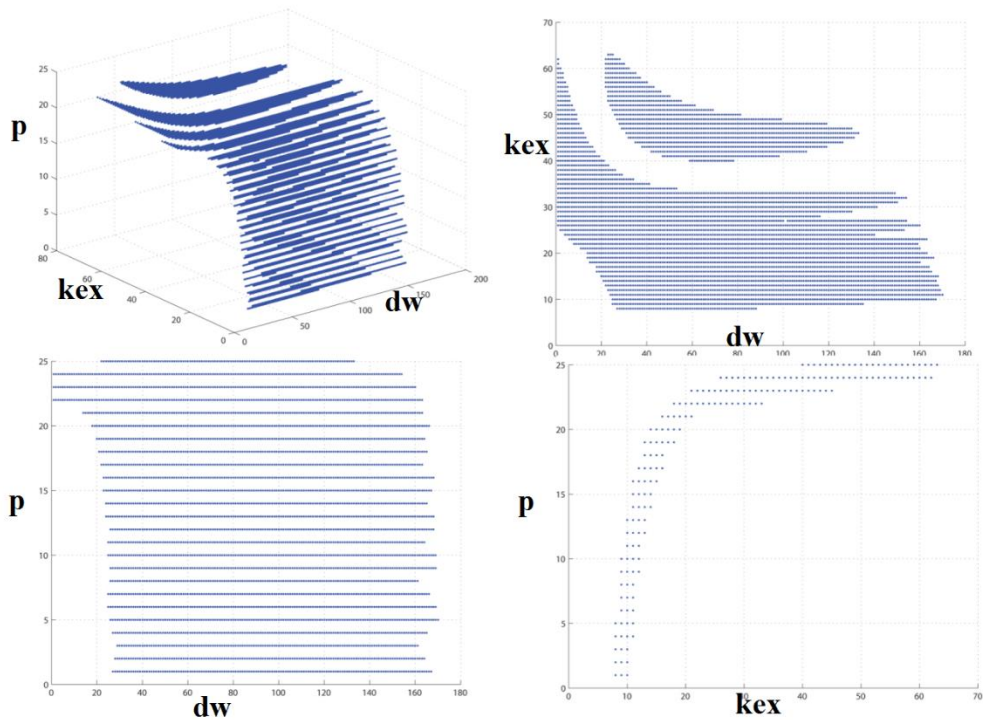


Figure 3.4.3.2 Parameter space output projection on p , k_{ex} and d_w dimension. The range of each parameter is indicated by blue dots, areas covered by blue dots represent range of the coupled parameters. The distance between adjacent dots depends on the spacing of the corresponding dimension in the grid search. Theoretically, the parameter space is continuous.

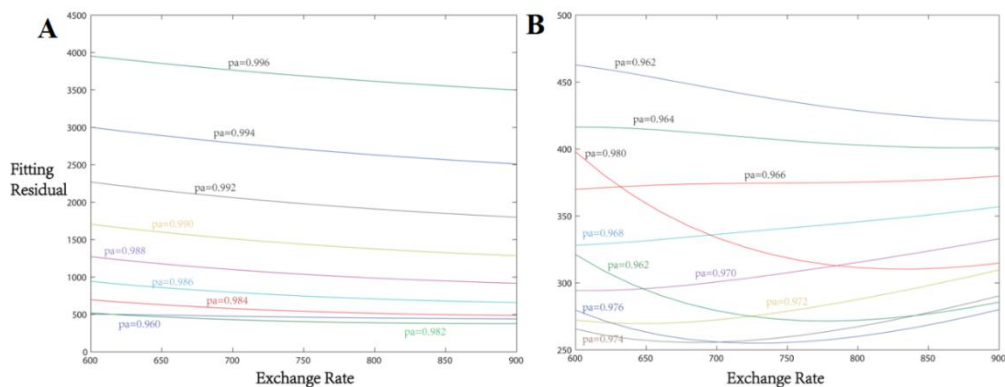


Figure 3.4.3.3 Grid search minimization in global fitting for apo-hIFABP. (A) Grid search output in stage one, fractional population of the major form was probed by examining the fitting residual based on initial values from individual fitting. (B) In stage two, fractional population of the major state was located between 97.4% and 97.6%.

Grid search aided minimization for parameter space mapping was performed in a manner described above. In the initial run, p_N was set between 0.7 and 1.0, with a spacing of 0.005; δ_ω was set between 20 and 2000 rad/s, with a spacing of 20 rad/s; k_{ex} was set between 0 and 10000 /s with a spacing of 100 /s; the $R_{1(2)}$ was set between $R_{1(2)}^0 - 1$ and $R_{1(2)}^0 + 1$ with a spacing of 0.2. After the initial run, kinetic parameters were limited to: $0.76 < p_N < 0.99$, $80 < \delta_\omega < 1480$ rad/s, $200 < k_{ex} < 1790$ /s and $R_{1(2)}^0 - 0.1 < R_{1(2)} < R_{1(2)}^0 + 0.1$. This process went on until p_N was limited between 0.974 and 0.976 (Figure 3.4.3.3) and the final minimization was performed with nonlinear least square fitting.

3.4.3.2 Global fitting result for two-state exchange model

After the grid search, parameters were limited in a relatively small range. The global fitting was then carried on with the output of the parameter space mapping as initial values. For the selected 108 residues, the resultant k_{ex} was 708.5 s^{-1} and the fractional population for the major form is 97.51%.

However, the fitting residual for the global fitting was not significantly increased compared to the total chi square of individual fittings (253.24 for global fitting and 166.79 for overall individual fittings). As shown in figure 3.4.2.1, the k_{ex} of many residues from individual fitting are quite far from the k_{ex} extracted from global fitting. Moreover, the errors of these residues were very small according to Monte Carlo simulations. This result proved that for CPMG data in a two-state exchange model, the target function had a very low sensitivity to the change of the parameters, indicating high flexibility of the

fitting, even for global analysis. In order to check whether a third conformation exists, we thus performed ^{15}N CEST and CPMG.

3.5 CEST data analysis

^{15}N CEST experiments were performed with ~ 2.5 mM, pH 7.0 hIFABP sample on a 800 MHz spectrometer at 307K as described in section 2.1.12. Peak intensities were extracted by nmrPipe software package and were calculated as the peak intensity of each residue divided by the corresponding peak height in the reference spectrum. 134 assigned residues, including four residues on the tag (V-4, G-3, T-2, Q-1) and two sidechain ^{15}N (W6 and W82) were selected for the initial analysis.

In the CEST data, 76 residues have only one major dip in the profile among which 10 display broad dips, 53 residues have an additional dip while residue E51, N71, F2, S4, and sidechain W6N_{ε1} show two minor dips (for E51 and N71, their broad and sharp dips convolute). Among the selected residues, two residues (F2 and S4) show obvious two minor dips, which correspond to two intermediates that exchange with the major form at a fast and slow exchange rates, respectively (Figure 3.5.1). This proves directly that the hIFABP, in the absence of ligands, exists in a more than two states' equilibrium.

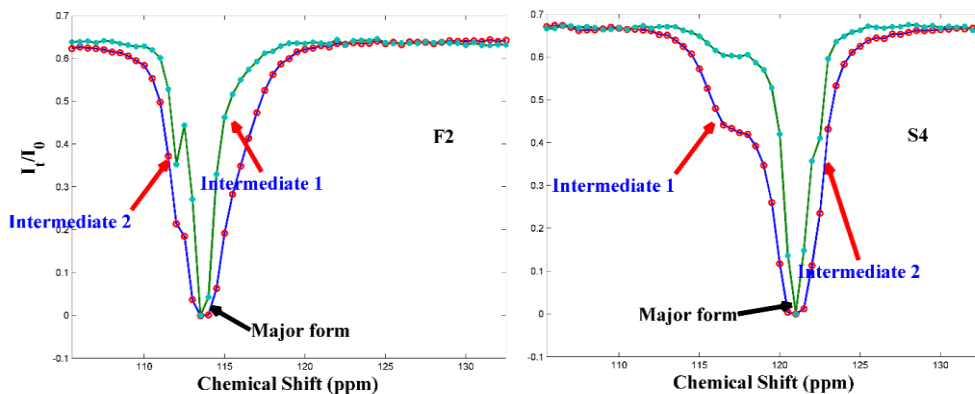


Figure 3.5.1 Representative residues that show obvious two minor dips. Cyan asterisks and red circles represent data recorded in the 30 Hz and 15 Hz weak rf fields, respectively. Solid lines are for visual inspection of the dips and are not for fittings. The location of the minor dips are indicated by the red arrows. Intermediate 1 represents a distinct conformation that interconverts with the major state with an exchange rate at several hundred s^{-1} ($>500 s^{-1}$). Intermediate 2 stands for an ensemble that exchanges with the major form with a rate constant at around $100 s^{-1}$. The CEST profiles provide direct experimental evidence that the apo-hIFABP has more than two states in equilibrium.

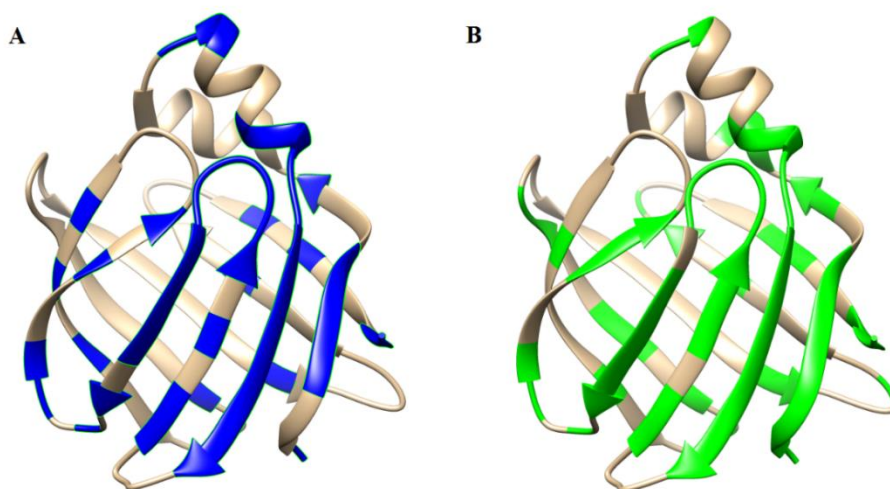
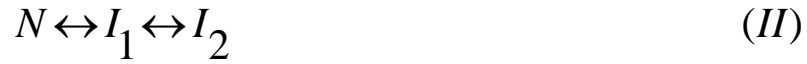


Figure 3.5.2 Residues with obvious chemical exchange effect plotted on ribbon structure of apo-hIFABP. (A) Residues that have minor dip in CEST data. (B) Residues with $R_{ex} > 3 s^{-1}$ in CPMG data measured in 18.8 T static magnetic field.

The 53 residues with obvious minor dips distribute across the whole protein, indicating the global feature of a slower exchange process. In addition, 83 and 93 residues have R_{ex} larger than $2 s^{-1}$ for the RD data at 500 MHz and 800 MHz

magnetic fields, suggesting the existence of a faster exchange process and more importantly, its global nature (Figure 3.5.2).

Since CEST experiment is sensitive to slow exchange processes and can probe chemical shifts of the minor states accurately. We chose to do an initial analysis on the CEST data first so that the chemical shifts can be limited in a relatively small range in the combined analysis of CEST and CPMG data later on. Three-state exchange models were proposed for apo-hIFABP:



In the three-site exchange model, the magnetization evolution of one ^{15}N spin can be described by:

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_{21} + kk_1 & -\delta\omega_1 & \omega_y & -k_{21} & 0 & 0 & -k_{31} & 0 & 0 \\ 0 & \delta\omega_1 & R_{21} + kk_1 & -\omega_x & 0 & -k_{21} & 0 & 0 & -k_{31} & 0 \\ -2 * M_{10} * R_{11} & -\omega_y & \omega_x & R_{11} + kk_1 & 0 & 0 & -k_{21} & 0 & 0 & -k_{31} \\ 0 & -k_{12} & 0 & 0 & R_{22} + kk_2 & -\delta\omega_2 & \omega_y & -k_{32} & 0 & 0 \\ 0 & 0 & -k_{12} & 0 & \delta\omega_2 & R_{22} + kk_2 & -\omega_x & 0 & -k_{32} & 0 \\ -2 * M_{20} * R_{12} & 0 & 0 & -k_{12} & -\omega_y & \omega_x & R_{12} + kk_2 & 0 & 0 & -k_{32} \\ 0 & -k_{13} & 0 & 0 & -k_{23} & 0 & 0 & R_{23} + kk_3 & -\delta\omega_3 & \omega_y \\ 0 & 0 & -k_{13} & 0 & 0 & -k_{23} & 0 & \delta\omega_3 & R_{23} + kk_3 & -\omega_x \\ -2 * M_{30} * R_{13} & 0 & 0 & -k_{13} & 0 & 0 & -k_{23} & -\omega_y & \omega_x & R_{13} + kk_3 \end{bmatrix} \quad (3.11)$$

In which M_{10} , M_{20} and M_{30} are the respective initial magnetizations of the three species at time 0; k_{ij} is the exchange rate between state i and j ; $kk_i = \sum(k_{ij})$ and $i=1, 2, 3, j=1, 2, 3, i \neq j$; R_{2i} and R_{1i} are the intrinsic transverse and longitudinal relaxation rates of the species i ; $\delta\omega_i$ is the ^{15}N chemical shift difference between species i and the major form.

The magnetization of the system at time t is calculated by:

$$\text{Magnetization} = \exp m(-P \bullet t) \times M_0$$

$$M_t = \text{Magnetization}(jk,1) \tag{3.12}$$

Where M_0 is the magnetization of all forms at time 0 and equal the fractional population of each distinct macrostate; jk is the position of the magnetization along z axis for a specific species in the evolution matrix P.

The above three-state models were tested with the CEST data. The results showed that the 3-state model 1 and 2 could explain our data very well while model 3 failed for most of the residues (Figure 3.5.3).

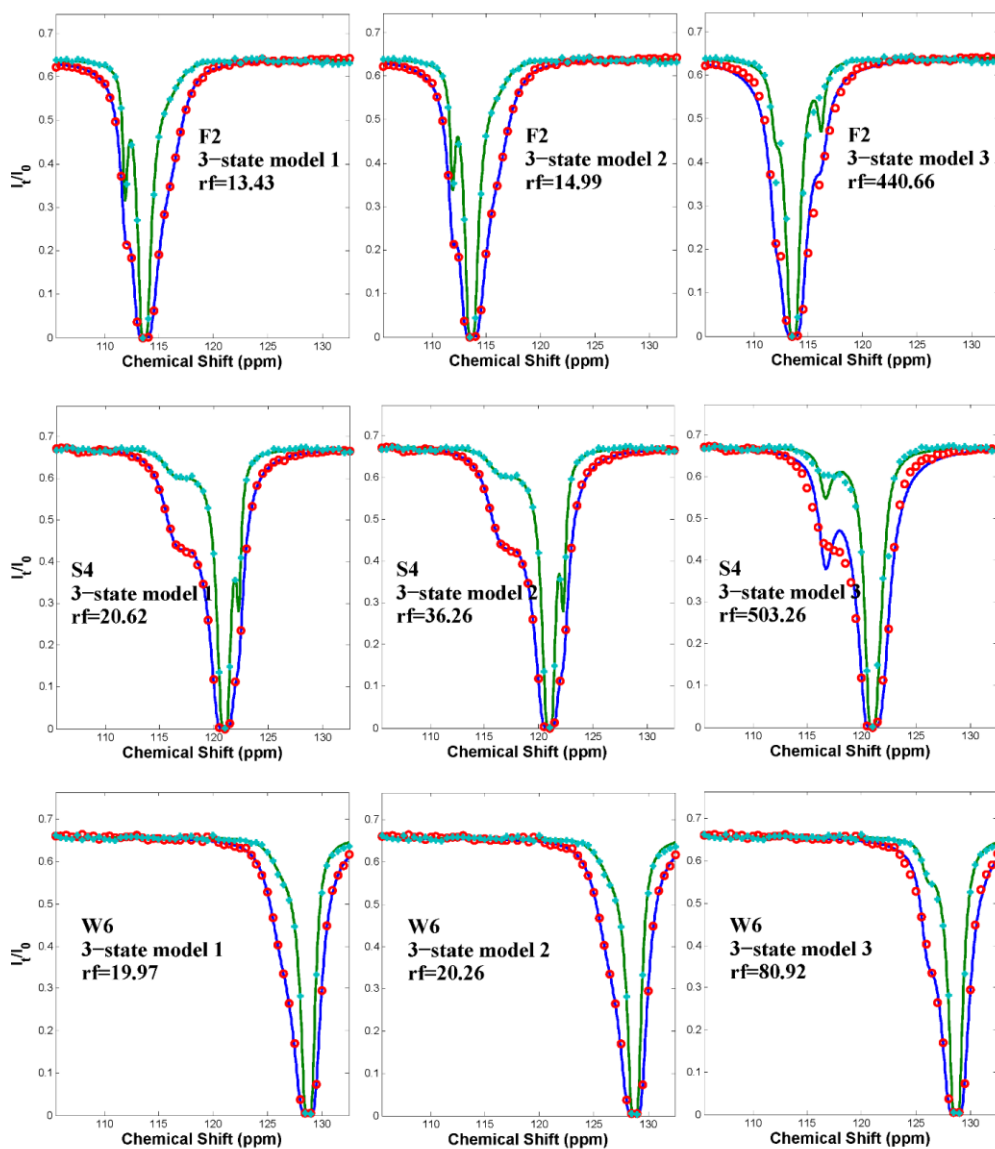


Figure 3.5.3 Representative residues for the three-state model tests. Cyan asterisks and red circles represent data recorded in the 30 Hz and 15 Hz weak rf fields, respectively. Solid lines are the best fits. Clearly, model 1 and 2 explained the data well while model 3 failed.

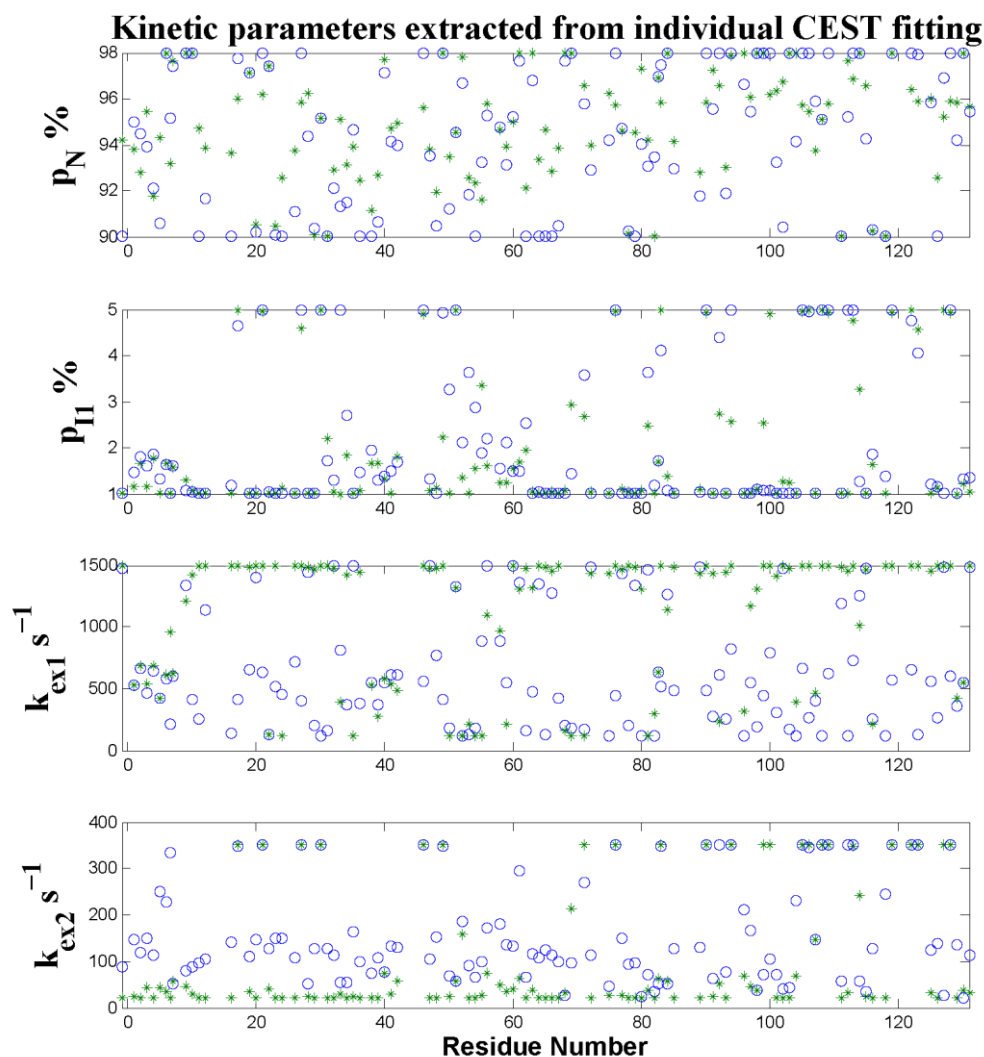


Figure 3.5.4 Kinetic parameters extracted from 3-state individual fitting with CEST data. In the fitting, lower and upper boundaries were set as: $0.90 \leq p_N \leq 0.98$, $0.01 \leq p_{II} \leq 0.05$, $120 \text{ s}^{-1} \leq k_{ex1} \leq 1500 \text{ s}^{-1}$, $20 \text{ s}^{-1} \leq k_{ex2} \leq 350 \text{ s}^{-1}$. Green asterisks and blue circles represent for parameters extracted by three-state model 1 and 2, respectively.

Although fitting residuals were relatively small, the extracted parameters from individual analysis distribute in a very broad range (Figure 3.5.4). To test the sensitivity of the CEST data to exchange rates in different ranges, Monte Carlo simulations were carried out on a residue basis. As can be seen from figure 3.5.5, for CEST data, only residues with obvious broad dips are sensitive to fast exchanges.

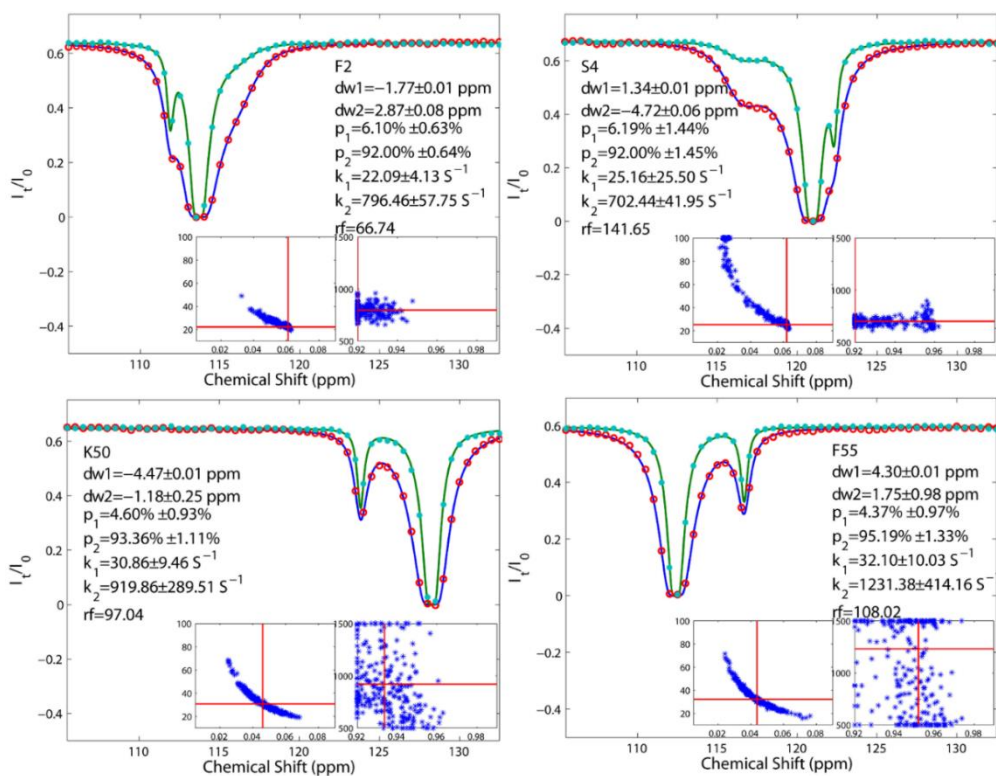


Figure 3.5.5 Sensitivity test for CEST data to fast and slow exchange processes using three-state model 1. For CEST data, only residues with obvious broad dips are sensitive to fast exchange. The subplots are the distribution of kinetics parameters from Monte Carlo simulations, the cross point of red line is the best fit of the original data. 250 sets of data was included in the MC simulation for each residue, and 2% of random error was injected.

Since most residues in apo-hIFABP only have sharp minor dips which correspond to the slower exchange process, the faster conformational rearrangement could not be characterized accurately by CEST data alone. This should also be true for studies using other proteins because for the fast exchange to be detected accurately by CEST experiments, the chemical shifts of the corresponding minor states relative to the major state should be large (typically, >1.5 ppm), and usually there will only be a very small number of residues that meet this condition. Therefore, the best way to probe the above mentioned two exchange processes accurately is through combined analysis of the CEST and CPMG data. Additionally, it is noteworthy that an initial analysis

on CEST data is always preferable since chemical shift plays a vital role in both CEST and CPMG data fitting. By accurately limiting δ_ω in a relatively small range, the minimization process could be substantially faster and hence saving much computation time.

As can be seen in figure 3.5.4, kinetic parameters extracted from individual fittings differ in a broad range. In this case, a global analysis was performed in order to obtain converged initial values for the subsequent combined fitting. Eleven residues whose minor dips are separated from the major dips by more than 2 ppm and one residue with obvious two minor dips were selected and subjected to the global fitting. The resultant global kinetic parameters were $p_{I1}=1.83\%$, $p_N=92.96\%$, $k_{ex1}=27\text{ S}^{-1}$, $k_{ex2}=720\text{ S}^{-1}$, $\chi^2=332$ for three-state model 1 (Figure 3.5.6) and $p_{I1}=1.94\%$, $p_N=93.16\%$, $k_{ex1}=105\text{ S}^{-1}$, $k_{ex2}=672\text{ S}^{-1}$, $\chi^2=378$ for three-state model 2. In the fitting, the intrinsic longitudinal relaxation rates for state N, I₁ or I₂ were considered equal since they have little effect on the result as long as $0.1 < R_1 < 4\text{ s}^{-1}$ (Bouvignies and Kay 2012), while the transverse relaxation rates R_{2i} for the three states were also considered the same. When R_{2i} values were not fixed to the same, the obtained parameters were $p_{I1}=1.86\%$, $p_N=93.05\%$, $k_{ex1}=27\text{ s}^{-1}$, $k_{ex2}=722\text{ s}^{-1}$, $\chi^2=309$ for three-state model 1 and $p_{I1}=1.97\%$, $p_N=93.16\%$, $k_{ex1}=103\text{ s}^{-1}$, $k_{ex2}=688\text{ s}^{-1}$, $\chi^2=359$ for three-state model 2. The minimization favors the latter one, yet it yielded a substantial increase in the degree of freedom to the fitting of the RD data, we thus chose to treat R_{2i} values of the three states as equal. The rest of the residues were then fitted with the global parameters fixed to extract the residue-specific parameters (δ_ω , R_{10} , R_{20}) according to the corresponding exchange models, respectively.

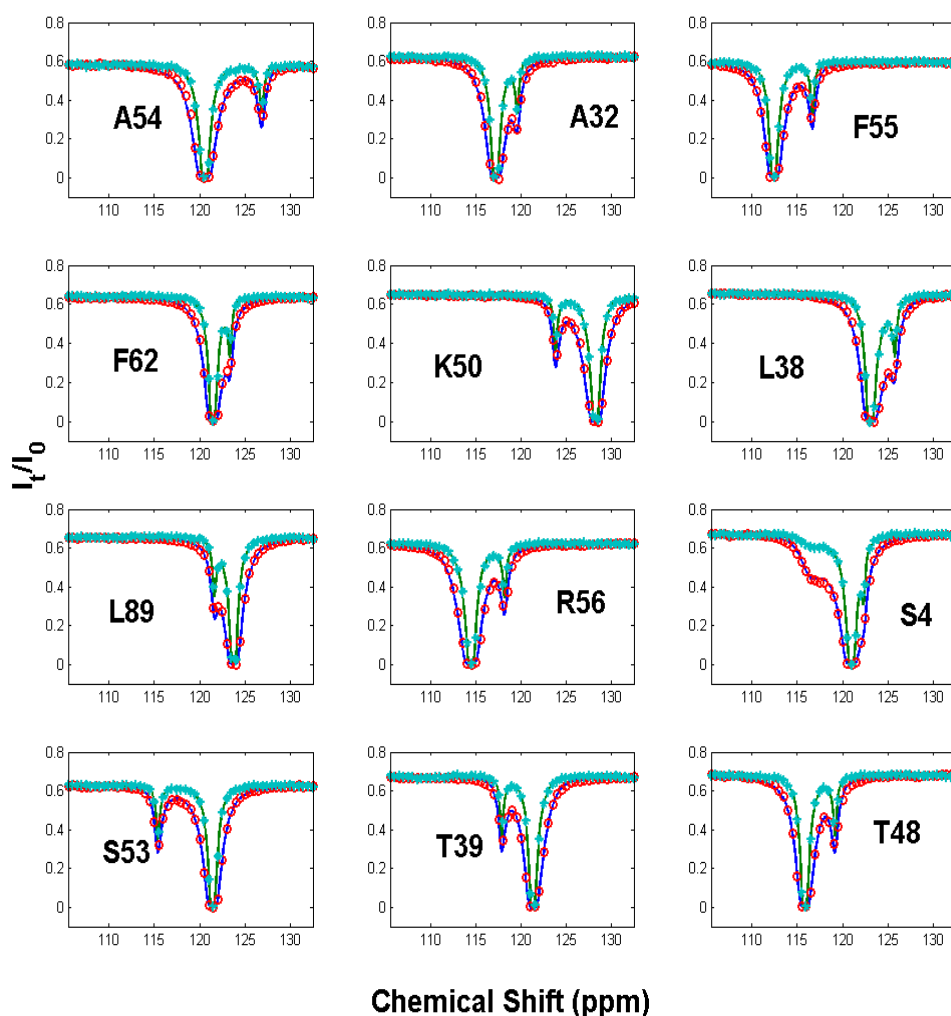
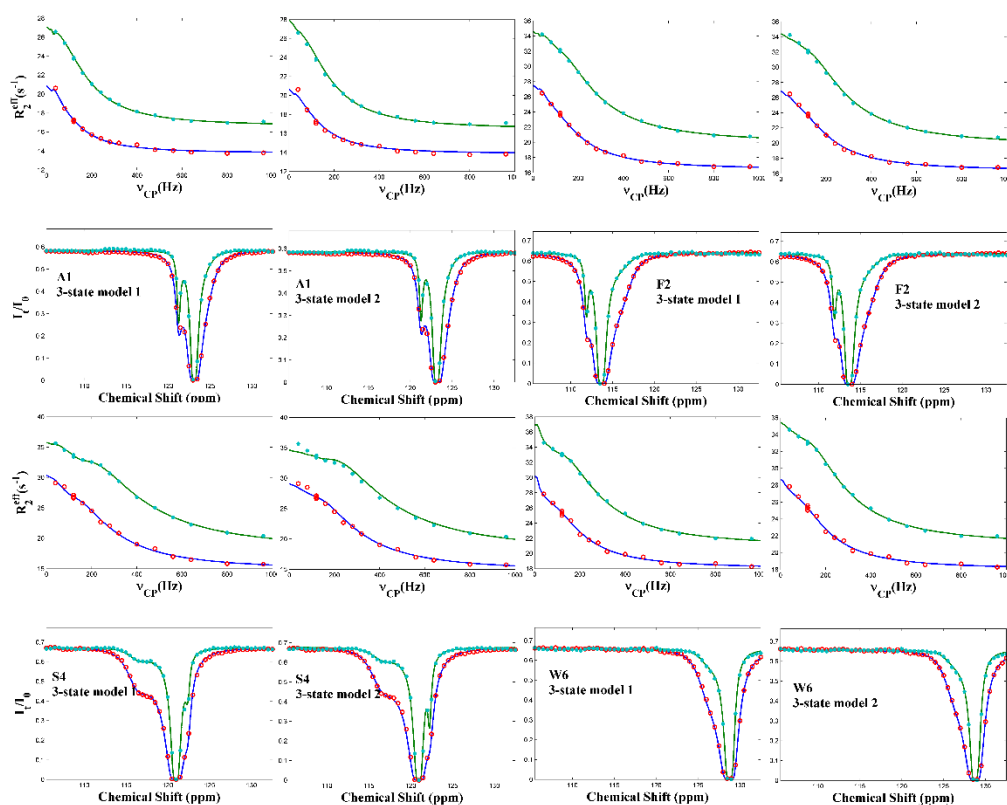


Figure 3.5.6 Global fitting of CEST data with three-state model I. 11 residues that have larger than 2 ppm and one residue with obvious two minor dips were selected and subjected to 3-state global fitting. The resultant global kinetic parameters $p_H=1.83\%$, $p_N=92.96\%$, $k_{ex1}=27\text{ s}^{-1}$, $k_{ex2}=720\text{ s}^{-1}$, were used to extract residue-specific parameters (e.g., δ_ω , R_1 and R_2). These obtained values were then used in the initial estimation for the combined analysis of CEST&CPMG data.

3.6 Combined analysis of CEST and CPMG data

For the RD data, nearly all residues have observable R_{ex} . It is obvious that one ought to choose residues only with sizeable R_{ex} and small random errors for data analysis, especially when RD data is the sole source of information. With

CEST data, however, we now can enhance the applicability of RD data by limiting the $\delta\omega$ in a very small range and accurately locate the exchange rate in the range of 20-200 s^{-1} , should there be any. In this scenario, all the 111 sets of the RD data together with the corresponding CEST data were subjected to residue-specific analysis. Our result showed that most of the residues could be well explained by both the three-state model I and II with $k_{ex1} \sim 50-60 s^{-1}$ and $k_{ex2} \sim 600-800 s^{-1}$ (Figure 3.6.1).



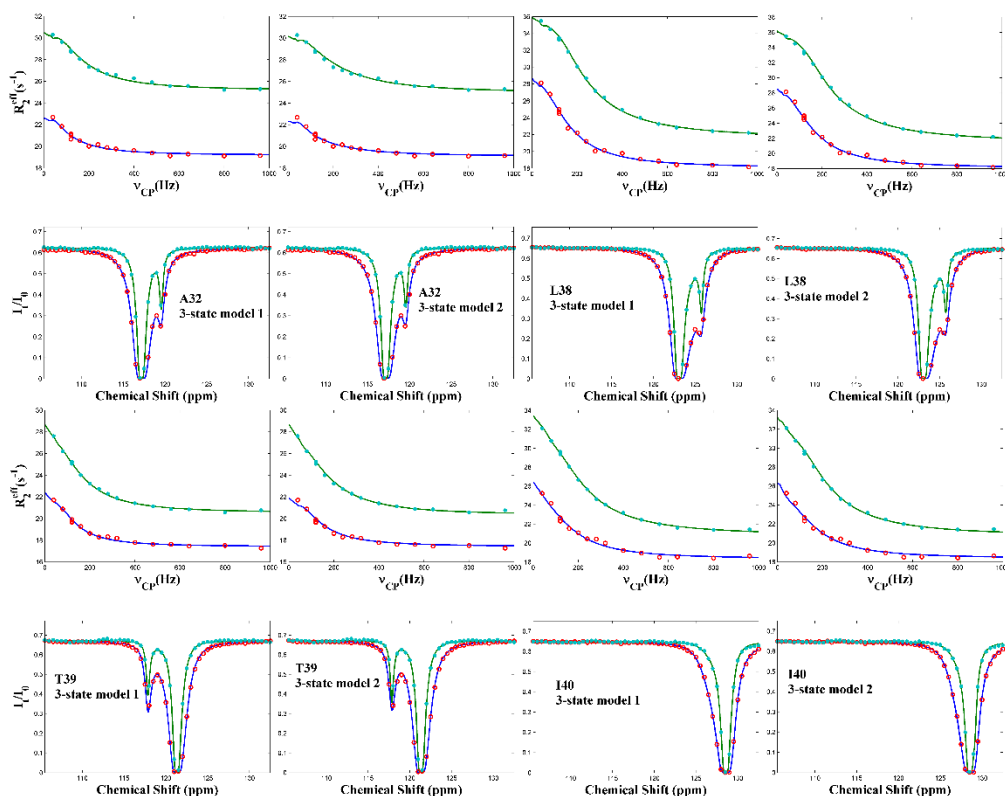


Figure 3.6.1 Representative residues fitted individually with three-state model I and II by combined analysis of CEST and CPMG data simultaneously. For CPMG data, green asterisks and red circles represent transverse relaxation rates measured in 500MHz and 800MHz magnetic fields, respectively. For CEST data, green asterisks are the peak intensities in 30 Hz rf field while red circles represent peak intensities measured in 15 Hz rf field, both in 18.8T static magnetic field. Solid lines are the best fits. Fittings of model 1 are on the left while those of model 2 are on the right for each residue.

However, a significant number of residues (T48, V49, S53, A54, F55, E59, V60, V61, L64, G65, V66) could be fitted well only with $k_{ex2} > 2000 \text{ s}^{-1}$, which is much larger than the exchange rates of other residues (Figure 3.6.2). As can be seen from the fitting plots, when k_{ex1} is limited below 900 s^{-1} , obvious deviations from experimental data were observed (Figure 3.6.3). Interestingly, these residues are predominantly located at β -strands C, D, E and the loops between them. Based on previous studies, these areas could be described as the gap region surrounded by β -strands D and E and two loops of the putative portal

region (Figure 3.6.4).

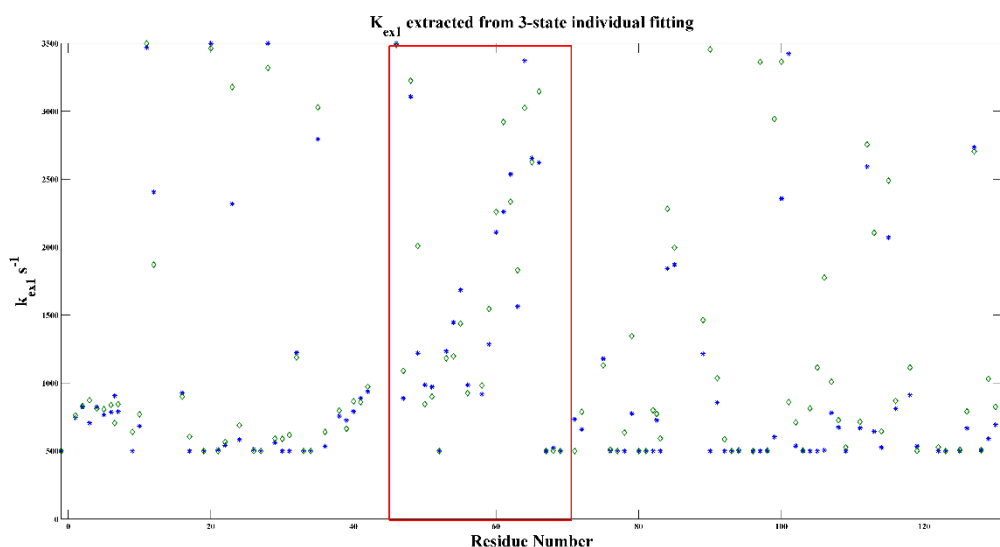
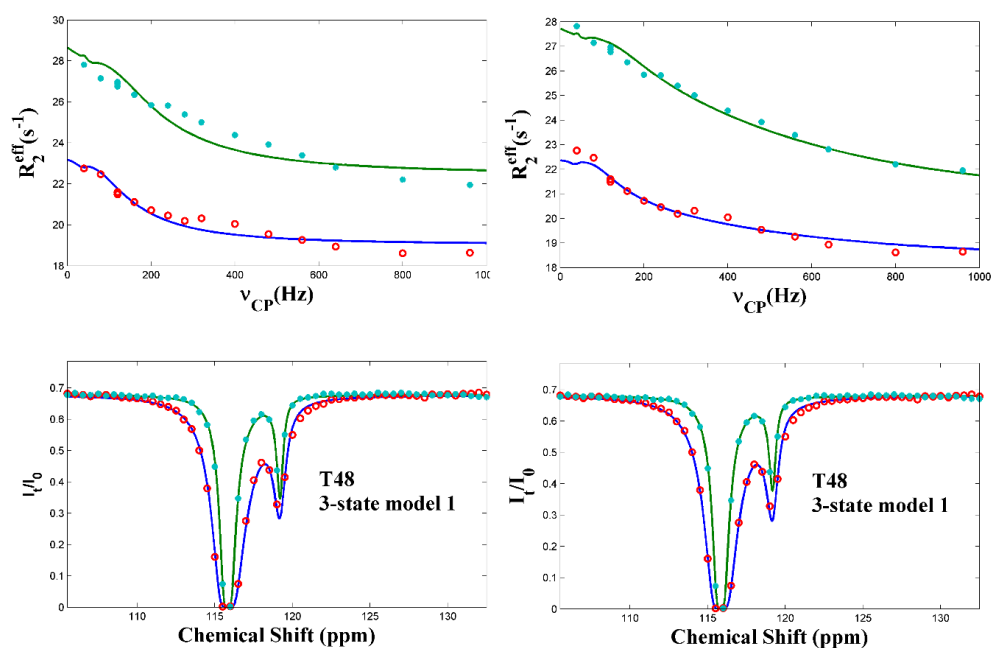
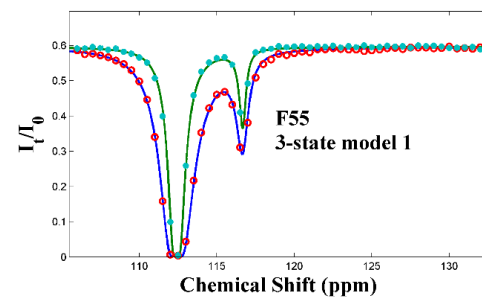
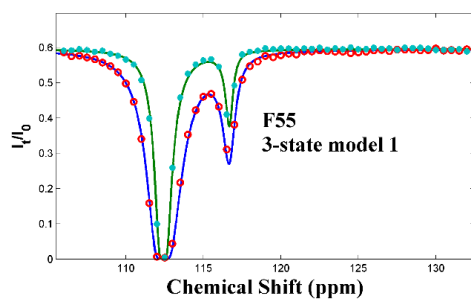
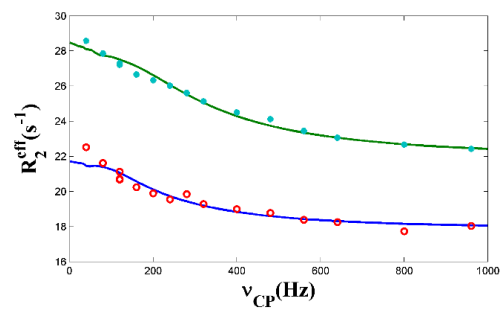
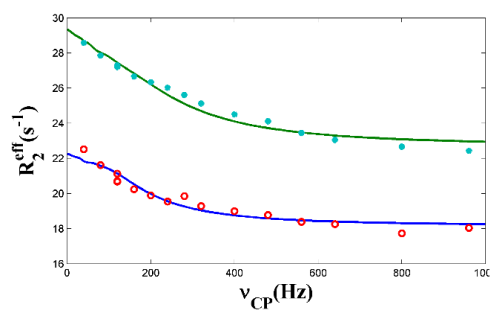
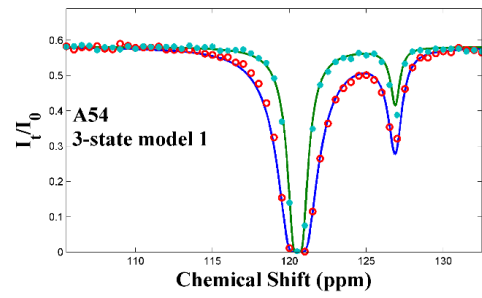
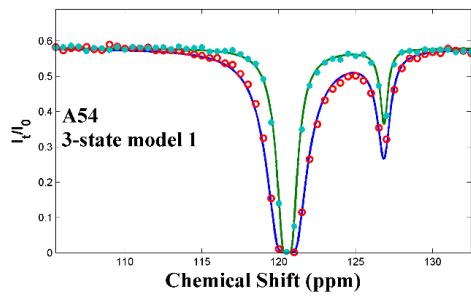
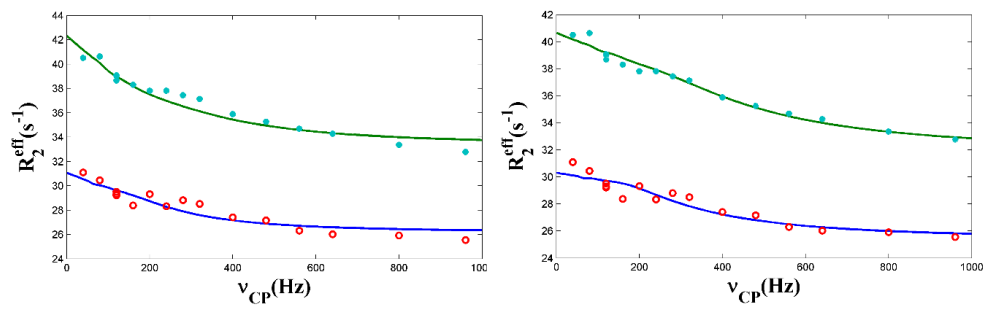
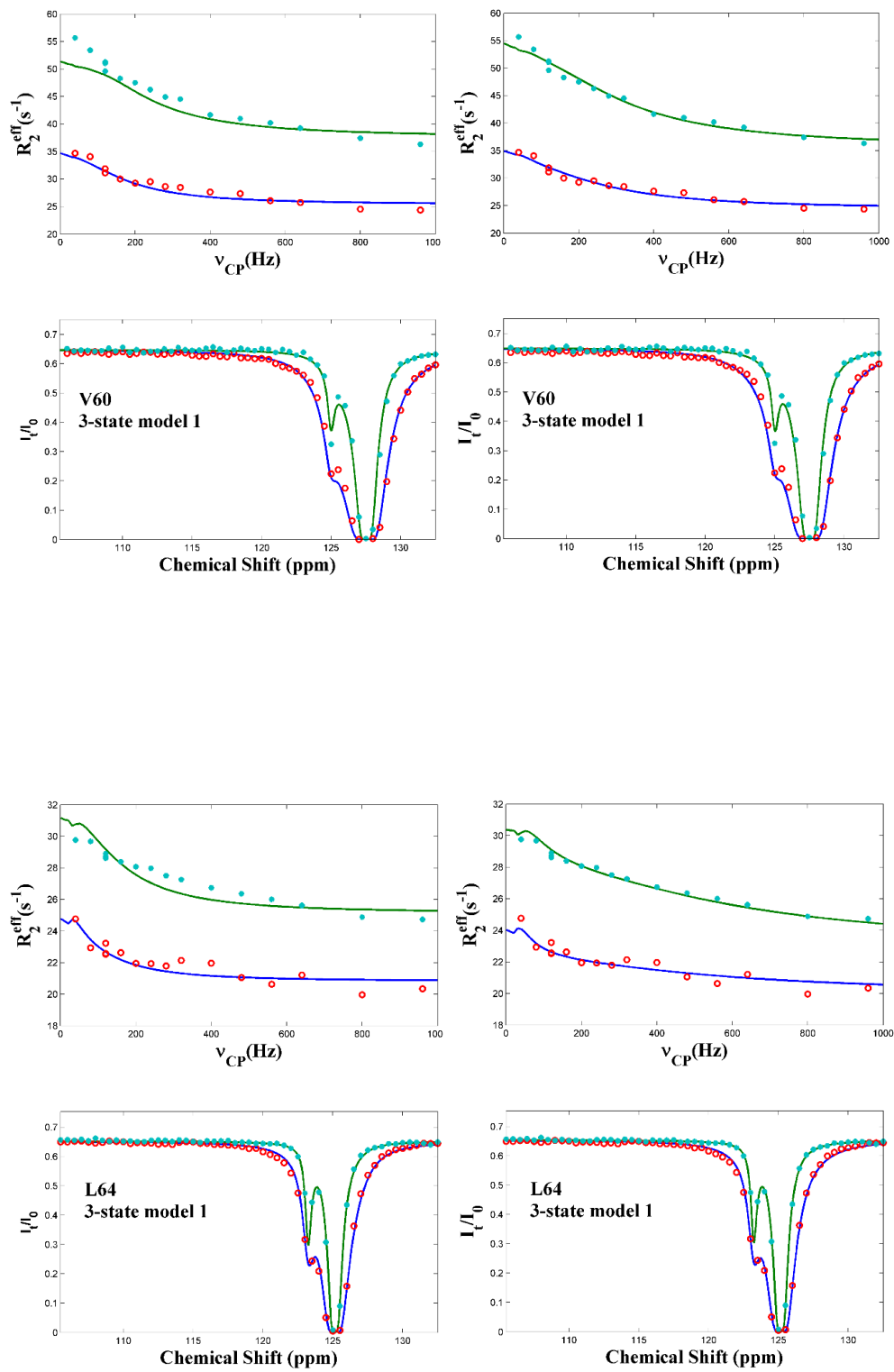


Figure 3.6.2 k_{ex1} distribution of individual fitting by 3-state model I. In the fitting, the exchange rate was limited between 500 s^{-1} to 3500 s^{-1} , and $20 \text{ s}^{-1} \leq k_{ex2} \leq 350 \text{ s}^{-1}$, $0.90 \leq p_N \leq 0.97$, $0.01 \leq p_{II} \leq 0.05$. The red rectangle indicates the area where residues that could be fitted only with $k_{ex1} \geq 2000 \text{ s}^{-1}$. It should be noted that other residues that produced k_{ex1} larger than 1500 s^{-1} could be fitted very well with $k_{ex1} \leq 900 \text{ s}^{-1}$ and hence are significantly different with those in the above mentioned area.







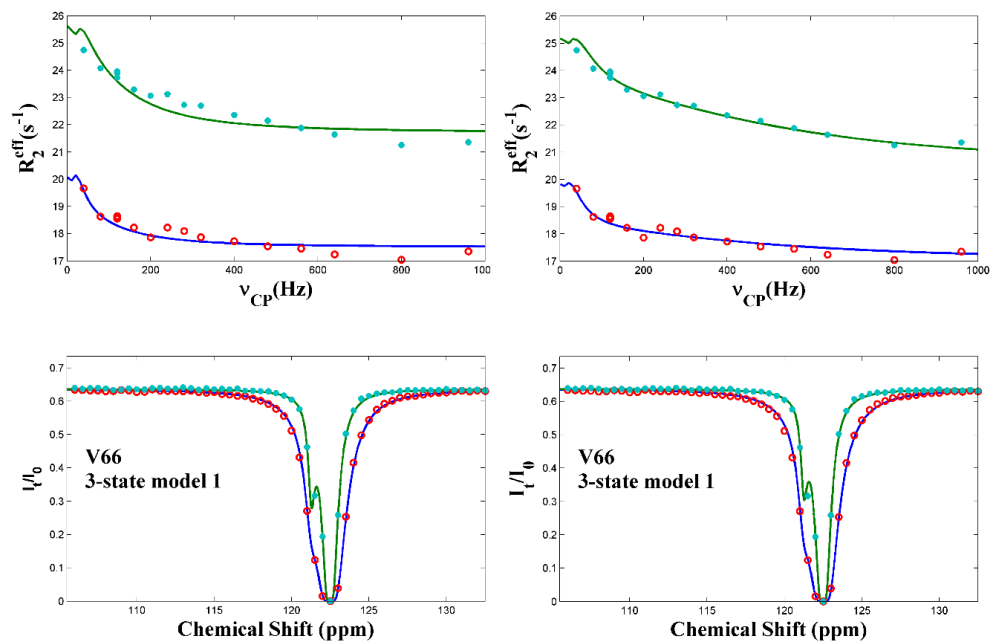


Figure 3.6.3 Representative residues that could be fitted well only with $k_{ex1} > 2000 \text{ s}^{-1}$. For CPMG data, green asterisks and red circles represent transverse relaxation rates measured in 500MHz and 800MHz magnetic fields, respectively. For CEST data, green asterisks are the peak intensities in 30 Hz rf field while red circles represent peak intensities measured in 15 Hz rf field, both in 18.8T static magnetic field. Solid lines are the best fits. For each residue, the plot on the left represents fitting with $k_{ex1} \leq 900 \text{ s}^{-1}$ while plot on the right side shows the fitting with k_{ex1} unlimited. All fittings are for three-state model 1, the results for model 2 were similar.

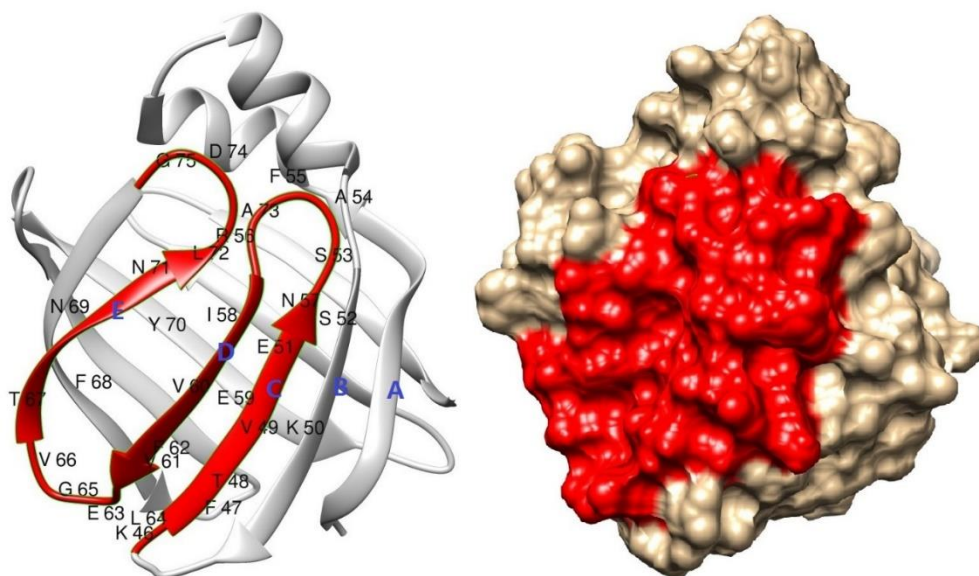


Figure 3.6.4 Gap region of hiFABP. From left to the right is the ribbon representation and the surface plot of apo-hiFABP, respectively. The gap region is highlighted in red. The distance between β -strands D and E is substantially

larger than that of any other two adjacent strands.

The fitting result for the above mentioned residues in the gap region is not a coincidence since we have performed three sets of CPMG experiments with different pulse sequences as well as temperature calibrations and the result is reproducible. Moreover, the faster exchange rate is directly suggested by the ascendant tendency of peak intensities even in the fast limit of the rf fields. i.e., if the fast exchange process experienced by these residues has a rate constant of $\sim 800 \text{ s}^{-1}$, which is the same as that of the other residues, the CPMG profile should become nearly flat when $\nu_{\text{cp}} \geq 800 \text{ Hz}$. On the contrary, the peak intensities of these residues kept increasing during the course of CPMG experiments and the signal recovery did not stop at the highest ν_{cp} . Theoretically, R_{ex} contributed by chemical exchange is proportional to δ_{ω} and k_{ex} . For residues mentioned above, no obvious broad dips were seen in their CEST profiles, therefore the respective chemical shift differences for each ^{15}N spin between the minor state and the major form are small. In this case, signal increase at the highest ν_{cp} indicates that the refocusing pulses used in the CPMG experiments could not suppress the contribution to R_{ex} from chemical exchange processes even in the fast limit of the rf fields. Another possible explanation is that an extra fast conformational rearrangement exists in the region where the 11 residues situated in aside from the two global exchange processes. This could be tested simply by fitting the corresponding data sets to a four-state exchange model with the global parameter (p_{N} , p_{I1} , p_{I2} , k_{ex1} and k_{ex2}) extracted from other residues that are not located in the same region fixed. In addition to the two scenarios described above, the protein might experience multiple

conformational rearrangements on different timescales and some of these motions are much faster than the two exchange processes we have identified and are thus beyond the detection limits of both the CEST and CPMG experiments. In this case, dynamic properties on different timescales should be combined for a genuine conclusion to be drawn. In the following sections, the three hypotheses will be tested thoroughly by robust model selection and comparison of intrinsic transverse relaxation rates using power spectrum density function.

3.7 Model selection for residues with exceptional RD and CEST profiles

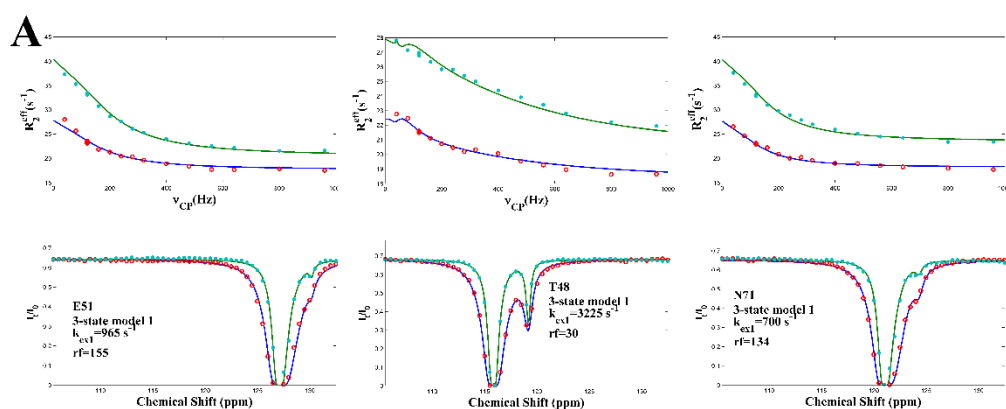
As has been described in the last section, the relaxation dispersion profiles of the residues in the gap region are significantly different with those of residues on other part of the protein. There are three hypotheses for this result:

- 1) Exchange rate for the gap region is much faster than that of the rest of the protein.
- 2) In the presence of the two global exchange processes, there is another fast exchange in the gap region.
- 3) The protein exists in at least a four-state equilibrium, the gap region undergoes multiple conformational rearrangements on different timescales.

3.7.1 Regional collective motions in local area, faster exchange rate based on three-state exchange model

In section 3.6, we showed that the fittings for residues in the gap region were much better when an exchange rate which is two to three folds larger than those

extracted from other residues (Figure 3.6.2 and 3.6.3). Since protein movements, either regional or global, are collective motions. This means residues that are located in the same area, are bound to move in a cooperative way. If the gap region undergoes a faster motion compared to the rest of the protein, residues in this region should be able to be fitted well with a global k_{ex1} . In this scenario, 25 residues in the gap region were subjected to the three-state model 1 and 2 for global analyses. Surprisingly, residue E51 and N71 could not be fitted well together with other residues. As figure 3.7.1 shows, all residues could be individually fitted very well with three-state model 1. However, neither E51 nor N71 could be fitted with other residues in the gap region while other residues in this region could be fitted well simultaneously. The three-state model 2 produced similar result. Therefore, the three-state models are not suitable for the gap region.



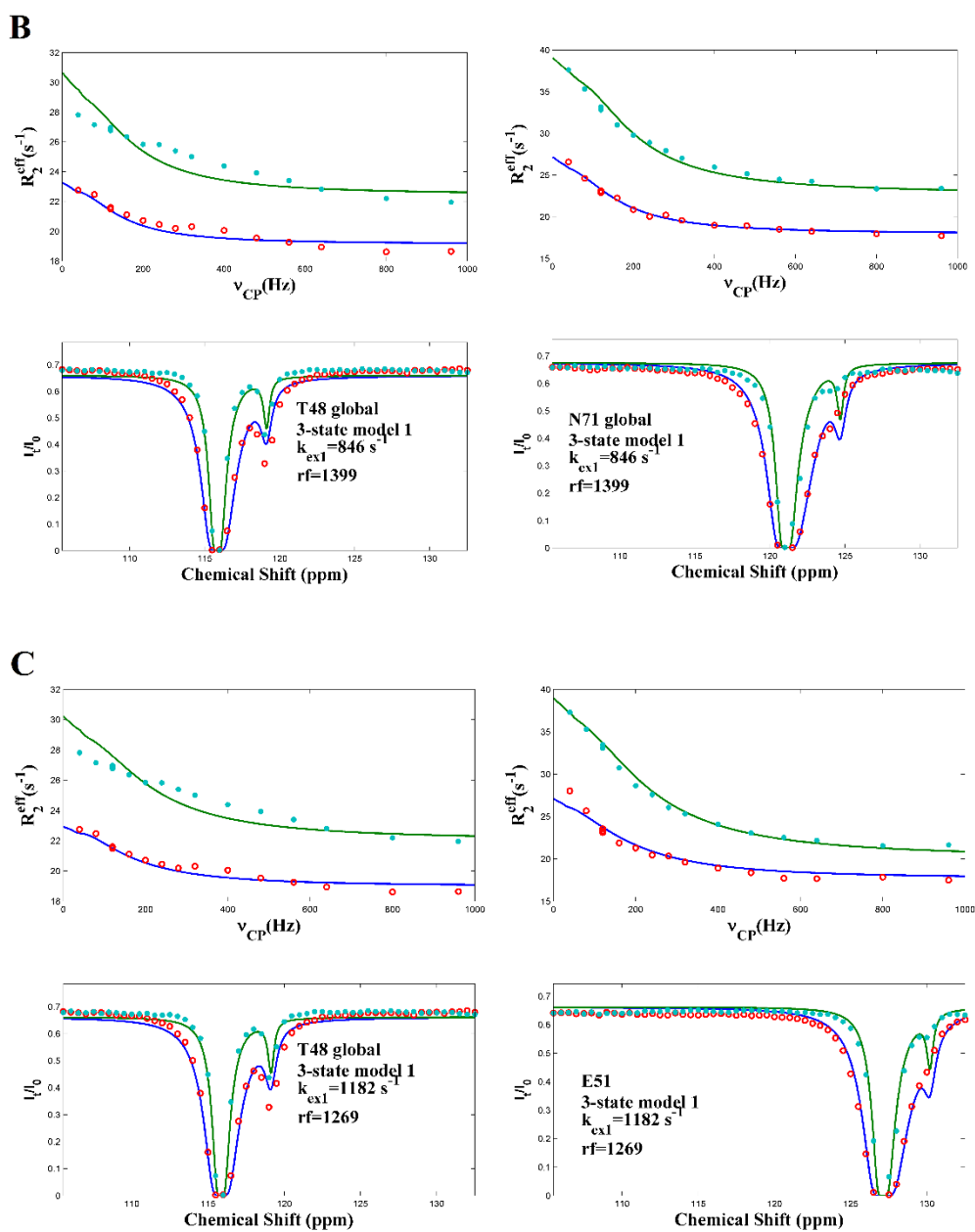


Figure 3.7.1 Three-state model 1 selection for the gap region. (A) Individual fitting of T48, E51 and N71. (B) Fitting of T48 and N71 simultaneously. (C) Fitting of T48 and E51 at the same time. Apparently, the three-state models are not suitable for the gap region. Result for model 2 was similar.

3.7.2 A third exchange process in the presence of the two global motions

This hypothesis could be tested simply by fitting the gap region with the parameters for the two global motions (p_N , p_{I1} , p_{I2} , k_{ex1} and k_{ex2}) fixed. Four-state exchange models on the basis of the three-state model 1 and 2 were proposed:



In which the third intermediate I_3 , interconverts with either one of the existing species.

The magnetization evolution matrix for four-state exchange system was derived similarly with that of the three-state and the effective relaxation rate could be computed according to:

$$A_1 = \text{conj}(A)$$

$$M_t = [\exp m(A \times d) \times \exp m(A_1 \times d) \times \exp m(A_1 \times d) \times \exp m(A \times d)]^n \times M_0$$

$$R_{2c}^{eff} = -\log[\text{real}(M_t(os)) / \text{real}(M_0(os))] / T_{cp}$$

Where R_{2i2} is the intrinsic transverse relaxation rate for the species i ; ω_{i0} is the chemical shift of the species i with the unit rad/s; k_{ij} is the exchange rate constant between species i and j ; A_1 is the conjugate matrix of the evolution matrix A ; $2n$ is the number of the 180° refocusing pulses; d is the delay between two successive 180° pulses; M_0 is the magnetization of the system at time 0; real means the real part of the magnetization for the phase sensitive detection; T_{cp} is the constant time delay and is 50 ms.

The magnetization evolution matrix for a ^{15}N spin in the CEST experiments for a four-state equilibrium is analogous to that of the three-site exchange. Similarly, the result is calculated according to:

$$\text{Magnetization} = \exp m(-P \bullet t) \times M_0$$

$$M_t = \text{Magnetization}(jk,1)$$

In which M_{i0} is initial magnetization for species i ; R_{1i} and R_{2i} is the intrinsic longitudinal and transverse relaxation rates for species i , respectively; k_{ij} is the overall exchange rate between species i and j , where $i \neq j$; other parameters are similar with those in the CPMG evolution matrix A . The magnetization of a species i at time t (M_t) is calculated in which jk represents the position of the longitudinal magnetization of the corresponding species in the evolution matrix P .

The residues in the gap region were then tested with the above four-state exchange models with the global parameters for the pre-existing motions fixed. The result showed that residue E51 and N71 could be explained reasonably well while other residues which produced faster exchange rates in the three-state fittings failed (Figure 3.7.2). The input for four-state model I and IV were: $p_N=94.73\%$, $p_{II}=2.07\%$, $k_{ex1}=807\text{ s}^{-1}$, $k_{ex2}=60\text{ s}^{-1}$; and $p_N=91.52\%$, $p_{II}=2.46\%$, $k_{ex1}=701\text{ s}^{-1}$, $k_{ex2}=95\text{ s}^{-1}$, respectively.

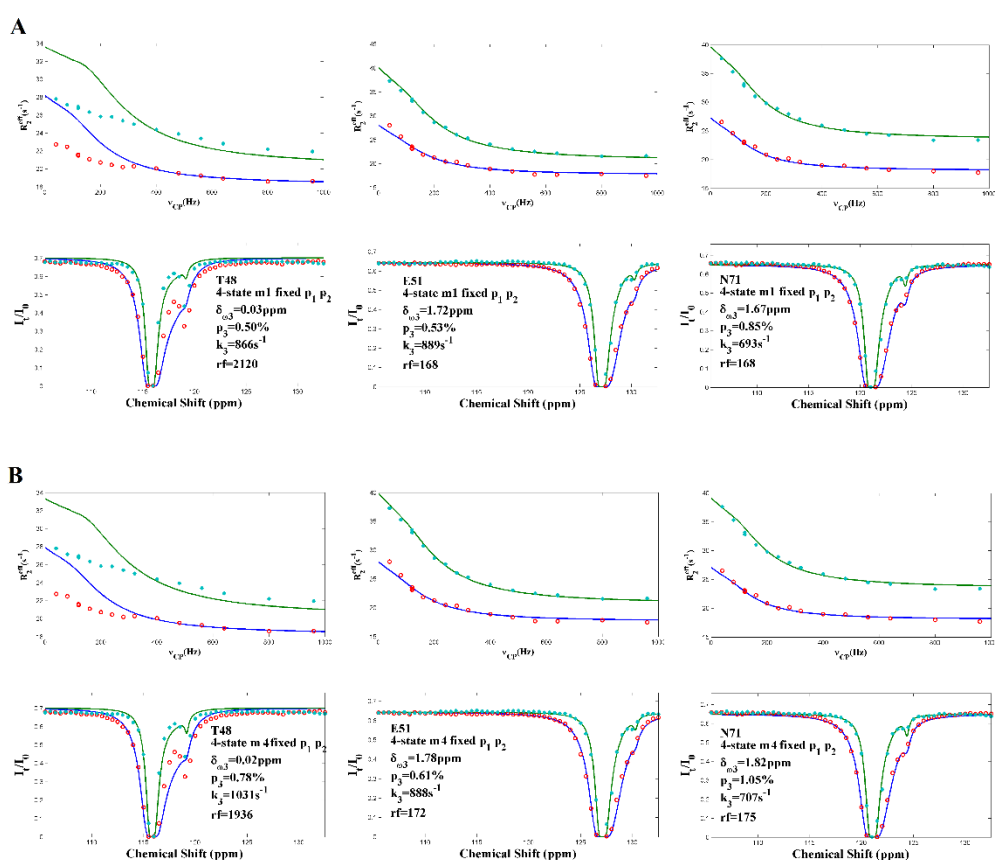


Figure 3.7.2 Representative residues fitted with four-state model when the global parameters for the other two exchange processes are fixed. (A) Fitting with four-state model I for residue T48, E51 and N71. (B) Result for four-state model IV.

3.7.3 Four-site exchange for apo-hIFABP

The four-state exchange models in which all parameters participate in the

minimization were carried out for residues in the gap region. For the six models listed above, 13 residues could be fitted well by all the models, 11 residues could not be explained by model III and model V, 5 residues could not be fitted by model II, 4 residues could not be fitted by model VI while model I and IV explained all residues well (Figure 3.7.3). For the representative residues in the model tests (Figure 3.7.4), differences in fitting quality were very clear according to values of the target functions or simply by visual inspection since discrepancies between data and fitting were self-explanatory.

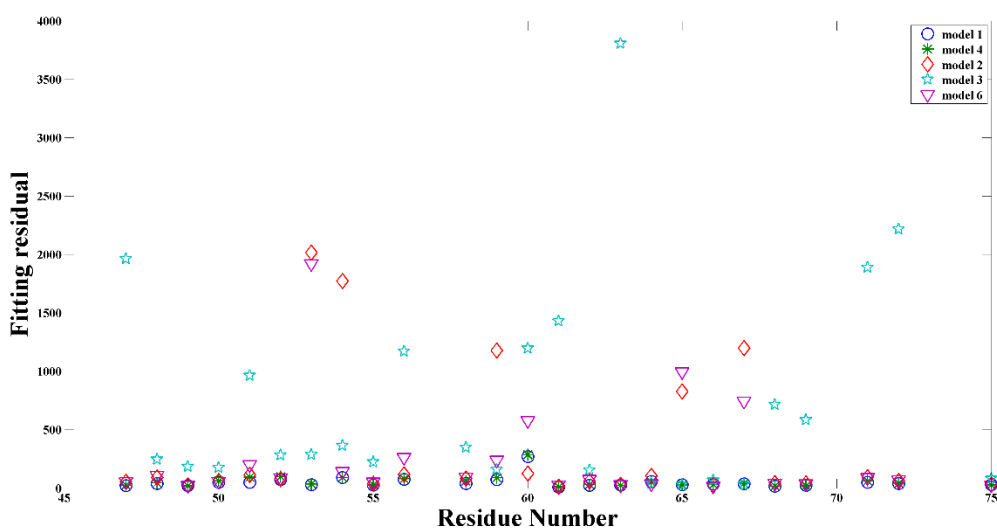
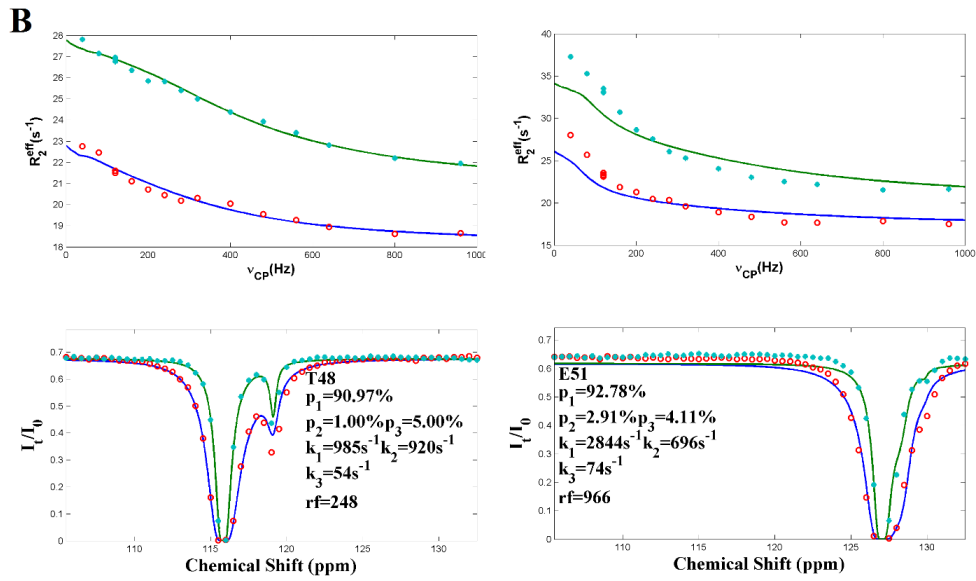
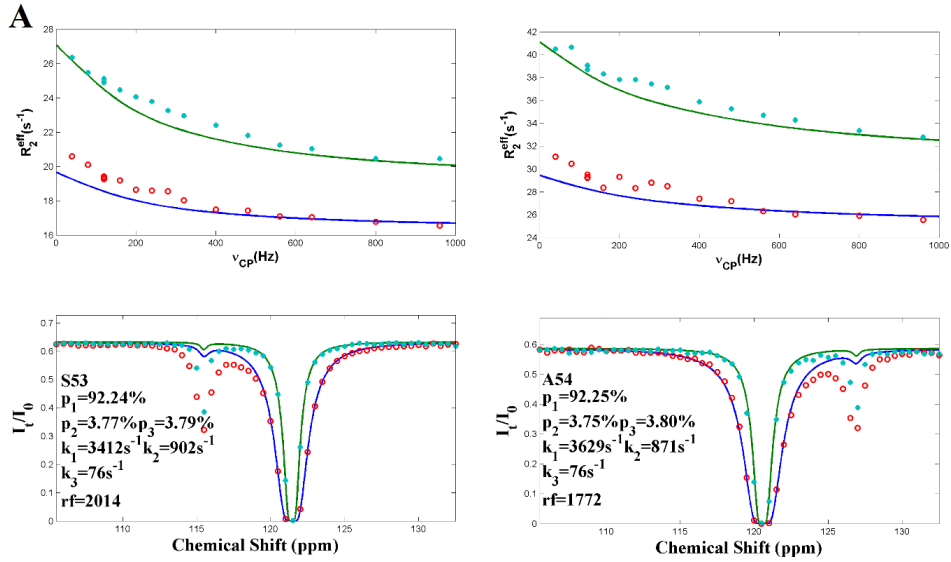


Figure 3.7.3 Fitting residuals for the six four-state models. Model 3 and 5 are equivalent.



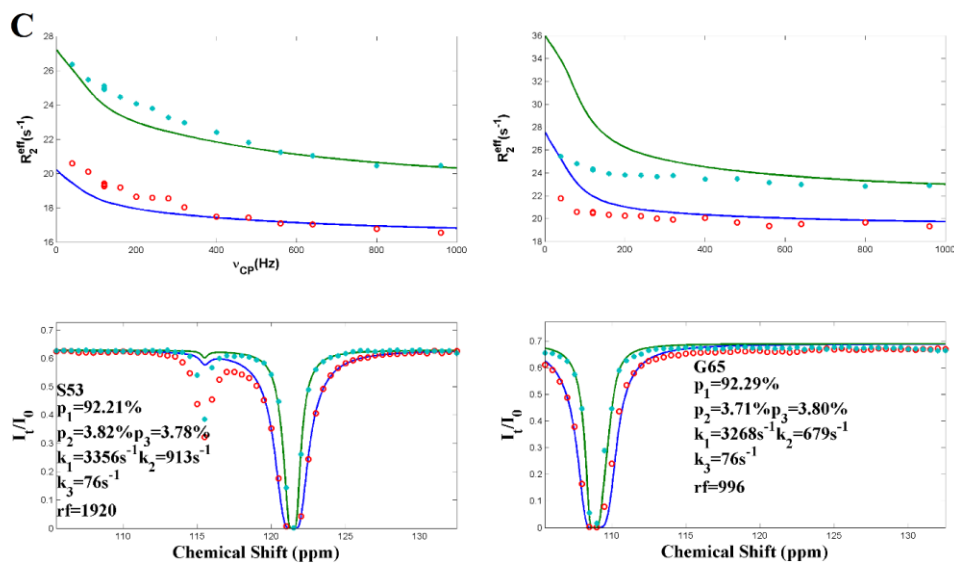


Figure 3.7.4 Representative residues in the four-state model test. (A) Four-state model 2; (B) Four-state model 3; (C) Four-state model 6. Model 3 and model 5 are equivalent.

Significant reductions of χ^2 (more than ten times) were observed for the four-state fittings compared with the three-state fittings. As can be seen in table 3.7.1, for most residues, the improvements of fitting from three-state to four-state model were obvious. Residue F47, K50, S52, E63 and T76 showed no obvious changes in chi squares between the tested models. For residue V60, the corresponding R_2 is 27.9 s^{-1} which is much larger than the average R_2 of the protein. The fitting residues for V60 were always large in any exchange model by CEST and/or CPMG data analysis, therefore it was excluded in the F-test.

Residue	F ratio	α
F47	NA	NA
T48	5.69	0.00
V49	8.64	0.00
K50	NA	NA
E51	111.68	0.00
S52	NA	NA
S53	29.00	0.00
A54	5.93	0.00
F55	38.97	0.00

R56	51.35	0.00
I58	106.13	0.00
E59	13.19	0.00
V60	NA	NA
V61	13.62	0.00
F62	11.45	0.00
E63	NA	NA
L64	5.66	0.00
G65	1.39	0.03
V66	7.45	0.00
T67	NA	NA
F68	61.64	0.00
N69	9.85	0.00
N71	77.28	0.00
L72	10.44	0.00
G75	12.21	0.00

Table 3.7.1 F-test for residues in the gap region between three-state model I and four-state model I. As can be seen, most residues produced much smaller fitting residuals by the four-state model. Residue F47, K50, S52, E63 and T76 have no obvious change in chi square between the tested models. Residue V60 was excluded in the F-test due to large error as described in text.

As described in section 3.7.1 and 3.7.2, residue E51 and N71 are different with other residues in the gap region and could not be fitted with them in global analysis. The corresponding F tests showed that the F factors for both E51 and N71 are larger than 70 (i.e., $\alpha=0$), indicating the four-state exchanging nature of the two residues. As indicated in table 3.7.1, this result also applies to other 17 residues in the gap region. Since E51 is located in the same area with the 11 residues that have substantial larger R_{ex2} , and N71 is on the loop of β -turn D, E which is part of the putative portal region, we believe these residues with exceptional RD and CEST profiles undergo motions that are much different to the rest of the protein. Therefore, E51 and N71, together with the eleven residues and those nearby that form the gap region, were subjected to four-state model I and IV. The individual fittings explain the NMR profiles of these

residues very well by the two selected four-state models. Surprisingly, even by four-state models, residue E51 and N71 still could not be fitted together with other residues in the gap region. The individual fittings of E51 and N71 suggest that the exchange rate for this two residues which fits the data should be around 1500 s^{-1} while some residues in this region (T48, V60, V66, etc.) produced exchange rates at around 3000 s^{-1} . If the gap region moves in a cooperative way, extracted exchange rates for adjacent residues should be the same. The discrepancies in E51, N71 and the other residues in the gap region suggest that there might be other conformational rearrangements occur in this structural segment and they could be on different timescales with CEST and CPMG.

3.7.4 Evaluation of the R_{ex} contributed by potential fast exchange process

By theory, the transverse relaxation rate of a ^{15}N spin in the CPMG experiments could be described by:

$$R_2 = R_2^{DD} + R_2^{CSA} + 0.5 \times R^{ext} \quad (3.13)$$

In which

$$\begin{aligned} R_2^{CSA} &= c^2 (3J(\omega_N) + 4J(0)) / 6 \\ R_2^{DD} &= 0.5d^2 (8J(0) + 3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H + \omega_N)) \\ c &= 2\pi\omega_N CSA_{NH} \\ d &= \mu_0 \gamma_H \gamma_N \hbar / r_{NH}^3 \end{aligned} \quad (3.14)$$

R_2^{DD} and R_2^{CSA} are the HN- ^{15}N dipolar relaxation and chemical shift anisotropy

relaxation, respectively. R^{ext} represents the dipolar longitudinal relaxation between HN and its proximal protons and it is largely independent of the static field. ω_N and ω_H are the lamor frequencies of ^{15}N and proton. γ_N and γ_H are the gyromagnetic ratios of ^{15}N and proton. \hbar is the plank's constant divided by 2π . CSA_{NH} is the chemical shift anisotropy of the ^{15}N and has a value of -160 ppm. r_{NH} is the bond length of the NH bond and is 1.02 Å.

$J(\omega)$ is the spectral density function(Lipari and Szabo 1982a; Lipari and Szabo 1982b):

$$J(\omega) = \frac{\tau_m S^2}{1 + (\tau_m \omega)^2} + \frac{(1 - S^2)\tau_e}{1 + (\omega \tau_e)^2} \quad (3.15)$$

In which τ_m and τ_e are the rotational correlation time and effective correlation time, respectively; S^2 is the generalized order parameter and ω represents the corresponding frequency at which the spectral density function is calculated.

The above three parameters in the spectral density function were extracted by subjecting R_1 , R_2 and NOE data to the Model Free Analysis on a residue basis.

Fast conformational rearrangements whose exchange rates range from 10^5 to 10^6 s^{-1} , could contribute to the R_{ex} (contribution to transverse relaxation rate from chemical exchange processes). Since the k_{ex} of these motions are relatively large, this effect could not be suppressed even at the fast limit of the rf field (typically 1000 Hz) or at the spin-lock $R_{1\text{rho}}$ experiments (1600 Hz). However, the presence of the fast exchange processes could be evaluated by comparing the intrinsic transverse relaxation rates at different static magnetic fields.

According to equation 3.13, the change of R_{ex} in two fields is dominated by the second term while the first term has a minor effect which shifts down the ΔR_{ex} .

The change of R_{ex} in different static fields is calculated by:

$$\begin{aligned}
\Delta R_2^{eff} &= R_2^{eff} (v_{cp}=960Hz, \omega_H=800MHz) - R_2^{eff} (v_{cp}=960Hz, \omega_H=500MHz) \\
&= R_{2(800MHz)} - R_{2(500MHz)} \\
&= \Delta R_2^{DD} + \Delta R_2^{CSA} + \Delta R_2^{ext}
\end{aligned} \tag{3.16}$$

Since the third term is independent of the external static field, the above equation gives:

$$\Delta R_2^{eff} \approx \Delta R_2^{DD} + \Delta R_2^{CSA} \tag{3.17}$$

In the presence of R_{ex} , τ_m and S^2 extracted from the Model Free Analysis will be overestimated. By ignoring the insignificant contributions to the spectral density function at frequency ω_H , $\omega_N + \omega_H$ and $\omega_H - \omega_N$, the experimental relaxation data can be approximated by:

$$\begin{aligned}
R_1 &= \frac{d^2}{4} [3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_N + \omega_H)] + \frac{c^2}{3} J(\omega_N) \\
&\approx \frac{3d^2}{4} J(\omega_N) + \frac{c^2}{3} J(\omega_N) = \left(\frac{3d^2}{4} + \frac{c^2}{3}\right) J(\omega_N) \\
&= \left(\frac{3d^2}{4} + \frac{c^2}{3}\right) J^{oc}(\omega_N) \\
R_2 &= \frac{d^2}{8} [4J(0) + 3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_N + \omega_H)] + \frac{c^2}{18} [4J(0) + 3J(\omega_N)] + R_{ex} \\
&\approx \frac{d^2}{2} J(0) + \frac{3d^2}{8} J(\omega_N) + \frac{2c^2}{9} J(0) + \frac{c^2}{6} J(\omega_N) + R_{ex} \\
&= \left(\frac{d^2}{2} + \frac{2c^2}{9}\right) J(0) + \left(\frac{3d^2}{8} + \frac{c^2}{6}\right) J(\omega_N) + R_{ex} \\
&\approx \left(\frac{d^2}{2} + \frac{2c^2}{9}\right) J^{oc}(0) + \left(\frac{3d^2}{8} + \frac{c^2}{6}\right) J^{oc}(\omega_N)
\end{aligned}$$

(3.18)

In which $J^{oc}(\omega_N)$ is the spectral density function used in the model free analysis.

Therefore, the extent to which $J^{oc}(\omega_N)$ is overestimated can be expressed by:

$$J^{oc}(0) = J(0) + \frac{R_{ex}}{\left(\frac{d^2}{2} + \frac{2c^2}{9}\right)} \quad (3.19)$$

The dominant part in ΔR_2^{eff} is given by:

$$\begin{aligned} \Delta R_2^{eff} &\approx \Delta R_2^{CSA} + \Delta R_2^{DD} \\ &\approx \Delta R_2^{CSA} \\ &\approx \Delta\left[\frac{2c^2}{3} J(0)\right] \\ &= 1.04c^2 J(0) \end{aligned} \quad (3.20)$$

Put 3.7.4.7 into 3.7.4.8 we have the calculated ΔR_2 :

$$\Delta R_2^{cal} \approx 1.04c^2 J(0) + \frac{1.04c^2 R_{ex}}{\left(\frac{d^2}{2} + \frac{2c^2}{9}\right)} \approx 1.04c^2 J(0) + 0.3R_{ex} \quad (3.21)$$

The experimentally measured change in R_{ex} should be:

$$\begin{aligned} \Delta R_2^{eff} &= R_2^{eff}(v_{cp(max)}, B=18.8T) - R_2^{eff}(v_{cp(max)}, B=11.7T) \\ &= R_{2(18.8T)} - R_{2(11.7T)} \\ &= \Delta R_2^{DD} + \Delta R_2^{CSA} + \Delta R_2^{ext} \\ &\approx \Delta R_2^{CSA} \approx 1.04c^2 J(0) + 1.56R_{ex} \end{aligned} \quad (3.22)$$

Thus, by comparing 3.7.4.9 and 3.7.4.10, we have:

$$\Delta R_2^{eff} - \Delta R_2^{cal} \approx 1.26R_{ex} \quad (3.23)$$

Therefore, the differences between theoretical and experimental ΔR_2 values indicate the presence of fast exchange processes that are responsible for the additional contribution to R_{ex} .

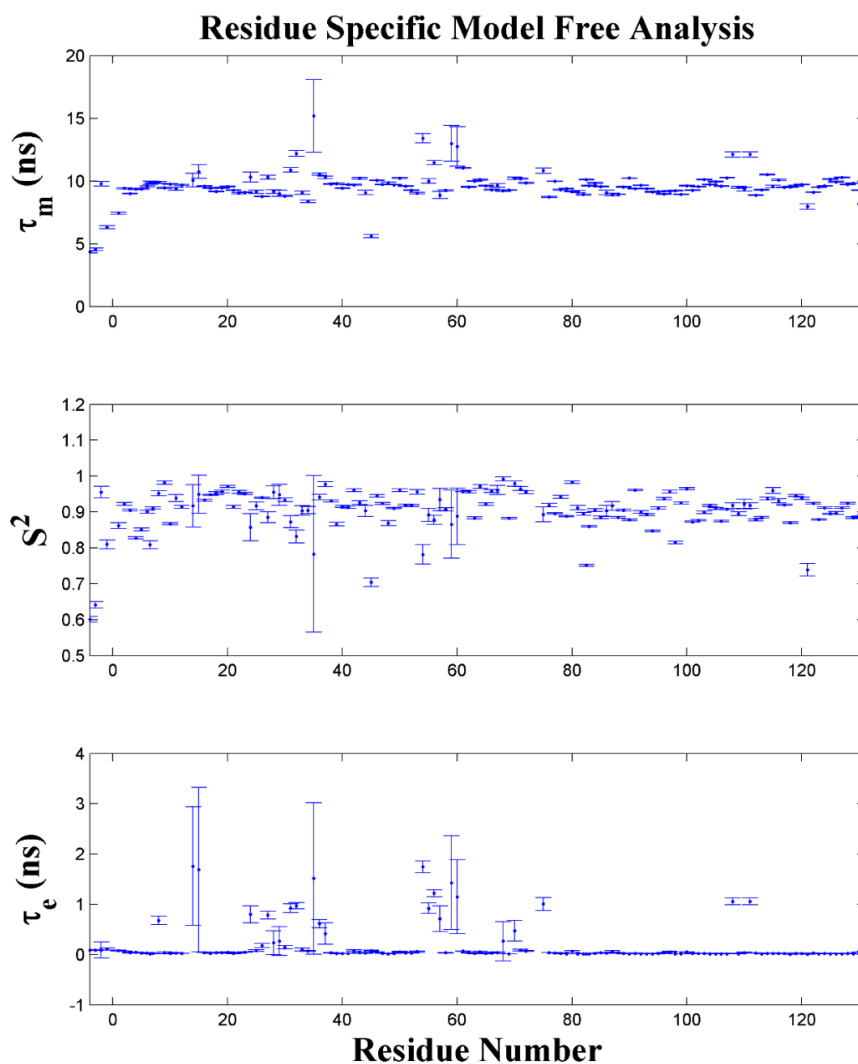


Figure 3.7.5 Residue-specific Model Free Analysis of apo-hIFABP. Residue N57, V60 and N87 were fitted using the R_2 values of their preceding or succeeding residues as described in the text. Error bar represents the fitting error.

The experimental relaxation data (R_1 , R_2 and NOE) were then subjected to the Model Free Formulism on a residue basis to extract residue-specific τ_m , S^2 and τ_e values (Figure 3.7.5). In the Model Free Analysis, residue N57, V60 and N87

could not be fitted due to their large R_2 values. The calculations were done using the R_2 of their preceding or succeeding residues. The differences in the theoretical values of R_{ex} were performed as described earlier in this section. It should be noted that the derived equation 3.22 only serves as an estimation in evaluating the result by relating the differences in the two set of data directly to R_{ex} . The calculation of the ΔR_2^{cal} was done with the full sized equation (3.14).

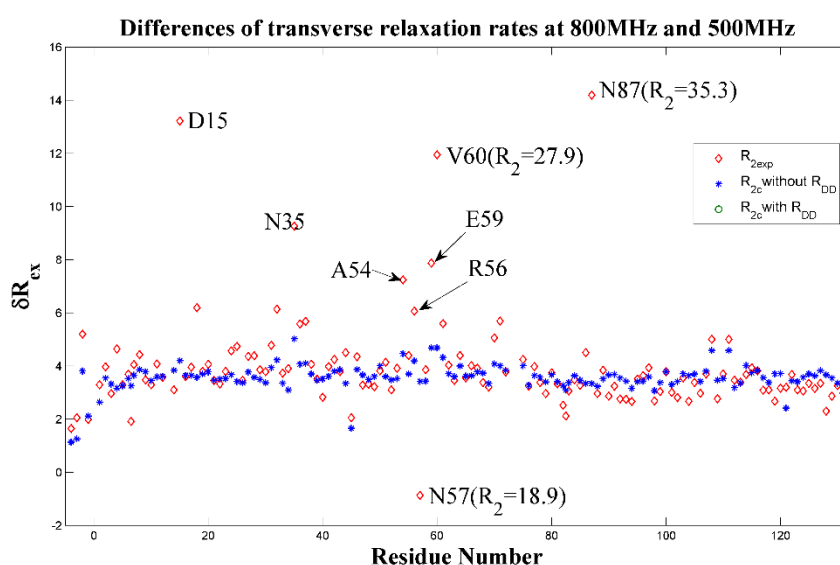


Figure 3.7.6 Comparison of ΔR_2^{eff} and ΔR_2^{cal} . Red diamonds represent the experimental data. Blue asterisks are the calculated value.

In addition, the effect of ΔR_2^{DD} was also evaluated. As described earlier, the contribution from R_2^{DD} was minor and only shifts down the value by less than 1 Hz (Figure 3.7.7).

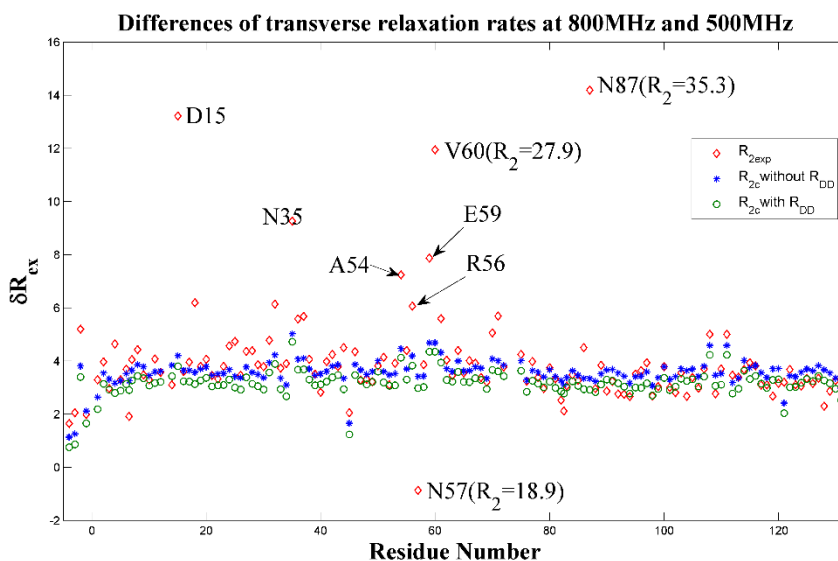


Figure 3.7.7 Comparison of ΔR_2^{eff} and ΔR_2^{cal} in the presence of ΔR_2^{DD} . Red diamonds represent the experimental data. Blue asterisks are the calculated value in the absence of ΔR_2^{DD} and green circles stand for results derived when ΔR_2^{DD} was put into consideration.

As can be seen in figure 3.7.6, a significant number of residues showed sizable R_{ex} contributed by fast timescale exchange processes. Residue N57 showed no difference between the two data sets, this might be attributed to the low peak intensity in spectra due to fast H₂O exchange. Two residues, V60 ($R_2=27.9$ s⁻¹) and N87 ($R_2=35.3$ s⁻¹), have much larger R_2 values than the average. The contribution to R_{ex} by fast chemical exchanges could not be compensated by other parameters in the residue-specific Model Free Analysis. Therefore, the results for these two residues in the above figure were not applicable. However, the presence of the fast exchange processes is directly suggested by their transverse relaxation rates, since their R_2 values are much higher than the average of those from the rest of the residues. In addition, some residues, D15 and N35 for instance, suggest R_{ex} from fast chemical exchanges as well, albeit their peak intensities were weak. More importantly, residues in the gap region,

like A54, R56 and E59, also indicate presence of faster conformational rearrangements. Therefore, we believe that there are additional structural fluctuations in the gap region and these motions are much faster than the three exchange processes reported by CEST and CPMG experiments. The presence of the faster motions also explains discrepancies between E51, N71 and other residues in the gap region when subjected to global fitting. Since chemical shift difference is crucial for NMR experiments to distinguish one macrostate from another, if chemical shifts of E51 and N71 from the assumed “distinct macrostate” in which these two residues are involved in the fast timescale motions, are very close to the major state or any of the pre-existing intermediates, their relaxation profiles should not be explainable with a much larger k_{ex1} , especially when the chemical shifts of the minor states were probed (or at least limited) accurately. On the other hand, when the chemical shift difference is large, it explains the reason that for certain residues in the gap region, why their relaxation profiles could not be fitted well when k_{ex1} is small.

3.8 Final analysis of relaxation data for apo-hIFABP

After the model selection in the previous sections. Collective analysis of CEST and CPMG data were performed based on three and four-state exchange models.

3.8.1 Three-state exchange models

Thirty eight residues whose R_{ex} ($R_{2(40Hz)} - R_{2(960Hz)}$) $> 3 \text{ s}^{-1}$ in 18.8 T static field were subjected to global fittings by three-state model I and II, respectively

(Figure 3.8.1). Error estimation of the global fittings was done by randomly select eighty percent of the residues in the global analysis, this process was repeated 50 times, and the errors were calculated as their standard deviations of the respective parameters. The rest of the residues were fitted by keeping the global parameters fixed to extract the residue-specific parameters. Error estimation for these residues were done with 250 sets of Monte Carlo simulations.

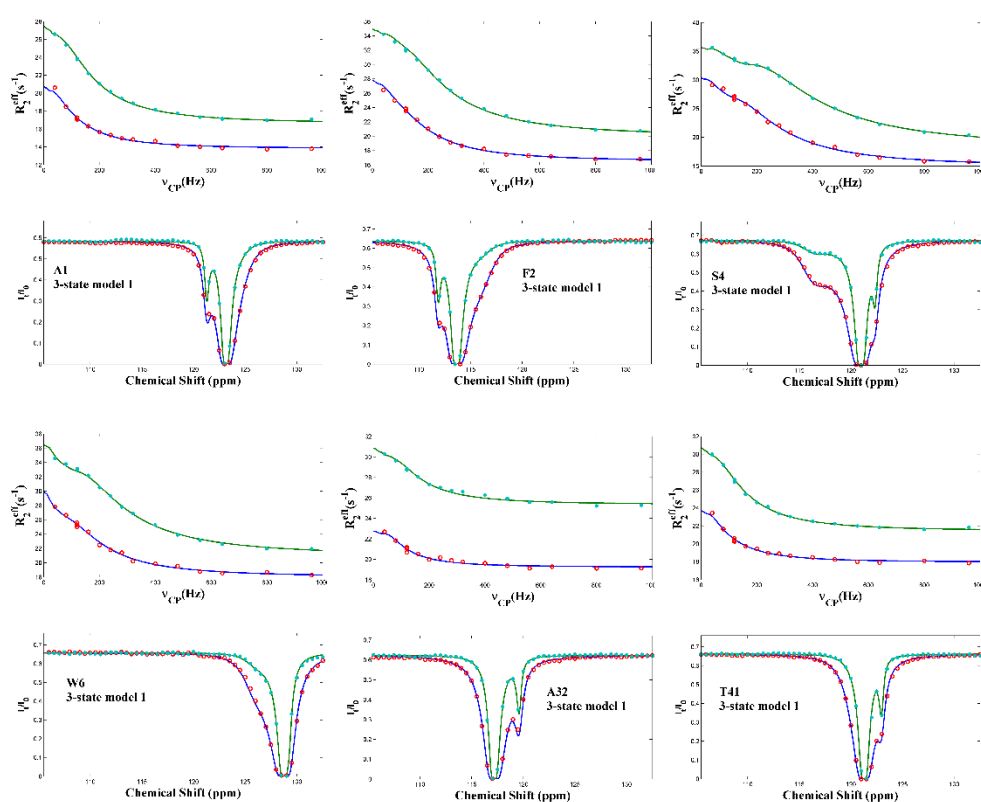


Figure 3.8.1 Representative residues in the global fitting with three-state model I. The plots for the three-state model II are similar.

The resultant global kinetic parameters are: $p_N=94.73\%$, $p_{II}=2.07\%$, $k_{ex1}=807\text{ s}^{-1}$, $k_{ex2}=60\text{ s}^{-1}$, $\chi^2=1761$ for three-state model I and $p_N=91.52\%$, $p_{II}=2.46\%$, $k_{ex1}=701\text{ s}^{-1}$, $k_{ex1}=95\text{ s}^{-1}$, $\chi^2=2844$ for three-state model II.

3.8.2 Four-state exchange models

Twenty five residues in the gap region were subjected to four-state model I and IV, respectively. Seventeen residues with $R_{ex} (R_{2(40\text{Hz})} - R_{2(960\text{Hz})}) > 3 \text{ s}^{-1}$ in the 18.8 T static field were put into global analyses (Figure 3.8.2). Error estimation was carried out by randomly select eighty percent of these residues and repeat fifty times as described in previous sections.

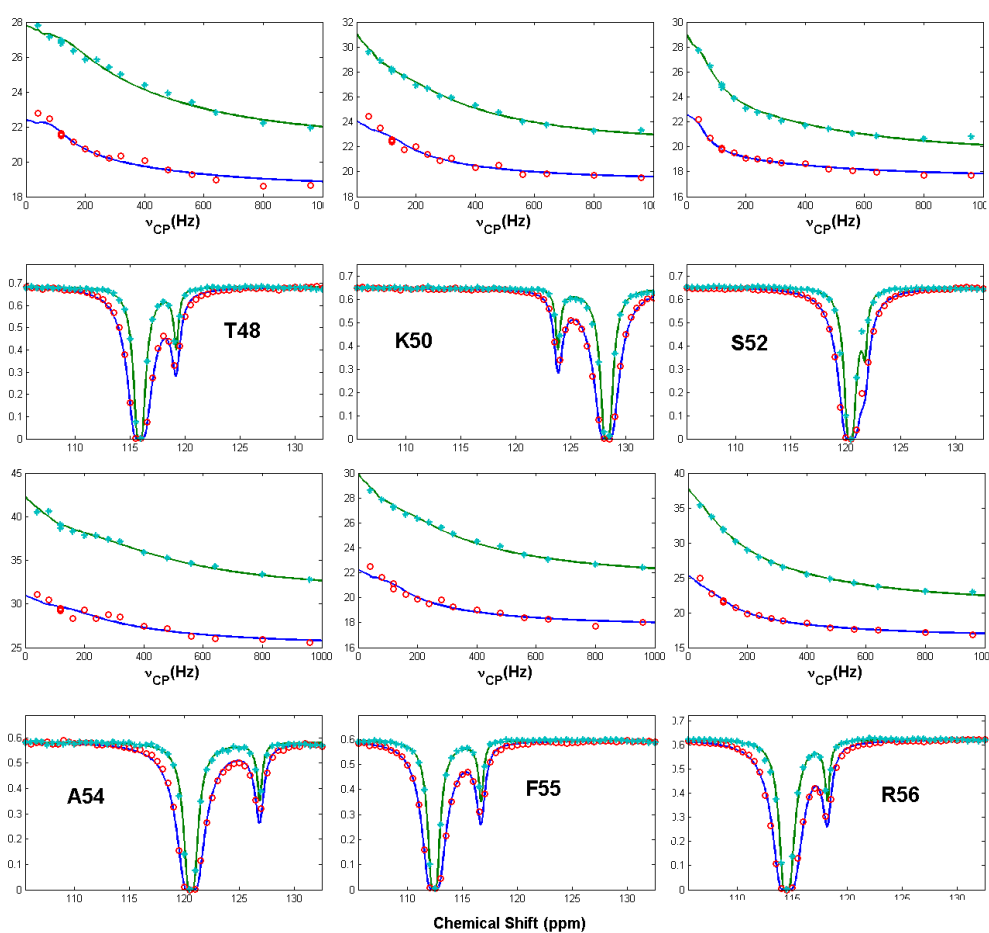


Figure 3.8.2 Representative residues in the global fitting with four-state model I. The plots for the four-state model II are similar.

The resultant kinetic parameters were: $p_N=90.73 \%$, $p_{I1}=3.20 \%$, $p_{I2}=3.33 \%$, $k_{ex1}=3405 \text{ s}^{-1}$, $k_{ex2}=750 \text{ s}^{-1}$, $k_{ex3}=69 \text{ s}^{-1}$, $\chi^2=1759$ for four-state model I and $p_N=90.70 \%$, $p_{I1}=3.39 \%$, $p_{I2}=3.33 \%$, $k_{ex1}=3031 \text{ s}^{-1}$, $k_{ex2}=815 \text{ s}^{-1}$, $k_{ex3}=76 \text{ s}^{-1}$,

$\chi^2=2137$ for four-state model IV. The other eight residues in the gap region were fitted with the global parameters fixed to extract the residue-specific values. The error estimation for these residues were carried out by 250 sets of Monte Carlo simulations.

CHAPTER 4

RESULT AND DISCUSSION

Chapter 4 Result and discussion

4.1 Method validation: robust extraction of kinetic parameters compared to data analysis on single NMR experiment.

As has been described in chapter 3, the CPMG experiments are insensitive to slow exchange processes ($k_{\text{ex}} < 200 \text{ s}^{-1}$). In the presence of a faster exchange process (e.g., $k_{\text{ex}} \sim 800 \text{ s}^{-1}$), the exchange with smaller rate constant often could not be identified. On the other hand, in a system where two exchange processes coexist, kinetic parameters for the two exchanges are still difficult to extract accurately since the change of one parameter could be compensated by the change of other parameter(s). This flexibility is generally attributed to the uncertainty of chemical shifts of the sparsely populated states. To test the reliability of the fitting when CPMG data is the sole source of information, simulations have been performed in which several different conditions were examined. The result of two-state and three-state CPMG fitting together with that of the combined CEST and CPMG analyses will be compared to show the flexibility of CPMG data fitting and hence, the advantage of the combined fitting over data interpretation on a single experiment.

4.1.1 CPMG data fitting

4.1.1.1 Three-site exchange processes approximated by two-site model

To test the performance of the two-state fitting using CPMG data, theoretical

data was generated for 10 residues. The system was assumed to be in a three-site exchange equilibrium in which fractional populations for the major form and minor form are 0.93 and 0.03, respectively. The respective rate constants for the two exchange processes are 850 s^{-1} and 60 s^{-1} . Chemical shift differences of the two minor forms with the major form range from -2.71 ppm to 4.5 ppm . The generated data was injected with 2% and 1.3% random error for CPMG and CEST respectively before fitting. The results showed that the two-state model could explain the data very well both individually and globally. The exchange rates extracted by individual fitting range from 300 s^{-1} to 801 s^{-1} with the corresponding p_N range from 96.11% to 98.81%. For global fitting, the p_N is 96.50% while $k_{\text{ex}}=749\text{ s}^{-1}$.

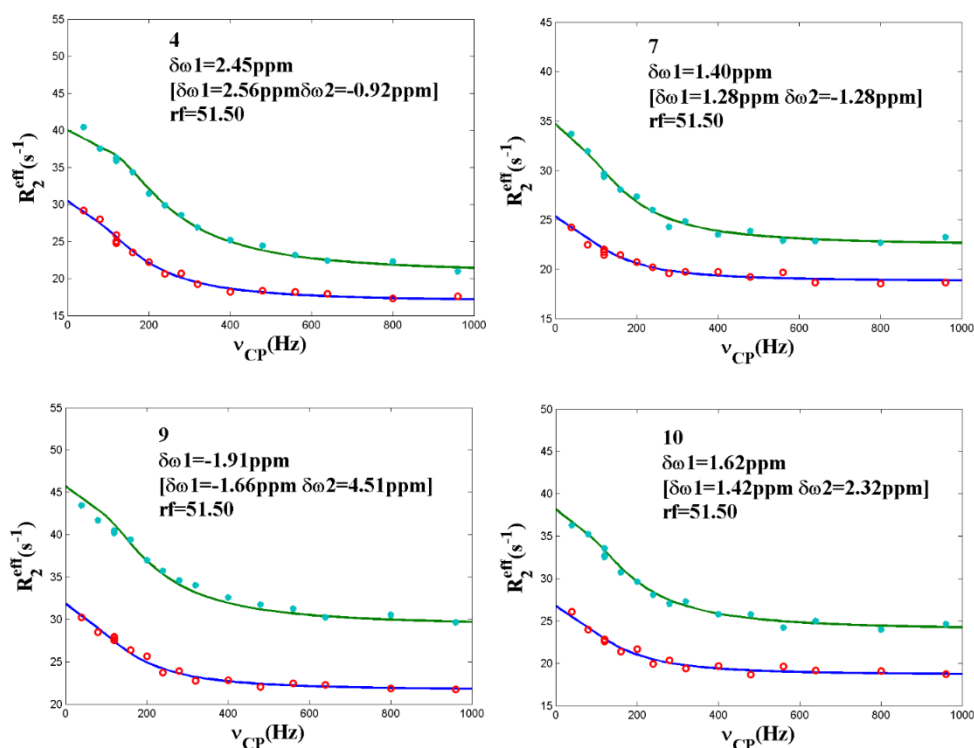


Figure 4.1.1 Representative residues in the two-state global fitting of the theoretical data. Data in the square brackets are the input δ_{ω} values for each residue.

As can be seen in figure 4.1.1, the two-state exchange modeled the data well. The extracted chemical shifts of the minor state were close to that of the input value for the fast exchange process while the absence of the other exchange process was compensated by slight changes of the other parameters and therefore was not identified. This explains the reason why for the apo-hIFABP, certain residues in the gap region produced substantially larger k_{ex} in the individual analyses yet had little effect on the global fitting residual. The change in k_{ex} was compensated by the change of δ_{ω} , and to a lesser extent, the population. Consequently, although individual fittings distributed in a very broad range, the global fitting which mimicked the complex motions with a simple two-site exchange was still reasonable.

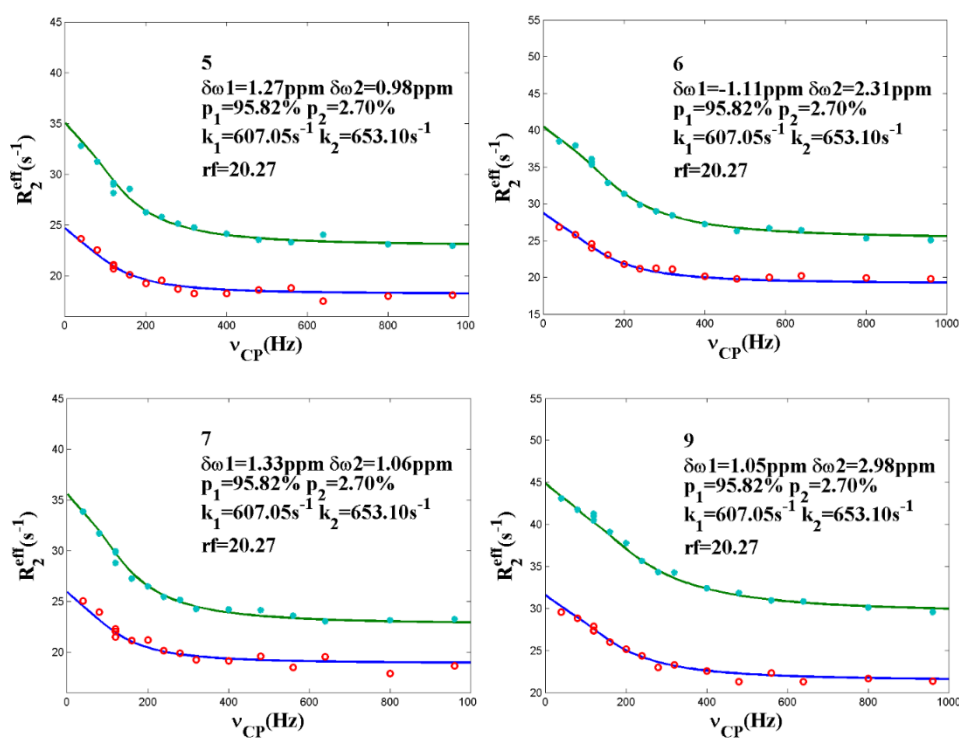


Figure 4.1.2 Residues in the global fitting of the theoretical data by three-state model I.

4.1.1.2 Three-site exchange processes examined by three-site models

The generated data was also subjected to fittings by three-state models. The result indicated that unlike two-state fitting, the three-state models are very sensitive to the initial values. Since the relaxation profiles are the overall effects of two motions, it is hence not applicable to precisely track the two motions at the same time solely by CPMG data analyses, especially without any pre-knowledge of the chemical shifts of the two invisible states.

As shown in figure 4.1.2, the global fitting produced better result compared to that of the two-state analysis. The initial values for δ_{ω} were set between -1 ppm to 1 ppm while the populations were set at 94.5% and 2.2 for the major form and the minor form. However, it is obvious that the result only represents one local minimum where the contribution to the relaxation from the slower exchange process was overestimated while the faster exchange shifted away slightly from the true value so that the target function was not compromised.

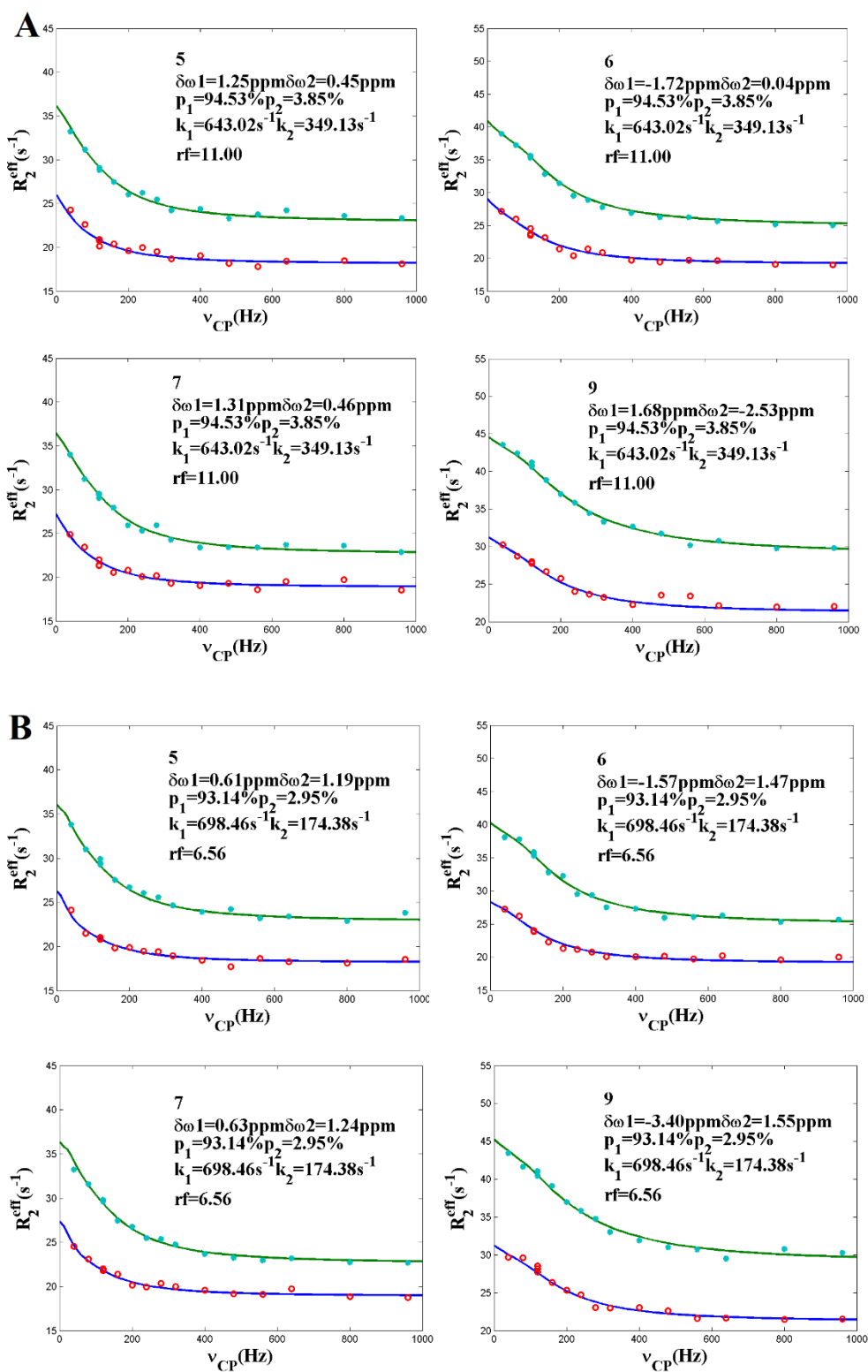


Figure 4.1.3 Representative residues in the global fitting by three-state model II and III. (A) Three-state model II. (B) Three-state model III.

The other two three-state exchange models were also tested. Surprisingly, both

the two models explained the synthetic data very well. The initial values for exchange rates were both set at 300 s^{-1} while other parameters were the same with those used for model I. As demonstrated in above figures (4.1.1 to 4.1.3), these fittings by different models could hardly be distinguished by visual inspections. The actual input model was three-state model I, yet the minimizations favored model II and III. Although one might still find the correct model by locating the global minimum with grid search aided parameter space mapping as introduced in section 3.4.3.1. It is noteworthy that in practice, the global minimum in the parameter hyper surface is not necessarily the one that represents the parameter set which accurately describes the experimental data. This could be partially attributed to the possible systematic errors which are always kept at minimum by spectroscopists at all costs and yet are still inevitable in reality. More importantly, for macromolecules such as proteins, multiple motions coexist on different timescales, further complicating relaxation dispersion profiles on which data analyses are based. As our simulations show, even for a simple three-site exchange system, data analyses on CPMG experiments alone, especially in the absence of any pre-knowledge about the corresponding chemical shifts of the minor states, are neither able to precisely extract two motions nor differentiate the actual model from other assumptive models. Therefore, probing chemical shifts of the invisible states, or at least limiting the chemical shifts in a relatively small range, is crucial for CPMG data analysis.

4.1.2 CEST data fitting

CEST experiments are very sensitive to motions with their lifetimes range from 5 ms to 50 ms. When slow motions on the above time scale exist in the system, CEST experiments can accurately extract the exchange rates and fractional populations of the sparsely populated conformers and more importantly, locate the corresponding chemical shifts. However, when motions that are in the optimum detection range of CPMG experiments (typically from 0.5 ms to 5 ms) coexist with slower motions mentioned above, CEST experiments alone will not be able to model these motions well at the same time due to the detection limits, especially when δ_ω is relatively small (e.g., < 2 ppm). As has been demonstrated in section 3.5, for the apo-hIAFBP, in the presence of a faster exchange process (~ 800 s⁻¹), CEST experiments could only extract kinetic parameters of the slower exchange (~ 60 s⁻¹). The obtained result for the faster exchange usually has large uncertainties due to the decreased sensitivity to motions faster than 200 s⁻¹. In this scenario, CEST experiment could be used as a complement for experiments that are sensitive to motions beyond its limit, which in our case, the CPMG experiment.

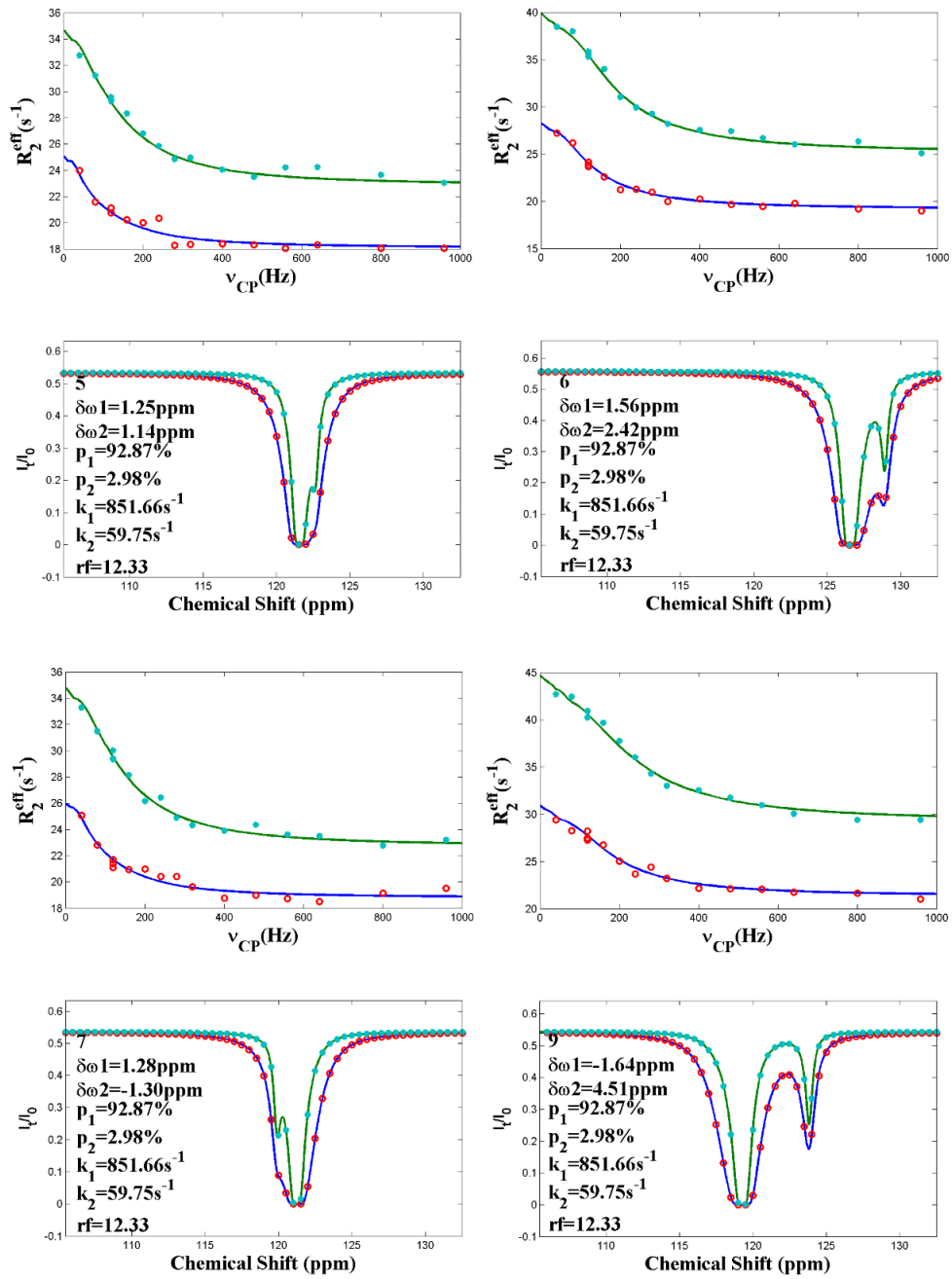
4.1.3 Combined CPMG and CEST analyses

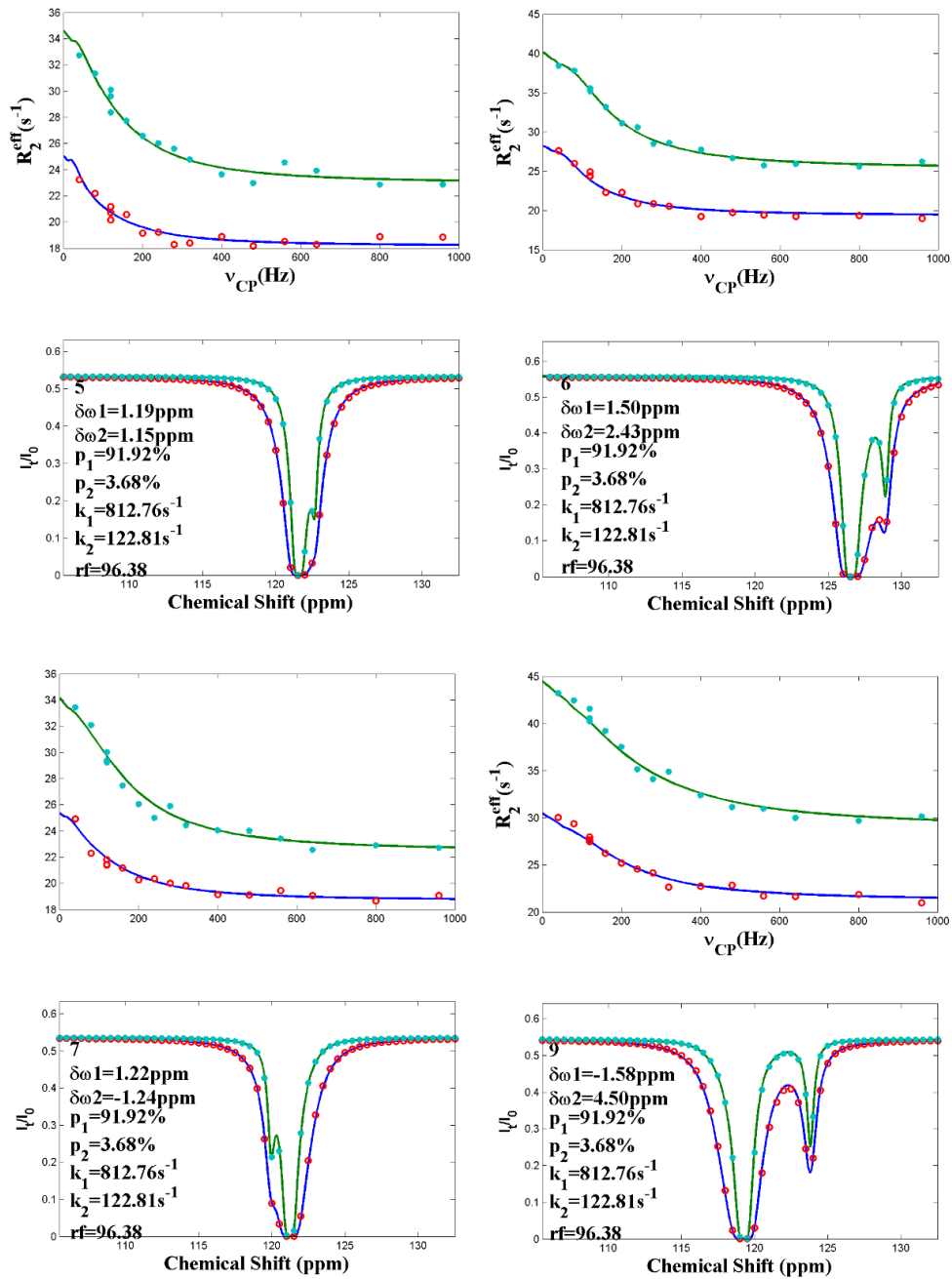
The presence of a slow motion with k_{ex} in the range of 20 s⁻¹ to 200 s⁻¹ was proved by CEST data with sharp minor dips. Meanwhile, a faster motion was suggested by CPMG profiles by sizeable R_{ex} (typically larger than 3 s⁻¹). This is obvious as we have shown in section 4.1.1, although the two motions could not be extracted precisely by CPMG fitting alone, the existence of a faster exchange process with k_{ex} around 600 s⁻¹ to 800 s⁻¹ was suggested by both two-state and

three-state fittings. In this case, the synthetic data set used in the tests for two-state and three-state CPMG fittings was subjected to the examination of the combined CEST and CPMG analyses.

As shown in figure 4.1.4, the results are self-explanatory. Model I describes the data best and model II produced a fitting residual seven times larger than that of model I, and model III failed to explain the data. It should be noted that CEST data is very sensitive to the initial values of δ_{ω} . For the slower exchange, δ_{ω} for each residue (if a sharp minor dip is present) is typically limited in ± 0.1 ppm range, and the chemical shift for the other state is limited between ± 1 ppm if there is no broad minor dip. Therefore, the combined analyses of CEST and CPMG data are less dependent on the initial δ_{ω} values and are easier to locate the global minimum.

A



B

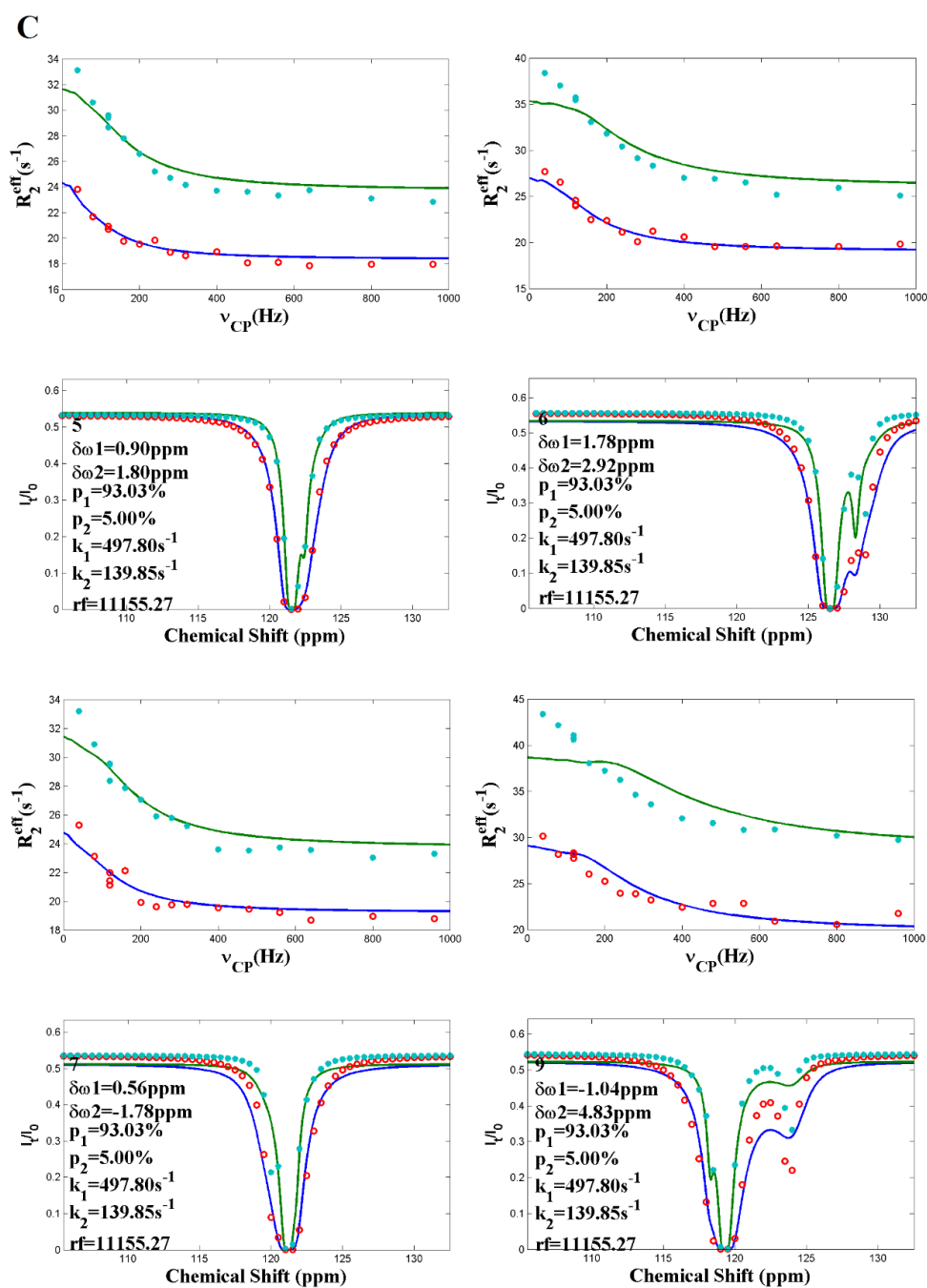


Figure 4.1.4 Global fitting of synthetic data with combined analysis of CEST and CPMG data. (A) Global fitting by three-state model I. (B) Global fitting by model II. (C) Global fitting by model III. Apparently, model III failed to explain the data while model I describes the data best with 7 times smaller fitting residual.

Data analyses for systems where more than three conformers coexist are similar with those of the three-state system. Searching ranges for chemical shifts of the sparsely populated states are set according to the CEST profiles. In practice, it

is preferable to do an analysis on CEST data first for an initial estimation of the possible slower exchange process and the corresponding chemical shifts of the related minor states and those of the faster exchanging conformers, should there be any.

4.2 NMR titration experiments

Chemical shift perturbation is a simple and straightforward experiment for locating the ligand binding site. Since chemical shift changes are very sensitive to structural changes and can be precisely detected, this means almost any genuine ligand binding will produce CSPs. This technique is widely used in drug discovery for high-throughput screening for potential drugs as binding partners to certain proteins in interest. Additionally, ligand titration experiments could be used to extract disassociation constant values even without assignments of any spectrum. In a typical CSP data analysis, the change of chemical shift is fitted along different ligand/protein concentration ratios to extract the K_d value of the specific binding. The major factor for the determination of the disassociation constant is the k_{off} rate for the ligand to leave the binding site. When k_{off} is much bigger than the frequency differences between the bound and unbound protein the exchange is fast in the chemical shift timescale and the resulting peaks are population-averaged signal. In this case, the K_d could be determined by monitoring the position changes of peaks in spectra. On the other hand, when k_{off} is not greater than the frequency differences of the two forms, the corresponding peaks become broadened, making it very difficult to determine the K_d . In addition, direct ligand binding

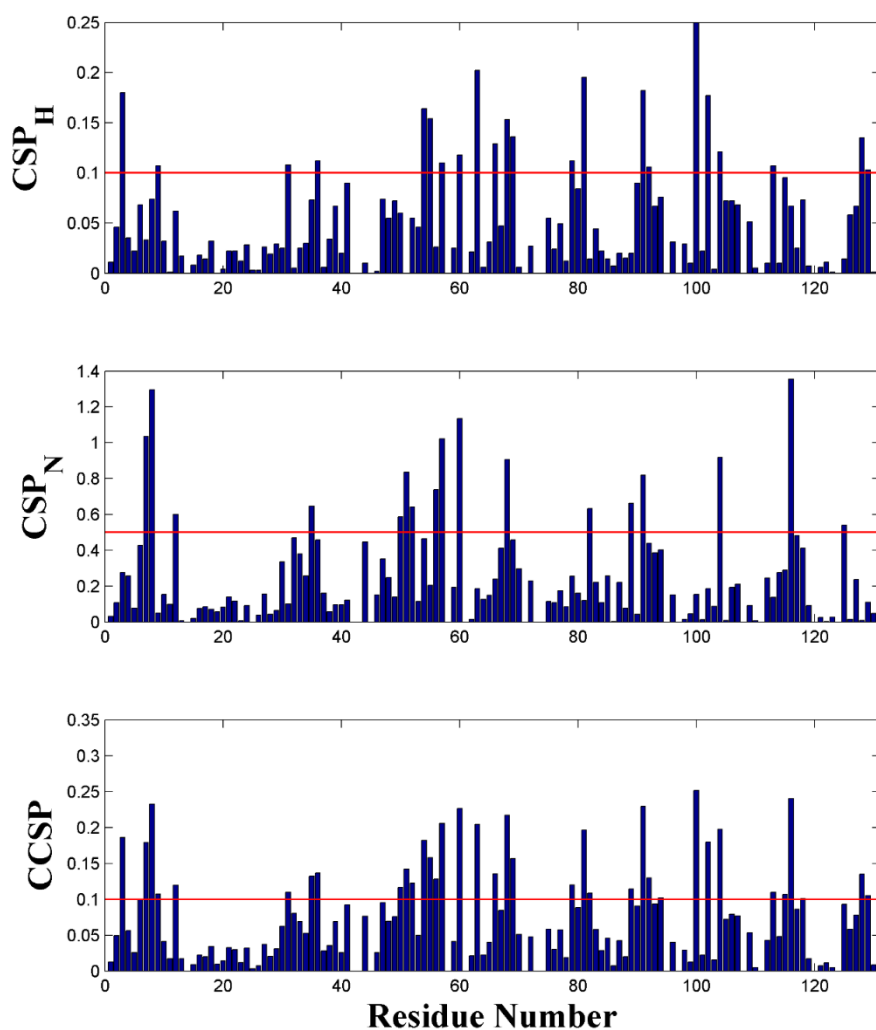


Figure 4.2.1 Chemical shift perturbation of apo-hIFABP upon binding with 1,8-ANS.

and allosteric effects as well as the non-specific surface interactions could produce CSPs and it is still impractical to differentiate these origins currently. In this work, the major goal is to study the dynamic properties of the hIFABP in the absence of ligand. The chemical shift walking directions of each residues were monitored and residues with significant changes above the standard deviation will be used to compare with those show large perturbations in the excited state in the absence of ligands so that valuable insight could be obtained

on the structure-function relationship.

As described in section 3.2, a large number of residues were perturbed upon binding with 1,8-ANS. The chemical shift changes were tracked on each stage of the binding and the CSP of the corresponding ^1H and ^{15}N and CCSP were calculated according to formula 3.2. It is well known that the chemical shift

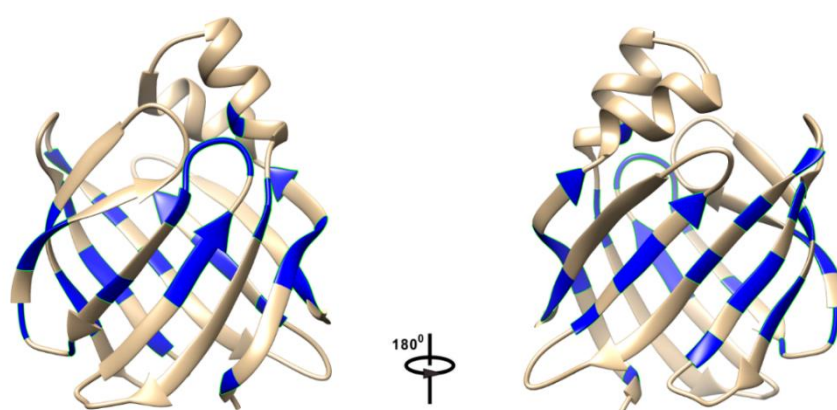


Figure 4.2.2 Residues with CCSP larger than 0.1 plotted on the ribbon structure of apo-hIFABP.

of a specific spin generally depends on the atom type and its chemical environment. For instance, the amide HN in proteins resonates in the range of 6 ppm to 10 ppm while the protons of a methyl group resonate in the range of 0 ppm to 3 ppm. Meanwhile, the chemical shifts of ^{15}N and ^{13}CO group could be very hard to predict since they are affected by many factors. The sidechain conformation of the preceding residue could contribute to the secondary shift of the ^{15}N for up to 8 ppm and this effect could be as large as 22 ppm when the sidechain type participates as well. The hydrogen bonding has effects to the chemical shifts of protons and nitrogen and even to the carbonyl carbon attached to the nitrogen as well. Therefore, the calculation of CCSP nowadays could only

serve as a method to help locating the binding site of a protein to its ligands and sometimes, to extract the dissociation constant when the protein concentration ranges from $0.2K_d$ to $5K_d$. As figure 4.2.1 shows, nearly half of the residues display sizable CSPs for ^1H and ^{15}N or the CCSPs. The locations of these residues spread around the β -barrel (Figure 4.2.2), which agrees well with the fact that the binding pocket for FABPs is inside of the β clam in which the ligand is surrounded by most residues in the β -strand area.

The CCSP profiles of 56 residues were selected to examine the binding pattern of hIFABP with 1,8-ANS (Figure 4.2.3). The data shows that the CCSPs increase with the ligand concentration, when normalized by the values of the highest ligand concentration, all CCSP curves converge at the end of the titration while differ significantly in other stages of binding. However, when subjected to normalization by the corresponding values at the stoichiometry binding stage, a clear pattern shows that there are two phases for the binding of 1,8-ANS to the hIFABP. Within the stoichiometry limit (ligand/hIFABP=1:1), binding to the protein generally follows the same pattern suggested by a nearly straight line of CCSP profiles. When the ligand become excessive, CCSPs of some residues diverge to different ranges as to represent the non-specific surface interactions of ANS with hIFABP. This was also evidenced by the changes of the chemical shift walking directions of certain residues (Figure 3.2.2).

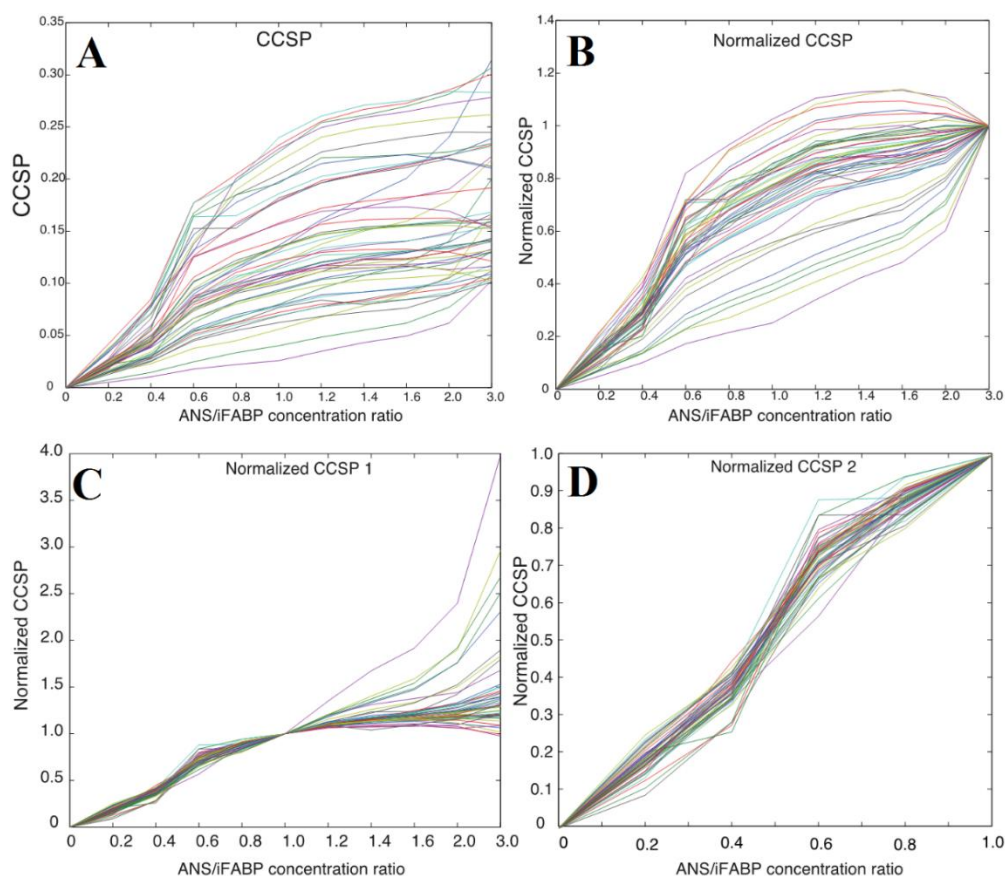


Figure 4.2.3 Combined chemical shift perturbation of 56 residues with CCSP larger than 0.1. (A) CCSP of hIFABP along molar ratios of ligand/protein. (B) CCSP normalized by the values when ligand/protein=3.0. (C) CCSP normalized with the values when ligand/protein=1.0. (D) Zoom in of CCSP profiles of plot C in the within the stoichiometry binding.

4.3 H-D exchange experiments

The analysis of the kinetics of hydrogen exchange in biological macromolecules could provide useful information on their structural and dynamic properties. It is well known that the major factor limiting the hydrogen exchange in proteins is the proton in the hydrogen bonding where a transient opening of the local structure is required for the corresponding proton to exchange with those in the solvent. On the other hand, hydrogen bonding properties are important in defining the protein structures and by determining the protection factor of amides one can obtain information on the backbone fluctuations and solvent

accessibility of local regions. As our time-resolved H-D exchange experiments show, the amide proton of 61 residues exchange relatively fast and the peak intensities drop to less than 2% (within 373 s) compared to the HSQC spectrum obtained in water.

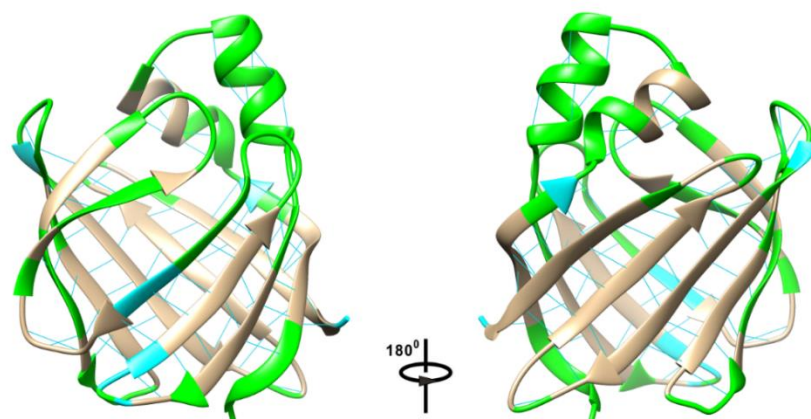


Figure 4.3.1 Fast exchanging residues in H-D exchange experiment plotted on ribbon structure of apo-hiFABP. Highlighted in green (cyan) are the residues whose peak intensities dropped to 2% (20%) and below till the end of the first spectrum (373 seconds). Hydrogen bonds are shown with lines.

These residues mainly located on the two helices, the loop areas and β -strand D, E (Figure 4.3.1). This result indicates that except the areas without hydrogen bonding, the helices and the β -strand D, E experience structural fluctuations in which the amide proton in hydrogen bond is prone to solvent exchange and therefore are loosely packed compared to the rest of the protein. The weak hydrogen bonding in the helical domain might be related to its function in which transportation of fatty acids to phospholipid membranes is mediated by collision of the structural segment with the target area. As for β -strand D and E, the structural fluctuation is also suggested by our relaxation experiments in which the gap region experiences conformational rearrangements that differ than those

in the other regions.

4.4 Four-state equilibrium: coexistence of multiple intermediates of apo-hIFABP in solution

For the NMR data, through robust model selection, our combined analyses of the longitudinal and transverse relaxation processes suggested that the apo-hIFABP exists in at least a four-state equilibrium. The extracted chemical shifts of the sparsely populated conformers were subjected to comparison with the predicted chemical shifts of the intrinsically disordered protein (IDP). The linear correlation indicates that the intermediates are well-structured and significantly different with IDP (Figure 4.4.1). This result differs from any of the previous NMR studies in which at least one unfolded form participates in the dynamical equilibrium together with the major form and intermediates.

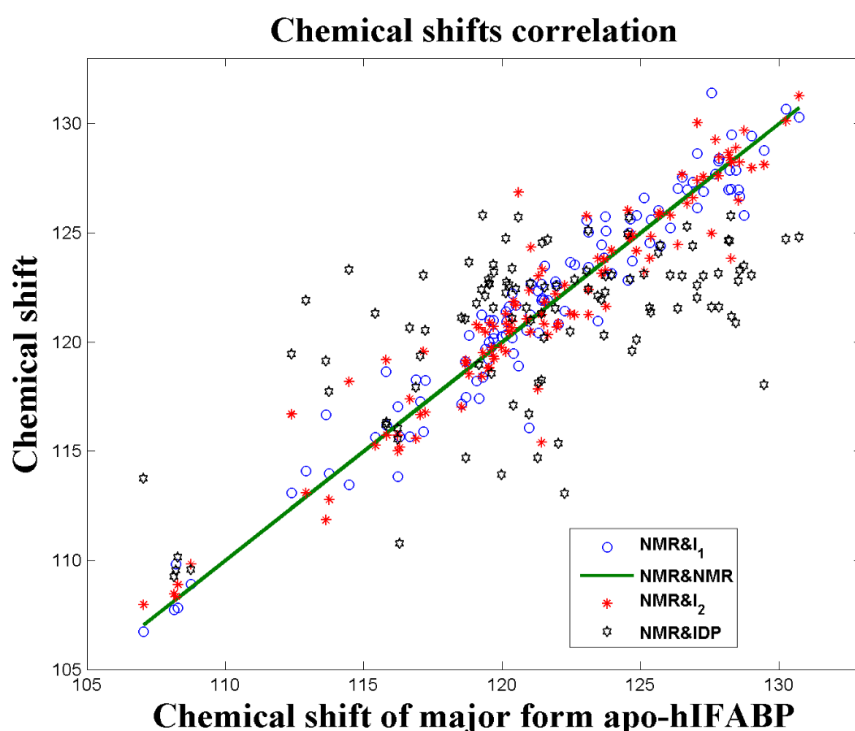


Figure 4.4.1 Chemical shifts comparison between minor states and IDP. NMR is the experimental data for the major state, I_x represents the intermediate, IDP is the data of the intrinsically disordered protein. The green line represents the linear correlation of the major form. The blue circles, red asterisks and black

pentagrams are the corresponding values for the two intermediates and IDP, respectively.

4.4.1 Regional motions and the statistical view of proteins in dynamical equilibrium

In a protein molecule, residues directly connected with each other are bound to move in a concerted way since protein motions do not involve breakage of covalent bonds. This kind of motions are therefore called concerted motions (or collective motions). In addition, there may be one or more regions that are flexible in a large protein molecule. Nonetheless, the two or more flexible parts are not necessarily present in a single molecule, hence the two structural segments move independently of each other. For instance, M is the protein sample in an NMR experiment, m_i represents one molecule in solution. The sample is an ensemble of n molecules and there are two subsets of M which we denote as M_1 and M_2 whose number of molecules are m and j ($m+j=n$), respectively. The molecules in M_1 have local structural fluctuations in the helical region while molecules in M_2 undergo a motion in the β -strand area. The respective fractional populations of the two minor states are m/n and j/n . In this case, the population of the major form should be $1-(m/n+j/n)$.

$$\begin{aligned}
 M &= \sum_1^n (m_i) \\
 M_1 &= \sum_1^m (m_i) \\
 M_2 &= \sum_1^j (m_i) \\
 M &= M_1 + M_2
 \end{aligned}
 \tag{4.1}$$

The two motions could be spotted by NMR as long as there are on the detectable timescales and the population of the two “excited” states are larger than 0.5%. It is clear that any molecule in the sample does not experience two motions at the same time, yet NMR will find the two minor “distinct macrostates” for the whole protein ensemble. In this case, these two motions do not comply with the collective manner since they are in different ensembles of molecules. In cases where the above motions occur in a same molecule, the situation becomes a bit complicated since whether the two motions are cooperative depends on the distance of the two structural segments where local fluctuations occur or in other words, the size of the protein molecule. It is now easy to understand such a behavior for a large protein with two or more domains where domains are connected with long loops. If the above two structural segments are not in the same domain for such a protein, the forces which restrict motions of adjacent residues simply do not apply. When the size of the molecule becomes smaller, for example, a small globular protein where only a single domain is present, cooperativity of the two motions will increase as the distance of the related two areas decrease until they become fully cooperative, or put differently, become a single collective motion. In this scenario, residues that connect the two structural segments share similar but not the same dynamic properties with the “distinct regions” and form a transition area. Therefore, a statistical view is required when it comes to the properties of a protein.

Variability in flexibility provides structural basis for different motions, i.e., different amplitudes and lifetimes, etc. In this context, structural segments in a protein could display different features in terms of motions, albeit the transition in a small protein tends to be less obvious. For a protein in solution, if “distinct”

conformations exist in equilibrium, chemical shift is the key element for NMR to differentiate them. Chemical shift, on the other hand, are predominantly determined by the atom type and its chemical environment. The latter one is often used to decide whether two conformations have the same structure in dynamic studies, albeit it is not necessarily true vice versa. In fact, from studies on excited states, the chemical shifts of a significant number of residues are almost identical within experimental error. This means regions where these residues situated have the same local structure in both the ground and excited state(s). As also evidenced by x-ray studies, most part of proteins with or without the ligand(s) are superimposable, namely, have the same local structure. As discussed above, there could be multiple flexible structural segments in a protein and statistically, these segments might be present in a single molecule or not while other parts of the molecule are generally the same in every molecule. Alternatively, only part(s) of the molecule differ in structure and the number of the part(s) defines how many excited states are in equilibrium.

The above theory explains our data very well. Most part of the protein undergoes two exchanges with lifetimes at 17ms and 1.2ms, respectively. The gap region delimited by β -strand C, D, E and the corresponding loop areas undergoes different motions, evidenced by significantly faster exchange rate from both three- and four-state fittings. Most residues in the gap region could not be fitted with $k_{ex1} < 2000 \text{ s}^{-1}$, E51 and N71 could only be fitted well with other residues in the gap region by any model. This indicates that the structural rearrangements in this area are much more complicated than a simple three- or four-site equilibrium and this conclusion is also supported by the presence of additional fast exchange processes according to the R_{ex} calculation based on

residue-specific Model Free analysis. According to our data, we suggest that the apo-form hiFABP undergoes motions differ from one region to another. The majority of the protein could be explained well with three-state exchange model I and II while the gap region could only be modeled with four-state exchange model I and IV. The resultant kinetic parameters are: $p_N=94.73\pm 0.17\%$, $p_{II}=2.07\pm 0.04\%$, $k_{ex1}=807\pm 9\text{ s}^{-1}$, $k_{ex2}=60\pm 4\text{ s}^{-1}$, for three-state model I and $p_N=91.52\pm 0.19\%$, $p_{II}=2.46\pm 0.06\%$, $k_{ex1}=701\pm 13\text{ s}^{-1}$, $k_{ex1}=95\pm 7\text{ s}^{-1}$ for three-state model II; $p_N=90.73\pm 0.16\%$, $p_{II}=3.20\pm 0.26\%$, $p_{I2}=3.33\pm 0.14\%$, $k_{ex1}=3405\pm 116\text{ s}^{-1}$, $k_{ex2}=750\pm 22\text{ s}^{-1}$, $k_{ex3}=69\pm 4\text{ s}^{-1}$ for four-state model I and $p_N=90.70\pm 0.09\%$, $p_{II}=3.39\pm 0.18\%$, $p_{I2}=3.33\pm 0.14\%$, $k_{ex1}=3031\pm 220\text{ s}^{-1}$, $k_{ex2}=815\pm 61\text{ s}^{-1}$, $k_{ex3}=76\pm 4\text{ s}^{-1}$, $\chi^2=2137$ for four-state model IV.

As discussed above, although different models were used to explain the data, when chemical shifts are put into consideration for the multistate system, the actual result should be four-state equilibrium for the apo-hIAFBP. The majority of the protein are well modeled by three-site model while the gap region requires a more complicated four-site model to explain. This is because for the gap region, the chemical shifts of the third intermediate are large so that the contributions to the R_{ex} from the corresponding exchange process are obvious. That is to say, residues aside from the gap region in the third intermediate generally have very small δ_ω with respect to the major form. According to our simulation, when δ_ω is close to zero (e.g., ± 0.01 ppm) the effect of the chemical exchange process to the relaxation rate will be infinitesimal. Consequently, apart from the gap region, most part of the protein could be very well approximated with three-site models. For a chemical exchange process to exist,

the very basic condition is that there are at least two species in the system. For NMR, the key parameter to differentiate one species from another one is the chemical shift. When put in the extreme limit where δ_ω of a residue is zero, actual chemical exchange is not present for that probe. As for a system described by formula 4.1, residues that are not involved in the two motions, produce same chemical shifts in subset M_1 and M_2 , and yet, chemical exchanges do exist for ensemble M .

To put in another way, for the system in formula 4.1, residue A has a relatively rigid conformation and hence produces same chemical shifts in subset M_1 and M_2 , the RD data of A should have $R_{ex}=0$ while no minor dips should be detected in CEST data and the exchange model for residue A is 0-state exchange which means no chemical exchange exists. For residue B in the helical region of the protein, it undergoes a motion with k_{ex} around 80 s^{-1} , the corresponding exchange model for residue B suggested by data fitting should be 2-state exchange (the excited state of residues B is in subset M_1). For residue C in the β -sheet, it is involved in a motion with k_{ex} around 650 s^{-1} , then the exchange model for C is 2-state exchange as well (the excited state of residues C is in subset M_2). However, since exchanges detected for residue B and C are different, the exchange model for the whole protein should be a three-state equilibrium. The same logic applies to the apo-hIFABP sample so that three intermediates coexist with the major form in the four-state equilibrium. Moreover, the kinetic parameters extracted from the three- and four-state fittings support this explanation since the rate constants of the two global motions are very close: $k_{ex1}=807 \text{ s}^{-1}$, $k_{ex2}=60 \text{ s}^{-1}$ (3-state model I) versus $k_{ex2}=750 \text{ s}^{-1}$, $k_{ex3}=69 \text{ s}^{-1}$ (4-state model I); $k_{ex1}=701 \text{ s}^{-1}$, $k_{ex1}=95 \text{ s}^{-1}$ (3-state model II) versus $k_{ex2}=815 \text{ s}^{-1}$, $k_{ex3}=76$

s^{-1} (4-state model IV). Therefore, there are three intermediates coexist with the major form of apo-hIAFBP, intermediate I_3 interconverts with intermediate I_1 (four-state model I) or the major form N (four-state model IV) with a rate constant of $3405 s^{-1}$ or $3031 s^{-1}$. The major structural difference of I_3 and N is in the gap region.

4.4.2 Physical meaning of the four-state equilibrium

As described in method validation section in section 4.1, the synthetic data set was examined by different models. The incorrect three-state model III was excluded according to combined relaxation data analyses, the actual model (three-state model I) could not be differentiated from model II, although the minimization did not favor the later one. The situation is similar in the three-state model test for the apo-hIFABP. Both model I and II fit the data well although model II always produce larger fitting residual. Since the four-state models are based on the corresponding three-state models in which the third intermediate interconverts with either one of the existing species, and the fitting results of the four-state models are consistent with those of the three-state models. In the following sections, the corresponding results for the two models (3-state model I and 4-state model I versus 3-state model II and 4-state model IV) will be discussed in detail.

4.4.2.1 Three-state model I and four-state model I

If these two models represent the actual exchange pathway for the apo-hIFABP in solution. According to the calculated kinetic parameters, we have:

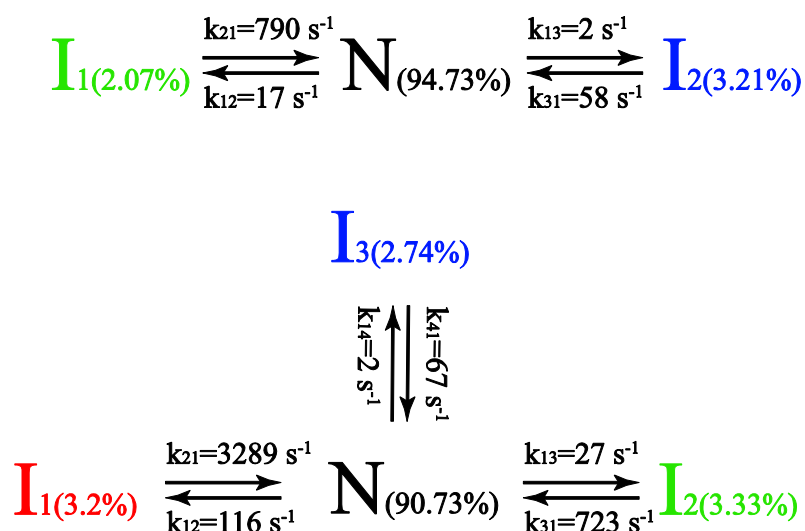


Figure 4.4.2 Four-state exchange pathway I for apo-hiFABP. Minor states are highlighted in different colors for comparison between the three-state and four-state model. Subscript of the intermediate represent position of the species in the magnetization evolution matrix and does not mean the name of an intermediate.

In the above model, the major state directly interconverts with the three intermediates. Highlighted in green, blue and red are the three excited states, respectively (Figure 4.4.2). Since the system is in a thermodynamic equilibrium, according to Boltzmann distribution, we have:

$$\frac{N_2}{N_1} = e^{-\frac{(E_2 - E_1)}{k_B \times T}} \quad (4.2)$$

In which N_i is the number of molecules of the state i ; E_i is the energy of the state i ; k_B is the Boltzmann constant and T is the temperature in Kelvin. The energy differences of the three excited state with respect to the major form are 8537.4 J/Mol (red), 8437.4 J/Mol (green) and 7068.5 J/Mol (blue), respectively. From the extracted chemical shift differences, we know that major structural fluctuations occur in the gap region and β -strand A, B nearby (Figure 4.4.3).

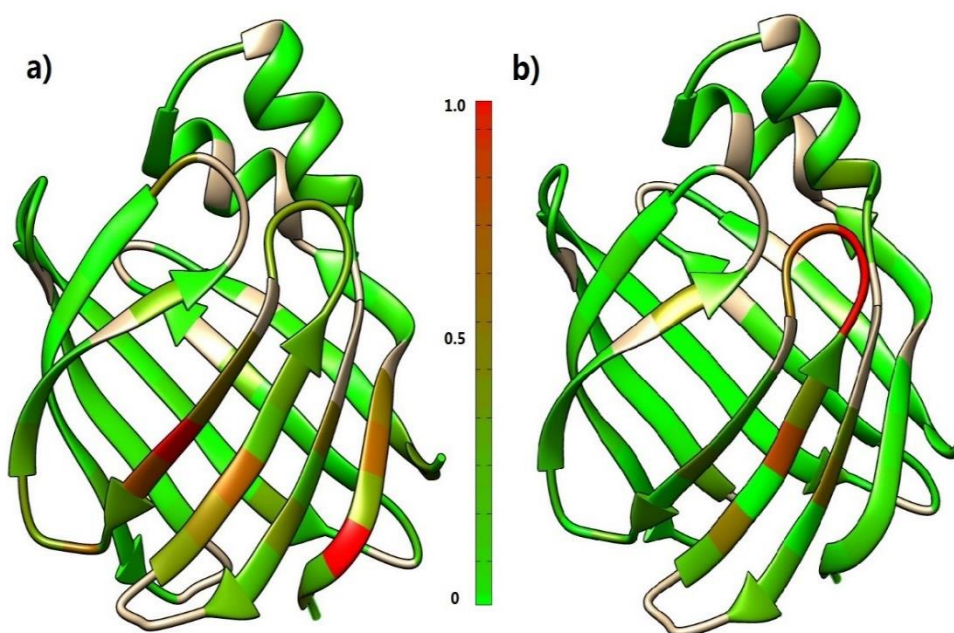


Figure 4.4.3 $\delta\omega$ plotted on the ribbon structure of apo-hiFABP. **(a):** $\delta\omega$ of the fast ($k_{ex} \sim 800 \text{ s}^{-1}$) exchanging intermediate. **(b):** $\delta\omega$ of the slow ($k_{ex} \sim 80 \text{ s}^{-1}$) exchanging intermediate. Residues are highlighted in RGB color from green (minimum) to red (max) based on normalized $\delta\omega$ value. $\delta\omega$ of the third intermediate ($k_{ex} \sim 3000 \text{ s}^{-1}$) is not shown since it is obvious that structural fluctuations occur in the gap region.

4.4.2.2 Three-state model I and four-state model IV

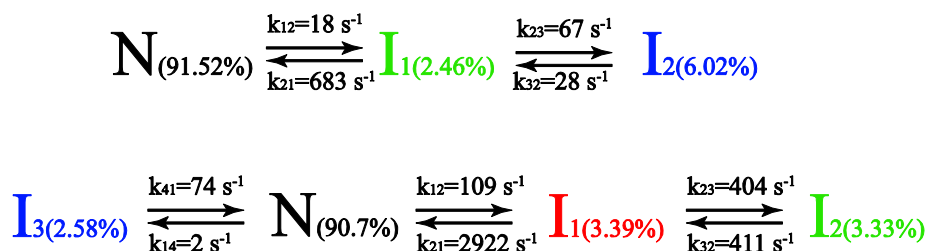


Figure 4.4.4 Four-state exchange pathway II for apo-hiFABP. Minor states are highlighted in different colors for comparison between the three-state and four-state model. Subscript of the intermediate represent position of the species in the magnetization evolution matrix and does not mean the name of an intermediate.

As shown in figure 4.4.4, in this model, the major form interconverts with two

minor forms with exchange rates at $\sim 80 \text{ s}^{-1}$ and 800 s^{-1} , respectively. The third intermediate becomes the transient state on the pathway to the minor form with $k_{\text{ex}}=815 \text{ s}^{-1}$. The corresponding energy differences of the three excited state with respect to the major form are 8389.5 J/Mol (red), 8432.3 J/Mol (green) and 6946.2 J/Mol (blue), respectively.

Although both the two pathways could explain the data well, the actual model is more likely to be pathway I in which three intermediates directly interconverts with the major state. As discussed in the method validation in section 4.1.3, the combined relaxation data analyses did not reject model II, nonetheless, the minimization favored the actual input model.

To summary, the apo-form hiFABP exists in at least a four-state equilibrium in which three intermediates directly convert with the major form. The resultant kinetic parameters for the three exchange processes are $p_{\text{N}}=90.73\pm 0.16 \%$, $p_{\text{I1}}=3.20\pm 0.26 \%$, $p_{\text{I2}}=3.33\pm 0.14 \%$, $k_{\text{ex1}}=3405\pm 116 \text{ s}^{-1}$, $k_{\text{ex2}}=750\pm 22 \text{ s}^{-1}$, $k_{\text{ex3}}=69\pm 4 \text{ s}^{-1}$, respectively. Moreover, apo-hiFABP experiences additional fast exchange processes with k_{ex} in the range of $10^5\sim 10^6 \text{ s}^{-1}$. The energy differences of the excited states with the major form range from 6 kJ/Mol to 8 kJ/Mol. According to the extracted chemical shifts of the minor states, major structural variations of the sparsely populated conformers with respect to the majority of the population occur mainly in the gap region. The structural fluctuations should be relatively small in amplitude and do not involve obvious openings of the putative portal for ligands to enter or exit. Since binding of ligands to FABPs almost certainly require an opening of the structure, the conformation selection mechanism is thus not likely for FABPs. More importantly, the reaction rate

constant for the major state to convert into the intermediates ($<120 \text{ s}^{-1}$) are much smaller than that of the FABP-ligand binding process ($\sim 23 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$) (G V Richieri, Ogata, and Kleinfeld 1996), making it impossible for ligands to select or interact with an existing intermediate in order to bind to the protein. Based on our observations, we propose that the excited states in equilibrium are irrelevant to the ligand binding process. The FABP-ligand binding should be via a ligand-induced binding mechanism in which the ligand interacts with the native state directly. In the absence of any unfolded states, the three intermediates should represent conformations in the last step of the folding pathway. In addition, since they all interconvert with the native state, they are of course not product of any misfolding process. This is thus an experimental evidence, and to the best of our knowledge, the first one, for the multiple independent protein folding pathways. The distinguishing dynamic properties displayed in the gap region shall draw new attention to this structural segment of this protein family. Last but not the least, for the first time, the combined analysis of CEST & RD data used in this work could extend the applicability of these two experiments and extract the kinetic parameters with better accuracy.

BIBLIOGRAPHY

- Abraham, A. 1961. *The Principles of Nuclear Magnetism*. Oxford University Press, London.
- Abumrad, N A, J H Park, and C R Park. 1984. "Permeation of Long-Chain Fatty Acid into Adipocytes. Kinetics, Specificity, and Evidence for Involvement of a Membrane Protein." *Journal of Biological Chemistry* 259 (14): 8945–53.
- Alpers, D H, N M Bass, M J Engle, and K DeSchryver-Kecskemeti. 2000. "Intestinal Fatty Acid Binding Protein May Favor Differential Apical Fatty Acid Binding in the Intestine." *Biochimica et Biophysica Acta* 1483 (3): 352–62.
- Bai, Yawen, John S Milne, Leland Mayne, and S Walter Englander. 1993. "Primary Structure Effects on Peptide Group Hydrogen Exchange." *Proteins* 17 (1): 75–86. doi:10.1002/prot.340170110.
- Baier, L J, J C Sacchettini, W C Knowler, J Eads, G Paolisso, P a Tataranni, H Mochizuki, P H Bennett, C Bogardus, and M Prochazka. 1995. "An Amino Acid Substitution in the Human Intestinal Fatty Acid Binding Protein Is Associated with Increased Fatty Acid Binding, Increased Fat Oxidation, and Insulin Resistance." *The Journal of Clinical Investigation* 95 (3): 1281–87. doi:10.1172/JCI117778.
- Besnard, Philippe, Isabelle Niot, H é è è Poirier, Lionel Cl é ment, and Andr é Bernard. 2002. "New Insights into the Fatty Acid-Binding Protein (FABP) Family in the Small Intestine." *Molecular and Cellular Biochemistry* 239 (1-2): 139–47.
- Bloembergen, N., E. M. Purcell, and R. V. Pound. 1948. "Relaxation Effects in Nuclear Magnetic Resonance Absorption." *Phys. Rev.* 73 (7): 679–712. doi:10.1103/PhysRev.73.679.
- Bouvignies, Guillaume, and Lewis E Kay. 2012. "A 2D ¹³C-CEST Experiment for Studying Slowly Exchanging Protein Systems Using Methyl Probes: An Application to Protein Folding." *Journal of Biomolecular NMR* 53 (4): 303–10. doi:10.1007/s10858-012-9640-7.
- Burnett, D A, N Lysenko, J A Manning, and R K Ockner. 1979. "Utilization of Long Chain Fatty Acids by Rat Liver: Studies of the Role of Fatty Acid Binding Protein." *Gastroenterology* 77 (2): 241–49.
- Carver, J P R. E. Richards. 1972. "A General Two-Site Solution for the Chemical Exchange Produced Dependence of T₂ Upon the Carr-Purcell Pulse Separation." *Journal of Magnetic Resonance* 105: 89–105.

- Cavanagh, John, Wayne J. Fairbrother, Arthur G. Palmer III, Mark Rance, and Nicholas J. Skelton. 2007. *Protein NMR Spectroscopy*.
- Cham, B E, and B R Knowles. 1976. "A Solvent System for Delipidation of Plasma or Serum without Protein Precipitation." *Journal of Lipid Research* 17 (2): 176–81.
- Chen, Jianhan, Charles L Brooks, and Peter E Wright. 2004. "Model-Free Analysis of Protein Dynamics: Assessment of Accuracy and Model Selection Protocols Based on Molecular Dynamics Simulation." *Journal of Biomolecular NMR* 29 (3): 243–57.
doi:10.1023/B:JNMR.0000032504.70912.58.
- Chen, Raymond F. 1967. "MACROMOLECULES : Removal of Fatty Acids from Serum Albumin by Charcoal Treatment Removal of Fatty Acids by Charcoal Treatment from Serum Albumin." *Journal of Biological Chemistry* 242 (25): 173–81.
- Clore, G. Marius, Attila Szabo, Ad Bax, Lewis E. Kay, Paul C. Driscoll, and Angela M. Gronenborn. 1990. "Deviations from the Simple Two-Parameter Model-Free Approach to the Interpretation of Nitrogen-15 Nuclear Magnetic Relaxation of Proteins." *Journal of the American Chemical Society* 112 (12): 4989–91. doi:10.1021/ja00168a070.
- Cohn, Steven M, Theodore C Simon, Kevin A Roth, Edward H Birkenmeier, and Jeffrey I Gordon. 1992. "Use of Transgenic Mice to Map Cis-Acting Elements in the Intestinal Fatty Acid Binding Protein Gene (Fabpi) That Control Its Cell Lineage-Specific and Regional Patterns of Expression along the Duodenal-Colonic and Crypt-Villus Axes of the Gut Epithelium." *The Journal of Cell Biology* 119 (1): 27–44.
- Corsico, B, D P Cistola, C Frieden, and J Storch. 1998. "The Helical Domain of Intestinal Fatty Acid Binding Protein Is Critical for Collisional Transfer of Fatty Acids to Phospholipid Membranes." *Proceedings of the National Academy of Sciences of the United States of America* 95 (21): 12174–78.
- Córsico, Betina, Heng Ling Liou, and Judith Storch. 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 (12): 3600–3607. doi:10.1021/bi0357356.
- Davis, D. G., M. E. Perlman, and R.E. London. 1994. "Direct Measurements of the Dissociation-Rate Constant for Inhibitor-Enzyme Complexes via the T1 Rho and T2 (CPMG) Methods."
- Delaglio, F, S Grzesiek, G W Vuister, G Zhu, J Pfeifer, and a Bax. 1995. "NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes." *Journal of Biomolecular NMR* 6 (3): 277–93.

- Der Vusse, Ger J Van, and Robert S Reneman. 1995. "Chapter 3 Substrate Utilization in Mammalian Cells." In *Cell Chemistry and Physiology: Part 1*, edited by E Edward Bittar and Neville Bittar, 4, Part 1:45–75. Principles of Medical Biology. Elsevier.
doi:http://dx.doi.org/10.1016/S1569-2582(06)80005-X.
- Distel, R J, G S Robinson, and B M Spiegelman. 1992. "Fatty Acid Regulation of Gene Expression. Transcriptional and Post-Transcriptional Mechanisms." *The Journal of Biological Chemistry* 267 (9): 5937–41.
- Ernst, Richard R., Geoffrey Bodenhausen, and Alexander Wokanu. 1987. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. New York: Oxford University Press.
- Falomir-Lockhart, Lisandro J, Lisandro Laborde, Peter C Kahn, Judith Storch, and Betina Córscico. 2006. "Protein-Membrane Interaction and Fatty Acid Transfer from Intestinal Fatty Acid-Binding Protein to Membranes. Support for a Multistep Process." *The Journal of Biological Chemistry* 281 (20): 13979–89. doi:10.1074/jbc.M511943200.
- Gary V. Richieri, Alan M. Kleinfeld. 1994. "Equilibrium Constants for the Binding of Fatty Acids with Fatty Acid-Binding Proteins from Adipocyte, Intestine, Heart, and Liver Measured with the Fluorescent Probe ADIFAB." *The Journal of Biological Chemistry* 269 (39): 23918–30.
- Georgopoulos, Angeliki, Hanna Bloomfield, Dorothea Collins, Margaret E Brousseau, Jose M Ordovas, John J O'Connor, Sander J Robins, and Ernest J Schaefer. 2007. "Codon 54 Polymorphism of the Fatty Acid Binding Protein (FABP) 2 Gene Is Associated with Increased Cardiovascular Risk in the Dyslipidemic Diabetic Participants of the Veterans Affairs HDL Intervention Trial (VA-HIT)." *Atherosclerosis* 194 (1): 169–74. doi:10.1016/j.atherosclerosis.2006.07.022.
- Glatz, J F, T Borchers, F Spener, and G J van der Vusse. 1995. "Fatty Acids in Cell Signalling: Modulation by Lipid Binding Proteins." *Prostaglandins, Leukotrienes, and Essential Fatty Acids* 52 (2-3): 121–27.
- Glatz, JFC, and GJ van der Vusse. 1996. "Cellular Fatty Acid-Binding Proteins: Their Function and Physiological Significance." *Progress in Lipid Research* 35: 243–82.
- Hallden, G. 1997. "Evidence for a Role of the Gut Hormone PYY in the Regulation of Intestinal Fatty Acid-Binding Protein Transcripts in Differentiated Subpopulations of Intestinal Epithelial Cell Hybrids." *Journal of Biological Chemistry* 272 (19): 12591–600.
doi:10.1074/jbc.272.19.12591.
- Hauerland, N H, X Chen, P Andolfatto, J M Chisholm, and Z Wang. 2013. "Developmental Changes of FABP Concentration, Expression, and

- Intracellular Distribution in Locust Flight Muscle.” *Molecular and Cellular Biochemistry* 123 (1-2): 153–58. Accessed December 18.
- Hertzel, A V, and D A Bernlohr. 1998. “Regulation of Adipocyte Gene Expression by Polyunsaturated Fatty Acids.” *Molecular and Cellular Biochemistry* 188 (1-2): 33–39.
- Hodsdon, M E, and D P Cistola. 1997. “Ligand Binding Alters the Backbone Mobility of Intestinal Fatty Acid-Binding Protein as Monitored by ¹⁵N NMR Relaxation and ¹H Exchange.” *Biochemistry* 36 (8). American Chemical Society: 2278–90. doi:10.1021/bi962018l.
- Hodsdon, Michael E, and David P Cistola. 1997. “Discrete Backbone Disorder in the Nuclear Magnetic Resonance Structure of Apo Intestinal Fatty Acid-Binding Protein: Implications for the Mechanism of Ligand Entry.” *Biochemistry* 2960 (6). American Chemical Society: 1450–60. doi:10.1021/bi961890r.
- Arthur G Palmer III, John Cavanagh, Peter E Wright, and Mark Rance. 1991. “Sensitivity Improvement in Proton-Detected Two-Dimensional Heteronuclear Correlation {NMR} Spectroscopy.” *Journal of Magnetic Resonance (1969)* 93 (1): 151–70. doi:http://dx.doi.org/10.1016/0022-2364(91)90036-S.
- Kaikaus, R M, W K Chan, P R Ortiz de Montellano, and N M Bass. 2013. “Mechanisms of Regulation of Liver Fatty Acid-Binding Protein.” *Molecular and Cellular Biochemistry* 123 (1-2): 93–100.
- Kay, Lewis, Paul Keifer, and Tim Saarinen. 1992. “Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity.” *Journal of the American Chemical Society* 114 (26): 10663–65. doi:10.1021/ja00052a088.
- Kneller, T. D. Goddard and D. G. n.d. “Sparky.” San Francisco: University of California.
- Kurian, E, W R Kirk, and F G Prendergast. 1996. “Affinity of Fatty Acid for (r)rat Intestinal Fatty Acid Binding Protein:further Examination.” *Biochemistry* 35 (12): 3865–74. doi:10.1021/bi952589y.
- LiCata, V J, and D a Bernlohr. 1998. “Surface Properties of Adipocyte Lipid-Binding Protein: Response to Lipid Binding, and Comparison with Homologous Proteins.” *Proteins* 33 (4): 577–89.
- Lin, Zhi. 2006. “Structural Studies on DDCAD-1: A Ca²⁺ Dependent Cell-Cell Adhesion Protein.” National University of Singapore.
- Lipari, Giovanni, and Attila Szabo. 1982a. “Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity.” *Journal of the*

- American Chemical Society* 104 (17): 4546–59.
doi:10.1021/ja00381a009.
- . 1982b. “Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 2. Analysis of Experimental Results.” *Journal of the American Chemical Society* 104 (17): 4559–70. doi:10.1021/ja00381a010.
- Long, Dong, Maili Liu, and Daiwen Yang. 2008. “Accurately Probing Slow Motions on Millisecond Timescales with a Robust NMR Relaxation Experiment.” *Journal of the American Chemical Society* 130 (8): 2432–33. doi:10.1021/ja710477h.
- “MATLAB 8.1 and Statistics Toolbox 8.1.” 2013. Natick, Massachusetts: The MathWorks, Inc., Natick, Massachusetts, United States.
- McConnell, Harden M. 1958. “Reaction Rates by Nuclear Magnetic Resonance.” *The Journal of Chemical Physics* 28 (3): 430.
doi:10.1063/1.1744152.
- McConnell, J. 1987. *The Theory of Nuclear Magnetic Relaxation in Liquids*. Cambridge University Press.
- Ockner, R K, and J A Manning. 1976. “Fatty Acid Binding Protein. Role in Esterification of Absorbed Long Chain Fatty Acid in Rat Intestine.” *The Journal of Clinical Investigation* 58 (3): 632–41. doi:10.1172/JCI108510.
- Ogata, R. T. 1996. “Thermodynamic and Kinetic Properties of Fatty Acid Interactions with Rat Liver Fatty Acid-Binding Protein.” *Journal of Biological Chemistry* 271 (49): 31068–74. doi:10.1074/jbc.271.49.31068.
- Osborne, M J, J Schnell, S J Benkovic, H J Dyson, and P E Wright. 2001. “Backbone Dynamics in Dihydrofolate Reductase Complexes: Role of Loop Flexibility in the Catalytic Mechanism.” *Biochemistry* 40 (33): 9846–59. <http://www.ncbi.nlm.nih.gov/pubmed/11502178>.
- Palmer, a G, Christopher D Kroenke, J Patrick Loria, and By Arthur G Palmer. 2001. “Nuclear Magnetic Resonance Methods for Quantifying Microsecond-to-Millisecond Motions in Biological Macromolecules.” *Methods in Enzymology* 339 (1997): 204–38.
- Peeters, R A, J H Veerkamp, and R A Demel. 1989. “Are Fatty Acid-Binding Proteins Involved in Fatty Acid Transfer?” *Biochimica et Biophysica Acta* 1002 (1): 8–13.
- Pérez-Bravo, F, M Fuentes, B Angel, H Sanchez, E Carrasco, J L Santos, L Lera, and C Albala. 2006. “Lack of Association between the Fatty Acid Binding Protein 2 (FABP2) Polymorphism with Obesity and Insulin Resistance in Two Aboriginal Populations from Chile.” *Acta Diabetologica* 43 (4): 93–98. doi:10.1007/s00592-006-0221-7.

- Piserchio, Andrea, Maria Pellegrini, Sunil Mehta, Scott M Blackman, Elizabeth P Garcia, John Marshall, and Dale F Mierke. 2002. "The PDZ1 Domain of SAP90: CHARACTERIZATION OF STRUCTURE AND BINDING." *Journal of Biological Chemistry* 277 (9): 6967–73. doi:10.1074/jbc.M109453200.
- Pohl, Jürgen, Axel Ring, Thomas Herrmann, Wolfgang Stremmel, and Robert Eehalt. 2004. "New Concepts of Cellular Fatty Acid Uptake: Role of Fatty Acid Transport Proteins and of Caveolae." *The Proceedings of the Nutrition Society* 63 (2): 259–62. doi:10.1079/PNS2004341.
- Poirier, H, I Niot, P Degrace, M C Monnot, A Bernard, and P Besnard. 1997. "Fatty Acid Regulation of Fatty Acid-Binding Protein Expression in the Small Intestine." *The American Journal of Physiology* 273 (2 Pt 1): G289–95. <http://www.ncbi.nlm.nih.gov/pubmed/9277406>.
- Pratley, R E, L Baier, D a Pan, a D Salbe, L Storlien, E Ravussin, and C Bogardus. 2000. "Effects of an Ala54Thr Polymorphism in the Intestinal Fatty Acid-Binding Protein on Responses to Dietary Fat in Humans." *Journal of Lipid Research* 41 (12): 2002–8. <http://www.ncbi.nlm.nih.gov/pubmed/11108733>.
- Ran, Xiaoyuan, Hong-hua Miao, Fwu-Shan Sheu, and Daiwen Yang. 2003. "Structural and Dynamic Characterization of a Neuron-Specific Protein Kinase C Substrate, Neurogranin." *Biochemistry* 42 (17): 5143–50. doi:10.1021/bi0271751.
- Redfield, A.G. 1965. "The Theory of Relaxation Processes." In *Advances in Magnetic and Optical Resonance*, edited by John S. Waugh, Volume 1:1–32. Academic Press.
- Richieri, GV, and RT Ogata. 2000. "Fatty Acid Binding Proteins from Different Tissues Show Distinct Patterns of Fatty Interactions." *Biochemistry*, 7197–7204. <http://pubs.acs.org/doi/abs/10.1021/bi000314z>.
- Richieri, G V, R T Ogata, and a M Kleinfeld. 1996. "Kinetics of Fatty Acid Interactions with Fatty Acid Binding Proteins from Adipocyte, Heart, and Intestine." *The Journal of Biological Chemistry* 271 (19): 11291–300. <http://www.ncbi.nlm.nih.gov/pubmed/8626681>.
- Richieri, G V, R. T. Ogata, and A M Kleinfeld. 1995. "Thermodynamics of Fatty Acid Binding to Fatty Acid-Biding Proteins and Fatty Acid Partition between Water and Membranes Measured Using the Fluorescent Probe ADIFABP." *Journal of Biological Chemistry* 270 (25): 15076–84. doi:10.1074/jbc.270.25.15076.
- Sacchettini, JC, and SM Hautf. 1990. "Developmental and Structural Studies of an Intracellular Lipid Binding Protein Expressed in the Ileal

- Epithelium.” *Journal of Biological ...*
<http://www.jbc.org/content/265/31/19199.short>.
- Scapin, G, J I Gordon, and J C Sacchettini. 1992. “Refinement of the Structure of Recombinant Rat Intestinal Fatty Acid-Binding Apoprotein at 1.2-Å Resolution.” *The Journal of Biological Chemistry* 267 (6): 4253–69.
<http://www.ncbi.nlm.nih.gov/pubmed/1740465>.
- Schleucher, J, M Schwendinger, M Sattler, P Schmidt, O Schedletzky, S J Glaser, O W Sørensen, and C Griesinger. 1994. “A General Enhancement Scheme in Heteronuclear Multidimensional NMR Employing Pulsed Field Gradients.” *Journal of Biomolecular NMR* 4 (2). ESCOM Science Publishers B.V., Leiden/Kluwer Academic Publishers: 301–6.
 doi:10.1007/BF00175254.
- Schmider, W, a Fahr, R Voges, W Gerok, and G Kurz. 1996. “Irreversible Inhibition of Hepatic Fatty Acid Salt Uptake by Photoaffinity Labeling with 11, 11-Azistearate.” *Journal of Lipid Research* 37 (4): 739–53.
<http://www.ncbi.nlm.nih.gov/pubmed/8732774>.
- Schurr, J M, H P Babcock, and B S Fujimoto. 1994. “A Test of the Model-Free Formulas. Effects of Anisotropic Rotational Diffusion and Dimerization.” *Journal of Magnetic Resonance, Series B* 105 (3): 211–24. doi:<http://dx.doi.org/10.1006/jmrb.1994.1127>.
- Spector, a a, K John, and J E Fletcher. 1969. “Binding of Long-Chain Fatty Acids to Bovine Serum Albumin.” *Journal of Lipid Research* 10 (1): 56–67. <http://www.ncbi.nlm.nih.gov/pubmed/5773785>.
- Stone, M J, W J Fairbrother, a G Palmer, J Reizer, M H Saier, and P E Wright. 1992. “Backbone Dynamics of the Bacillus Subtilis Glucose Permease IIA Domain Determined from 15N NMR Relaxation Measurements.” *Biochemistry* 31 (18): 4394–4406.
- 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 (23): 13317–23.
 doi:10.1074/jbc.271.23.13317.
- Thumser, a E, and J Storch. 2000. “Liver and Intestinal Fatty Acid-Binding Proteins Obtain Fatty Acids from Phospholipid Membranes by Different Mechanisms.” *Journal of Lipid Research* 41 (4): 647–56.
- Tochio, Hidehito, Franki Hung, Ming Li, David S Bredt, and Mingjie Zhang. 2000. “Solution Structure and Backbone Dynamics of the Second {PDZ} Domain of Postsynaptic Density-95.” *Journal of Molecular Biology* 295 (2): 225–37. doi:<http://dx.doi.org/10.1006/jmbi.1999.3350>.

- Trigatti, B L, D Mangroo, and G E Gerber. 1991. "Photoaffinity Labeling and Fatty Acid Permeation in 3T3-L1 Adipocytes." *The Journal of Biological Chemistry* 266 (33): 22621–25.
- Vassileva, G, L Huwyler, K Poirier, L B Agellon, and M J Toth. 2000. "The Intestinal Fatty Acid Binding Protein Is Not Essential for Dietary Fat Absorption in Mice." *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology* 14 (13): 2040–46. doi:10.1096/fj.99-0959com.
- Veerkamp, J H, and H T van Moerkerk. 2013. "Fatty Acid-Binding Protein and Its Relation to Fatty Acid Oxidation." *Molecular and Cellular Biochemistry* 123 (1-2): 101–6. Accessed December 18.
- Waggoner, David W. 1990. "In Situ Labeling of the Adipocyte Lipid Binding Protein with 3-[125] Iodo-4-Azido-N-Hexadecylsalicylamide. Evidence for a Role of Fatty Acid Binding Proteins in Lipid Uptake." *The Journal of Biological Chemistry* 265: 11417–20.
- Wangsness, RK, and F. Bloch. 1953. "The Dynamical Theory of Nuclear Induction." *Physical Review* 89 (4): 728–39. doi:10.1103/PhysRev.89.728.
- Wilkinson, T C, and D C Wilton. 1987. "Studies on Fatty Acid-Binding Proteins. The Binding Properties of Rat Liver Fatty Acid-Binding Protein." *The Biochemical Journal* 247 (2): 485–88.
- Wu, F, B Corsico, C R Flach, D P Cistola, J Storch, and R Mendelsohn. 2001. "Deletion of the Helical Motif in the Intestinal Fatty Acid-Binding Protein Reduces Its Interactions with Membrane Monolayers: Brewster Angle Microscopy, IR Reflection-Absorption Spectroscopy, and Surface Pressure Studies." *Biochemistry* 40 (7): 1976–83.
- Xia, Junchao, Nan-jie Deng, and Ronald M Levy. 2013. "NMR Relaxation in Proteins with Fast Internal Motions and Slow Conformational Exchange: Model-Free Framework and Markov State Simulations." *The Journal of Physical Chemistry. B* 117 (22): 6625–34. doi:10.1021/jp400797y.
- Yang, D, Y K Mok, J D Forman-Kay, N a Farrow, and L E Kay. 1997. "Contributions to Protein Entropy and Heat Capacity from Bond Vector Motions Measured by NMR Spin Relaxation." *Journal of Molecular Biology* 272 (5): 790–804. doi:10.1006/jmbi.1997.1285.